Regional Selective Assistance in Scotland: does it make a difference to plant performance?

John Duncan Moffat

Submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy

College of Social Sciences
Business School
Economics
University of Glasgow

November 2010

Abstract

Regional Selective Assistance (RSA) is the largest and oldest business support scheme currently operating in Scotland. It provides grants to firms undertaking capital investment projects in economically deprived EU designated 'Assisted Areas'. As a component of regional policy, the scheme is principally designed to safeguard and generate employment in the Assisted Areas. Many of the grants are given to help foreign firms to set up in Scotland. The aim of this thesis is to estimate the impact of receipt of these grants on plant performance as measured by productivity and survival.

The main econometric problem to be confronted when estimating the impact of grants is self-selection bias. Because plants self-select into the treated group, the treated group will have different characteristics from the untreated group which would lead to differences in performance had neither group received treatment. This creates difficulties in estimating the impact of treatment as a simple comparison of a variable across treated and untreated groups will not measure the causal impact of treatment. This problem was dealt with using propensity score matching and instrumental variables.

The dataset was created by linking a register of plants that received an RSA grant into the longitudinal ARD which contains the necessary range of financial variables for empirical analysis. This part of the thesis was crucial as failure to identify a high percentage of plants that received a grant in the ARD would seriously undermine the empirical analyses. In the end, a higher proportion of plants that received a grant were linked with the ARD than has been previously achieved using these databases.

In the first empirical chapter, the growth of labour productivity and TFP between 1994 and 2004 in Scottish manufacturing plants was decomposed to reveal the contribution of plants that receive an RSA grant. This showed that RSA-assisted plants made a small but positive contribution to both measures of productivity growth.

The latter two empirical chapters showed that receipt of an RSA grant had no statistically significant impact on either the TFP or the survival probability of Scottish manufacturing plants between 1984 and 2004 in any of the industries considered. This is a major concern as it casts doubt on whether the jobs created and safeguarded by an RSA grant will endure.

Table of Contents

List of Appendices

List of Tables

List of Figures

Acknowledgement

There are many people to whom I owe thanks for assisting me with my thesis. However, the person to whom I am most indebted is my supervisor, Professor Richard Harris, who has always been willing to take time from his schedule to answer my questions and offer comments on my work. From the start, he has been very kind and friendly to me and this has contributed in no small part to my enjoyment of the PhD.

I am also grateful to the Department of Economics at Glasgow University for allowing me to present at departmental seminars, providing the opportunity to teach and generally making me feel very welcome. In particular, I am grateful to my second supervisor, Dr Francesca Flamini, for proof reading my thesis and Nicola Birkin for her speed in dealing with any administrative issues I may have had.

I am also very grateful to the Economic and Social Research Council and the Scottish Government for generously funding this work. In particular, thanks are due to Dr Gill Clark and Dr Aileen McIntosh whose efficient administration of the scheme made the PhD much easier. I am also grateful to my Scottish Government supervisors, Lynn Graham and Marina Curran, for organising my annual placement at the Scottish Government.

I would also like to thank the staff at the Office for National Statistics Virtual Micro-Data Laboratory in London, Newport and Glasgow for allowing access to the data and handling my clearance requests promptly.

Also deserving of thanks are my office mates who have helped to keep me fit with frequent and furiously competitive games of table football. I would also like to thank my parents for giving me the opportunity to study for a PhD. Finally, I am exceptionally grateful to my girlfriend for looking after and putting up with me.

Declaration

I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

This work contains statistical data from ONS which is Crown copyright and reproduced with the permission of the controller of HMSO and Queen's Printer for Scotland. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

Abbreviations

ABI	Annual Business Inquiry
ARD	Annual Respondents Database
ATE	Average Treatment Effect
ATNT	Average Effect of Treatment on the Non-Treated
ATT	Average Effect of Treatment on the Treated
BERR	Business, Enterprise and Regulatory Reform
CSO	Central Statistical Office
EU	European Union
FDI	Foreign Direct Investment
GB	Great Britain
GBI	Grant for Business Investment
GDP	Gross Domestic Product
GVA	Gross Value Added
GMM	Generalised Method of Moments
IDBR	Inter-Departmental Business Register
IFG	Invest for Growth
LATE	Local Average Treatment Effect
MTE	Marginal Treatment Effect
NVQ	National Vocational Qualification
OLS	Ordinary Least Squares
R&D	Research and Development
RSA	Regional Selective Assistance
SAMIS	Selective Assistance Management Information System
SEK	Swedish Krona
SIC	Standard Industrial Classification
SFA	Selective Financial Assistance
SMART	Small Firm Merit Awards for Research and Technology
SME	Small and Medium Enterprise
SPUR	Support for Products under Research
TFP	Total Factor Productivity
UK	United Kingdom

1. Introduction

1.1. Introduction

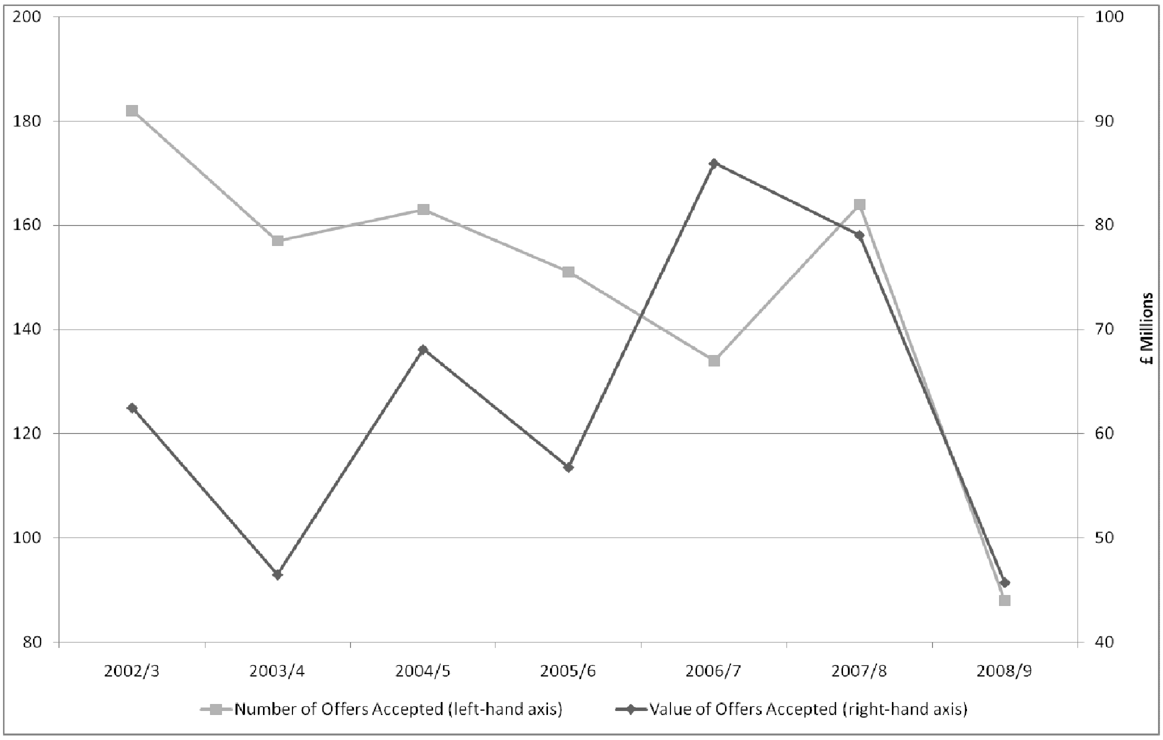
Regional Selective Assistance (RSA) is the largest and oldest business support scheme currently operating in Scotland. It provides grants to firms undertaking capital investment projects in economically deprived European Union (EU) designated ‘Assisted Areas’. As a component of regional policy, the scheme is principally designed to safeguard and generate employment in the Assisted Areas. As such, many of the grants are given to help foreign firms to set up in Scotland. The amount that can be offered is determined by a number of factors including the location and size of the project and the number of jobs it will create or safeguard. In order to receive an RSA grant, an additionality criterion must be satisfied which requires that awards will only be made if the project could not have proceeded in the same form without the grant. A displacement criterion must also be met which demands that the jobs created by the project must not be offset by job losses in other parts of the Assisted Areas. The aim of this thesis is to estimate the impact of receipt of these grants on plant performance.

This chapter is an introduction to the thesis. The next section provides a motivation for the thesis. The third section justifies the choice of variables upon which the impact of RSA will be analyzed in the chapters 7 and 8. The fourth section will describe the contents of each chapter. The final section concludes.

1.2. Motivation

Figure 1.1 shows the number and value of grants offered and accepted by plants in Scotland over the last six years. These figures are taken from the Scottish Government’s annual reports on RSA.

Figure 1.1: Number and Value of Grants Accepted in Scotland (2003 prices), 2002/3-2008/9

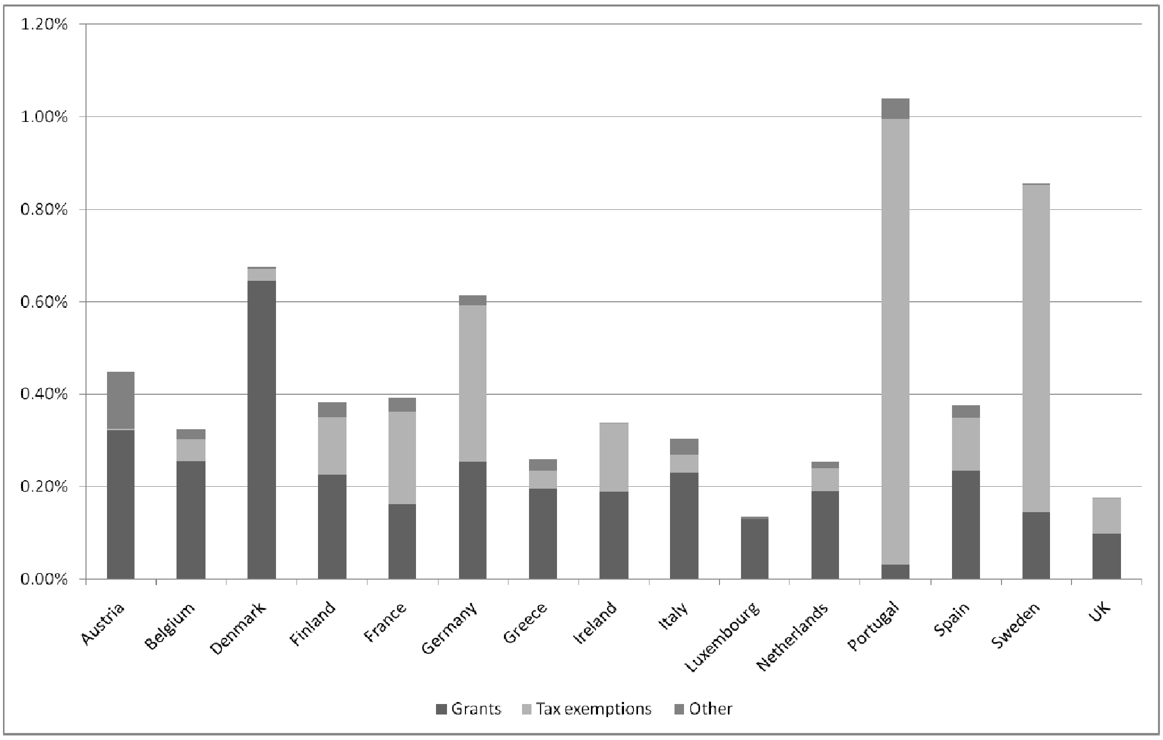


Source: Scottish Government (2003-2007a, 2008-2009a)

The number of grants accepted fell between 2002/3 and 2006/7 from over 180 to less than 140 before rising to over 170 in 2007/8. However, in 2008/9, only 90 grants were accepted. The value of grants accepted does not generally mirror the number of grants accepted as the value of grants accepted rose between 2002/3 and 2006/7 from slightly over £60 million to around £85 million. This implies that the average value of grant rose significantly during this period. However, in 2007/8, the value of grants fell to under £80 million before falling precipitously to just over £45 million in 2008/9. It remains to be seen whether these recent falls in both the number and value of grants are the start of serious cut-backs in funding for the RSA scheme or merely a short-term fluctuation caused by difficult budgetary circumstances. Regardless of which is the case, it is clear that large amounts of public money are currently spent on the RSA scheme. The size of the amount spent is emphasised by comparison with the other business support schemes that currently operate in Scotland. In 2008/9, less than £5 million was awarded under the Small Firm Merit Awards for Research and Technology (SMART): Scotland scheme (Scottish Government, 2009b) while less than £13.5 million was awarded under the R&D grants scheme (Scottish Enterprise, 2009). Data from the Selective Assistance Managements Information System (SAMIS) showing the extent to which expenditure on RSA dwarfed that of other schemes between 1972 and 2003 is presented in chapter 5.1. Research into the impact of receipt of an RSA grant on plant performance is therefore essential to allow policy-makers to understand whether this money is being well-spent.

However, the results obtained in this thesis concerning the effectiveness of the RSA scheme will be of interest throughout Europe. Figure 1.2 provides a breakdown of the various types of state aid that are provided in each country of the EU-15.[1]

Figure 1.2: State Aid by Type as a Percentage of Gross Domestic Product (GDP) in the EU-15, Annual Average 2006-2008



Source: EU (2010)

Crisis measures and coal sector are excluded.

'Other' consists of equity participations, soft loans, tax deferrals and guarantees.

As is shown in figure 1.2, in most countries of the EU-15, the largest component of state aid takes the form of grants. Unfortunately, disaggregated data is not available which shows how much is spent on different types of grants. It is therefore not possible to state how much of the money that is spent on grants are spent on investment grant schemes of the type embodied by the RSA scheme. Nevertheless, the website of the European

[1] The EU defined state aid as 'a form of state intervention used to promote a certain economic activity' (EU, 2009).

Commission shows that most European countries operate some form of investment grant scheme, some of which are discussed due to their being the subject of empirical analysis in chapter 4, so the analysis contained in this thesis may be of interest not only to Scottish policy-makers but to policy-makers throughout Europe.[2]

1.3. Choice of Dependent Variables for Econometric Chapters

The dataset that will be used in the empirical analysis allows the impact of receipt of an RSA grant on numerous measures of plant performance to be analysed. Perhaps the most obvious measure to choose is employment or labour demand given that the main objective of the RSA scheme is to promote and safeguard employment and the important role of employment in determining living standards. This was not done here for two reasons. Firstly, the impact of receipt of an RSA grant on employment has been examined extensively in the past and a statistically significant and positive impact has invariably been found (see, for example, Criscuolo, Martin, Overman and Van Reenen, 2007; Hart, Driffield, Roper and Mole, 2008a). It can therefore be regarded as an established fact that RSA has a positive impact on employment. Secondly, as a 'clawback' clause requires that RSA grants are repaid if a specific job target is not achieved, a positive impact on employment is hardly surprising as firms are unlikely to apply for RSA grants unless they fully intend to meet the jobs target (see chapter 2.6 for a description of the eligibility criteria for the RSA scheme). For these reasons, the impact of receipt of an RSA grant on employment was not analysed in this thesis.

Another obvious measure of plant performance that could have used for analysis is investment. This was recently done by Criscuolo, Martin, Overman and Van Reenen (2007) who found that receipt of an RSA grant had a positive and statistically significant impact on investment. However, given the existence of the additionality criterion mentioned above and the fact that payment of the grant is only made after the plant has actually purchased the capital, a positive impact of receipt of an RSA grant on investment ought to be ensured. Analysis of the impact of receipt of an RSA grant on investment or the capital stock is therefore not particularly interesting.

---

[2] The European Commission website provides a description of the various types of financial support that are available in each of the member states (European Commission, 2009).

The impact of receipt of an RSA grant on output could also be analysed. However, given what has been said concerning the likely impact of receipt of a grant on employment and the capital stock, it is very unlikely that receipt of a grant will not also have a positive impact on output. For no impact to be found would require either that the larger stocks of employment and capital that result from receipt of a grant do not increase the productive capacity of the plant or, if there is an increase in productive capacity, there to be no demand for the greater output that can be produced by the plant. The former is unlikely because more labour and capital will generally increase the productive capacity of the plant. The latter will only occur if the demand curve is sufficiently inelastic that the firm does not profit maximise by lowering its price and producing more output. However, this is also unlikely as a firm would not choose to undertake a project which enhanced its productive capacity if it did not intend to increase its output. Output is therefore not a particularly interesting aspect of firm performance upon which to estimate the impact of receipt of an RSA grant.

Instead, the impact of receipt of an RSA grant on productivity will be analysed. The importance of productivity in determining living standards is well documented. According to Krugman (1997), in the determination of living standards, ‘productivity isn’t everything but in the long run, it is almost everything’. Similarly, Baumol (1984) states that ‘it can be said without exaggeration that in the long run probably nothing is as important for economic welfare as the rate of productivity growth’. Empirical evidence showing the truth of these statements is provided by the OECD (2003).

The size of the contribution to the growth of GDP per capita made by a firm which experiences an increase in productivity as a result of receiving an RSA grant will depend upon the shape of the demand curve it faces. With a sufficiently elastic demand curve, a grant-induced increase in productivity will allow the firm to increase revenues and therefore make greater profits or pay higher wages to its employees. An increase in productivity may also boost GDP per capita by inducing the firm to employ additional workers over and above the levels that the firm is obliged to employ under the terms of the grant. On the other hand, a sufficiently inelastic demand curve raises the possibility that the firm responds to the higher productivity by increasing profits and reducing employment. However, this should not happen because the rules of the RSA scheme stipulate that the grant must be used to safeguard or promote employment. Regardless of the shape of the demand curve, an increase in productivity will help firms to survive that may otherwise close. This is of particular relevance here because the rules of the RSA scheme stipulate that grants cannot be provided in cases where the necessary finance for the project could have been obtained from the private sector and that, for those projects which seek to safeguard rather than create employment, the applicant firm would have had to make redundancies without a grant. It is therefore reasonable to suggest that RSA grant recipients are a poor subset of firms which may be at higher risk of closure than the average.

In light of the potential impact on GDP per capita of productivity improvements caused by receipt of an RSA grant and also the inclusion of a target to improve productivity performance in the Government Economic Strategy (Scottish Government, 2007b), productivity is an obvious measure that can be used to analyse the impact of receipt of an RSA grant. A positive impact would arise if grants are used to acquire capital which allows the plant to produce existing products more efficiently or to produce new products that can be produced more efficiently than older products. As will be discussed in the literature review in chapter 4.4, there is no clear consensus as to what impact receipt of an RSA grant has on productivity. Furthermore, those studies that have analysed the existence of a causal relationship between receipt of an RSA grant and productivity have used inferior datasets to that which will be employed here and have employed different methodologies. For these reasons, further research into this relationship is required.

The impact of receipt of an RSA grant on survival probability will also be analysed. Only Harris and Robinson (2005) have performed such an analysis. They found that receipt of an RSA grant increased the probability of survival. The lack of studies into this subject is surprising given the close relationship between plant survival and the security of the jobs created and safeguarded by RSA. In addition to using an inferior dataset to that employed in chapter 8, their study made no allowance for the consequences of self-selection into the group of plants that receive a grant. For this reason, further analysis of the causal relationship between receipt of an RSA grant and survival is needed.

1.4. Chapter Summaries

This thesis consists of eight chapters in addition to this introduction. The next chapter begins by providing an explanation for why governments attempt to reduce disparities in unemployment rates across regions. It then proceeds to discuss various theories to explain regional variations in unemployment rates which imply different forms of intervention to support firms in poorly performing regions. These are the neoclassical theory, the dynamic capabilities theory and the evolutionary theory. Although not supported by all of the theories considered, the most frequently used type of support has been labour and, more often, capital subsidies. Given that the main aim of regional policy is to promote employment, it might appear strange that capital subsidies have been used more frequently than labour subsidies. An explanation for this apparent paradox is provided. A description of the evolution of regional policy since its origins during the depression is then given. This section will focus in particular on the 1960s and 1970s as this was the time when the largest amounts of money were spent on regional policy. It will also describe the development of the RSA scheme in detail. The final section will describe the grant schemes that currently operate in Scotland with an emphasis on the RSA scheme.

Chapter 3 will describe the econometric problems that arise when trying to analyse the impact of assistance because plants that receive a grant are a self-selected sample of the population of plants with different characteristics to those plants that did not receive a grant. It will then proceed to discuss four estimators that purport to allow the researcher to obtain consistent estimates of the impact of receipt of an RSA grant under these circumstances: the fixed effects, matching, instrumental variables and control functions estimators. Having described these estimators, a motivation for the choice of the matching and instrumental variables estimators rather than the fixed effects and control function estimators in chapter 7 and 8 is given.

Chapter 4 will review the theoretical and empirical literature that offers predictions as to what results may be expected from the empirical analyses of chapters 6, 7 and 8. This chapter is heavily weighted towards the empirical literature because the theory which explains why RSA should have an impact on total factor productivity (TFP) and survival is straightforward and relatively few papers have been written on the macroeconomic impact of business support programmes such as RSA. The empirical literature, by contrast, is voluminous and the review will be organized according to the nature of the data and the method employed to overcome the problems caused by self-selection into the group of plants that receive assistance. Although the results are diverse, receipt of a business support grant is generally found not to lead to a statistically significant increase in productivity. Of those papers that analyse the impact of receipt of an RSA grant, Harris and Robinson (2004) find a positive and statistically significant effect while Criscuolo, Martin, Overman and Van Reenen (2007) and Hart, Driffield, Roper and Mole (2008a) do not. Only two papers (Girma, Görg and Strobl, 2007b; Harris and Trainor, 2007) have been written on the impact of receipt of business support grants on the probability of survival and both find that receipt of a grant has a positive impact. However, these papers do not satisfactorily deal with the consequences of self-selection into the treatment group.

Chapter 5 begins by providing descriptive statistics on the schemes that currently operate in Scotland using data from the SAMIS database which is the register of plants that received support under the various grant schemes that operate in Scotland. This showed that the RSA scheme offers the largest number of and the most generous grants of all the business support schemes operating in Scotland. It then proceeds to describe how plants that received an RSA grant were identified in the dataset to be used for the empirical analysis. This involved linking the SAMIS database to the Annual Respondents Database (ARD) which contains the necessary information on inputs and outputs for econometric analysis. This process managed to link a higher proportion of plants that received a grant than has been managed before in analyses of RSA using these datasets (Harris and Robinson, 2004; Criscuolo, Martin, Overman and Van Reenen, 2007) which allows greater confidence in the empirical results. A description of the variables in the created dataset which will be used in the empirical analyses is then given. Finally, the chapter provides a comparison of the characteristics of plants that received an RSA grant with those that did not. This shows that RSA recipients possess different characteristics to those plants that do not receive a grant and therefore that self-selection bias will be an issue in the econometric analysis.

Chapter 6 decomposes the growth of aggregate labour productivity and aggregate TFP in Scottish manufacturing plants between 1994 and 2004 in order to reveal whether plants that received an RSA grant during this period contributed positively or negatively to aggregate productivity growth and the channels through which this contribution was made.

This is accomplished using the Haltiwanger decomposition. This shows that RSA-assisted plants made a small positive contribution to both measures of aggregate productivity growth, primarily because RSA-assisted plants that existed in both 1994 and 2004 tended to increase both their productivity and their market share. However, the decomposition also suggested a far greater contribution could be made if entrants were more and better targeted.

Chapter 7 will examine whether receipt of an RSA grant has a causal impact on plant TFP. In order to do so, two sources of bias must be overcome: the first arises due to the aforementioned self-selection into the group of plants that receive an RSA grant and the second arises due to the endogeneity of the factor inputs in the production function. The consequences of self-selection are tackled by creating a matched sample using propensity score matching and by using the instrumental variables estimator. The endogeneity of the factor inputs is dealt with using the system generalised methods of moments (GMM) estimator. The results using both the matched sample and instrumental variables shows that receipt of an RSA grant had no statistically significant impact on TFP.

Chapter 8 will investigate whether a causal relationship exists between receipt of an RSA grant and the probability of survival. This is done using a Cox proportional hazard model. The problems that arise due to self-selection into the treatment group are tackled by estimating the model on a matched sample created by propensity score matching. The results obtained using the matched sample showed that receipt of an RSA grant had no statistically significant impact on the probability of survival. It is worth noting that when no control for the consequences of self-selection into the treatment group was employed, the results implied that receipt of an RSA grant led to a statistically significant reduction in the probability of closure for some industries. This demonstrates that self-selection into the treatment group, when ignored, can generate biased estimates of the treatment effect.

The final chapter is a conclusion. This will begin by setting out the contribution to the literature made by this thesis. It will then summarise the results from the empirical chapters. Some policy recommendations will then be made on the basis of these results. These are essentially ways of changing the scheme so that it becomes more focused on improving the productivity of plants that receive grants. Finally, some suggestions for future work will be provided.

1.5. Conclusion

This chapter has provided a motivation for this thesis on the grounds of the levels of public resources put into the RSA scheme. Evidence was also provided to show that many European countries also spend large sums of money on grant schemes so the results obtained in this thesis should be of wider interest than merely to policy-makers in Scotland. The next section justified the choice of productivity and survival as the variables upon which the impact of receipt of an RSA grant will be estimated in the chapter 7 and 8. This was on the basis that gaining an understanding of the impact of receipt of an RSA grant on these variables is interesting and important and the literature has yet to reach a consensus on these questions. The last section gave an outline of each of the chapters of the thesis.

2. Regional Industrial Assistance and its Rationale

2.1. Introduction

Regional industrial assistance is primarily aimed at reducing disparities in unemployment across regions. However, different theories provide different explanations for the existence of such disparities. For instance, the neoclassical model suggests that externalities lead to differences in economic performance across regions. By contrast, the evolutionary model would emphasise a lack of innovation as the cause of relative underperformance. These differences are important because different understandings of the cause of the problem imply different solutions. This chapter will therefore discuss three theories which offer different explanations of the existence of disparities in unemployment across regions and support different types of government intervention.

Although not necessarily supported by the theories considered, the most popular form of intervention has historically been capital subsidies in the United Kingdom (UK). Given that the main aim of regional policy is to reduce disparities in unemployment across regions, it may seem strange that capital subsidies have been used more frequently than labour subsidies given that the standard analysis shows that, because capital subsidies lead to a substitution of capital for labour, employment subsidies should be more effective in boosting employment. As will be discussed in greater detail below, the standard analysis is static and consequently does not take account of improvements in the technology embedded in capital over time and the fact that capital provides a stream of productive services over many periods. A proper consideration of the nature of capital explains the favouring of capital over labour subsidies.

Regional industrial assistance has been provided in some form in the UK since the depression. However, the 1960s and 1970s were the heyday of regional policy. This was also the time at which the RSA scheme was born. In the course of giving a history of regional industrial assistance, and in particular the grant schemes on which the largest amounts of money have been spent, this chapter will also provide a description of the evolution of the RSA scheme. Industrial assistance in Scotland now takes the form of investment or innovation grants. Investment grants are provided under the RSA scheme while innovation grants are provided under the SMART: Scotland and the Research and Development (R&D) Grants schemes. A detailed description of the current operation of these schemes will be given below.

The chapter is structured as follows: the next section will describe why reducing regional disparities in employment is important; the third section will look at different explanations of why unemployment rates differ across regions and the policies implied by these different explanations; the fourth section will provide a comparison of capital and labour subsidies in order to determine which is best suited to increasing employment; the fifth will look at the history of regional policy in the UK and the sixth section will give a detailed description of the RSA scheme and the innovation grants schemes that currently operate in Scotland; the final section concludes.

2.2. Justification for Intervention

The main objective of regional industrial policy is to create employment in areas with persistently high unemployment. As such there are, according to Taylor and Wren (1997), four main arguments which support the use of regional industrial policy. The first argument contends that the reduction of unemployment in areas of high unemployment has direct social and economic benefits. The latter includes higher income for the unemployed, higher income for others through multiplier effects, lower expenditure on transfer payments, higher tax revenues and better prospects for areas with high unemployment due to the association between unemployment and poor educational attainment, a poorly skilled workforce and competitiveness. The social benefits accrue to those that would be unemployed from the avoidance of the demoralisation and poor health associated with unemployment. Others would also benefit from the lower crime rates and the improved physical and social environment that are associated with lower rates of unemployment.

The second advantage of reducing disparities in unemployment across regions is that it should ease inflationary pressures in the economy. This is based upon the notion that national inflation can be generated when only one region is experiencing excess demand for labour. The resultant wage inflation is transmitted to other labour markets even if they have high unemployment rates through national industry-wide wage agreements, inter-plant wage-setting in multi-plant firms and wage-setting on the principle that workers in similar occupations ought to be paid a similar amount. This wage inflation then leads to price inflation.

Creating employment in areas of higher unemployment is also desirable on the grounds that unbalanced economic growth leads to the persistence and intensification of economic problems. This occurs because disparities in economic performance will lead to the migration of workers from areas with high unemployment to areas with low unemployment. As the workers that migrate tend to be more skilled than those that remain, this will improve the skills of low unemployment regions at the expense of high unemployment regions and thereby exacerbate differences in regional economic performance.

Finally, reducing unemployment in areas of high unemployment is politically necessary. This is the result of the feeling of unfairness created by persistent differences in economic performance across regions.

2.3. Sources of Regional Variations in Unemployment

The previous section has explained why the reduction of regional variations in unemployment is thought to be desirable. This section will discuss different explanations of why differences in unemployment rates arise across regions. This is necessary as different understandings of the source of disparities in regional unemployment rates imply different types of government intervention to counteract it. The explanation provided by the government for the RSA scheme is based on the notion of market failure that arises from the neoclassical model. This section will therefore begin by discussing the notion of market failure. The next sections will discuss the dynamic capabilities and evolutionary theories of the firm which provide more nuanced explanations of why some firms succeed and therefore how firms in areas of high unemployment should be supported. The policies that could be employed to remove the source of variations in regional unemployment will be discussed and the extent to which capital grants schemes such as RSA could perform this role will be discussed. The other main type of support offered in Scotland is grants to support innovation and the extent to which these theories support this type of grants will also be considered.

Neoclassical Theory

In accordance with standard textbook theory (see, for example, Myles, 1995), the Green Book of HM Treasury (2003) states that government intervention is normally justified either by market failure or equity considerations. Market failure refers to a situation in which the unregulated market fails to deliver an efficient outcome. An efficient outcome exists when nobody can be made better off without someone else being made worse off. Intervention on the grounds of equity arises from government and (and to the extent that government represents society) societal preferences for equity. It is noted in the Green Book that intervention can incur costs and economic distortions which have to be taken into account when determining whether intervention is desirable.

As our interest here is primarily in a grant scheme which seeks to reduce disparities in unemployment across regions, it is clear that equity considerations play a role in motivating the use of these grants. However, because disparities in economic performance can also be argued to represent an inefficient outcome caused by market failures that are present to different extents across regions, such intervention can also be justified on the grounds of efficiency. The following discussion will therefore be framed in terms of market failure. The advantage of this approach is that specific types of market failure call for specific types of policy whereas equity considerations can be used to justify any form of intervention which aims to secure a more equitable outcome.

Annex 1 of the Green Book gives four situations in which the market fails: when goods have the characteristics of public goods; when externalities arise from some activity; when agents have imperfect information and when firms have market power. The first of these calls for the provision of public goods; the third demands laws to alleviate the information asymmetry while the last requires stronger competition policy. They therefore do not offer a justification for the existence of grant schemes. It is also not clear why any of these sources of market failure would explain differences in economic performance across regions and, as a result, they do not offer a justification for regional industrial assistance in general. The existence of externalities can however be regarded as a justification for the existence of such support.

An externality exists ‘whenever some economic agent’s welfare (utility or profit) includes real variables whose values are chosen by others without particular attention to the effect upon the welfare of the other agents they affect’ (Myles, 1995: 313). In other words, the agent’s welfare is partly determined by the actions of other agents who do not consider the impact of their own behaviour on the welfare of other agents. In this case, either the costs incurred by the agent of taking a particular action do not equal the social costs of the action or the private benefits from taking the action do not equal the social benefits. This leads to a socially inefficient outcome as private agents undertake a socially inefficient level of the action. When externalities are positive, too little of the activity is undertaken and when externalities are negative, too much of the activity is undertaken.

There are many types of externalities but the most relevant when seeking to explain disparities in economic performance is agglomeration externalities. These exist when there are cost reductions that accrue to firms that are situated in the vicinity of other firms. Such externalities may arise in many forms. According to Hart, Driffield, Roper and Mole (2008a), they include ‘collaboration and networking opportunities, technological externalities (e.g. spillovers, linkages), information transfer, the freeing up of internal human and financial capital (which can then be utilised in other innovative actions within the firm), the leverage of additional private sector financial support, or the range of perceived or actual benefits associated with large urban labour markets (e.g. skill sets)’. Agglomeration externalities will arise to different extents across regions because of differences in the concentration and behaviour of firms. Regional subsidies to capital and labour will remove the cost advantage to the richer areas which benefit from agglomeration externalities and thereby attract more firms to the disadvantaged areas. This should in turn generate the externalities, the lack of which caused the cost disadvantage in the first place. However, this is not automatic as firms must behave in certain ways to create and benefit from externalities. For instance, they must be involved in technological development if there are to be technological externalities. Therefore, this source of market failure implies support for schemes that specifically aim to support activities which give rise to externalities rather than merely trying to attract more firms to a given area. This provides support for innovation grant schemes that encourage R&D rather than capital grants schemes as this is an activity commonly thought to create externalities (see, for example, Gertler, 2003).

Another type of externality which would seem to offer a strong justification for the existence of capital grant schemes is that generated by capital investment. According to De Long and Summers (1991), equipment investment 'is a natural place to expect external economies and linkages to be important'. They provide empirical evidence supporting the existence of such an externality but do not specify precisely how it may arise. The idea is formalised by Keuschnigg (1998). The features of his model are monopolistic competition, product differentiation and free entry. These create an equilibrium in which the level of capital accumulation is lower than the socially optimum level. This occurs because each agent takes the investment level of other agents, and therefore the number and price of capital goods, as given when making decisions concerning whether to make an investment. However, each investment creates new firms which increases product diversity and leads to greater specialisation in production. This lowers the price of capital goods and, by doing so, increases the profitability of investment projects for all agents. Although it is not profitable for one agent to make the marginal investment, a coordinated increase in investment from all agents increases social welfare. Such an externality appears to provide a strong basis for capital grants schemes as, by reducing the price of capital, these should improve upon the socially inefficient level of investment that occurs in the absence of intervention. However, it should be noted that, for this externality to arise requires that capital investment creates new firms and products. The extent to which capital grant schemes create new firms and products depends upon the type of capital that is bought using the grant. If capital grants are used simply to fund replacement investment, this will not enhance product diversity. Furthermore, because the market for capital goods is likely to be quite integrated across the country, it is not obvious why this externality would arise differentially across regions. This externality would therefore appear to support a nationwide capital grant scheme rather than a regional grant scheme such as RSA.

A market failure not considered in the Green Book but suggested by Hart, Driffield, Roper and Mole (2008a) is that of incomplete markets. When markets are complete, firms are able to borrow the money they require at an interest rate reflecting the lenders' perceptions of the riskiness of the loan. However, financial markets may not be complete. As a result, the private sector may not provide loans to start-up firms or firms seeking to expand their operations because they regard them as too risky. This prevents such firms from obtaining private sector finance and prevents the achievement of an efficient outcome. Because the private sector will perceive that loans to firms in poorly performing regions carry a greater risk, firms in these areas will find it more difficult to obtain loans than firms in richer areas. Overcoming this externality requires the provision of finance to such firms. As regional capital grant schemes can be viewed as doing precisely this, the RSA scheme would appear to be well justified by this market failure. Innovation grant schemes can also be viewed in this way although the inherent riskiness of R&D projects, the lengthy period of time that it may take for returns to materialise and frequently large amount of finance required suggests that the difficulties involved in obtained private sector finance for this sort of project will be greater than those encountered when obtaining finance for other projects (Scottish Government, 2009b). Therefore, this market failure provides a stronger rationale for innovation grant schemes than capital grant schemes.

The rationale provided by the government for intervention is based on the notion of market failure which arises from the neoclassical model of the economy. Of the various different types of market failure, regional capital grant and innovation grant schemes can be justified by the existence of externalities and incomplete markets. These rationales for intervention are however stronger for innovation grant schemes than capital grant schemes. However, the neoclassical model from which the notion of market failure arises is a static model that depends on assumptions such as perfect information and complete mobility of factors of production which do not hold in reality (Harris and Robinson, 2001a). More importantly, the theory of the firm contained in this model abstracts from many aspects of the firm which are important when considering why disparities in economic performance may exist between regions and how the government could act to reduce them. What follows in the next two sections is a discussion of two theories of the firm which provide further insight into how the government may act to reduce disparities in economic performance across regions. The following draws heavily from Harris and Robinson (2001b).

Dynamic Capabilities Theory

The notion of dynamic capabilities, which has been developed in a series of papers by Teece (see, for example, Teece and Pisano, 1994; Teece, 1996) is a means of explaining how firms acquire and sustain a competitive advantage in a rapidly changing business environment. According to this view, the strategic dimensions of a firm are its managerial and organisational processes, its present position and the paths available to it. These dimensions cover the firm's capabilities or competencies. When it is difficult to replicate or imitate these capabilities or competencies, they can be thought of as distinctive competencies. Another characteristic of such competencies is that they cannot be bought on the market. If they could be bought on the market, they would only provide a transitory competitive advantage as competitors could simply buy them, thereby removing the competitive advantage from the firm that was the first to develop that competence. Dynamic capabilities are the 'subset of the competences/capabilities which allow the firm to create new products and processes, and respond to changing market circumstances' (Teece and Pisano, 1994: 541). In other words, dynamic capabilities are the driver of innovation which is, in turn, the driver of technological change. As such, it is reasonable to suppose that firms in poorly performing areas lack such dynamic capabilities and their development should therefore be the focus of policy. The following section will provide a description of different types of dynamic capabilities. It will be structured by the strategic dimensions mentioned above.

Managerial and organisational processes (henceforth referred to simply as processes) simply refer to the manner in which things are done in the firm. They encompass the state of three aspects of the firm which, depending on their condition, may constitute a dynamic capability. Firstly, processes determine the efficiency with which activity is coordinated by managers both internally and externally. Examples of external coordination include strategic alliances, buyer-supplier relations and technological collaboration. Another part of the firm's processes which may constitute a dynamic capability is the way in which the employees of the firm learn. This is an important source of competitive advantage as fast learning will allow a firm to be flexible and employ new processes quickly. In addition to improving individual skills, learning improves organisational skills which lead to improved processes so the relationship between learning and processes is bidirectional. Finally, the firm's ability to recognise the need to reconfigure its asset structure and then to accomplish the requisite reorganisation is another key part of its processes and an important dynamic capability. This requires an awareness of the market and the technological environment and a willingness and ability to change.

The second dimension of the firm and potential source of dynamic capabilities is the firm's position. This refers to its current stock of business assets. One type of asset that cannot be regarded as a dynamic capability is generic plant and equipment which can be bought on the market. By contrast, technological assets can be a source of dynamic capabilities.

Although some technological assets can be traded on the market, much technology cannot be bought because the owner is unwilling to sell or because of difficulties in transactions involving technology. Another source of dynamic capabilities is the firm’s stock of complimentary assets. These assist in the production and delivery of new products. In the presence of liquidity constraints, a firm’s financial assets are another potential source of dynamic capabilities. However, the importance of liquidity constraints should not be overstated as they are unlikely to be a long-term hindrance. Locational assets can also be a dynamic capability because, although property markets are well established, building and environmental restrictions can prevent trade in such assets. Such restrictions provide durable advantages in the form of lower transport costs and market access.

The final part of the firm’s dimensions, its paths, refers to the strategic options available to the firm and are therefore an obvious determinant of its ability to react to changing market conditions. The inclusion of this dimension indicates an acceptance of the notion of path dependency whereby the options available to a firm is a function of its past decisions (key papers in the development of this theory are David, 1985; Arthur, 1989). Path dependency is particularly important in determining the technological opportunities available to the firm. These are a lagged function of scientific activity external to the firm and also of innovative activity commissioned, either internally or externally, by the firm itself. The implication of the latter is that technological opportunities are unique to a firm that has undertaken innovative activities in the past.

In terms of policy, this theory implies that government policy should be aimed at improving the dynamic capabilities of firms in areas of high unemployment. Assistance should be provided directly to the firm instead of trying to build up the technology infrastructure of the region since the latter will merely facilitate the sharing of generic knowledge. Since this type of knowledge is easy to obtain, the latter type of policy will not provide firms with a source of competitive advantage. However, due to their nature, helping firms to develop dynamic capabilities is not easy. Capital or labour subsidies will help only to the extent that they help to overcome liquidity constraints, as generic plant and equipment are not a source of dynamic capability due to their being purchasable on the market. Policies aimed at encouraging innovation, such as innovation grant schemes, could lead to the development of dynamic capabilities as these will create new technological assets and opportunities. They should therefore be regarded more favourably by proponents of this theory.

Evolutionary Theory

The evolutionary theory, which offers rather different policy conclusions to the dynamic capabilities theory, takes much from the work of Schumpeter and thereby represents a radical break from standard approaches based on the notion of equilibrium. Standard theory shows that, with a fixed set of products and technologies, a competitive market structure will result in the lowest price to consumers whereas a monopolistic market structure will lead to higher prices and lower output. It is competition based on price which produces this superior outcome under a competitive market structure. Schumpeter regarded this analysis as missing the point because, due to its holding of products and processes constant, it places too much emphasis on price competition (Diamond, 2004). Schumpeter argued that, in reality, competition occurs primarily through the innovation process, rather than through prices, as entrepreneurs seek to develop new products or technologies which provide them with a cost or quality advantage over their rivals. This allows them to survive and make large profits but reduces the profits of or destroys their rivals. The innovation process can therefore be summarised as a process of 'creative destruction' (Schumpeter, 1943). As this innovation process occurs continuously, the notion of equilibrium is incompatible with the capitalist economy because the economy never reaches an equilibrium state but remains in a constant state of flux.

Although debate surrounds whether Schumpeter can properly be classified as an evolutionary economist, it is undeniable that his thinking has deeply influenced the development of evolutionary economics (Metcalfe, 2000). The following summary of the evolutionary theory of technical change borrows heavily from Metcalfe and Georghiou (1997). The model involves three distinct stages, each of which will be described in turn.

The first is the innovation stage which generates diversity. As with the resource-based theory, innovations are understood to be a function of the characteristics of the firm. However, it is also emphasised that innovations are 'guided and constrained by cognitive frameworks and the embedding of those frameworks in institutional rules and practices' (Metcalfe and Georghiou, 1997: 78). In other words, external factors are important as innovations are developed using information gleaned from outside sources. Radical innovations are rare because the direction of innovation is driven and constrained by existing intellectual and institutional capital built for research in a particular area. Ex post, the innovation process appears to be wasteful as innovations are usually the consequence of trial and error experimentation. ‘Errors’ however provide valuable information on where not to direct research effort in the future and are an unavoidable element of the innovation process.

From the diversity created by innovation, the selection of which technologies will be diffused throughout the economy occurs. This is the second stage in the evolutionary model and is accomplished primarily by the market, where suppliers and users of innovations meet. The market fulfils this role by co-ordinating ‘the development of demand, investment and the growth of productive capacity together with the processes of learning which take place jointly between users and suppliers’ (Metcalfe and Georghiou, 1997: 79). Diffusion may also occur outside the market through imitation of existing practice. The extent of imitation will depend, among other factors, upon the stringency and coverage of intellectual property laws.

The final stage is feedback of knowledge from the selection to the innovation process. This stage generates path dependencies in the innovation process as the selection of technologies in one period influences the nature of innovations in the next. This raises the possibility that firms and markets could go down technology trajectories that are not optimal in the long-run (see Altman, 2000 for a model showing how this can happen).

The importance of innovation is made clear by considering what would happen in its absence. Without the first stage in the model, each firm would adopt the existing best technology, the market would be homogenised and the economy would not grow. Innovation is therefore crucial as it replenishes the system with the new innovations which lead to technological progress and growth.

Due to its importance in this theory, the model would appear to regard a lack of innovation as the source of differences in economic performance across regions and therefore provides support for government intervention that bolsters innovation in the poorer regions. Policy is advocated that encourages innovative activities from all firms. This is preferable to the more targeted approach supported by the dynamic capabilities theory because the outcomes from the innovation process are too unpredictable for policymakers to identify which firms to support. The evolutionary theory does not support capital or labour subsidies because they do not directly aim to bolster the innovation stage of technological development. Indeed by supporting plants that would otherwise be forced to make redundancies, these subsidies may impede the Schumpeterian process of ‘creative destruction’ that creates growth in the economy. Discretionary schemes, in general, are regarded unfavourably as, by providing support for individual plants, such schemes try to ‘pick winners’ and this is seen as difficult due to the uncertainty of innovation outcomes. Innovation grant schemes are better founded with regards to the evolutionary theory of the firm as these do aim to bolster the innovation stage. However, as they are also discretionary grant programmes, they can also be criticised on the grounds that they try to ‘pick winners’.

2.4. Capital or Labour Subsidies?

The previous section has offered a number of different explanations for the existence of disparities in regional unemployment. Each explanation implies a particular policy to remove the source of the disparities in unemployment. Although the existence of market failure caused by incomplete markets or the existence of externalities can be regarded as providing support for subsidies to the factors of production, neither the dynamic capabilities or the evolutionary theories imply support for such subsidies as a first-best policy. Nevertheless, although they may not directly tackle the suggested causes of disparities in unemployment across regions, capital and labour subsidies may still perform a function in reducing them. Indeed, they have been the most frequently used instrument of regional industrial assistance, as will be shown in the next section. Capital grants have been used more frequently than labour subsidies which may seem surprising given that the main aim of regional policy is to reduce disparities in unemployment rates across regions. The following analysis will discuss the impact of both types of subsidy and thereby explain whether labour or capital subsidies are superior in terms of creating employment.

According to a standard static analysis, the impact of a capital or labour subsidy on the employment level of the recipient firm depends upon the induced substitution and output effects. As a capital (labour) subsidy lowers the price of capital (labour), the substitution effect on employment will be negative (positive). For a fixed level of subsidy, its magnitude will depend upon two factors. The first is the supply elasticities of labour and capital which determine the extent to which changes in demand for factors of production translate into changes in inputs. The second is the shape of the firm's isoquant as this governs the size of the induced change in demand for each factor of production.

Both capital and labour subsidies will lead to a reduction in the cost of producing a unit of output. Assuming the recipient firm responds to this induced reduction in costs by cutting prices, there will be an increase in demand for its output. In order to satisfy this increased demand, firms will employ more inputs of both capital and labour. The output effect on labour of both capital and labour subsidies is therefore positive. For labour subsidies, the impact on employment is unambiguously positive but for capital subsidies, only when the induced output effect of the subsidy is larger than the substitution effect will the subsidy have a positive impact on employment. The magnitude of the output effect will be determined by a number of factors: the size of the reduction in production costs precipitated by the subsidy; the extent to which firms lower their prices in response to the subsidy; the elasticity of demand for output with respect to price; and the technical conditions of production.

The analysis above implies that the firm is capable of continuing in operation with or without the subsidy. However, some firms may have costs that are so high that they cannot operate without some form of subsidy. For such firms, both capital and labour subsidies of sufficient size will allow it to continue trading which will save jobs. However, in accordance with the analysis above, the capital subsidy will lead to a substitution from labour to capital while a labour subsidy will lead to a substitution of capital for labour. There will, of course, also be an output effect which will increase the levels of both capital and labour. However, it is clear that the analysis implies that the number of jobs saved would be larger with labour subsidies.

A serious problem with the simple analysis presented above is that it does not take account of the general equilibrium effects of subsidies. Both capital and labour subsidies will create multiplier effects which will increase employment in other sectors in the surrounding area and beyond. As workers will tend to spend a high proportion of their wages in the region, this would imply that the multiplier from labour subsidies will be large. By contrast, the capital that is bought using the subsidy is likely to be purchased from outside the region which implies that the capital subsidy multiplier may be lower than the labour subsidy multiplier.

The analysis has so far presented a strong case for labour rather than capital subsidies as the best form of grants for reducing unemployment. However, taking account of dynamics calls this conclusion into question. The analysis has thus far been static and therefore has had nothing to say on the fact that the technology embodied in capital improves over time. This means that, when new capital is purchased, it will be more productive than older capital. Furthermore, the standard analysis does not properly capture the nature of capital. Capital, once bought, provides productive services over many years unlike labour which must be paid in each period in order to procure productive services. When making the decision as to the quantity of inputs to buy, firms will therefore have a stock of existing capital to take into consideration.

Holding the level of output constant, a capital subsidy leads a firm to buy more capital than it would otherwise have bought. As the new capital will be more modern than the existing capital, this will increase the productivity of the firm. The costs of producing a unit of output will therefore fall for two reasons: because the cost of capital is subsidised and because the more modern capital is more productive. Holding output constant, a labour subsidy would lead to the employment of more labour and less of the more technologically advanced capital. The impact of a labour subsidy on the costs of producing a unit of output is therefore ambiguous: by reducing the cost of employing labour, the labour subsidy generates a reduction in the cost of producing a unit of output but this effect is counterbalanced to some extent by the lower levels of the more technologically advanced capital that is bought. Therefore, the capital grant is more effective in reducing the costs of producing a unit of output and is consequently more likely to assist in securing the viability of the firm in the long-run. It will therefore also generate a larger output effect. This is the main reason why capital subsidies are preferred over labour subsidies.

The advantage of discretionary capital subsidies over automatic capital subsidies in terms of generating employment is clear. As the analysis above has shown, it is not obvious whether the provision of a capital grant will lead to an increase in employment in the recipient firm. To avoid the danger that grants are provided to firms which use them to cut employment, scheme administrators can ensure that discretionary capital subsidies are only provided to firms that promise to increase their employment. This allows scheme administrators to ensure that maximum cost-per-job conditions are satisfied (see Swales, 1997 for a discussion of the UK government's cost-per-job ceilings that are applied to the RSA scheme).

2.5. History of Regional Policy in the UK

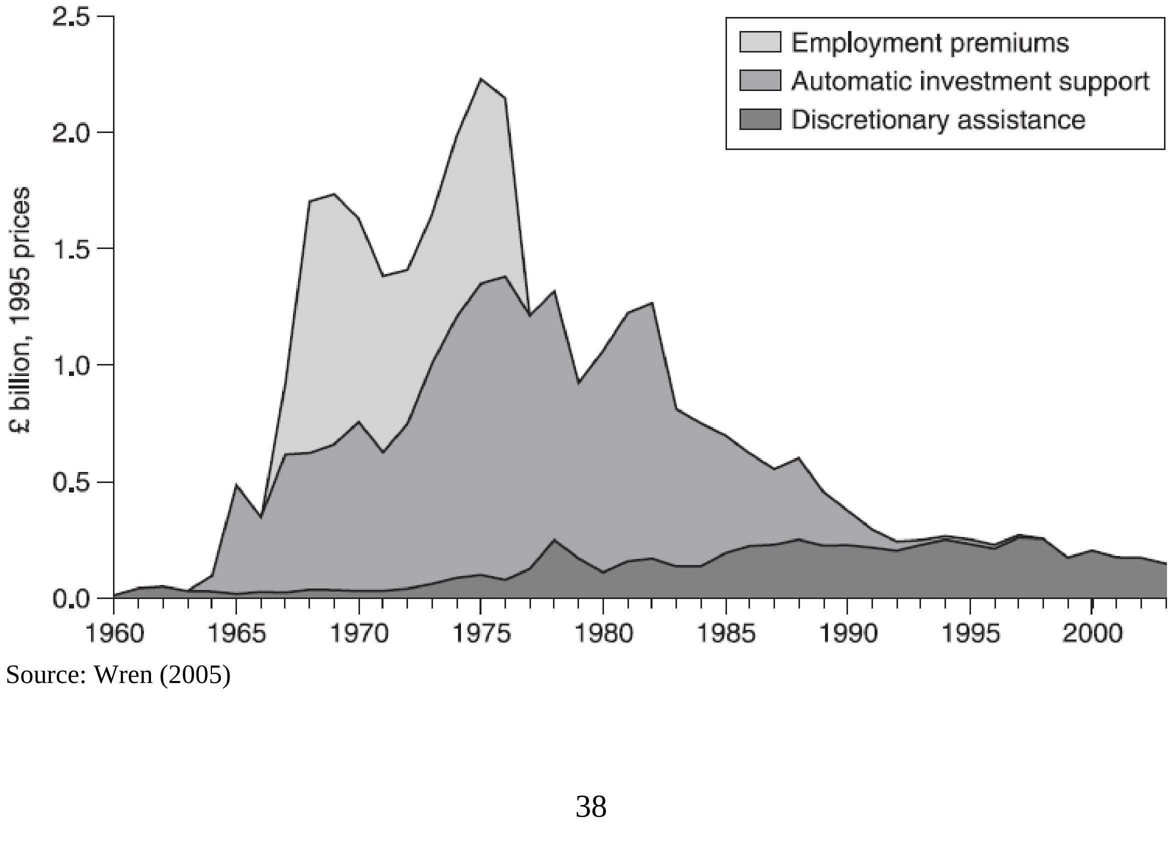
This section will provide a description of the evolution of regional policy in the UK. The focus will be on regional policy conducted by the national government which has traditionally been the main actor in this sphere. This is because the RSA scheme began and remained until recently part of national regional policy. It should be noted though that, in recent years, local government and the EU has taken on a greater role in the provision of regional industrial assistance. The section will begin by giving a general description of regional policy before focusing on three specific components of regional policy on which expenditure has been greatest: automatic capital assistance, employment premiums and discretionary assistance. The following general description of the early evolution of regional policy is based upon Armstrong and Taylor (2000).

Regional industrial policy began in the late 1920s when, in response to high unemployment in areas dependent on staple export industries, money was provided to workers to allow them to migrate from high to low unemployment regions. By 1938, more than 200,000 workers had received financial assistance to move under this scheme. However, given that unemployment reached three million in 1933, the impact of this scheme on aggregate unemployment was small. Further aid was provided under the Special Areas Acts of 1934 and 1937. Unlike the earlier policy, this aid attempted to create employment in areas of high unemployment rather than encouraging workers in high unemployment areas to move to areas of low unemployment. The impact of the aid provided under these acts was, however, also small due to the low level of expenditure.

The 1944 White Paper on Employment Policy signalled a shift towards a greater commitment to regional policy. This committed the post-war government to reducing unemployment in the depressed regions of the country. To achieve this, the Distribution of Industry Act of 1945 introduced many measures including loans and grants to firms, the power to build factories and the power to provide services to industries in the deprived areas. The most powerful instrument, however, was the imposition of a system of controls on the location of industry. But this enthusiasm for regional policy did not last and these measures were little used during the 1950s.

However, by the 1960s, serious concern was being expressed about the relative performance of Britain compared to other industrialised nations. It was recognised that the underperforming regions contained large supplies of labour which, if productively employed, could contribute significantly to the improvement of national economic performance. Secondly, there were worries about the negative externalities generated by the relative growth of the Greater London area. Therefore the amount spent on regional policy increased sharply as a number of measures were introduced to encourage firms to locate in poorly performing areas. Figure 2.1 shows expenditure on different types of regional assistance between 1960 and 2003. It includes assistance provided in the form of grants, loans and investment tax allowances. The figures are ‘grant equivalents’ which is the grant amount that has equal net present value to the subsidy. If a similar national scheme exists, the extra value provided to the recipient firm by virtue of its being located in the Assisted Areas is given.

Figure 2.1: Expenditure on Regional Industrial Assistance in the UK, 1960-2003



Source: Wren (2005)

As shown in figure 2.1, the total amount spent on regional assistance rose dramatically from the middle of the 1960s. The increase in expenditure was achieved through the introduction of employment premiums and automatic investment support, which will be discussed below. Partly in response to recession, expenditure on regional industrial policy rose again in 1972. However, this increased level of expenditure did not last as the amount of investment eligible for automatic assistance fell as a result of the recession of the late 1970s. Expenditure continued to decline in the early 1980s as the government attempted to increase the cost-effectiveness of regional policy. As part of this, the proportion of the country with assisted area status was severely cut back. Location controls, which had been used frequently during the 1960s, were also abolished in 1982. The reduction in expenditure on regional industrial assistance was to some extent counteracted by increased expenditure on regional policy by the EU which is not included in figure 2.1.

The role of regional policy was radically changed in 1988. The government now regarded the poor performance of the regions as the effect of 'economic inefficiency due to supply-side rigidities and a deficiency of entrepreneurial activities' (Armstrong and Taylor, 2000: 219). Regional policy therefore became focused upon the removal of these supply-side rigidities and the stimulation of indigenous entrepreneurship in a bid to promote indigenous growth. This differed from previous regional industrial policy which aimed to attract inward investment to the assisted areas rather than promoting development from within the regions. Automatic capital grants provided to all firms in the assisted areas were ended with the abolition of the Regional Development Grants scheme. These changes in the focus of regional policy dovetailed with the government's increasing focus on improving the competitiveness of British industry.

So far, a very general history of regional policy has been provided. The following will discuss the evolution of employment premiums, automatic investment support and discretionary support on which the largest amounts of money have been spent. The latter includes the RSA scheme, the evolution of which will be described in detail. Information on the older schemes and the evolution of the RSA scheme is primarily taken from Wren (1996) and Wren (2005).

The largest component of regional industrial assistance in this period was automatic or nondiscretionary capital support on which £18.8 billion (1995 prices) was spent between

1963 and 1997. This was introduced through regionally differentiated tax allowances and the grant equivalent reached almost £500 million by 1965. Due to the erosion of the real value of these allowances caused by the introduction of corporation tax, these allowances were replaced by Investment Grants in 1966. However, Investment Grants were then replaced in 1970 by regionally differentiated first-year writing down allowances before being reintroduced again in 1972 under the guise of Regional Development Grants. Expenditure on Regional Development Grants rose to over £1.2 billion by 1976 so, in an effort to reduce the cost of the scheme, construction and mining were made ineligible for Regional Development Grants in 1977. Expenditure on Regional Development Grants fell during the recession because of a drop in investment but rose again afterwards until 1982. The entire scheme was revised in 1984 so that the amount received was linked to the number of jobs created by the capital investment project. The amount spent then fall steadily until the scheme was replaced by the Regional Enterprise Grants scheme in 1988.

Regional Enterprise Grants took the form of either investment or innovation grants. The former was a simple capital subsidy while the latter were provided to assist firms to introduce a new product or production process using the latest technology. However, the Regional Enterprise Grants scheme only provided support to small firms and the expenditure was consequently comparatively small compared to what had gone before. The scheme was abolished in 1997, ending automatic investment support in the UK.

The second largest component of regional industrial assistance was employment premiums on which £7.8 billion was spent between 1960 and 2003. These were introduced by the Selective Employment Tax in 1966. This was a surcharge on the National Insurance contributions which was refunded to firms outside of the construction and services sector in a bid to shift the structure of the economy more towards exporting industries. A Selective Employment Premium was also created at this time for firms outside of construction and services. In 1967, this was made available only to firms in the Assisted Areas. The Regional Employment Premium was also introduced in 1967 which paid a given amount for each employee eligible for the Selective Employment Premium. Spending on employment premiums had reached over £1 billion by 1968. However, the real value of the premiums declined in real terms from 1967 to 1974 when they were doubled to restore their value. However, in 1976, employment premiums were abolished altogether as part of expenditure cuts.

Prior to the introduction of the RSA scheme, discretionary loans and grants were provided under various Local Employment Acts. In 1972, they were replaced by the RSA scheme which was introduced alongside the Regional Development Grants scheme under the 1972 Industry Act. At first, RSA projects were divided into two categories. Category A projects included new projects and expansions that generated additional employment. Assistance was provided to such projects in the form of cheap loans, interest-relief grants and grants that helped to pay the costs of moving to an Assisted Area. Firms in the service sector that did not primarily serve local markets were eligible to apply for assistance under this category. Category B contained projects that maintained or safeguarded existing employment. For these projects, assistance was offered as a loan at commercial rates although funding was only provided in cases where it could not be obtained on reasonable terms from the private sector. For both types of project, the level of assistance provided was related to the number of jobs created or safeguarded. Furthermore, applicants for either category of assistance had to demonstrate that the firm was viable and that the bulk of funding would come from the private sector. Assistance could also be given in any form, including the provision of guarantees or the buying of share capital, to firms that were about to shed large numbers of jobs. However, this was only done in exceptional circumstances and the amount of finance provided was related to the number of jobs at risk. As is clear from figure 2.1, expenditure on RSA was small throughout the 1970s compared with the expenditure on employment premiums and automatic capital grants. From 1972 to 1976, most of the money spent under the RSA scheme took the form of loans and equity rather than grants. However, from 1977, the amount spent on loans and grant decreased and grants became the predominant source of funding.

In an effort to reduce its cost and improve its effectiveness, the newly elected Conservative Government undertook a review of industrial aid in July 1979 which led to the tightening of the eligibility requirements for the RSA scheme. A ‘Proof of Need’ condition was introduced which stated that assistance could only be given where funding could not be obtained elsewhere on the required terms. The provision of RSA must therefore lead to a major alteration in the nature or scale of a project, an advance in timing or a change in location to an Assisted Area. This condition had previously existed for Category B projects but was now extended to Category A projects, of which there were far more. Another significant change was the imposition of a ‘Regional and National Benefit’ condition which required that assisted projects strengthen the regional and national economy. As a result of this condition, possible job displacement caused by the provision of assistance was now taken into account when determining whether applications were successful. This gave RSA a bias towards firms operating in international markets whose competitors are located outside of the UK. Projects were also now asked to provide more 'productive and secure jobs' (Wren, 2005: 253) indicating that productivity enhancement had joined employment promotion as an aim of the scheme. Furthermore, in an effort to reduce the administration costs of the scheme, loans were made only in exceptional circumstances so assistance was, from 1980, provided almost exclusively in the form of capital grants. From 1983 onwards, no loans were provided and RSA became purely a grant scheme with the amount provided determined by the fixed and working capital involved in the project and the number of jobs created or maintained.

Further changes to the scheme were made in 1984. Firstly, relocation projects that did not lead to a net increase in employment were made ineligible for the scheme. Secondly, a limit on the size of grants that could be offered per job created or safeguarded was introduced. Lastly, a 'clawback' clause was introduced which dictated that the grant was repayable if the capital assets bought using the RSA grant were not retained for three years or if the jobs created or safeguarded by the grant did not last for 18 months.

Responsibility for administration of the RSA scheme shifted to the newly devolved Scottish Parliament and the Welsh assembly in 1999. This heralded the end of the uniform provision of regional grants across the assisted areas of Great Britain (GB).[3] In 2000, RSA in Scotland was made unavailable to projects which involved capital expenditure of less than £500,000. For such projects, the Invest for Growth (IFG) scheme was created although this only provided support to small and medium enterprises (SMEs). The Enterprise Grants and the Assembly Investment Grant were the equivalent schemes created at this time in England and Wales respectively. All of these schemes have since been abolished. In Scotland, the IFG scheme was disbanded and RSA grants are once more available to businesses of all sizes. In England, both the RSA scheme and the Enterprise Grants scheme were replaced by the Selective Finance for Investment in England scheme in 2004. This scheme included in its eligibility criteria an explicit requirement that projects

[3] Northern Ireland has always had its own version of the RSA scheme called Selective Financial Assistance (SFA).

should create high productivity jobs which does not exist for the equivalent schemes in Scotland and Wales. According to the latest evaluation, the 'new scheme was designed to be more in line with BERR's (the Department for Business, Enterprise and Regulatory Reform's) overall objectives on regional policy which sought to develop the competitive strengths of the region through more sustainable forms of industrial development' (Hart, Driffield, Roper and Mole, 2008b: 11). The Selective Finance for Investment in England scheme was in turn replaced in 2008 by the Grant for Business Investment scheme which retained the productivity requirement. In Wales, both the RSA scheme and the Assembly Investment Grants were replaced in 2008 by the Single Investment Fund. Although there are slight differences across Scotland, England and Wales in the schemes that currently exist, the essence of the schemes remain the same in that they aim to promote employment in the assisted areas through the provision of capital grants.

2.6. Current Grant Schemes in Scotland

Having provided a history of regional policy since the 1960s, this section will describe the grant schemes currently operating in Scotland. The first part will discuss the RSA scheme. The second section will describe the innovation grant schemes which currently operate in Scotland. It should be noted that the innovation grants schemes are not a part of regional policy as firms throughout Scotland can receive these grants. Descriptive statistics on the RSA scheme and other discretionary grant schemes are provided in chapter 5.2.

RSA

According to the latest annual report, 88 RSA grants worth a total of £52.1 million (current prices) were offered and accepted in Scotland in 2008/09 (Scottish Government, 2009a). As noted in the introduction, this represents a large fall on the value of grants accepted in recent previous years. These offers were provided to support capital investment of £518.4 million and to safeguard or create more than 5,000 jobs. 11 of these offers were larger than £1 million and the largest was £10 million. The extent to which RSA is used as an instrument to attract foreign direct investment (FDI) is also clear from the annual report. 20 of the offers of RSA accepted in Scotland in 2008/9 were made to foreign owned firms but the total value of these offers was over £20 million. This implies that the average size of offer accepted by foreign owned firms was considerably larger than the average size of offer accepted by UK-owned firms. This is unsurprising since the size of the planned capital expenditure associated with offers accepted by foreign owned firms was almost £260 million - over half of the planned capital expenditure associated with all the offers of RSA accepted in the year. The average number of planned jobs associated with accepted offers of RSA is also far larger for foreign-owned than UK-owned firms (see Harris, 2010: section 2 for some descriptive statistics on the use of RSA to attract FDI).

There are currently seven criteria that must be satisfied for a project to be eligible to receive an RSA grant (Scottish Government, 2009b). Most of these have already been discussed above but greater detail on them will now be provided.

The first is that the project must take place in an assisted area. The assisted areas are divided into three tiers which have different EC specified limits concerning the maximum proportion of the project costs that can be covered by grants (the current map of the Assisted Areas in Scotland is shown in the Appendix A2.1). Tier 1 is the area where the highest proportion of the costs of the project can be covered by an RSA grant. The area covered by this tier is currently the Highlands and Islands of Scotland and, as such, the scheme is administered by Highlands and Islands Enterprise. Their website does not provide specific details on the limits of funding available to businesses of different sizes (Highlands and Islands Enterprise, 2009). Tiers 2 and 3 are administered by the Scottish government. Large businesses in tier 2 may receive a maximum of 15% of their project costs from an RSA grant; medium sized businesses may receive up to 25% of project costs while small businesses can receive a maximum of 35% of their project costs.[4] In tier 3, only small or medium sized businesses can receive RSA grants. Medium sized businesses may receive a maximum of 10% of their project costs covered by an RSA grant while small businesses can receive up to 20% of the costs of the project.

The second criterion is that the project must directly create or safeguard jobs within the recipient firm. The jobs can be either full or part-time but must be permanent posts. Furthermore, the jobs created or safeguarded must not be offset by jobs losses in some

[4] Small businesses are defined as businesses that employ fewer than 50 people and have turnover of less than £6.7 million or a balance sheet total of less than £6.7 million. Medium sized businesses are businesses that employ fewer than 250 people and have turnover of less than £34 million or net assets of less than £29 million.

other part of the assisted areas. This is the displacement criterion. This desire for gains in net employment in Scotland leads to a preference for businesses that sell to markets outside of Scotland and whose competitors are mainly situated outside the assisted areas.

The fourth requirement is that projects must involve an element of capital investment. This includes expenditure on land, buildings, plant, machinery, software and the acquisition of intellectual property.

The next criterion is that the project and applicant business has to be financially viable and that projects must also receive most of their funding from the private sector. Finally, the project must require a grant for it to proceed. This is the so-called additionality criterion. Grants can be awarded in cases where the grant increases the size of the project, improves the project in some way or accelerates the project; they cannot be awarded if the project would go ahead in the same form regardless of whether a grant is provided. Grants should also not be provided if the business has already committed to carrying out the project.

Although the rules of the scheme stipulate that all projects must be additional, studies have shown that this criterion is not always satisfied. This is an important issue when estimating the impact of receiving an RSA grant because projects supported by an RSA grant may have an impact on aspects of plant performance but, if the project would have gone ahead in exactly the same form without the RSA grant, attributing the change in plant performance to receiving an RSA grant is problematic.

For Scotland, Hart, Driffield, Roper and Mole (2008a) report that only 1.9% of firms say that their project would have gone ahead in exactly the same form without an RSA grant over the period from 2000 to 2004. This suggests that additionality may not be much of a problem. However, 19.7% of firms report that receiving an RSA grant had no effects beyond speeding up the project. 21% of firms report that, without the grant, they would have achieved only some of the business outcomes and 28.7% report that they probably would not have achieved the business outcomes. 28.7% report that they would definitely not have achieved the same business outcomes had they not received an RSA grant.

Given that it is not possible to identify in SAMIS which projects were non-additional, little can be done to tackle the problems posed by non-additionality. It must therefore be recognised that if there is a differential impact of additional and non-additional projects, the existence of non-additional projects will lead to biased estimates of the impact of receiving an RSA grant. Fortunately, the magnitude of the problem is not as great as in other parts of the UK. In Northern Ireland between 1998 and 2004, almost 10% of firms report that their projects were entirely non-additional while around 38% report that receiving an SFA grant had no effect apart from speeding up the project (Hart, Driffield, Roper and Mole, 2008c). In England for the period between 2000 and 2004, the corresponding figures for the RSA scheme are 5% and 26.3% while, for the SFIE scheme, the corresponding figures are 4.9% and 22.8% (Hart, Driffield, Roper and Mole, 2008b). This shows that the RSA scheme in Scotland has comparatively high levels of additionality.

There are two eligibility criteria for the Grant for Business Investment scheme in England not used by the RSA scheme in Scotland (House of Commons, 2009). The first of these is that supported projects have to generate an improvement in productivity. Whether a project satisfies this criterion is determined by a comparison of gross value added (GVA) per employee for the jobs associated with the project with the industry and national averages. Secondly, the majority of jobs associated with applicant projects for the Grant for Business Investment (GBI) scheme must be at National Vocational Qualification (NVQ) 2 level or above. There is therefore a clearer focus on productivity enhancement for the GBI scheme than for the RSA scheme.

The size of grant that the Scottish Government can offer depends upon the location of the project and the size of the applicant business. Other factors taken into consideration when determining the size of grant awarded is the size of the project, the number of jobs created or safeguarded, their quality, type and how much the Scottish government believes is required for the project to go ahead. Payments are made once previously agreed levels of capital expenditure and employment have been met.

Innovation Grants

There are two types of innovation grant offered by the Scottish Government. The first that will be discussed is the SMART: Scotland scheme. The SMART scheme began throughout the UK as a pilot in 1986 and two years later the full scheme was launched. Unlike RSA which is available only in the assisted areas, SMART: Scotland is available throughout Scotland. It is a programme which provides discretionary grants to individuals planning to start a business or SMEs for projects which 'represent a significant technological advance for the UK sector or industry concerned' (Scottish Government, 2009b). In particular, grants can be provided for technological and commercial feasibility studies which involve early stage R&D or R&D projects that are attempting to develop a 'pre-production prototype of a new product or process' (Scottish Government, 2009b). Successful applicants for support for feasibility studies will receive 75% of the costs of the project, which must last between 6 and 18 months, as long as this does not exceed the maximum grant of £70,000. Successful candidates for grants for R&D projects receive 35% of the costs of the project up to a maximum of £600,000. Such projects must last between 6 and 36 months and involve costs of at least £75,000. Grants for R&D Projects were previously provided by the Support for Products under Research (SPUR) scheme which has effectively been subsumed into the SMART: Scotland scheme.

A separate scheme is the R&D Grants scheme. This began in 2008 and replaced the R&D Plus, the Small Company Innovation Scheme and the SME Collaborative Research schemes. R&D grants are given to support 'businesses developing new products, processes and services to improve company competitiveness and to benefit the Scottish economy' (Scottish Government, 2009b). Unlike the SMART: Scotland scheme, R&D grants are available to firms of all sizes. There are two types of project which are eligible for R&D grants: industrial research and experimental development. The former type of project is defined as 'the planned research or critical investigation aimed at the acquisition of new knowledge and skills for developing new products, processes or services or for bringing about a significant improvement in existing products, processes or services' (Scottish Government, 2009b). The latter type of project involves 'the acquiring, combining, shaping and using existing scientific, technological, business and other relevant knowledge and skills for the purpose of producing plans and arrangements or designs for new, altered or significantly improved products, processes or services' (Scottish Government, 2009b). R&D grants can be provided to SMEs at a maximum rate of 35% of project costs for grants up to £40,000 and at a maximum rate of 25% of project costs for grants above £40,000. For larger firms, 25% is the maximum rate at which grants can be provided. To be eligible for grants larger than £40,000, applicants must prove that the project will raise the number of R&D jobs in Scotland. Projects are expected to last between 6 and 36 months.

2.7. Conclusion

This chapter began by providing an explanation of why the government attempts to reduce disparities in unemployment across regions using regional policy. It then sought to assess the extent to which capital grants schemes such as RSA and innovation grants schemes are supported by different theories of the firm. It was found that while capital and innovation grant schemes can be justified on the basis of market failure, the latter type of support is better justified as liquidity constraints will be more severe for firms seeking to undertake R&D due to the riskiness of this activity and because innovation grants specifically attempt to encourage an externality generating behaviour rather than merely encouraging firms to locate together in the hope that these firms will generate externalities. It has been difficult to reconcile capital grants schemes with the dynamic capabilities or the evolutionary theories of the firm. The former supports policies that create dynamic capabilities but capital grant schemes merely provide financial support to buy capital which does not constitute a dynamic capability. The evolutionary theory of the firm supports policies to strengthen innovation but capital grants schemes are not designed to do this. Innovation grants schemes are better justified in relation to these theories of the firm as these should help to create dynamic capabilities by creating technological assets and are direct attempts to support the innovation stage.

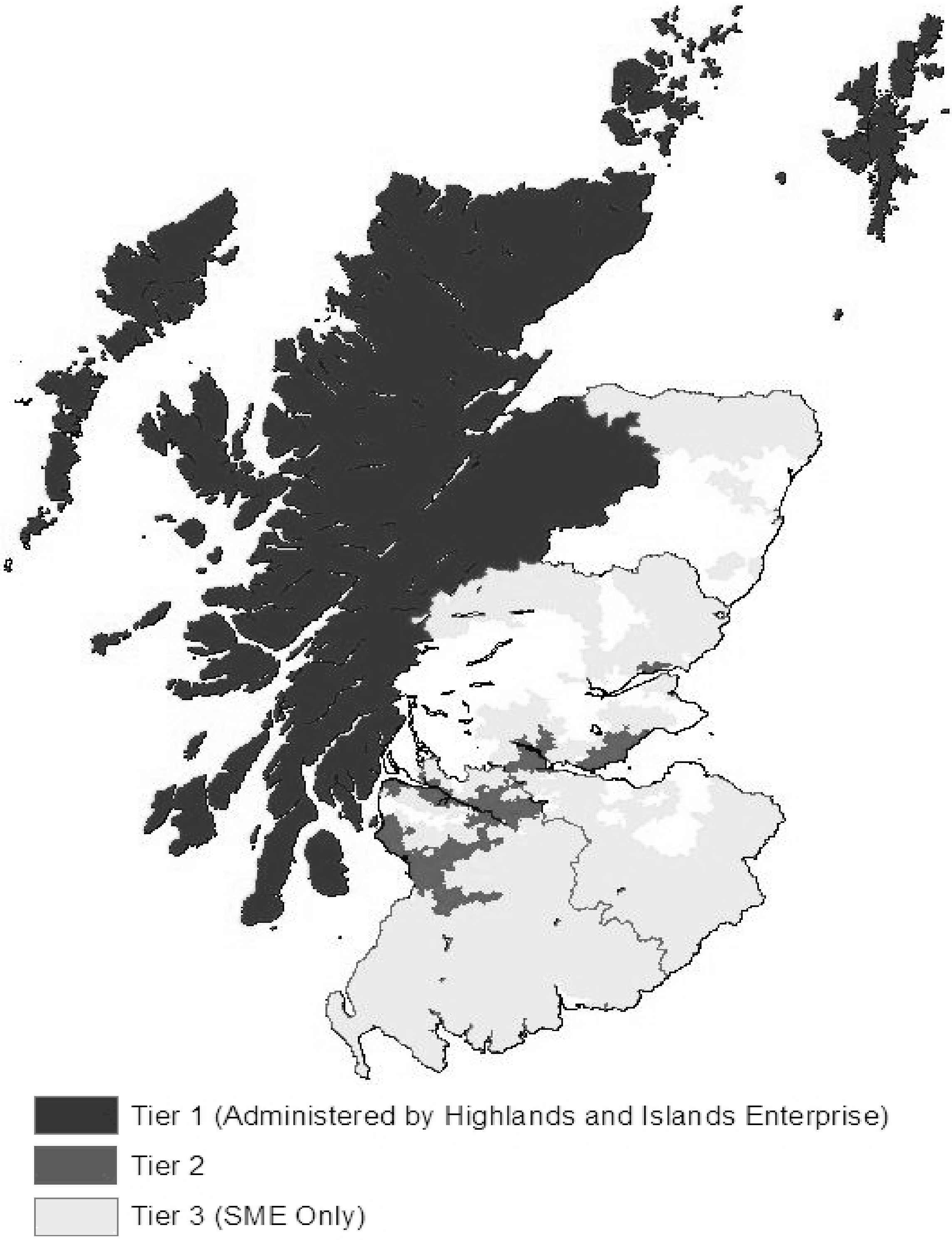
The following section offered a comparison of capital and labour subsidies in an effort to explain the apparent paradox that, although the aim of regional policy is to promote employment in areas of high unemployment, capital subsidies have been more popular than labour subsidies.

The next section provided a description of the evolution of regional policy since its introduction during the depression. The focus was on the grant support schemes introduced during the 1960s and 1970s as this was the time at which the largest sums of money were spent on regional policy and this was also the time at which the RSA scheme was introduced. The development of the RSA scheme was then described in detail. The chapter finished with a description of the grants schemes that currently operate in Scotland. This set out the present eligibility criteria for receipt of an RSA grant. An interesting detail that emerged was that the equivalent scheme to RSA in England now has an eligibility criterion that projects must enhance productivity but no such criterion exists for the RSA scheme in Scotland.

A2.1. Assisted Areas Map

Figure A2.1 shows the current map of assisted areas in Scotland.

Figure A2.1: Assisted Areas in Scotland, 2007-2013



Source: Scottish Government

3. Methodological Literature Review

3.1. Introduction

This chapter will set out the econometric problem of self-selection which arises when estimating the impact of receiving an RSA grant on an outcome variable such as TFP or survival probability. It will then proceed to discuss four different estimation methods that can be employed to overcome this problem. The structure and notation are the same as in the review of the methods available to analyse the impact of education on earnings by Blundell, Dearden and Sianesi (2005).

The parameter of interest is the average effect of treatment on the treated (ATT). Other measures that are discussed in the literature are the average treatment effect (ATE), the average effect of treatment on the non-treated (ATNT), the marginal treatment effect (MTE) and the local average treatment effect (LATE) (see Blundell and Costa Dias, 2008: 10-11 for a more detailed discussion of different estimands). The ATE is the effect of treatment on a randomly assigned group of plants. The ATNT measures what the impact of treatment would be on plants that did not receive treatment. The MTE is the effect of treatment on plants at the margin of receiving treatment (key papers in the development of the MTE are Bjorklund and Moffitt, 1987; Heckman and Vytlacil, 1999). The LATE is the ATE for a specific subpopulation of the treatment group and will be discussed in greater detail later. The ATT is the most relevant when assessing the effectiveness of a voluntary programme which is what will be done in chapters 7 and 8. The other treatment effects become of interest when considering whether or not a programme ought to be extended. Clearly, when returns to treatment are homogenous, all four parameters of interest are identical but, in the presence of heterogeneous returns, the distinction is no longer trivial.

The next section of this chapter will describe the problem of self-selection. The third will discuss the four main estimators that purport to overcome this problem: fixed effects, matching, instrumental variables and control functions. The final section concludes by stating which estimators will be employed in the empirical analyses of chapters 7 and 8 and explaining why these were chosen.

3.2. The Self-Selection Problem

The group of observations that received treatment[5] are said to be self-selected when the decision of whether or not to receive treatment is taken by the plant. In such a situation, the treatment group is not a random sample of the population and will have characteristics that would lead to better or worse performance (in terms of the outcome variable) than observations in the untreated group, in the event that neither group received treatment. This is because the decision to seek treatment will be taken on the basis of an assessment of the benefit that will accrue to the plant from treatment and this benefit will itself be a function of the characteristics of the plant. A comparison of the mean of an outcome variable across the treated and untreated groups will then not provide an unbiased estimate of the ATT because the estimate will be contaminated by the difference in performance between treated and untreated groups that arises due to the differences in characteristics across the two groups that are unrelated to treatment status.[6]

The self-selection problem will now be shown more formally using two approaches: Rubin's potential outcomes approach (developed in a series of papers such as Rubin, 1973; Rubin, 1974; Rubin, 1977) and the standard econometric approach.

Using potential outcomes, the ATT is given by:

$$\beta_{ATT} = E\left[y_{it}^{1} - y_{it}^{0} \mid D_{it} = 1\right] \tag{3.1}$$

where y_{it}^{1} is the outcome variable for plant *i* at time *t* in the event that it received treatment; y_{it}^{0} is the outcome variable for plant *i* at time *t* in the event that it did not receive treatment

[5] In what follows, observations that received treatment are referred to as comprising the treatment or treated group while observations that did not receive treatment comprise the untreated group. The control group is a sub-set of the untreated group that is created for the difference-in-difference and matching estimators. The conditions it must satisfy will be discussed later.

[6] Note that the decision of whether or not a grant application is successful is taken by a governmental body. However, this does not alter the fact that the treated and untreated group will have different characteristics. Indeed, if the government tries to choose 'winners', it will increase the likelihood of treated and untreated groups having different characteristics.

and D_{it} is a dummy variable taking the value of one if plant i receives treatment at time t and zero otherwise.[7]

Equation (3.1) says that the ATT is the difference between the mean of the outcome variable for those observations in the treatment group and the mean of the outcome variable for the same group of observations, had they not received treatment. As the value of the outcome variable for observations that received treatment, in the event that they did not receive treatment, y_{it}^{0}, is unobserved, the problem is one of missing information. y_{it}^{0} is often referred to as the missing counterfactual (see, for example, Blundell and Costa Dias, 2008: 31).

To solve this problem, consider trying to estimate the ATT by using the mean of the outcome variable for observations that did not receive treatment in place of the unobserved mean of the outcome variable for observations in the treatment group, in the event that they did not receive treatment, as follows:

$$\hat{\beta}_{ATT} = E\left[y_{it}^{1} \mid D_{it} = 1\right] - E\left[y_{it}^{0} \mid D_{it} = 0\right]. \tag{3.2}$$

This approach would give an unbiased estimate of the ATT if the following assumption holds:

$$E\left[y_{it}^{0} \mid D_{it} = 1\right] = E\left[y_{it}^{0} \mid D_{it} = 0\right]. \tag{3.3}$$

This assumption simply states that the mean of the outcome variable for those observations that received treatment, in the event that they did not receive treatment, is equal to the mean of the outcome variable in the untreated group.

Equation (3.2) can be rewritten as:

$$\hat{\beta}_{ATT} = E\left[y_{it}^{1} - y_{it}^{0} \mid D_{it} = 1\right] + \left(E\left[y_{it}^{0} \mid D_{it} = 1\right] - E\left[y_{it}^{0} \mid D_{it} = 0\right]\right). \tag{3.4}$$

The first term is the ATT. The second term (after the addition sign) is a bias term that arises when the expected value of the outcome variable for observations in the treatment group, had they not received treatment, differs from the expected value of the outcome variable for observations in the untreated group (see Heckman, Ichimura, Smith and Todd, 1998 for a decomposition of this bias).

[7] In the empirical analysis of later chapters, plants are regarded as receiving treatment from the time at which they receive an RSA grant.

If the treatment group are a random sample from the population of observations, as in properly designed social experiments, this bias term will equal zero as assumption (3.3) will hold (see, for example, Angrist, Imbens and Krueger, 1999). Unfortunately, social experiments are rare in economics primarily due to the expense of conducting them (see, for example, Lalonde, 1986 or Card and Robins, 1998 for evaluations of social experiments). In non-experimental settings such as the analysis of a business support scheme, the bias term will not equal zero because of differences in the characteristics of observations in the treatment and untreated groups. As discussed above, this is the result of treatment status being determined by non-random decisions on behalf of the plant, which are based upon an assessment of the benefit which will arise from treatment. This estimate of the benefit will itself be influenced by the characteristics of the plant and, as a result, the characteristics of the observations in the treated and untreated groups, which determine the outcome variable, will differ.

It is possible to speculate on the direction of the bias from estimating equation (3.2). According to the rules governing the distribution of RSA grants, they should only be provided to plants that cannot receive funding for their project from other sources. It is therefore reasonable to suggest that these plants may possess characteristics that would lead to relatively poor performance, in terms of the outcome variable, had they not received treatment so it would be expected that:

$$E\left[y_{it}^{0} \mid D_{it} = 1\right] < E\left[y_{it}^{0} \mid D_{it} = 0\right]. \tag{3.5}$$

When this holds, the estimate of the ATT will be biased downwards.

The problem of self-selection will now be set out in the usual econometric way. This exposition follows Angrist and Pischke (2009). Consider employing the following simple model, directly analogous to equation (3.2), to estimate the ATT:

$$y_{it} = \alpha + \beta_{ATT} D_{it} + \varepsilon_{it}. \tag{3.6}$$

In this model, y_{it} represents the value of the outcome variable for plant i at time t; α is an intercept term that equals the mean of the outcome variable for plants that did not receive treatment, $E\left[y_{it}^{0} \mid D_{it} = 0\right]$; β_{ATT} is intended to measure the ATT as set out in (3.1) and ε_{it} is an error term.

If the following assumption holds, estimation of equation (3.6) using ordinary least squares (OLS) will give an unbiased estimate of the ATT:

$$Cov(D_{it}, \varepsilon_{it}) = 0. \tag{3.7}$$

This assumption is the analogue of assumption (3.3) and states that there is no correlation between the treatment variable and the error term.

The problem with estimating equation (3.6) is essentially one of omitted variables (see, for example, Heckman (1979); Angrist and Krueger, 2001). Variables that determine both treatment status and the outcome variable are omitted and this generates a correlation between the treatment variable and the error term. It is therefore desirable to include all variables of this type that are observed so that, instead of estimating equation (3.6), the following is estimated:

$$y_{it} = m(X_{it}) + \beta_{ATT} D_{it} + \varepsilon_{it}, \tag{3.8}$$

where $m(X_{it})$ is assumed to be the correct specification for the observed variables, X_{it}, that determine both treatment status and the outcome variable.[8] Assuming all variables with these properties are observed, assumption (3.7) will hold and equation (3.8) will yield an unbiased estimate of the ATT.

However, finding the correct specification for the observed variables, X_{it}, is not easy as there is no way of knowing beforehand what this will be. Generally, a linear specification for the X_{it} variables is assumed. But if this is not the correct specification for the X_{it} variables, this will lead to biased estimates of the ATT as a result of the differences in the distribution of the X_{it} variables across treated and untreated groups that arise due to self-selection into the treatment group. This sensitivity to the specification of the X_{it} variables arises because, for those observations in the treated group for which there is no observation in the untreated group with the same value of the X_{it} variables, the OLS regression depends entirely on the specification of the X_{it} variables for its estimate of what the outcome variable would have been in the event that they did not receive treatment (see Imbens and Wooldridge, 2009). Misspecification will therefore lead to a biased estimate of the ATT.

Another problem arises with the estimation of equation (3.8) if unobservable variables that determine both treatment status and the outcome variable exist. This will mean that

[8] The X_{it} can, of course, also be correlated with the error term and this also generates a bias in the estimate of the ATT. See Frölich (2008) for a discussion of the implications of such a correlation. Assume for now that X_{it} are exogenous.

assumption (3.7) does not hold and equation (3.8) will then fail to provide an unbiased estimate of the ATT.

Implicit in equation (3.8) is the assumption that the impact of treatment on the outcome variable is the same across observations; in other words, that returns to treatment are homogenous. When the impact of treatment on the outcome variable differs across observations, returns to treatment are said to be heterogeneous (Heckman, Smith and Clements, 1997 present evidence showing that heterogeneity in returns to treatment can be empirically important), and equation (3.8) becomes:

$$\begin{gathered} y_{it} = m(X_{it}) + E[b_{it} \mid X_{it}, D_{it}=1]D_{it} + b(X_{it})D_{it} + \varepsilon_{it}, \\ \varepsilon_{it} = \alpha_{it} + (b_{it} - E[b_{it} \mid X_{it}, D_{it}=1])D_{it}, \end{gathered} \tag{3.9}$$

where b_{it} is the observation specific returns to treatment, $b(X_{it})$, consists of interactions between X_{it} and the treatment variable to capture observable heterogeneous returns to treatment and α_{it} is the observation specific error. The ATT is given by:

$$\beta_{ATT} \equiv E[b(X_{it}) + b_{it} \mid D_{it} = 1]. \tag{3.10}$$

Equation (3.10) shows that returns to treatment are heterogeneous as they depend upon X_{it} which vary across plants and time and upon b_{it}, the observation specific returns to treatment.

OLS estimation of equation (3.9) will give unbiased estimates of the ATT when:

$$Cov(D_{it}, \alpha_{it}) = 0. \tag{3.11}$$

This simply states that there is no correlation between the treatment variable and the observation specific error. No assumption need be specified concerning the unobservable returns to treatment as $(b_{it} - E[b_{it} \mid X_{it}, D_{it}=1])D_{it}$ cannot be correlated with the treatment variable. This is only the case when parameter of interest is the ATT. When the parameter of interest is the ATE, MTE or the ATNT, heterogeneous returns to treatment present greater difficulties and further assumptions must be specified for OLS to provide unbiased estimates (Blundell and Costa Dias, 2008 give the further assumptions that are required for estimation of the alternative treatment effects). For the remainder of the chapter, the heterogeneous returns model will only be set out in situations where the estimated parameters have different interpretations under homogeneous and heterogeneous returns.

The next section discusses estimation methods that purport to allow unbiased estimation of the ATT when the treatment group is self-selected.

3.3. Estimation Methods

A number of approaches have been used to estimate the ATT. Here, discussion is limited to the following estimators: fixed effects; matching; instrumental variables and control functions. A review of applications of these methods is provided in chapter 4.4.

The other method that is frequently used in the literature but is not discussed below is the discontinuity design estimator (see Imbens and Lemieux, 2008; Van Der Klaauw, 2008 for a detailed discussion of this estimator). This estimator requires a discontinuous change in treatment probability at a threshold of a continuous variable. As the eligibility rules for receipt of an RSA grant do not create such a discontinuity, this estimator could not be implemented in the empirical analyses and this explains why it is not discussed here.

Difference-in-Difference/Fixed Effects Estimator

The simplest method that is used to estimate the ATT is the difference-in-difference estimator which can be used when a natural experiment is identified. It has a long history having been first used in 1855 by the physician John Snow (1855) to show that cholera was a water-borne rather than an air-borne disease. As will be shown below, it is a special case of the fixed effects estimator. It is estimated by subtracting the difference in the mean of the outcome variable between a start year and an end year for a control group from the difference in the mean of the outcome variable between the start and end year for the treatment group. The validity of the difference-in-difference estimator depends upon the assumption that those plants that received treatment would have performed in exactly the same way as the control group, had they not received treatment. If this does not hold, the estimated treatment effect will be contaminated by the difference in performance which is unrelated to treatment status.

In terms of conditional outcomes, the following assumption must be satisfied (Harris, 2005b):

$$E\left[y_{it}^{0} \mid D_{it}=1\right]-E\left[y_{it'}^{0} \mid D_{it}=1\right]=E\left[y_{it}^{0} \mid D_{it}=0\right]-E\left[y_{it'}^{0} \mid D_{it}=0\right], \tag{3.12}$$

where t is a period of time later than t'. Assumption (3.12) says that the difference between the mean of the outcome variable between period t and t' for the treatment group, in the event that they did not receive treatment, is equal to the difference in the mean of the outcome variable between t and t' for the control group. In other words, the treatment and control groups would have had the same increase or decrease in the outcome variable, had the treatment group not been treated.

Having found a control group that satisfies assumption (3.12), the proxy for the outcome variable for those observations in the treatment group, in the event that they did not receive treatment, is obtained by rearranging assumption (3.12) as follows:

$$E\left[y_{it}^{0} \mid D_{it} = 1\right] = E\left[y_{it'}^{0} \mid D_{it} = 1\right] + E\left[y_{it}^{0} \mid D_{it} = 0\right] - E\left[y_{it'}^{0} \mid D_{it} = 0\right]. \tag{3.13}$$

The proxy is the mean of the outcome variable for the treatment group in the initial period plus the difference between the mean of the outcome variable for the control group at time t and t'.

The ATT can then be estimated using the formula:

$$\beta_{ATT} = \left\{E\left[y_{it}^{0} \mid D_{it} = 1\right] - E\left[y_{it'}^{0} \mid D_{it} = 1\right]\right\} - \left\{E\left[y_{it}^{0} \mid D_{it} = 0\right] - E\left[y_{it'}^{0} \mid D_{it} = 0\right]\right\}. \tag{3.14}$$

It is the difference between the growth of the outcome variable for plants in the treatment group between periods t and t' and the growth of the outcome variable for plants in the control group between periods t and t'.

The difference-in-difference estimate of the ATT can also be obtained from the following econometric model (see, for example, Blundell and MaCurdy, 1998):

$$y_{it} = \beta_{ATT} D_{it} + \eta_i + t_t + \upsilon_{it}, \tag{3.15}$$

where η_i is a time-invariant fixed effect, t_t is a time effect common to each plant in time t, and ε_{it} is an error term.

The econometric analogue to assumption (3.12) is then:

$$Cov\left(D_{it}, \upsilon_i \mid \eta_i, t_t\right) = 0. \tag{3.16}$$

Assumption (3.16) says that there is no correlation between the treatment variable and the error term, having controlled for the time-invariant effects and the time effects.

The difficulty with this estimator is finding a control group that satisfies assumption (3.12). To guarantee that equation (3.14) gives an unbiased estimate of the ATT, the treatment and control group must perform in exactly the same way in terms of the outcome variable, had neither group received treatment, so that the estimated parameter does not capture influences on the outcome variable that are unrelated to treatment. However, observations in the treatment group apply for treatment because they calculate that they will benefit from it sufficiently for it to be worthwhile applying while observations in the control group calculate that the benefit is not sufficiently large for it to be worthwhile applying. It is reasonable to expect that the estimate of the benefit from treatment is a function of the characteristics of the observation. As these characteristics also determine the performance of the observation, the performance of plants in the treatment and control group are expected to differ and this violates assumption (3.12) (see Harris, 2005b).

The general fixed effects estimator does not rely upon such demanding assumptions as the difference-in-difference estimator. When returns to treatment are homogeneous, it can be written as follows:

$$y_{it} = m(X_{it}) + \beta_{ATT} D_{it} + \eta_i + \upsilon_{it}. \tag{3.17}$$

The difference between equations (3.15) and (3.17) is that the former must contain time effects while the latter does not necessarily, although these could easily and often are contained in X_{it}. More importantly, equation (3.17) includes observable variables, X_{it}, that determine both treatment status and the outcome variable. According to assumption (3.12), having controlled for time-invariant or ‘fixed’ effects and time effects, no correlation exists between the error term and the treatment variable and X_{it} does not exist. As this assumption will generally not hold for the reasons set out in the previous paragraph, the fixed effects estimator is more broadly applicable than the difference-in-difference estimator.

It is necessary to give a brief description of how parameters of the fixed effects model are estimated. In both equations (3.15) and (3.17), the treatment variable may be correlated with the time-invariant effects but uncorrelated with the error term as follows:

$$Cov(D_{it}, \eta_i \mid X_{it}) \neq 0, \tag{3.18}$$

$$Cov(D_{it}, \upsilon_{it} \mid X_{it}) = 0. \tag{3.19}$$

Intuitively, the most obvious way to estimate equation (3.17) is by including a dummy variable for each plant to control for these time-invariant effects. However, this creates a large number of parameters and dramatically reduces degrees of freedom. Furthermore, when the number of plants is large and the number of time periods small, there is the problem of incidental parameters (Neyman and Scott, 1948). This is that when the number of time periods for which plants are observed is fixed and the number of plants goes to infinity, the coefficients on the dummy variables are inconsistent because the number of parameters rises as the number of plants rises (see, for example, Baltagi, 2005).

Instead, the time-invariant effects are removed by the within-transformation. Demeaning the variables in equation (3.17) eliminates the time-invariant effects (along with any time-invariant variables) so that estimation of the following model by OLS provides unbiased estimates of the ATT:

$$\begin{aligned} \tilde{y}_{it} &= m\left(\tilde{X}_{it}\right) + \beta_{ATT}\tilde{D}_{it} + \tilde{\upsilon}_{it}, \\ \tilde{y}_{it} &= y_{it} - \bar{y}_i, \\ \tilde{X}_{it} &= X_{it} - \bar{X}_i, \\ \tilde{D}_{it} &= D_{it} - \bar{D}_i, \\ \tilde{\upsilon}_{it} &= \upsilon_{it} - \bar{\upsilon}_i, \end{aligned} \tag{3.20}$$

where a bar is used to denote the mean calculated over time (see, for example, Wooldridge, 2007; Baltagi, 2005 for more detail on the fixed effects estimator). In addition to removing the time-invariant effects, this transformation also removes any variables that are constant over time. In some applications, this may be a considerable disadvantage.

When the treatment variable is a dummy variable taking the value of one when the observation receives treatment and zero otherwise, the same benefit of eliminating correlation between the error term and time-invariant effects can be gained in a simpler way. This is the approach taken by Harris and Robinson (2004) and Harris and Trainor (2005). Rather than eliminating the time-invariant effects through the within-transformation, this approach involves adding a dummy variable that equals one throughout time for plants that receive treatment at any time to equation (3.8) or (3.9). When the treatment variable is a dummy, correlation between the time-invariant effects and the error term is removed because the larger or smaller time-invariant effects for those treatment group observations are controlled for by the dummy variable, and having done this, the treatment dummy, cannot be correlated with the time-invariant effects. But when the treatment variable is continuous, this approach is inadequate to guarantee the removal of correlation between the treatment variable and the time-invariant effects as these effects may be correlated with the ‘amount’ of treatment. The main advantage of this strategy is that it does not necessitate the within-transformation.

The major problem with this approach and the fixed effects estimator is that, while dealing with the problem of correlation between the treatment variable and the time-invariant effects, it requires the error term to be uncorrelated with the treatment variable, as shown by assumption (3.19). If plants tend to receive treatment when they are performing either better or worse than normal due to unobserved characteristics, there will be a correlation between the treatment variable and the error term in (3.20). This is likely if, as with RSA, grants are given to promote or safeguard employment. This is because plants may seek grants when they are performing relatively well due to unobserved factors in order to expand or seek grants to safeguard employment when they are performing relatively poorly due to unobserved factors (Criscuolo, Martin, Overman and Van Reenen, 2007).

Matching Estimator

The early development of the matching estimator owes much to Rubin (see, for example, Rubin, 1973; Rubin, 1979). Essentially, it involves the construction of a control group which is as similar as possible, in terms of observed characteristics, to the treatment group. Differences in the mean of the outcome variable between the treatment and control groups are then attributed to treatment. It can therefore be regarded as an attempt to recreate the conditions of a social experiment. Its main weakness is that it simply assumes that there are no differences in unobserved characteristics that determine treatment status and the outcome variable. If this assumption is not satisfied, the estimates of the ATT obtained using this method will be biased.

The first part of this section will look at the assumptions that must be satisfied for the matching estimator to provide unbiased estimates of the ATT. The second will discuss the various ways of implementing the matching estimator. The final part will discuss its limitations.

The first step involves constructing a control group of observations from amongst the untreated observations such that the selected group of observations is as similar as possible, in terms of observed characteristics, to observations in the treatment group. The conditional independence or unconfoundedness assumption which underpins the matching estimator is given by (Rosenbaum and Rubin, 1983):

$$y^0 \perp D \mid X. \tag{3.21}$$

Assumption (3.21) states that having created y^0 conditional upon X, the distribution of the outcome variable across treatment and control groups, in the absence of treatment, is independent of treatment status. One way of conditioning upon X is to create a matched sample using X. Therefore, assumption (3.21) states that, within the matched sample, the distribution of the outcome variable, in the event that neither group received treatment, is the same for those observations in the treatment group as for those observations in the constructed control group. When this assumption holds, the difference in the mean value of the outcome variable across treatment and control groups is entirely attributable to treatment and so is an unbiased estimate of the ATT.[9] Judgement on whether assumption (3.21) will hold should be made with reference to the availability in the dataset of variables that determine both treatment status and the outcome variable.

It is worth noting that another way of conditioning upon X is to estimate equation (3.8) because if the conditional independence assumption holds, then OLS estimation of equation (3.8) will provide unbiased estimates of the ATT. This is because it has been assumed that equation (3.8) contains the correct specification for those variables that determine both treatment and outcome variable. However, as discussed earlier, it is difficult to find the correct specification for these variables and the wrong specification will lead to biased estimates of the ATT because the estimate of the ATT is sensitive to the choice of specification when the X_{it} variables are not balanced across treatment and untreated groups.

The matching estimator avoids this problem as it balances the X_{it} variables across treatment and control groups by removing from the sample those observations that cannot be well matched to an observation with similar characteristics but a different treatment status. When the X_{it} are balanced satisfactorily, it is no longer necessary to include X_{it} variables in an outcome regression so the problem of specification is sidestepped. As will be discussed later, when the matched sample is not perfectly balanced, it is advisable to include the X_{it} variables in the outcome regression to control for remaining differences in these variables

[9] This is true regardless of whether returns to treatment are homogeneous or heterogeneous.

across treated and untreated groups. The problem of misspecification will then not be as severe as when an unmatched sample is employed because the distribution of X_{it} will be better balanced across treatment and control groups.

It is important to consider whether assumption (3.21) will hold when variables that are determined by treatment status are used to create the matched sample such that:

$$X^0 \neq X^1. \tag{3.22}$$

where X^0 is the X variables in the absence of treatment and X^1 is the X variables when the observation is treated.

Observation *it* would be matched to observation *ju* if:

$$F\left(X_{it}^1 \mid D_{it} = 1\right) = F\left(X_{ju}^0 \mid D_{ju} = 0\right), \tag{3.23}$$

where F denotes any function. However, because X is a function of treatment, X_{it} and X_{ju} would differ, had neither group received treatment:

$$F\left(X_{it}^0 \mid D_{it} = 1\right) \neq F\left(X_{ju}^1 \mid D_{ju} = 1\right). \tag{3.24}$$

The implication of equation (3.24) is that a matched sample created using X will consist of a treatment group that is different from the control group in terms of observed characteristics, if both groups were in the same state of not receiving treatment. As shown by Lechner (2008), because the determinants of the outcome variable are distributed differently, the outcome variable will also be distributed differently across treatment and control groups, if neither group received treatment, which is a violation of assumption (3.21). In practice, this means lagging variables that determine treatment status and the outcome variable and that are influenced by treatment before using them to create the matched sample. In what follows, it is assumed that X only consists of variables that are not determined by treatment status.

A second assumption needed to ensure that the matching estimator is feasible is:[10]

$$P(D_{it} = 1 \mid X) < 1. \tag{3.25}$$

Assumption (3.25) says that X cannot be a perfect predictor of treatment. By implication, this means that for all values of X, there are both treated and untreated observations. This is

[10] Matching assumptions (3.21) and (3.25) are sufficient for estimation of the ATT but not for estimation of the ATNT or MTE. Stronger assumptions that achieve this are given by Heckman and Navarro-Lozano (2004): 32.

necessary because if for some values of X, there were only treated units, it would not be possible to find matches for such plants in the untreated group. When this is the case, matching can only be performed for observations that satisfy assumption (3.25) and the estimated treatment effect must be defined as the ATT for those observations only. Clearly, this is only an issue when the impact of treatment is heterogeneous.

Having created the matched sample, the ATT can be estimated using the following equation:

$$\hat{\beta}_{ATT} = \frac{1}{N^*} \sum y_{it} - \hat{y}_{ju}, \tag{3.26}$$

where N^* is the number of treated observations in the matched sample. Equation (3.26) states that the ATT is estimated simply as the average difference between the outcome variable for observations in the treatment and control group.

The value of the outcome variable for the control group is given by:

$$\hat{y}_{ju} = \sum W_{itju} y_{ju}, \tag{3.27}$$

where W_{itju} is the weight attached to observation y_{ju} in the control group.

There are many ways of generating W_{itju} and hence $\hat{y}_{ju}$. The simplest way is nearest neighbour matching whereby each observation in the treatment group is matched to the single observation in the untreated group with the most similar value of X. The only difference with caliper matching is that the 'caliper' excludes treated observations for which there are no close matches in order to ensure that the treated and control groups are not too dissimilar (see Cochran and Rubin, 1973). These two approaches are examples of one-to-one matching where $W_{ju} = 1$.[11] One-to-many matching can also be performed whereby each treated observation is matched to many untreated group observations with a higher weight assigned to those untreated group observations with more similar values of X (see, for example, kernel based matching as proposed by Heckman, Ichimura and Todd, 1997). A description and comparison of different matching estimators is provided by Zhao (2004).

[11] Note that, when one-to-one matching is employed, equation (3.6) provides the same estimate of the ATT as equation (3.26).

When many variables are being used to create the matched sample, it becomes difficult to find well matched observations. This is the 'curse of dimensionality' (Zhao, 2004: 91). The most popular means of avoiding the problem of dimensionality is to use a balancing score.[12] A balancing score is defined as a function of the observables that guarantees the following condition is satisfied:

$$X \perp D \mid q(X). \tag{3.28}$$

This states that having created a sample that is matched upon the balancing score, $q(X)$, the distribution of the observables that determine both treatment status and the outcome variable is independent of treatment status. The propensity score is a frequently used balancing score. It measures the probability of being in the treatment group given the values of X_{it} and is usually estimated using either a logit or a probit model regression:

$$p(X_{it}) \equiv P(D_{it} = 1 \mid X_{it}). \tag{3.29}$$

Equation (3.29) shows that propensity score is equal to the probability of being in the treatment group given observed characteristics.

Rosenbaum and Rubin have shown that condition (3.28) allows the conditional independence assumption, (3.21), to be rewritten as (Rosenbaum and Rubin, 1983):

$$y^0 \perp D \mid p(X). \tag{3.30}$$

This states that the distribution of the outcome variable in the absence of treatment is independent of treatment status, having matched upon the propensity score. The significance of assumption (3.30) is that the conditional independence assumption holds even when the matching procedure is performed using the scalar variable $p(X)$ instead of using X. The propensity score can then be used in place of X to construct the matched sample (Dehejia and Wahba, 2002 demonstrate the effectiveness of propensity score matching by comparing their results with those obtained in an experimental setting).

The most obvious criticism of matching is the fact that the conditional independence assumption (3.21), which underpins its ability to provide unbiased estimates of the ATT, is very demanding. It requires that every variable that determines both treatment status and the outcome variable is observed. As a result, its plausibility depends crucially upon the richness of the available dataset in relation to such variables (Harris, 2005b). If relatively few of these variables are contained in the dataset, it is unlikely that the conditional

[12] Another means is by employing a metric such as that proposed by Mahalanobis (1936).

independence assumption will hold and the matching estimator will therefore be unable to provide unbiased estimates of the ATT.

One approach often taken in the literature is to create a matched sample and then employ the difference-in-difference estimator discussed above (see Pellegrini and Centra, 2006; Ankarhem, Daunfeldt, Quoreshi and Rudholm, 2009). The idea behind such an approach is that the matching process creates the conditions under which assumption (3.12) will be satisfied. While this approach is preferable to the simple difference-in-difference estimator, a matched treatment and control group may still experience differences in performance (had neither group received treatment) for two reasons. Firstly, the treatment and control group may differ in terms of unobserved characteristics that cause them to perform differently. Secondly, matching on observed characteristics in the period before the treatment group receives treatment does not mean that these observed characteristics are also matched across treatment and control groups in the period in which the treatment group receives treatment. This is therefore another potential source of bias.

A better approach is instead to construct a matched sample and use it to estimate equation (3.8). Unlike the approach just described, such an approach will control for differences in observed characteristics at the time at which treatment occurs. The advantage over estimating equation (3.26) is that estimating equation (3.8) will control for differences in the distribution of variables that determine the treatment and outcome variable across treatment and control groups in the matched sample. While these differences will be much reduced in the matched sample compared to the full sample, depending on the type of matching used and the number of the X variables, such differences may still be significant.[13]

Instrumental Variables Estimator

While the matching method assumes that all variables that determine both treatment status and the outcome variable are observed in the dataset, the instrumental variables provides a means of obtaining consistent estimates of the treatment effect when variables of this type are unobserved. In other words, treatment status is allowed to be determined by

[13] This approach is recommended by Imbens & Wooldridge (2009)

unobserved determinants of the outcome variable. An instrumental variable is correlated with the treatment variable but uncorrelated with the error term. When such a variable can be identified, it can be used to purge the treatment variable of its correlation with the error term and hence provide consistent estimates of the treatment effect.

The first application of instrumental variables was by Wright (1928) in an attempt to overcome the endogeneity that arises when estimating simultaneous equations. The two-stage-least squares method, which allows more efficient estimation when more than one instrument is available, was developed by Theil (1953). Angrist and Krueger (2001) provide a good overview of the origins and the uses of instrumental variables.

Because the nature of the instrumental variables estimator differs for homogeneous and heterogeneous returns to treatment, the discussion will address each in turn.

Suppose that the following model is estimated:

$$y_{it} = m(X_{it}) + \beta_{ATT} D_{it} + \varepsilon_{it}, \tag{3.31}$$

where $m(X_{it})$ is the correct specification for the observed variables that determine both treatment status and the outcome variable so there is no bias from differences in the distribution of observed characteristics across treated and untreated groups.

Assume that differences exist in the distribution of unobservable characteristics that determine the outcome variable such that there is correlation between the treatment variable and the error term:

$$Cov(D_{it}, \varepsilon_{it} \mid X_{it}) \neq 0. \tag{3.32}$$

Under these circumstances, instrumental variables are the standard method of recovering consistent estimates of the ATT.[14] For a variable to qualify as a valid instrumental variable it must be a non-trivial determinant of treatment status and must not determine the outcome variable directly. Formally, an instrument, Z_{it}, must satisfy the following assumptions:

$$Cov(Z_{it}, \varepsilon_{it} \mid X_{it}) = 0, \tag{3.33}$$

$$Cov(Z_{it}, D_{it} \mid X_{it}) \neq 0. \tag{3.34}$$

[14] Instrumental variables estimates are consistent but not unbiased because they involve a ratio of random quantities. It is therefore advisable to use large samples when using this method.

When such an instrument is available, consistent estimates of the ATT can be obtained using two-stage-least-squares estimation. This is performed by substituting the fitted values from a regression of treatment status on variables that satisfy assumptions (3.33) and (3.34) in place of the treatment variable in equation (3.9) (see Angrist and Pischke, 2009: 121-127 for a more detailed discussion of how the two stage least squares estimator is implemented). Intuitively, the instrument removes from the treatment status variable that part which is correlated with the error term. In the case of heterogeneous returns, the situation is more complex.

It should be noted that the error term may take the form of the error term in equation (3.17). An obvious solution is therefore to employ the fixed effects estimator to remove the time-invariant effects, η_i. However, it is assumed here that assumption (3.19) does not hold so that there is also correlation between the treatment dummy and υ_{it}. In this case, the fixed effects estimator is insufficient to provide unbiased estimates of the ATT. Nevertheless, it can be useful in removing the time-invariant effects so that prospective instrumental variables need only be uncorrelated with υ_{it} rather than with both η_i and v_{it}. This combination of fixed effects estimation and instrumental variables is often used (see Criscuolo, Martin, Overman and Van Reenen, 2007 for an example).

Suppose now that there are heterogeneous returns to receiving treatment so that equation (3.31) becomes:

$$\begin{gathered} y_{it} = m(X_{it}) + E[b_{it} \mid X_{it}, D_{it}=1]D_{it} + b(X_{it})D_{it} + \varepsilon_{it}, \\ \varepsilon_{it} = \alpha_{it} + (b_{it} - E[b_{it} \mid X_{it}, D_{it}=1])D_{it}, \end{gathered} \tag{3.35}$$

where $b(X_{it})$ captures observable heterogeneous returns to treatment, b_{it} is the unobserved observation specific return to treatment and α_{it} represents the unobserved no-treatment component. The ATT is given by equation (3.10).

In addition to satisfying assumption (3.34), when returns to treatment are heterogeneous, an instrument must satisfy the following assumption of not being correlated with the observation specific error which is the equivalent of assumption (3.33):

$$Cov(Z_{it}, \alpha_{it} \mid X_{it}) = 0, \tag{3.36}$$

Equation (3.36) is, however, insufficient to recover consistent estimates of the ATT when returns to treatment are heterogeneous due to the existence of the $(b_{it} - E[b_{it} \mid X_{it}, D_{it}=1])D_{it}$ term.

One solution to this problem is to assume the following:

$$Cov(Z_{it}, b_{it} \mid X_{it}, D_{it} = 1) = 0. \tag{3.37}$$

This states that, having conditioned upon X_{it}, Z_{it} is uncorrelated with the unobserved observation specific return to treatment for those observations in the treatment group. If this holds, consistent estimates of the ATT can be obtained. However, since assumption (3.34) demands that the treatment variable is determined by the instrument, assumption (3.37) does not allow plants to be influenced by their returns to treatment when making their choice of treatment status because this would mean that the instrument is correlated with the error term which violates assumption (3.33). This implies either that plants are irrational or that they are ignorant about their unobserved return to treatment. It is therefore an unattractive assumption to make.

Without invoking this assumption, it is not possible to estimate the ATT using instrumental variables estimation. However, Imbens and Angrist (1994) show that it is possible to estimate a LATE (see Angrist and Imbens, 1995 for an application). The LATE is the ATE for those observations that would change their treatment status in response to a change in the value of the instrumental variable.

In the following exposition, assume that the instrument is a dummy variable. The dummy takes the value of 1 for observations for which, because of government policy, it is more attractive to receive a treatment than for other observations, and zero for all other observations. Observations can react in four different ways to the instrument changing in value from zero to one and can be disaggregated into groups accordingly. The first group of observations will be in the treatment group regardless of the value of the instrument and can be called the always-takers. For this group, the change in the value of the instrument makes no difference to their treatment status. By contrast, the second group of observations will be in the untreated group regardless of the value of the instrument. This group cannot be induced to receive treatment by a change in the value of the instrument and are the never-takers. The third group of observations are induced to enter the treatment group by the dummy variable taking the value of one but would not receive treatment otherwise. This group is the compliers. The final group act in a perverse way and leave the treatment group when the dummy is equal to one but are part of it when the dummy equals zero. This final group are the defiers. Define the events:

$$\begin{aligned} E_{1it} &\equiv \{D_{it} \mid Z_{it} = 1\} \\ E_{0it} &\equiv \{D_{it} \mid Z_{it} = 0\}, \end{aligned} \tag{3.38}$$

and, in addition to assumptions (3.34) and (3.36), assume that:

$$\text{either} \left[E_{1it} \geq E_{0it}\right] \text{or} \left[E_{1it} \leq E_{0it}\right] \text{ for all } it. \tag{3.39}$$

This is known as the 'monotonicity' assumption. This requires that the change in treatment status in response to a change in the value of the instrument from zero to one is unidirectional throughout the sample. In other words, if there are some observations that belong in the group of compliers, there are no defiers. This assumption is important as it precludes the possibility that the treatment effect would be positive for all observations but that the size of the groups of compliers and defiers is such that the estimated treatment effect is zero or even negative. Note also that assumption (3.39) guarantees that the instrument actually alters the treatment status of at least some observations and thus strict inequality holds for some *it*.

If (3.39) holds, Imbens and Angrist (1994) show that the following parameter is obtained using two stage least squares:

$$b(X_{it}) + E[b_{it} \mid X_{it}, E_{1it} > E_{0it}] = E\left[y_{it}^{1} - y_{it}^{0} \mid X_{it}, E_{1it} > E_{0it}\right]. \tag{3.40}$$

Equation (3.40) shows that the LATE is the average returns to treatment amongst those observations that are induced to receive the treatment by the change in the value of the instrument. Intuitively, the only group of observations identified above that is observed in both the treated and untreated group is the compliers. The always-takers are never in the untreated group while the never-takers are never in the treated group. The defiers are assumed not to exist. As a result, the data is only informative about the compliers so it is only possible to estimate a treatment effect for this group (see Imbens and Wooldridge, 2009 for a more formal exposition).

So far, it has been assumed that there is only one instrument and that it is continuous. When more than one variable is included in the instrument set, the estimated coefficient is simply a weighted average of the individual LATE coefficients with the weights determined by the size of the effect that each instrument has upon the treatment dummy. When the instrument is a continuous variable, the LATE measures the impact on the outcome variable for those observations that are induced to change their participation status as a result of variation in the instrument within a specified range.

The main difficulty with the instrumental variables approach in general is finding an instrument that satisfies the criterion of being correlated with the treatment variable but which can be legitimately excluded from the outcome equation. Almost all of the variables that have been used as instrumental variables are open to criticism because it is difficult to justify fully the restriction that the instrument does not directly determine the outcome variable. If a variable is used that is correlated with the error term in the outcome equation, the estimated treatment effect can be more biased than the OLS coefficient so it is essential that any instrument truly satisfies the orthogonality assumptions (Angrist and Krueger, 2001).

Furthermore if an instrument is used that is only weakly correlated with the treatment variable, the two stage least squares estimates tend to be centred on the corresponding biased OLS estimate (Bound, Jaeger & Baker, 1995). The instrument must therefore be a strong determinant of whether or not an observation is in the treatment group whilst still being legitimately excluded from the outcome equation.

The fact that observations are likely to experience heterogeneous returns to treatment introduces another layer of complexity. As discussed, under these circumstances, the instrumental variables model estimates the ATE among those observations that are induced to change their participation status by a change in the value of the instrument - the LATE. The LATE estimate will vary depending on which instrument (or instruments) is used. As a result, great care must be taken when interpreting the estimates obtained using instrumental variables.[15]

Control Functions

The control functions approach is a generalisation of the Heckman selection model (Heckman, 1979). It is the most sophisticated means of obtaining unbiased estimates of the ATT when the treatment group is self-selected as it incorporates information from a treatment status model into the outcome variable regression. When there are differences in the distributions of the unobservable variables that determine the outcome variable across treated and untreated groups, there is correlation between the treatment variable and the

[15] The instrument employed in chapter 7 is such that all treated observations are compliers. Therefore, this issue is not as problematic as it is in most applications.

error term. Self-selection bias can therefore be seen as a form of omitted variables bias. The control function approach removes this bias by including additional terms, estimated from a treatment status model, in the outcome regression which removes from the error term that part that is correlated with treated status and so permits consistent estimates of the ATT (see Maddala, 1993: 257-290 for a broader discussion of the control function approach).

Consider the homogeneous returns to treatment model:

$$y_{it} = m(X_{it}) + \beta_{ATT} D_{it} + \varepsilon_{it}, \tag{3.41}$$

Once more, assume that $m(X_{it})$ is the correct specification for the observed variables that determine both treatment status and the outcome variable. However, there are unobserved variables that determine both treatment status and the outcome variable so there is a correlation between the treatment status dummy and the error term.

Treatment status is assumed to be determined by the following binary response model:

$$E_{1it} = 1\{m_D(Z_{it}, X_{it}) + v_{it} \geq 0\}, \tag{3.42}$$

where Z_{it} is a vector of variables that determine treatment status but are not included in the outcome equation and v_{it} is an error term that is assumed to be uncorrelated with both Z_{it} and X_{it}. Equation (3.42) shows that when $v_{it} \geq -m_D(Z_{it}, X_{it})$, the observation receives treatment.

The idea that underpins the control function method is that v_{it} is correlated with the error term in the outcome equation ε_{it}. The intuition for this is that the unobserved component that determines treatment may also have explanatory power for the outcome variable if these unobservable factors that determine treatment status also determine the outcome variable.

Taking expectations of equation (3.41), the following equation is obtained:

$$\begin{aligned} E[y_{it} \mid D_{it}, X_{it}, Z_{it}] = m(X_{it}) + \beta_{ATT} D_{it} + (1 - D_{it}) E[\varepsilon_{it} \mid v_{it} < -m_D(Z_{it}, X_{it})] \\ + D_{it} E[\varepsilon_{it} \mid v_{it} \geq -m_D(Z_{it}, X_{it})]. \end{aligned} \tag{3.43}$$

The terms $E[\varepsilon_{it} \mid v_{it} < -m_D(Z_{it}, X_{it})]$ and $E[\varepsilon_{it} \mid v_{it} \geq -m_D(Z_{it}, X_{it})]$ are the expected values of the error term when the observation is in the untreated and treatment group respectively. This illustrates well the problem of self-selection as when the treated and untreated groups have different characteristics that determine the outcome variable, these terms differ and, as a result, equation (3.41) will not produce consistent estimates of the ATT. When the form of $E[\varepsilon_{it} \mid v_{it} < -m_D(Z_{it}, X_{it})]$ and $E[\varepsilon_{it} \mid v_{it} \geq -m_D(Z_{it}, X_{it})]$ are known, equation (3.43) can be estimated using OLS.

The key assumption in the control function approach is the following:

$$\varepsilon_{it} \perp (D_{it}, Z_{it}) \mid v_{it}. \tag{3.44}$$

This states that the error term in equation (3.41) is uncorrelated with the treatment status dummy, having conditioned upon the error term from equation (3.42). It also states that, having conditioned upon v_{it}, the Z_{it} variables included in the treatment status equation are uncorrelated with ε_{it}. The implication of this is that the impact of the Z_{it} variables on the outcome variables is channelled entirely through the treatment status dummy or is controlled for by v_{it}. It is therefore similar to assumption (3.32) for instrumental variables which states that the entire impact of the instrumental variable is through its impact on the treatment status dummy (Blundell and Costa Dias, 2008). Such an assumption is not necessary for the X_{it} variables included in the treatment status equation as these are also included in the outcome regression and are therefore controlled for.

The conditional means of the error term in equation (3.43) can be written as:

$$\begin{aligned} E[\varepsilon_{it} \mid v_{it} \geq -m_D(Z_{it}, X_{it})] &= r\lambda_{1it}(X_{it}, Z_{it}), \\ E[\varepsilon_{it} \mid v_{it} < -m_D(Z_{it}, X_{it})] &= r\lambda_{0it}(X_{it}, Z_{it}), \end{aligned} \tag{3.45}$$

where λ_{0it} and λ_{1it} are control functions, the form of which are determined by $m_D(Z_{it}, X_{it})$ and the distribution of the error terms in the outcome and treatment equation. Both of these are unknown.

Assuming joint normality of the error terms in equations (3.41) and (3.42) as in the Heckman selection model, the control functions are inverse Mills ratios:

$$\begin{aligned} \lambda_{0it} &\equiv -\frac{\phi\{m_D(Z_{it}, X_{it})\}}{1 - \Phi\{m_D(Z_{it}, X_{it})\}}, \\ \lambda_{1it} &\equiv \frac{\phi\{m_D(Z_{it}, X_{it})\}}{\Phi\{m_D(Z_{it}, X_{it})\}}. \end{aligned} \tag{3.46}$$

where ϕ denotes the standard normal density function and Φ is the standard normal cumulative distribution function. The $m_D(Z_{it}, X_{it})$ terms can be calculated from the fitted values from a logit or probit model of the treatment equation.

These control functions therefore, λ_{0it} and λ_{1it}, allow consistent estimation of the ATT using the following equation:

$$y_{it} = \alpha_0 + m(X_{it}) + \beta_{ATT} D_{it} + r(1 - D_{it})\lambda_{0i} + rD_{it}\lambda_{1it} + \omega_{it}, \tag{3.47}$$

where ω_{it} is an error term. The inclusion of the control functions removes from the error term, ε_{it}, that part that is correlated with the treatment dummy so that ω_{it} is uncorrelated with the treatment variable. Under these assumptions, the coefficient on the control functions in equation (3.47), r, can be written as $r = \sigma_\varepsilon \rho_{\varepsilon v}$, which states that r is equal to the standard deviation of ε_{it} multiplied by the correlation between the error terms in equations (3.41) and (3.42).

An attractive feature of the control function approach is that a test of the statistical significance of the coefficient on the control functions, r, is a test of whether there is potentially any self-selection bias. A positive and statistically significant coefficient on λ_{1it} indicates that observations that receive treatment would perform better due to unobserved characteristics than observations in the untreated group, in the event that they did not receive treatment. A positive and statistically significant coefficient on λ_{0it} indicates that observations in the untreated group would have performed better than treated observations in terms of unobserved characteristics, had the treated observations not been treated.

There are a number of drawbacks to the control function approach (see Puhani, 2000 for a detailed critique of the control function approach). Unlike the instrumental variables estimator, it is not straightforward to control for time-invariant effects in the control function model. While the outcome equation can be estimated as a fixed effects model, the probit model that is estimated to give the inverse Mills ratios cannot be estimated as a fixed effects model due to the incidental parameters problem. As estimating the outcome equation as a fixed effects model but the treatment status equation as a random effects model is inappropriate (Zabel, 1992), other approaches have been developed but these sacrifice some of the simplicity of the control function approach (see, for example, Wooldridge, 1995).

Another problem is that the Heckman selection model, which is the most common version of the control function method, generally requires a variable in the equation determining treatment status that can be legitimately excluded from the outcome equation (an instrument) because the inverse Mills ratio is approximately linear over wide ranges of its argument. Without these variables, there will be collinearity between the regressors and the control function in the outcome equation so the estimated parameters will tend not to be efficiently estimated (Little and Rubin, 1987). As discussed in section 3.2, this instrument will typically be difficult to find.

Also problematic is that the estimated coefficients are sensitive to the assumed distributions of the error terms in both the outcome and treatment equations (Little and Rubin, 1987). Conventionally, normality is assumed (as above) but if this does not hold, the estimated parameters may not be consistent.

Finally, the control function method demands a full specification of the treatment equation while the instrumental variables approach only requires the identification of one variable that determines treatment status but that does not belong in the outcome equation (Blundell and Costa Dias, 2008). In view of this requirement, this approach compares unfavourably to the instrumental variables estimator outlined above.

3.4. Conclusion

This chapter has described the four main methods that are used to estimate the ATT in non-experimental settings. The two that will be employed in the empirical analysis are propensity score matching combined with multivariate regression and the instrumental variables estimator. The fixed effects estimator will not be used because assumption (3.19), which requires that having controlled for the fixed effects and the observable characteristics that determine treatment status and the outcome variable, there is no correlation between the treatment variable and the error term, will not hold if plants tend to apply for grant when they are doing relatively badly or relatively well as this is likely to be the case with RSA. The control function approach will also not be used because it possesses the greatest problem of the instrumental variables estimator - finding an instrument – in addition to the need for a full specification of the treatment equation.

4. Literature Review

4.1. Introduction

This chapter will review the theoretical and empirical literature on the effects of capital subsidies in order to build expectations of what will be found in the empirical analyses of chapters 6, 7 and 8. The part of the chapter which discusses the theoretical papers is far shorter than the part that describes the empirical papers simply because far fewer theoretical papers on the impacts of capital subsidies have been written.

This next section will set out what the theoretical literature suggests may be expected from the empirical analyses of chapters 6, 7 and 8. Chapter 6 decomposes the growth of aggregate productivity to calculate the contribution of RSA-assisted plants and the channels through which this contribution is made. A handful of theoretical models have been developed which provide predictions as to what the impact of schemes such as RSA may be at the macroeconomic level and these will be reviewed. The empirical analyses of chapters 7 and 8 test microeconomic predictions concerning whether receipt of an RSA grant has an impact on TFP and survival respectively. This section will review what impact the literature predicts will be found.

Unlike the theoretical literature, the empirical literature on the impact of government grants upon firm performance is voluminous. Many different methodologies and datasets have been employed and the conclusions regarding their effectiveness are diverse. The third and fourth sections of this chapter will review empirical papers that have analysed business support programmes using macroeconomic and microeconomic data respectively.

4.2. Theoretical Literature Review

This section will begin by reviewing theoretical papers which suggest what the impact of RSA at a macroeconomic level will be. This will provide guidance as to what may be expected from the productivity decomposition of chapter 6. It will then discuss the impacts of receipt of capital subsidies at the microeconomic level. This will be useful in giving a priori expectations for those chapters that investigate the impact of RSA on TFP and on survival probability.

Macroeconomic Impact of RSA

Fuest and Huber (2000) seek to explain why governments tend to use investment rather than employment subsidies in regions with high unemployment. They do so using a model in which firms are heterogeneous in the sense that they have different exogenously determined, random output or productivity shocks.[16] Bargaining between trade unions and firms raises the wage rate which leads to the closure of low productivity firms and a consequent inefficiently low level of employment.[17] There is also an inefficient low level of capital as plants realise that trade unions capture part of the benefit from capital investment through the bargaining process. The impact of an unfunded capital subsidy is that firms demand more capital and profits rise. This rise in profits attracts more firms to enter and this increases aggregate employment. That more firms operate in the market means however that the average level of the productivity shock is lower. However, when a tax is imposed on labour to fund the subsidy, this tax reduces capital investment and employment. However, the benefits of the capital subsidy outweigh the costs of the labour tax so the overall impact is an increase in the capital stock. The number of firms operating is higher than when there is no funded subsidy which means that the average level of productivity is lower.

Restuccia and Rogerson (2008) investigate the impact on aggregate output and TFP of policy induced heterogeneity in the price of inputs faced by heterogeneous plants. This is relevant as the RSA scheme can be regarded as a scheme which reduces the price of capital (chapter 2.6 gives a detailed description of the RSA scheme). They do so using a version of the neoclassical growth model in which half of the plants are subsidised and half are taxed. Again, plants are heterogeneous only in their level of TFP which is constant throughout time. The size of the subsidy is set so that the net effect on steady state capital accumulation of the distortion in prices is zero. The model is calibrated using US data and

[16] The output of firm i is given by: $Y(K_i, L_i) + z ì_i$, where K_i and L_i are capital and labour respectively, z is a positive constant and $ì_i$ is the random output or productivity shock. $Y(K_i, L_i)$ is a production function common to all firms.

[17] Note that no explanation is given for why this should happen in only some regions.

the implications of various policy-induced changes in the price of inputs are then studied. In the case where the probability of being subsidised or taxed is unrelated to productivity, a 50% tax on capital requires a 10% subsidy to keep steady state capital accumulation unchanged and this leads to a fall in aggregate output and TFP of 3% from the state in which there are no price distortions. When lower productivity plants are more likely to receive the subsidy than be taxed, a 50% tax in capital requires a 44% subsidy to keep steady state capital accumulation unchanged and leads to a fall in output and TFP of 11%.

A major problem with these models is that they do not allow firm technology to vary over time. A model that does is by Samaniego (2006). He seeks to understand the quantitative impact of industrial subsidies to failing firms using a general equilibrium model of establishment dynamics. As plants that apply for an RSA grant may do so in order to avoid redundancies, the notion of a failing plant is relevant here.

In his model, the firm's production function takes the Cobb-Douglas form, consisting of an exogenous productivity growth factor, an idiosyncratic productivity shock, capital and labour. The idiosyncratic productivity shock is assumed to follow a random walk. The vintage of technology is embodied in capital rather than directly entering the production function. Entry is costly with the cost an increasing function of the vintage of technology acquired. Entrants draw their initial idiosyncratic productivity shock from a distribution that is distributed entirely to the left of the distribution of productivity shocks for existing firms.[18] As firms age, they fall behind the technological frontier. In each period existing firms have the option of upgrading their vintage of technology at a cost (with the cost an increasing function of the vintage of technology purchased), falling further behind the technological frontier or closing. Firms close when their continuation value is less than their random draw from a cumulatively distributed continuation shock. The provision of a subsidy will cause a firm to stay open that would otherwise close if the value of the subsidy is greater than the gap between the continuation value of the firm and the continuation shock. Subsidies are funded by a tax on firm profits and a balanced budget condition is assumed to apply. Timing in the model is as follows: at the start of the period, plants draw a continuation shock and choose whether or not to stay in operation; assuming they choose to remain in operation, they decide whether or not to update their vintage of

[18] The significance of this specification of the productivity shocks is that it ensures that entrants are smaller than existing plants.

technology; finally, they observe the value of their idiosyncratic productivity shock and produce output.

Simulations calibrated using US data show that, for empirically reasonable levels of subsidy, the introduction of a subsidy leads to lower employment and lower consumption. This is the result of a fall in aggregate labour productivity. However, firm productivity actually rises because of two factors. Firstly, subsidies allow firms to reach a point where upgrading their vintage of technology becomes the optimal choice. Secondly, the average idiosyncratic shock is higher when subsidies are provided because the proportion of new plants in the economy falls and these have lower levels of the idiosyncratic shock. The fall in aggregate labour productivity is the result of a rise in the average plant size. This, in turn, is the consequence of fewer plants entering because of the tax on profits used to fund the subsidy because entrants are typically small.

Microeconomic Impact of RSA

The impact of receipt of a capital grant on the employment and capital stock of the firm has been discussed in chapter 2.4. This showed that the impact of receipt of a capital grant on the capital stock will be positive while the impact on employment depends upon whether the induced substitution effect is outweighed by the output effect. If the latter predominates, there will be a positive impact on the employment of the recipient firm. This is more likely to happen when the capital that is bought with the subsidy is relatively modern compared to the existing stock of capital because of the larger reduction in costs generated by the larger amount of new capital bought by the firm. Given that the primary purpose of the RSA scheme is to create and safeguard employment and that the scheme is discretionary, grants should only be provided in cases where the output effect is larger than the substitution effect. Furthermore, the RSA scheme has a ‘clawback’ clause that requires that the grant is repaid if agreed employment targets are not met. It is therefore expected that receipt of an RSA grant will have a positive impact on employment.

The relationship between receipt of an RSA grant and TFP is less clear. Although many papers test the prediction that business support programmes have an impact on TFP, they tend not to provide a clear explanation of the channels through which such an impact may be expected to occur (see, for example, Girma, Görg and Strobl, 2007a; Harris and

Robinson, 2004). The clearest explanation of how such an impact may arise comes from Harris (1991a). He argues that capital subsidies such as the RSA scheme are expected to have a positive impact on TFP through two main channels. The first is through replacing older capital with more modern capital which requires the plant to reorganise production along more efficient lines. This implies that the impact will be greater when the capital that is being replaced is older. The second is through net investment which allows the plant to create new products that can be produced with greater efficiency than older products.[19]

The prediction from the literature concerning the impact of government grants on survival is relatively straightforward. It is perhaps for this reason that, once again, papers that investigate this issue tend not to spell out precisely why it may be expected that government grants are expected to have an impact on survival (see, for example, Girma, Görg and Strobl, 2007b; Harris and Trainor, 2007). The decision of a firm to close depends fundamentally on expectations of future profits and the liquidation value of the firm. When the discounted expected profits over future periods are less than the liquidation value of the firm, the firm will optimally choose to cease production (see, for example, Jovanovic, 1982; Hopenhayn, 1992). As shown by Samaniego (2006), a subsidy increases discounted expected profits so that plants, that would otherwise close, choose to remain in operation.

4.3. Empirical Papers Using Macroeconomic Data

This section will discuss some of the empirical evaluations of business support programmes that have been conducted using macroeconomic data. The first part will describe the only previous productivity growth decomposition that has analysed the contribution of plants that have been supported by a business support programme. The second part will discuss two papers that have used shift-share analysis; the third section will review one paper that employed simulation and the final section will discuss an application of multivariate regression.

[19] It should be noted that the paper by Samaniego (2006) does not suggest a causal relationship between receipt of a grant and productivity. In his model, a grant allows a plant to avoid closure and then, at a specified time, it becomes optimal for the plant to update its technology.

Productivity Decomposition

The only paper that has examined the contribution of plants that received business support grants to aggregate productivity growth is by Harris and Robinson (2005).[20] Using a dataset created by merging SAMIS into the ARD, they employ the Haltiwanger method to decompose the growth of labour productivity and TFP between 1990 and 1998 in UK manufacturing plants. This allows them to identify the share of the growth of aggregate productivity attributable to plants that received an RSA grant and the channels through which this contribution is made. Their results show that plants that received an RSA grant made a large positive contribution to the growth of labour productivity but a negative contribution to the growth of TFP. The contribution to aggregate labour productivity comes primarily from RSA grant recipients that improve their productivity between 1990 and 1998 also increasing their market share while the negative contribution to the growth of aggregate TFP is mostly due to RSA grant recipients with low TFP in 1990 increasing their market share. While the contribution from entry and exit is large for plants that did not receive a grant, it is far smaller from entering and exiting plants that received an RSA grant.[21]

This paper reveals the proportion of aggregate productivity growth accounted for by plants that received support. It does not attempt to estimate the causal impact of a business support programme at the macroeconomic level. It therefore does not sit very comfortably alongside the rest of the papers in this section which attempt to estimate a causal impact of a business support programme.

---

[20] It is arguable that this paper belongs more properly under the heading of 'Empirical Papers using Microeconomic Data' as the aggregates used in the productivity decomposition are calculated from the microeconomic data of the ARD. However, as the decomposition is performed using aggregate data, it is placed in this section.

[21] This is the paper, referred to in chapter 1.3, which gives results from an analysis of the impact of receipt of an RSA grant on survival. However, only a paragraph is devoted to this so it is impossible to critically review this aspect of the paper in any depth. Instead, this is done briefly in chapter 8.1.

Shift-Share Analysis

In a seminal paper, Moore and Rhodes (1973) employ shift-share analysis to estimate the impact on manufacturing employment and investment of the moves made between 1960 and 1963 towards a more active regional policy in the UK (see chapter 2.5 for a history of regional policy in the UK). They address the problem of the missing counterfactual by constructing a series which purports to show what employment and investment would have been after 1963 in the Development Areas had there been no change in policy. This is achieved by applying the UK industry growth rates to the employment and investment levels of each industry in the Development Areas. The calculated figures are then summed to yield aggregate 'expected' estimates for employment and investment in the Development Areas for each year. These are then compared to the actual figures to give an estimate of the impact of policy. This approach is supposed to isolate the effects of policy on the outcome variable by controlling for the impact of industrial structure. Their estimates are that in 1971, employment was 12% higher in the Development Areas than it would have been had there been no move towards a more active regional policy. For Scotland, Wales and Northern Ireland, actual investment was found to be 30% higher in 1970 than expected investment.

In essence, this approach relies upon the same assumption that underpins the difference-in-difference estimator. What is being assumed is that the growth rate of each industry in the development areas would have been the same as the growth rate of the same industry outside the development areas, had those industries in the development areas not benefited from a more active regional policy. However, it may be hypothesised that those industries in the development areas would have performed differently without the more active regional policy than those industries outside the development area. This would be the case if the decision to introduce a more active regional policy is taken due to the characteristics of the industries in the development areas and these characteristics make them prone to relatively poor performance. In this case, the difference-in-difference assumption would not hold and the estimates obtained would be a biased estimate of the impact of the more active regional policy.

A more advanced version of the shift-share estimator is employed by Canning, Moore and Rhodes (1987). This version attempts to control for the fact that those industries that benefited from the change in policy may perform differently to those that did not benefit, had there been no change in policy, by making use of data that is available from before the change in policy to calculate the difference in trend growth rates across industries inside and outside the development areas. More specifically, the difference between actual and expected values of the variable of interest, calculated as above, is regressed on time using data from before the change in policy. The coefficient on time is then used to extrapolate this difference forward from the time of the change in policy. The gap between the difference between the actual and expected values and the extrapolated values of this difference is the estimate of the policy effect. Using this methodology, Canning, Moore and Rhodes find that between 1959 and 1971, regional policy in Northern Ireland created an extra 33,000 manufacturing jobs.

However, the estimates rely upon the assumption that the difference between the actual and expected values of the series would have continued to grow at the same rate over time, had there been no change in policy. This may not be the case if, for instance, the industries that were to benefit from the change in policy would have performed relatively worse than in previous periods if there was no change in policy. This would be likely if the government had changed policy because they had anticipated such relatively poor performance in the development areas and wanted to prevent it.

Furthermore, estimates obtained using this method are sensitive to the time period used to estimate the coefficient on the time trend which is then used to extrapolate the difference between actual and expected values of the variable of interest. This is a serious problem as it is impossible to know which time period will provide the most similar trend in the difference between the actual and expected series as that which would be observed after the change in policy, had there been no change in policy. Therefore, while this version of the shift-share estimator is superior to the basic version used in Moore and Rhodes (1973), it remains open to criticism.

Multivariate Regression

Beason and Weinstein (1996) use a panel of Japanese sectoral data that starts in 1955 and ends in 1990 to consider which industries benefit from government assistance in the form of tax relief, low interest rate loans, subsidies, tariffs and import quotas. They find that the targeted industries tended to be low-growth and have decreasing returns to scale. They then look at the impact of this assistance. Using fixed effects estimation, they find that these measures increased investment and growth in the targeted industries but they cannot find strong evidence of a major impact on TFP growth.[22]

The econometrics of this paper can be criticised on the grounds that, even though fixed effects are included in the model, there may be a correlation between the treatment variable and the error term that will bias the estimate of the treatment effect. This would arise if industries tended to receive more assistance in years when they are performing better or worse than average because of unobserved variables.

A similar paper by Lee (1996) uses a panel of sectoral data to investigate the impact of industrial and trade policy on GVA, capital growth and TFP in South Korea between 1963 and 1983. His empirical model for output is derived from the neoclassical growth model, allows for fixed and time effects and is estimated using both weighted least squares to correct for cross-equation heteroskedasticity and three-stage-least-squares to control for the possible endogeneity of the treatment variables. His results show that tax incentives had a positive and statistically significant impact on GVA and capital growth but no effect on TFP while low interest bank loans had no impact on any of the dependent variables considered. Trade protection was found to reduce the growth rates of both labour productivity and TFP.

Although each equation is estimated by three-stage-least-squares, and the possibility of correlation between the error term and the policy variables is thereby acknowledged, the way in which it was implemented does not guarantee an unbiased estimate of the impact of policy. This is because the instruments employed are the once-lagged policy variables. If the error term is autoregressive, these instruments are invalid as they will be correlated with the error term (Bond, 2002). It is therefore better to adopt a dynamic version of the equation so that the error term is serially uncorrelated and lags may be used as instruments.

Another criticism that can be levelled at this approach is that no effort is made to understand the interrelationships between the parameter estimates obtained using different

[22] Some estimates suggest that sectors that received a high proportion of low interest loans enjoyed higher TFP growth rates, ceteris paribus, but this effect is not found to be robust.

dependent variables. For instance, the estimate of the impact of policy on GVA growth will clearly be a function of the estimate of the impact of policy on the growth of both capital and TFP. On the other hand, the impact on the growth of capital will be partly determined by the impact on GVA as the demand for capital is a derived demand. Failure to acknowledge these interrelationships creates difficulties in understanding the channels through which policy has an impact.

Simulation

A more sophisticated means of policy evaluation that avoids this criticism is to estimate a structural model and use this to estimate the impact of policy. This is the method employed by Harris (1991b) to investigate the impact of automatic capital subsidies on employment in the Northern Irish manufacturing sector. His model incorporates an industry production equation, an industry demand equation and factor demand equations. The parameters of these equations are estimated by full-information maximum likelihood using data for the period from 1950 to 1983. The parameterised model is then used to generate estimates of what output, labour and capital would have been, had automatic capital grants not been provided. The results show that had the automatic capital grants been unavailable, output would have been around 3.9% lower, the capital stock would have been smaller by almost 23.8% and employment would have been higher by 26.1%. This shows that automatic capital grants created a large substitution of labour for capital and that this substitution effect outweighed the output effect so that the overall impact of the grants on employment was negative.

The use of industry level data in this paper was not problematic because the support provided was automatic. However, many papers use industry level data to analyse the impact of discretionary business support programmes and this is not the best unit of observation to use in the analysis of the impact of such programmes. This is because industry level data precludes the comparison of treatment and control groups as there is no such thing as a treated industry because some plants or firms within each industry will not have received treatment and there is also unlikely to be an untreated industry, as some plants or firms within each industry will generally have received treatment. To find genuine treatment and control groups when the programme is discretionary, it is necessary to use plant or firm level data. Such data permits the direct comparison of an outcome variable across treated and control groups and thereby facilitates the estimation of treatment effects.

4.4. Empirical Papers Using Microeconomic Data

Most of the recent papers on the impact of grants on firm performance have used microeconomic data which allows the construction of treatment and control groups in line with the approach advocated in the evaluation literature (see, for example, Blundell and Costa Dias, 2008). Such papers will be reviewed in this section. They are grouped according to the method employed to deal with self-selection and this section therefore links into chapter 3.3 where the methods available to control for self-selection are discussed.

Difference-in-Difference

One of the most popular approaches to evaluating business support programmes is by employing the difference-in-difference estimator. This is the approach taken by Hart and Scott (1994) who, as part of a broad analysis of the effectiveness of small firm policy in Northern Ireland, investigate the impact of receiving a SFA grant between 1984-5 and 1988-9 on employment growth in small manufacturing firms between 1986 and 1990. SFA is Northern Ireland’s equivalent of the RSA scheme. They offer four alternative control groups: small manufacturing firms that did not receive support in Northern Ireland, small manufacturing firms in Leicestershire, small manufacturing firms in Wearside, and small manufacturing firms in the Republic of Ireland. Results are reported for all four although, as acknowledged by the authors, the last two are not appropriate control groups because government assistance was also available in Wearside and in the Republic of Ireland and those plants that received such assistance could not be identified. They find that employment growth in SFA assisted firms was 19.1% higher than in non-assisted Northern Irish firms and 22.3% higher than in small firms in Leicestershire.

The most obvious problem with these estimates is the nature of the assumption required to obtain unbiased estimates of the treatment effect. This is that grant recipients in Northern Ireland would have experienced exactly the same employment growth as non-assisted Northern Irish firms or firms in Leicestershire, had they not received a grant. With respect to the former, as Northern Irish grant recipients are a self-selected group of the population, they are likely to differ from Northern Irish firms that did not receive treatment in ways that effect their employment growth. With respect to the latter, not only are firms in Northern Ireland likely to have different characteristics from plants in Leicestershire that determine employment growth, macroeconomic conditions may also differ between the two areas and this will also lead to differences in their rates of employment growth. As a result, this application of the difference-in-difference estimator is not convincing.

The difference-in-difference estimator is also used by Bronzini and de Blasio (2006) who provide an evaluation of the impact on investment of capital grants provided under Italy's Law 488/1992. These grants, intended to reduce regional inequalities in income, are awarded to manufacturing and extractive firms through auctions where applicant firms score points in relation to criteria such as the number of jobs that will be created by the project. As a control group they use those firms that applied for assistance but were rejected. Results show that in the year after receiving the first instalment (of three) recipients increase their investment levels relative to the control group. However, in the years after the last instalment is received, investment levels are lower for the treatment group which suggests that firms may have intertemporally substituted their investment in response to the grant. To test the robustness of this result to the control group they firstly use as an alternative control group firms which received scores in the auction process close to successful firms and secondly firms with investment profiles similar to treated firms prior to the provision of grants. Neither approach significantly changes the results.

In their paper, the assumption that has to hold for the main estimates to be unbiased is that firms that received a grant would have experienced exactly the same growth in investment as those plants that applied for assistance but were rejected, had they not received a grant. As the firms that did not receive treatment did not do so because they did not score sufficiently well in the auction, it is reasonable to expect that they are different to those plants in terms of characteristics. If these characteristics affect their investment growth, this violates the assumption outlined above. Although it is commendable that this paper experiments with other control groups, these can be similarly criticised.

Multivariate OLS

The difference-in-difference estimator can be performed with a regression of the first-differenced outcome variable on a treatment dummy variable using OLS. However, if there are observed variables that are likely to be correlated with the treatment variable and the error term, these should be included in the regression. This is what is done by Bergström (2000) in his analysis of the impact on TFP of Swedish selective capital subsidies. His dataset comprises firms that received a grant in 1989 and a control group randomly drawn from the population of Swedish firms. He begins by using a logit model to show that firms that received a subsidy in 1989 tended to have been situated in a support area, to have been younger and to have had lower labour productivity than those that did not receive a subsidy. In the main regression, he uses the growth of output between 1989 and dates ranging from 1990 to 1993 as the dependent variable. The treatment variable is the level of subsidy received by the firm in 1989. The equation is estimated by OLS and a bounded influence estimation technique which minimises the influence of outliers (see Maddala, 2001: 476-479 for an introduction to bounded influence estimation). Results obtained by selecting 1990 as the terminal date showed that subsidies raised TFP growth in the year after the subsidies were provided (although the coefficient is not statistically significant using the bounded influence estimator). However using 1991 to 1993 as terminal dates gave a negative, and often statistically significant, coefficient on the subsidy variable suggesting that subsidies may actually reduce TFP growth in the longer term.

This approach is an improvement on the simple difference-in-difference estimator as it controls for differences in observed characteristics across treatment and control groups. Omitting these variables would create a correlation between the treatment variable and the error term. However, the econometrics may still be unsatisfactory as no attempt is made to control for possible correlation between the variable measuring the size of the grant awarded and the error term arising from differences in unobserved characteristics across treatment and control groups (see Chapter 3.3 for more discussion of the shortcomings of the fixed effects estimator). As firms that receive assistance are a self-selected group, they may have unobserved characteristics that allow them to grow faster or slower than the control group and this will generate a bias in the estimated coefficient.

A further problem with this paper is that correlation between the error term and the growth of factor inputs is ignored. As discussed in chapter 7.3, this arises when firms have some knowledge about the realisation of the error term and use this information to make their choice of the growth of inputs. As discussed by Frölich (2008), failing to account for the endogeneity of control variables can bias the estimate of the treatment effect.

Fixed effects

A paper of particular relevance to this thesis is by Harris and Robinson (2004). As in chapter 7, they use a plant-level panel dataset, created by linking SAMIS with the ARD, to estimate the impact of receiving an RSA grant on TFP. Their model is also an augmented log-linear Cobb-Douglas production function in which they deal with the endogeneity of capital, employment and intermediate inputs using the system GMM estimator (see appendix A7.1 for a discussion of the system GMM estimator). Unlike in this thesis, they also look at the impact of the SMART and SPUR schemes which are discussed in chapter 2.6. In addition to a number of control variables, the model contains four dummy variables. The first is a dummy that takes the value of one in every period if a plant was assisted at any time by RSA and the second is the equivalent for the SMART and SPUR schemes. These dummies allow the intercepts to vary across treatment and control groups and therefore show whether RSA or SMART/SPUR recipients had high or low TFP levels prior to receiving support. The other two dummies are equal to one from the time that a plant receives an RSA or SMART/SPUR grant and their coefficients are intended to estimates the treatment effect. When the control group is drawn from plants throughout the whole of Britain, the results indicate that RSA recipients had initially lower levels but that SMART/SPUR recipients had higher levels of TFP. In the case of RSA, receiving support was found to increase TFP by 2.5% but receiving a SMART/SPUR grant led to no statistically significant change in TFP.

One criticism of this paper is that a larger proportion of the plants that received treatment were not linked to the ARD than in the dataset to be used in the empirical chapters here. As discussed in chapter 5.3, failure to identify RSA recipients can potentially generate biased estimates of the treatment effect.

A similar model is employed by Harris and Trainor (2005) to investigate the impact of SFA in Northern Ireland on TFP. Their dataset is created by linking the ARD with a dataset constructed from the records of the Industrial Development Board. Again, a GMM estimator is used to control for the endogeneity of factor inputs. Their model is more sophisticated than that estimated in Harris and Robinson (2004). It allows for the possibility that SFA grant recipients may differ from plants that do not receive a grant in their technologies by using interaction variables. Furthermore, they separate SFA into capital grants and other forms of grant assistance, test if a 1990 change in industrial policy changed the impact of either on TFP and investigate whether their impact differs according to the location of the plant owners. The model is also estimated by industry because of 'a strong a priori assumption that industries are likely to differ in terms of their underlying production functions, product life cycles, and thus the potential impact of SFA' (Harris and Trainor, 2005: 65). This approach is also taken in chapter 7. The results show that capital grants were more likely than other types of grant support to increase TFP and that the 1990 change in industrial policy towards improving competitiveness rather than promoting employment did lead to a greater impact of SFA on TFP.

This approach of both these papers is a specific case of the fixed effects estimator discussed in chapter 3.3. Instead of including a dummy variable for each plant, only a single dummy variable is included which equals one throughout time for plants that receive assistance. The problem with this approach is the same as that of the fixed effects in that it fails to deal with the potential existence of correlation between the error term and the treatment variable. This dummy variable will only control for self-selection bias if the mean of the error term is the same for plants that receive treatment before and after they receive treatment. If, for instance, plants tend to be performing better or worse than normal due to some unobserved characteristic and this leads to them applying for assistance, a correlation will exist between the treatment variable and the error term that is not controlled for using this method. In this way, this method is not an adequate control for self-selection bias.

A standard application of the fixed effects estimator is provided by Kangasharju and Venetoklis (2002). Their paper aims to calculate the effect of various types of subsidy provided by the Finnish government on firm employment between 1995 and 1998. They estimate a fixed effects panel model with control variables and time dummies in an attempt to mitigate the problem of self-selection bias. The key dependent variable, 'own payroll', is calculated by subtracting the value of the subsidy from the firm's payroll in order to avoid a trivial relationship between the value of the subsidy received and the dependent variable. Allowing subsidies to have an impact on employment in the year in which they are provided and in the following year, results show that receiving employment subsidies led to an average increase of 11% in 'own payroll'. Further analysis shows that an extra Euro of subsidies generates only an extra 34 cents in 'own payroll'. This suggests that subsidies are displacing employment as on average firms cover 60% of the costs of a subsidised job.[23] At the 95% level, Investment and Operation subsidies and R&D subsidies are not found to have a statistically significant impact on this measure.

The fixed effects estimator in its standard form possesses no advantage in terms of overcoming self-selection bias over the approach outlined in the Harris and Robinson (2004) and Harris and Trainor (2005). It is therefore subject to the same criticisms that it does not control for bias that arises when firms receive grants in years in which they are performing unusually well or badly due to some unobserved factor.

A criticism that is more specific to this paper is that only allowing the subsidies to have an impact in the year in which they are provided and the following year is restrictive and will lead to underestimates of the treatment effect if some of the impact occurs after this period. Another concern is that, unlike in Harris and Robinson (2004) and Harris and Trainor (2005), no attempt is made to tackle the endogeneity of other covariates in the model. Sales is included in the model and, as owners choose their level of output and factor inputs simultaneously, this will be correlated with the error term and this may lead to biased estimates of the impact of treatment (Gorter, Hassink, Nijkamp and Pels, 1997).

Matching and Multivariate Regression

Perhaps the most intuitive way of evaluating a business support programme with non-experimental data is with the matching estimator. Girma, Görg and Strobl (2007b) use this approach to investigate the impact of grants on plant survival using a Cox proportional hazard model. Their dataset is constructed by linking three datasets collected by Forfás, the

[23] For there to be no displacement, one Euro must lead to an increase in payroll of at least 1.5 Euros if firms cover 60% of the subsidised job.

agency responsible for industrial development, science and technology in the Republic of Ireland. The control group is constructed using propensity score matching. Their results indicate that receiving a government grant lowered the probability of closing. It is noted that the closure of a domestic plant is more likely to mean the closure of the entire enterprise than the exit of a foreign-owned plant as in the latter case the enterprise may only be shifting production to another country. With this difference in mind, tests for a differential effect on closure probability of grants received by foreign-owned plants were run but no significant difference in closure probability was found. The authors therefore conclude that the evidence that grants are successful at maintaining foreign investment is not as strong as that showing their effectiveness on domestic firms.

A strange feature of this paper is that the probit model which generates the propensity scores, which are then used to create the matched sample, contains variables that do not feature in the hazard model while the hazard model contains variables that are omitted from the probit model. If the strategy of estimating a multivariate outcome equation rather than one that only includes the treatment variable is taken, all variables that determine both treatment status and the outcome variable should be included in both the probit model and the hazard model. This is because, on the one hand, variables included in the probit but not the hazard model will lead to a relatively poorly matched sample in terms of those variables that are included in the hazard model because the sample will be matched on these variables in addition to those included in the hazard model. On the other hand, those variables included in the hazard model but not included the probit model will not be as well balanced as they would have been if they had been included in the probit model. This strategy will therefore create an unnecessarily poor balance of the observed characteristics across the treatment and control groups in the outcome regression. As creating a well balanced sample is the reason for creating a matched sample, this approach is problematic.

A more general point is the perennial issue with the matching estimator: whether there are any unobserved determinants of both the treatment and outcome variable that will generate a correlation between the treatment variable and the outcome variable. If such determinants exist, the estimated treatment effect will be biased.

Matching and Difference-in-Difference

Ankarhem, Daunfeldt, Quoreshi and Rudholm (2009) investigate whether Regional Investment Grants given to Swedish firms between 1990 and 1999 had a positive impact on firm performance, as measured by employment and returns to equity growth. Using a large panel dataset containing all limited firms in the two support areas of Sweden, they use the difference-in-difference estimator on a sample of firms created using propensity score matching. As it is difficult to say, a priori, how long it will take for the grant to impact upon the dependent variable, estimates were calculated over one, three and five year intervals following receipt of the grant. The authors also allow for a differential impact of the grants on firms of different sizes by calculating estimates using firms which have turnover of less than Swedish Krona (SEK) 100,000, firms with turnover greater than SEK 100,000 but less than 1 million and firms with turnover that is greater than SEK 1 million. Estimation is performed by year. No statistically significant impact of Regional Investment Grants on returns to equity was found for any of estimations performed. A positive and statistically significant impact is found for a few of the employment estimates but the vast majority show no significant effect.

As in Bronzini and de Blasio (2006), the effects of Italy's Law 488/1992 are the subject of a paper by Pellegrini and Centra (2006). They create a matched sample, using both kernel and nearest neighbour matching, to create a control group drawn from those firms that applied but were rejected for a grant and then apply the difference-in-difference estimator to control for time invariant differences in performance across treatment and control groups. Using either matching procedure they find that grant recipients experienced significantly higher rates of growth in turnover, employment and fixed assets than firms in the control group which is unsurprising given the nature of the grants provided. Importantly though, labour productivity grew slower in the treatment group by a statistically significant amount which casts doubt upon the sustainability of the jobs created by the scheme.

While the combination of matching and difference-in-difference represents an improvement on the simple difference-in-difference method, its validity can still be questioned. The fact that treatment and control groups have been matched on observed characteristics does not guarantee that they are well matched on unobserved characteristics that determine treatment status and the growth in the outcome variable. When such characteristics exist, the difference-in-difference assumption will not hold. Furthermore, even if the observed covariates are perfectly matched in the start period, differences may appear by the end period that are unrelated to treatment and which determine the value of the outcome variable. These differences ought to be controlled for and this is the motivation for using the matched sample to estimate a multivariate regression rather than using the difference-in-difference estimator.

Instrumental Variables

The standard econometric means of dealing with self-selection is instrumental variables estimation. Girma, Görg and Strobl (2007a) choose this method in their analysis of the effect of government grants on TFP. Their dataset is created by linking data from the Irish Economy Expenditure survey, an annual survey of manufacturing plants in the Republic of Ireland, with data collected by the Industrial Development Authority on grant payments.[24] TFP is regressed on grant levels and a number of control variables using the system GMM estimator. When the grant levels variable is constructed using all grants available from the IDA, there is found to be no statistically significant impact on TFP. However when the grant level variable is disaggregated into a variable that includes grants that are likely to enhance productivity and another containing all other grants, a positive and statistically significant effect is found for the former. To test the hypothesis that financially constrained plants benefit most from grants, they augment their model with interactions between the disaggregated grant level variables and debt to equity ratios and find that, for productivity enhancing grants, the hypothesis holds.

The major problem with this paper is the lack of good instrumental variables. The system GMM estimator tackles the endogeneity of factor inputs well but is not a good way of controlling for self-selection bias. Essentially, the instruments in this paper are lagged levels and first differences of the grant levels for the endogenous first differences and levels of the grant levels respectively. Such instruments will be very weak predictors of current grant levels and the estimates will therefore be centred on the biased OLS estimates (see chapter 3.2). Furthermore, they may not be exogenous as, in the levels equation of the

[24] These are the Irish equivalents of the ARD and SAMIS respectively.

system GMM estimator, there is no transformation which removes time-invariant heterogeneity and lagged differences of the grant levels could well be correlated with this time-invariant heterogeneity. Another issue is that TFP is calculated in a first stage using the Levinsohn and Petrin approach which, as discussed in appendix A7.1, is based on untenable assumptions.

Using the same dataset as in their paper on the effects on TFP of grants in Northern Ireland, Harris and Trainor (2007) find that SFA was mostly aimed at preventing redundancies in existing plants rather than creating employment through the provision of assistance to new plants. Using a Cox proportional hazards model, they find, using no method to control for self-selection bias, that assisted plants had on average a 24.1% lower probability of closure than plants receiving no assistance. Noting the potential existence of self-selection bias, they then re-estimate the model using the predicted values from a Tobit regression of value of SFA received upon relevant variables instead of the actual values of SFA received. This yields a corresponding figure of 15%.

A major problem with this application of IV estimation is the absence of a good IV. The only variable that appears as a determinant of grant level but does not appear in the proportional hazards model is the natural logarithm of firm age. However, firm age, in levels, does appear in the outcome equation. So identification depends only upon this transformation and the nonlinear relationship between the treatment variable and the explanatory variables used in the treatment regression induced by the Tobit model. As a result, there is likely to be substantial collinearity involving the treatment variable in the outcome equation which may lead to inefficient estimates. Another problem is that instrumental variables cannot be used in the Cox proportional hazards model in this manner (Bijwaard, 2008). The method of replacing the treatment variable with the predicted values from a regression of the treatment variable on exogenous determinants of treatment is only valid if the model is linear. The Cox proportional hazards model is nonlinear.

A more conventional application of IV estimation is found in Criscuolo, Martin, Overman and Van Reenen (2007). Their paper investigates the impact of grants provided under the UK’s RSA scheme between 1988 and 2003 on employment, investment and labour productivity using a dataset derived from the same datasets as are used here; namely, SAMIS and the ARD. As an instrument for treatment status they use dummies representing the different levels of support to which plants are entitled. It is argued that these are subject to exogenous variation because lags in the process determining the Assisted Areas Map lead to the entitlement map being determined by measures of economic deprivation that are between three and five years out-of-date by the time the new map comes into force. The Assisted Areas map is also affected by changes in EU-wide average GDP per capita and unemployment which are heavily influenced by exogenous factors such as the accession of new countries to the EU. During the period under consideration there were changes to the Assisted Areas map in 1993 and 2000. Using this strategy they find a large and statistically significant effect of RSA on employment and investment. The authors do not find a significant impact on labour productivity or TFP.

In the analysis of business support programmes, this is the best application of the instrumental variables estimator of which the author is aware. As such, a version of this instrumental variables strategy will be used in the empirical analysis later. The possibility that the instrument is not valid will therefore be discussed in chapter 7.3. Other criticisms may also be levelled at this paper. Firstly, the analysis is carried out at the reporting unit level. As discussed in chapter 5.3, RSA grants are given to individual plants so the reporting unit is not the most appropriate level at which to conduct analysis. Furthermore, many firms that received RSA have not been linked from SAMIS to the ARD. These firms therefore remain erroneously in the control group. Assuming a positive impact of RSA, this will lead to downwards biased estimates of the impact of RSA grants. While it is not possible to match all firms to the ARD, it is possible to link a higher percentage than is achieved here. Lastly, there is no attempt made to control for correlation between the error term and the control variables. As will be discussed in chapter 7.3, this may also lead to biased estimates of the treatment effect.

Control Functions

In all the papers discussed below, the dependent variable is a growth rate so that the analyses are cross-sectional. This may reflect the difficulties in extending the control function method to the panel data case (see chapter 3.3 for a discussion of this issue).

Faulk (2002) analyses the impact of the Georgia Job Tax Credit on employment growth. Using firm level data, a switching regression model is employed to allow the impact of the explanatory variables to vary across treatment and untreated groups. The inverse Mills ratio is estimated and included in the outcome equation for treatment and untreated groups to control for self-selection into the treatment group. The coefficients on the inverse Mills ratio are found to be insignificant in both equations indicating that firms did not self-select into the scheme. The results show that 23.5% to 27.6% of the employment created by firms that participated in the programme was attributable to the employment boosting effects of the Job Tax Credit. Attention is however drawn to the issue of additionality as these estimates imply that 72.4% to 76.5% of the jobs that were created by the participant firms would have been created without the assistance of the Job Tax Credit.

A very similar study is by Gabe and Kraybill (2002). Their paper examines the impact of five development programmes on the growth of 366 establishments in Ohio that undertook major expansions between 1993 and 1995. The largest share of assisted firms received assistance under the Ohio Job Tax Credit which is Ohio's version of the programme examined by Faulk (2002). Their focus is not only on the effect of these programmes on actual employment growth but also on the impact on the employment growth that firms announce they will achieve. The latter question is of interest because firms may overestimate the number of jobs they intend to create in order to secure more generous awards from the government. A switching model is employed with self-selection controlled for by the inclusion of inverse Mills ratios. They find that overall these programmes had little effect on actual employment growth but a large and positive impact on the amount of jobs that firms announce they will create.

An obvious problem with switching models is that although an estimate can be calculated of the impact of the scheme on the outcome variable, because separate equations are estimated for the treatment and untreated groups, the statistical significance of this estimate is not automatically provided as it is when a single equation approach is used. It should also be noted that the justification used for switching regressions – that they allow the impact of the explanatory variables to vary across treatment and control groups – does not necessitate separate equations as this could be achieved with one equation using interaction variables (such interaction variables are used in Harris and Trainor, 2005).

Control functions are used in a single equation by Roper and Hewitt-Dundas (2001) to investigate the impact of business support grants on turnover growth, employment growth and profit to asset ratios between 1991 and 1994 in small businesses operating in the Republic of Ireland and Northern Ireland. For the Republic of Ireland, the selection term was found to be not statistically significant using any of the variables mentioned above as dependent variables. For Northern Ireland, it was not significant when turnover growth and the profit to assets ratio were used. However when employment growth was the dependent variable, the selection term was negative and statistically significant at the 5% level suggesting that grants were being targeted at firms that, in the absence of assistance, would have experienced below average levels of employment growth. For the Republic of Ireland, the coefficient on the assistance dummy is not statistically significant at the 5% level using any of the dependent variables although it is positive and statistically significant at the 10% level in the employment growth equation. Results are the same for Northern Ireland with the exception that the assistance variable is now positive and significant at the 5% level when employment growth is the dependent variable. Another regression is run using the Northern Irish data to test the impact of grants on labour productivity. This yields a negative and statistically significant coefficient on the assistance indicator which shows that grants, while boosting employment growth, have a detrimental impact on labour productivity. This suggests that that the jobs created by the grants may not be sustainable in the long-run.

Roper and Hart (2005), using the same approach, evaluate the impact of participation between 1996 and 1998 in the Business Links programme on the growth of employment, turnover and labour productivity between 1996 and 2000 in small firms in England. The Business Links programme consisted primarily of subsidised consultancy and specialist advice and is thus an example of ‘soft’ support. Probit models provide little evidence that Business Links was being targeted at firms that had performed well in the past. In the main regression, selection terms are found to be not statistically significant when any of the dependent variables are used. The coefficients on the Business Links dummy are also never significant which suggests that the Business Links programme had no impact on the chosen indicators of performance. However, the authors also find that when the selection term is not included in the employment growth equation, a positive and statistically significant effect of Business Links participation is obtained. This highlights the importance of controlling for self-selection bias.

Control functions are also used by Hart, Driffield, Roper and Mole (2008a) in their wide-ranging report into the impact of RSA grants provided in Scotland between 2000 and 2004 on plant performance. Their dataset is gathered from a telephone survey of 157 RSA recipients and 157 non-recipients and includes a broad set of variables thought likely to influence performance and growth. It also includes what the authors argue to be good IVs: the existence of a published business plan, firm age, exporting behaviour and the degree of local R&D. Using the growth of employment between 2004 and 2006 as the dependent variable, they find that RSA grants had a positive and statistically significant impact.[25] The estimated coefficient on the inverse Mills ratio is negative indicating that RSA grants went to plants that would have experienced below average employment growth had they not received an RSA grant. Using sales and labour productivity growth as dependent variables, the treatment dummy was not significant and it is suggested that two years may have been an insufficient period of time to capture the entire effects of assistance on these variables.

As with most applications of the control function estimator, it is hard to argue that a complete specification of the treatment equation has been provided in these papers. As discussed in chapter 3.3, the requirement for a full specification of the treatment equation is one of the drawbacks of this approach. Furthermore, the instruments used in these papers may be invalid. For example, in Hart, Driffield, Roper, and Mole (2008) the following instruments are used: the existence of a published business plan, firm age, exporting behaviour and the degree of local R&D. However, an argument can be made that all of these instruments are not valid: the existence of a published business plan may reflect managerial skills; firm age may capture the vintage of technology (Samaniego, 2006); exporting may force a firm to become more efficient (Harris and Li, 2007) and the degree of local R&D may capture technology spillovers (Harhoff, 2000). If these variables are not valid instrumental variables, their usage as such will lead to biased estimates of the treatment effect. Finally, the criticism that the endogeneity of other covariates is ignored can also be levelled at all these studies.[26]

---

[25] Results from a two-stage-least-squares model are also reported. These are the same in terms of the sign and statistical significance of the treatment variable.

[26] These criticisms can also be made against the papers discussed in the control functions and switching regressions section.

4.5. Conclusion

This first part of this chapter reviewed what predictions are available from the theoretical literature concerning the empirical analyses of chapters 6, 7 and 8. In relation to the productivity decomposition, the theoretical literature appears to suggest that a scheme such as RSA will have a negative impact on the growth of aggregate productivity. In relation to chapters 7 and 8, receipt of an RSA grant is expected to have a positive impact on TFP and a negative impact on the probability of closure.

The review of empirical papers offered ambiguous evidence concerning what results may be expected from the chapter that examines the impact of RSA on TFP. Limiting the discussion to studies that look specifically at the RSA scheme, Harris and Robinson (2004) find a positive and significant impact of RSA on TFP. Criscuolo, Martin, Overman and Van Reenen (2007) find no significant impact on either TFP or labour productivity and Hart, Driffield, Roper and Mole (2008a) find no impact on the growth of labour productivity. The evidence on the impact of other business support programmes on productivity is also inconclusive although more studies find no statistically significant effect than studies that do. Fewer studies analyse the impact of business support programmes on survival. Of the two discussed here, Girma, Görg and Strobl (2007b) and Harris and Trainor (2007) find a significant reduction in closure probability. As the programmes analysed in these papers are similar to the RSA scheme, this provides further evidence in support of an a priori expectation that RSA will be found to reduce closure probability.

The approach taken to analyse the impact of RSA on TFP in chapter 7 differs from all the studies in that it not only takes account of bias arising from self-selection but also deals with the endogeneity the factor inputs. In the studies that examine the effectiveness of RSA using the same dataset as is used here, Criscuolo, Martin, Overman and Van Reenen (2007) deal with the former but not the latter while Harris and Robinson (2004) tackle the latter but not the former. Furthermore, the dataset used here is superior to that used in both of these studies as more of the plants that received RSA have been identified in the ARD.

The method adopted to examine the effectiveness of RSA in reducing the closure rate is very similar to that taken in Girma, Görg and Strobl (2007b) in that a Cox proportional hazard model is estimated on a matched sample. Aside from the obvious point that they were looking at the effectiveness of Irish grants rather than RSA, there are other differences. Most importantly, their specification of the probit model that determines treatment is not consistent with the hazard model because variables are included in the former that are excluded in the latter and excluded in the former but included in the latter. The approach adopted in chapter 8 is not open to the same criticism.

5. Data

5.1. Introduction

The dataset that will be used in the empirical analyses of chapters 6, 7 and 8 is created by merging the SAMIS into the ARD. The former is a register containing information on plants that received support under the following grant schemes: RSA, IFG, SMART and SPUR.[27] The ARD is a longitudinal database that contains a range of variables including factor inputs and outputs. Successfully linking the two databases is crucial as failure to identify a high proportion of plants which received an RSA grant in the ARD will undermine the results in the empirical analyses. The work undertaken to merge the two datasets builds on that of Harris (2005a) and Criscuolo, Martin, Overman and Van Reenen (2007). This resulted in over 91% of the plants that received RSA grants in Scotland being located in the ARD. This is a higher link than has been achieved in previous studies and therefore provides a firmer basis for analysis of the impact of RSA.

The chapter is structured as follows: the next section will provide descriptive statistics from SAMIS on the distribution of grants across year, industry and region; the third section describes the method used and the extent to which SAMIS was linked into the ARD; the fourth section gives a description of the variables in the merged dataset; the fifth gives a comparison of the characteristics of RSA plants and non-assisted plants and the final section concludes.

5.2. SAMIS Descriptive Statistics

The Department for Business, Enterprise and Regulatory Reform[28] maintains the SAMIS database which contains information on the plants and firms that receive assistance under various business support schemes in GB. Although the statistical analysis conducted later will focus exclusively on the largest of these schemes, RSA, this section will provide

[27] Information on firms that received support under the LINK and Benchmark schemes is also given in SAMIS but, as these are not grant schemes, they will not be discussed.

[28] Formerly the Department for Trade and Industry (DTI)

descriptive statistics on all the grant schemes contained in SAMIS for comparative purposes and to place the RSA scheme in the context of the different types of grant support available in Scotland.

Table 5.1 shows some basic descriptive statistics on the number and value of grants awarded under different schemes. Detailed information on the RSA and SMART schemes is given in chapter 2.6. As discussed in that chapter, the IFG and SPUR schemes have recently been subsumed into the RSA and SMART schemes respectively. The former was an investment grant scheme which provided grants to smaller projects than those assisted by RSA. The SPUR scheme was an innovation grant scheme that supported R&D projects. Further details on the IFG and SPUR schemes are given by Harris (2005a).

Table 5.1: Number and value of grants awarded by scheme in Scotland, 1972-2003

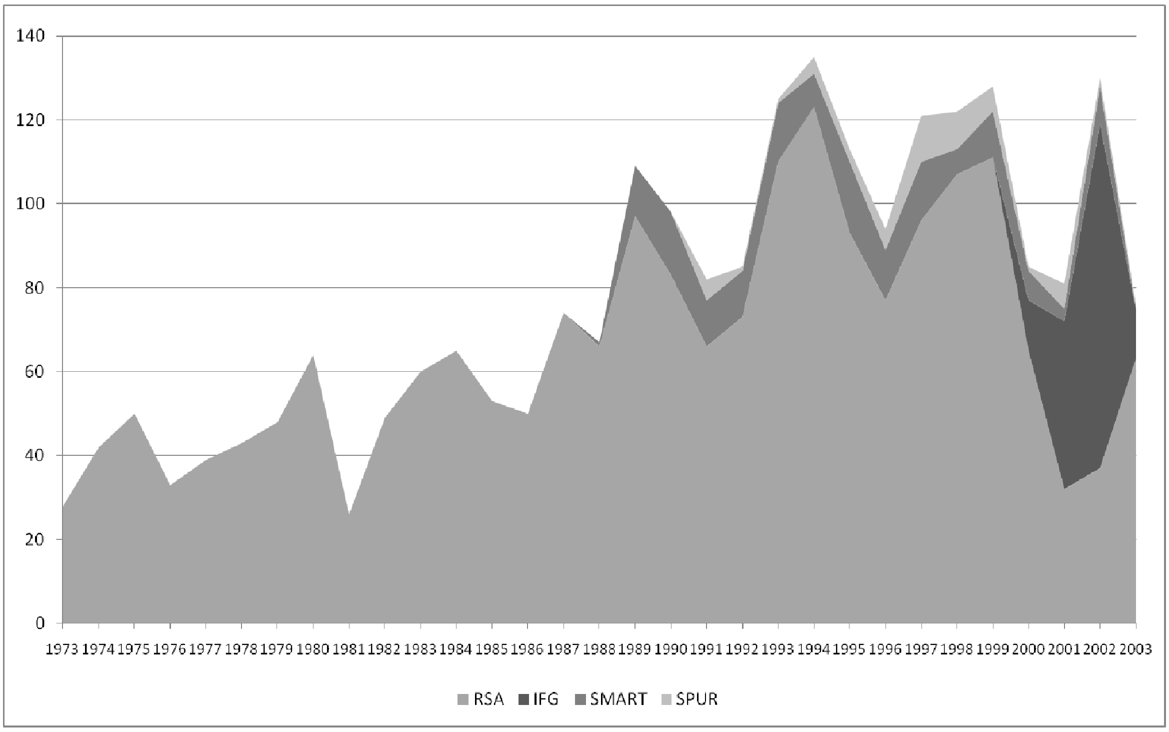
	Number of Grants	Mean Number of Grants Per Year	Mean Grant Value (2003 prices)
RSA	2,023	65	747,467
IFG	146	37	58,209
SMART	151	10	56,395
SPUR	55	4	191,265

Source: SAMIS

The first column shows that 2,023 grants have been awarded under the RSA scheme in Scotland since its initiation in 1972. The corresponding figures for IFG, SMART and SPUR are far smaller but these numbers are, in large part, a reflection of the different lengths of time for which the schemes have operated. To avoid this spurious comparison, the next column shows the average number of grants awarded each year during the period in which the schemes were in operation. This shows that, at 65 grants per year, many more RSA grants were awarded on average per year than any other type of grant. 37 IFG grants were provided on average each year while, for the innovation grants, the average number of SMART and SPUR grants was only 10 and 4 respectively. The final column gives the average value of grants awarded under each scheme. The RSA scheme is far more generous than the other schemes with an average award of almost £750,000. The average award made under the IFG scheme was under 8% of the value of the average award made under the RSA scheme which is the result of IFG grants being only awarded to smaller projects. The average award made under the SMART scheme was of a similar magnitude to that provided to IFG grant recipients while the average SPUR grant was almost £200,000. The difference between the average award made under the SMART and SPUR scheme is a consequence of the SPUR scheme having provided support for large R&D projects while the SMART scheme, prior to subsuming the SPUR scheme, gave support for less costly feasibility studies.

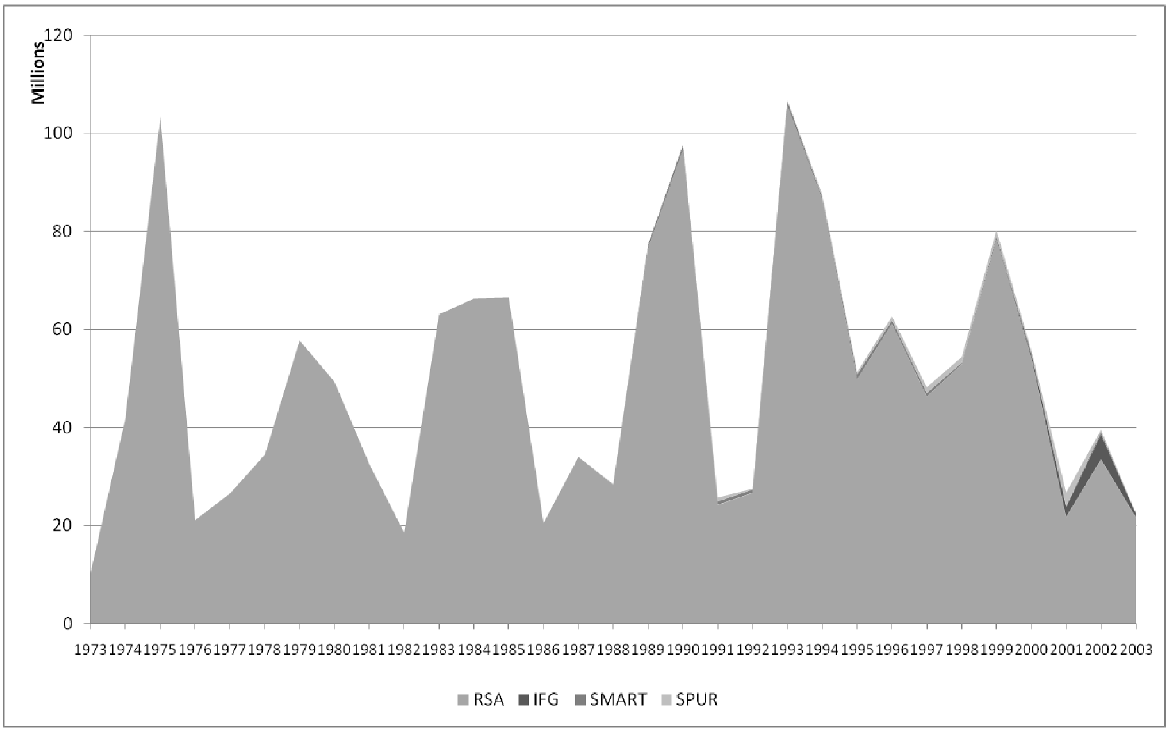
Figure 5.1 shows the number of grants received each year under each scheme in Scotland. The numbers for recent years should be treated with caution as, due to delays in updating SAMIS, the database may not include all of the more recent awards. Figure 5.2 gives the total value of grants provided each year under each scheme.

Figure 5.1: Number of grants received in Scotland by year, 1972-2003



Source: SAMIS

Figure 5.2: Value of grants received in Scotland by year, 1972-2003 (2003 prices)



Source: SAMIS

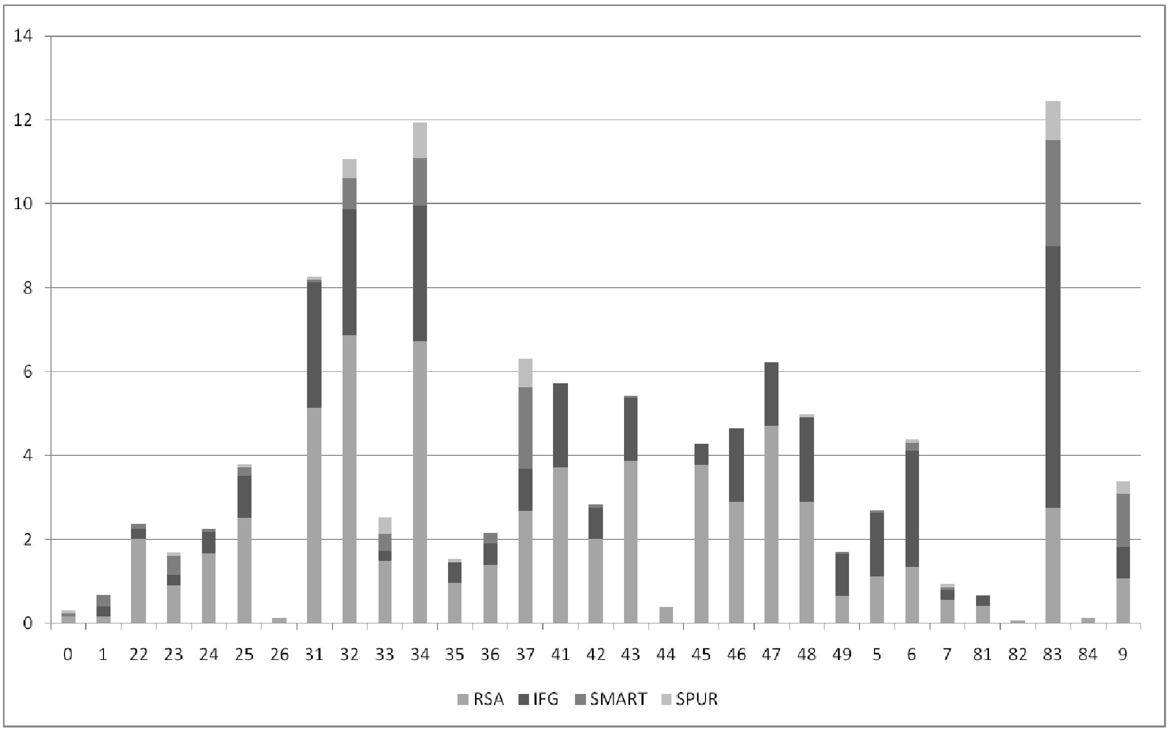
As is clear from figure 5.1, RSA is the scheme that has been in operation for the longest period of time, having begun throughout GB in 1972, although the first grants were awarded in Scotland in 1973. The number of awards made under the RSA scheme has fluctuated considerably over time but an upward trend is discernible until 1999. Dramatic falls were witnessed in 2000 and 2001 due to the introduction of the IFG scheme which provided investment grants to smaller projects that had previously been given under the RSA scheme. Turning to the innovation grants schemes, the first SMART grants were provided in 1988 in Scotland and the number of awards has displayed no obvious trend. The largest number of SMART grants awarded was 17 in 1995. The first award in Scotland under the SPUR scheme was made in 1991. Compared to other schemes, relatively few of these grants have been provided with a peak in 1997 of 11 awards.

The most obvious lesson from figure 5.2 is that the value of RSA grants has always dwarfed that of grants provided under different schemes. This is a reflection of both the larger number of RSA grants as well as their larger value. The value of RSA awards has fluctuated even more than their number but has displayed no discernible upward or downward trend. In 1975, 1990, 1993 and 1994 over £80 million of RSA support was awarded but in 1973, 1976, 1982, 1986, 1991, 2001 and 2003, less than £25 million was awarded under the scheme. As the number of RSA grants tended to rise until 2000, this implies that RSA grant levels have fallen over time.

Because of the relatively high value of the grants awarded under the RSA scheme, figure 5.2 is not particularly helpful in showing the value of grants awarded under other schemes. To better illustrate this, figure A5.1 in appendix A5 shows the value of grants awarded by year under the IFG, SMART and SPUR schemes only. This shows that the IFG scheme never awarded more than £6 million in spite of providing a larger number of grants than RSA in 2001 and 2002. As discussed earlier, this is a consequence of the smaller projects that it assisted. In terms of the innovation grant schemes, the value of SMART awards peaked at £1 million in 1995 and has since declined. The largest amount awarded under the SPUR scheme was £2.8 million in 2001. As discussed earlier, because SPUR awards were provided to assist more costly projects than SMART awards, this is consistent with expectations.

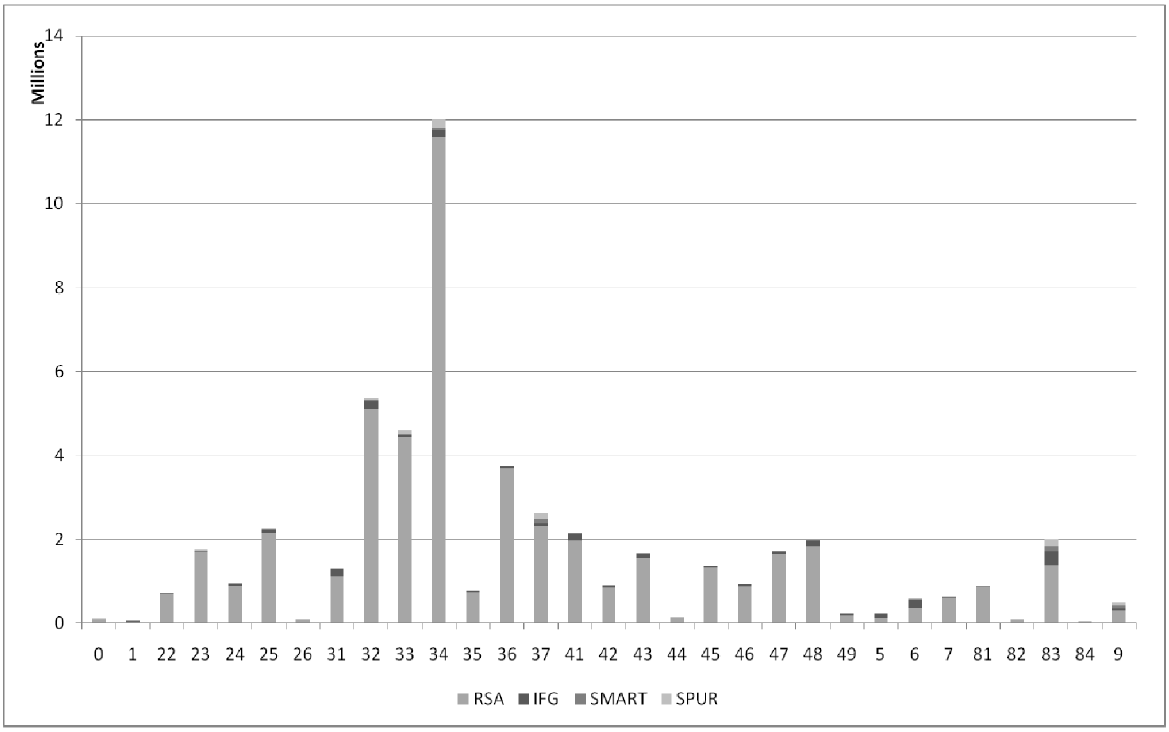
Figure 5.3 shows the number of grants awarded under each scheme for each 2-digit standard industrial classification (SIC) 80 code. In order to avoid making misleading comparisons due to differences in the length of time for which schemes have been operating, the total number of grants for each industry is divided by the number of years that each scheme has been in operation so that the figures represent the average number of grants received under each scheme for each year that the scheme operated. The industries represented by each code are listed in Appendix A2. Figure 5.4 shows the value of grants received by 2-digit SIC 80 industry on the same basis as the number of grants is shown in figure 5.3.

Figure 5.3: Average number of grants received per year in Scotland by 2-digit industry, 1972-2003



Source: SAMIS

Figure 5.4: Average value of grants received per year in Scotland by 2-digit industry (2003 prices), 1972-2003



Source: SAMIS

The following industries received an average of over three RSA grants per year between 1973 and 2003: the manufacture of metal goods (SIC 31), mechanical engineering (SIC 32), electrical and electronic engineering (SIC 34), food, drink and tobacco manufacturing industries (SIC 41), textile industry (SIC 43), footwear and clothing industries (SIC 45) and manufacture of paper & paper products; printing & publishing (SIC 47). That these are all manufacturing industries is a consequence of RSA being initially only available to manufacturing plants (Harris and Robinson, 2001a). The following industries received an average of at least three IFG grants per year during the years of that scheme's operation: the manufacture of metal goods (SIC 31), mechanical engineering (SIC 32), electrical and electronic engineering (SIC 34) and the business services (SIC 83) industries. That these industries, with the exception of business services, also receive a large share of RSA grants is unsurprising given that the latter is a spin-off from the RSA scheme.

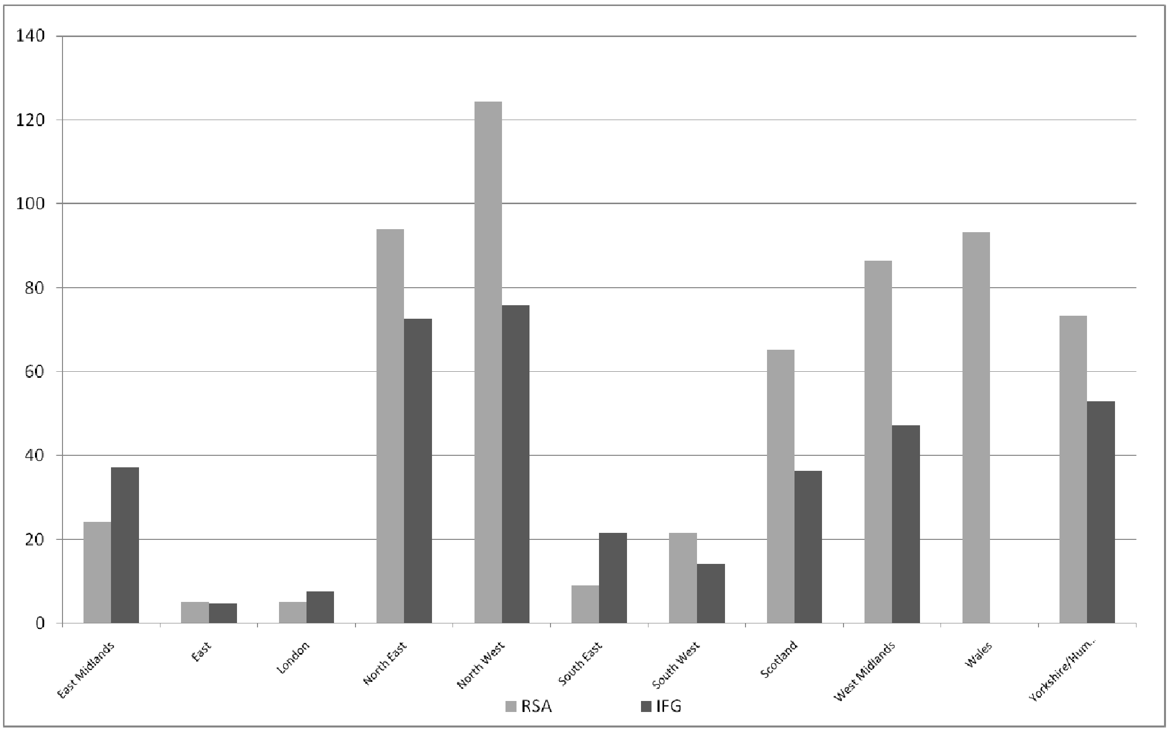
Turning to the innovation grant schemes, the following industries have received an average of over 0.7 SMART grants per year: mechanical engineering (SIC 32), electrical and electronic engineering (SIC 34), instrument engineering (SIC 37), business services (SIC 83) and R&D (SIC 94). The SPUR scheme has provided an average of over 0.65 grants to the electrical and electronic engineering (SIC 34), instrument engineering (SIC 37) and business services (SIC 83) industries. As SMART and SPUR are similar in that they both represent attempts to boost innovative activities, that industries receiving a high share of SMART grants generally also receive a large share of SPUR is consistent with expectations.

The instrument engineering (SIC 37), business services (SIC 83) and the R&D (SIC 94) industries all receive a large share of either or both SMART and SPUR grants and smaller shares of RSA grants. On the other hand, the food, drink and tobacco manufacturing (SIC 41), textile industry (SIC 43), footwear and clothing industries (SIC 45) and manufacture of paper and paper products; printing & publishing (SIC 47) industries receive large shares of RSA grants but smaller shares of SMART and SPUR grants. This is a result of the differing nature of these schemes: RSA, as a capital grant scheme, has tended to go to capital intensive industries whereas SMART and SPUR has tended to go to more innovative industries.

Turning to figure 5.4, the industries which received the largest value of RSA grants are the mechanical engineering (SIC 32), manufacture of office machinery & data processing equipment (SIC 33) and electrical and electronic engineering (SIC 34) and the manufacture of other transport equipment (SIC 37) industries. The latter received by far the largest amount of RSA grants at almost £12 million a year. The industries that received an average of over £160,000 of IFG grants are also those that received the largest number of IFG grants. Electrical and electronic engineering (SIC 34), instrument engineering (SIC 37) and business services (SIC 83) are the industries that receive the largest amount of SMART grants while manufacture of office machinery & data processing equipment (SIC 33), electrical and electronic engineering (SIC 34), instrument engineering (SIC 37) and business services (SIC 83) received the largest amounts of SPUR grants.

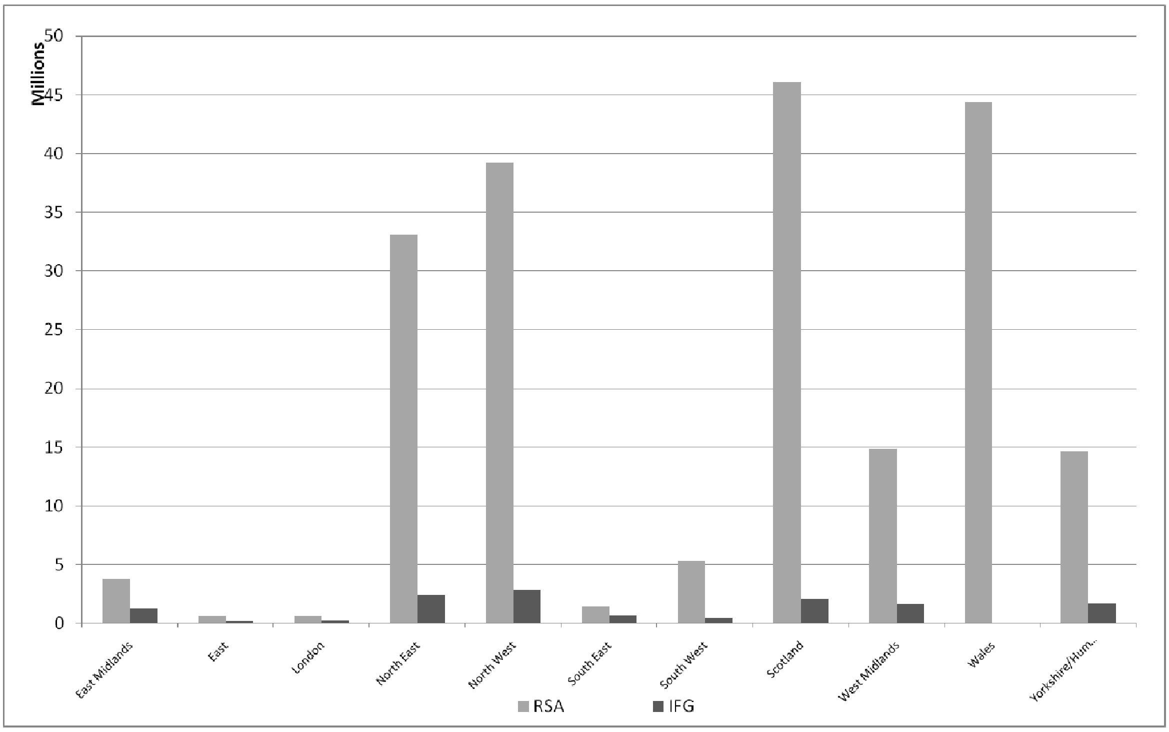
Figure 5.5 shows the average number of grants received per year during the years in which the scheme operated by region for the RSA and the IFG schemes. Figure 5.6 gives the average value of grants awarded each year on the same basis by region for the same schemes. Information on the location of the recipients of SMART and SPUR grants is not given because it is not readily available in SAMIS.

Figure 5.5: Average number of RSA and IFG grants received per year by region, 1972-2003



Source: SAMIS

Figure 5.6: Average value of RSA and IFG grants received per year by region (2003 prices), 1972-2003



Source: SAMIS

As Scotland's population represents 8.69% of the total population of GB (Office for National Statistics, 2009a), Scotland, by receiving 10.85% and 9.81% of RSA grants and IFG grants, is receiving a disproportionately large share of RSA and IFG grants. This is explained by the relatively large proportion of Scotland that has assisted areas status. More remarkable is the fact that plants located in Scotland received a higher value of RSA grants than plants in any other region. This amounts to 22.58% of the total value of RSA awarded and points to the fact that the average value of an RSA grant in the rest of GB was only £311,772 – less than half the average amount of RSA grant awarded in Scotland. This is a consequence of the assisted areas of Scotland generally being in higher tiers than those of other regions which allows higher proportions of project costs to be covered by RSA grants. In terms of IFG grants, Scotland received 15.39% of the total value of grants which is a result of the average value of IFG grants in the rest of GB being only 60% of the value of IFG grants awarded in Scotland.[29]

In summary, this section has shown that RSA has, in most years, offered the largest number of grants per year and that the total value of these grants has always been far larger than the total value of grants provided under other schemes. Ignoring the years after 2000 when the IFG scheme was introduced, the number of RSA grants provided has risen since its introduction in 1972 although this has not been matched by an increase in the total amount awarded, implying that the value of RSA grants has fallen slightly. The largest number of RSA grants has gone to engineering and manufacturing plants. Finally, Scotland has received a disproportionately large number of RSA grants due to its having a larger proportion of its territory designated as assisted areas. The value of RSA grants provided in Scotland is, on average, more than twice as large as the average value of an RSA grant provided in the rest of GB which is a consequence of Scottish assisted areas generally being in higher tiers than those of other regions.

5.3. Data Linking

In order to estimate the impact of RSA on plant performance it is, of course, necessary to be able to identify which plants received such support in the dataset to be used for the

[29] Note that the equivalent of the IFG scheme in England and Wales was the Enterprise Grants scheme.

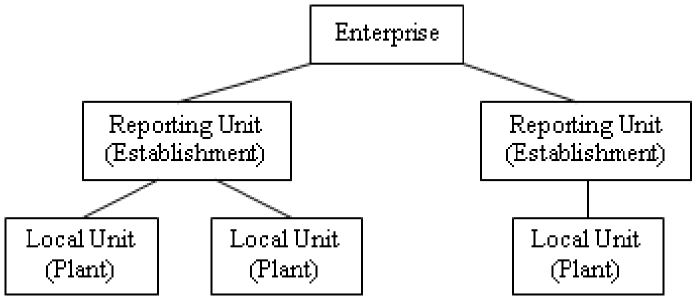
analysis. This section will describe how, for this purpose, the SAMIS database was linked to the ARD. The process also involved a third database: the inter-departmental business register (IDBR). Although the empirical analyses will examine the impact of RSA on Scottish plants, use will also be made of non-assisted plants in the rest of GB. Therefore, to avoid incorrectly classifying RSA recipients outside Scotland as plants that did not receive a grant, the linking process must be undertaken for RSA recipients throughout GB. The first part of this section provides a description of each of the three databases; the next section describes the method employed to link SAMIS to the ARD and the final section gives statistics showing the extent of the link achieved across time and industry.

Databases

SAMIS, from which statistics were provided in the previous section, has data on over 50,000 firms in GB that applied for an RSA grant between 1972 and 2003. This includes the postcode, SIC code, employment level of the applicant firm and the date in which the application was made. For successful applications, the date when the payment was made and the value of the grant are also recorded.

The IDBR provides the name, address, ownership structure, industrial classification and employment level of all plants in the UK (the description of the IDBR and the ARD is based on that provided by Oulton, 1997; Robjohns, 2006). Plants are organised into local units, reporting units and enterprise groups. Local units are plants or offices at a single geographical location. A reporting unit, or establishment, is the smallest unit which can provide the full range of data required for the Annual Business Inquiry (ABI), which is discussed in the next paragraph. When a local unit can provide the full range of information necessary for the ABI, it will report to the ABI. When it reports on behalf of itself only, it is a ‘single’ as the reporting unit consists of only one local unit. However, not all local units are able to provide the required information for the ABI and, for these plants, another local unit will report on their behalf. In this case, the local unit that reports is a ‘parent’ while those local units on whose behalf it reports are its ‘children’. The reporting unit then consists of both the parent and children local units. Enterprises consist of reporting units that share a common owner. Figure 5.7 shows a hypothetical enterprise consisting of two reporting units, one of which is a parent with one child while the other is a single.

Figure 5.7: Structure of Hypothetical Enterprise



Local units, reporting units and enterprises are all identified by unique reference numbers in the IDBR which allow them to be tracked through time. As the ARD also contains these reference numbers, an RSA recipient that has been found in the IDBR is also found in the ARD.

The ARD is a longitudinal dataset dating from 1970 (see Griffith, 1999 for more information on the ARD).[30] It is created by combining information from the IDBR, termed ‘indicative data’, with more detailed information collected at the reporting unit level by the ABI, referred to as ‘returned data’. The more detailed information collected in the ABI includes data on investment, intermediate inputs and gross output. In each year there is a ‘selected’ and a ‘non-selected’ file. The ‘selected’ file contains a combination of indicative and returned data on reporting units – the level at which the ABI is collected - which were selected for surveying in the ABI. The ‘non-selected’ file contains indicative data from the IDBR and covers establishments that were not selected for sampling in the ABI, the local associated with such reporting units and the local units associated with reporting units selected for inclusion in the ABI.

Reporting units are selected for surveying in the ABI based on employment data contained in the IDBR with the sampling frame skewed towards larger reporting units. At present, 25% of reporting units with fewer than 10 employees are surveyed in the ABI; 50% of reporting units with between 10 and 99 employees are surveyed; the proportion surveyed of reporting units with between 100 and 249 employees varies by industry from 100% to less than 50% while 100% of reporting units with 250 or more employees are surveyed (Robjohns, 2006).

[30] However, the data from 1970 to 1972 is incomplete.

As the more detailed 'returned' data in the ARD is generally required econometric analyses, many studies (see, for example, Criscuolo, Martin, Overman and Van Reenen, 2007) have used the reporting unit as their unit of analysis. However, when analysing the impact of RSA grants, this is not appropriate as RSA is awarded to support capital investment in specific plants rather than throughout the enterprise. Furthermore, the reporting unit is an accounting rather than an economic unit. As such, the number of plants covered by a reporting unit may change as enterprises open and close plants, buy and sell plants or simply because of changes in the way that an enterprise chooses to report to the ABI (Harris, 2005b). The consequences of using the reporting unit rather than the local unit to calculate measures of the capital stock are investigated by Harris (2005c). To permit econometric analysis at the more appropriate local unit level, it is therefore necessary to 'spread back' to the local unit those variables that are only collected in the ABI at the reporting unit. These include important variables such as gross output, intermediate inputs and investment. This is done using the plant level employment data collected in the IDBR using the assumption of constant labour-investment ratios and labour productivity levels within reporting units. This should be born in mind when interpreting the results as standard errors will be artificially reduced by this method.

Linking Process

The dataset used in the analysis here is built upon the dataset used in Harris (2005a). The following will describe the steps taken by Harris to construct his dataset. His starting point was a version of SAMIS provided by BERR containing 33,328 RSA applicants (successful and unsuccessful) which had been linked at either the local unit, reporting unit or enterprise level to the IDBR. This represented 61% of all RSA applicants listed in SAMIS. Using IDBR reference numbers common to both SAMIS and the ARD, he was able to link 14,649 RSA applicants at the local unit level.[31] Some could not be linked because the ARD only covers the manufacturing sector before 1997 and, at the time, was only available until 2001. Due to the limited length of time for which plants outside manufacturing had been observed, these were removed from his dataset. The use of Central Statistical Office (CSO)

[31] As discussed later, this is the most appropriate level at which to conduct analysis of RSA.

reference numbers[32] that go back to 1970 gave a sufficiently long time series for manufacturing plants. The removal of plants outside manufacturing reduced the number of RSA applicants linked to the ARD at the local unit level to 11,194.

Further work has been done in linking between SAMIS and the ARD by Criscuolo, Martin, Overman and Van Reenen (2007). In total, they were able to link 68% of the RSA applicants in SAMIS to the ARD. Of the additional links, some were made at the local unit level and these have been added to those made by Harris. Other links were made at the reporting unit and enterprise level. The main unit of observation in their paper was the reporting unit rather than the plant so they did not attempt to use these links to find additional links at the plant level. This was done here by listing the plants that fall under the linked reporting units or enterprises in the ARD and choosing the plant that best matches the description of the plant in SAMIS on the basis of postcode, SIC code, plant employment and year in which the application was made. Often there is no choice to be made as reporting units and enterprise groups frequently encompass only one plant.

The final stage in the linking process was the most laborious. It involved manually trying to locate all the other plants in SAMIS that had not yet been linked to the ARD. At this point, plants in SAMIS outside manufacturing and those whose applications were unsuccessful were removed from the list of plants that had to be linked. The former were excluded as the empirical analyses will be conducted using a sample consisting exclusively of manufacturing plants. The latter were excluded because all plants in the ARD not linked to an RSA recipient in SAMIS are assumed to be untreated and will therefore have the correct treatment status regardless of whether they are identified in the ARD. The approach taken was twofold. Firstly, a search was conducted in the ARD by postcode for the RSA receiving plant. If on the basis of SIC code, plant employment and the year in which the application was made it was judged that one of the plants displayed was the RSA recipient in SAMIS, the link was made. If this approach failed to provide a link, a second attempt was made. This involved searching by SIC code, plant employment and year in which the

[32] IDBR and CSO reference numbers are both local unit reference numbers. IDBR reference numbers were introduced in 1994 to replace CSO reference numbers. However, work carried out by Harris extended the CSO reference number series past 1993 so that a plant existing in both 1993 and 1994 (or any year after 1993) is identifiable as the same plant. This allows the researcher to use every year of data contained in the ARD rather than just 1973-93 or 1994-2004.

application was made. If this search gave a plant likely to be the RSA recipient in SAMIS, a link was made.

This latter stage in particular depends upon the judgement of the researcher and it is unavoidable that in some cases an incorrect link will have been made. It is important to consider the implications both of failure to link at all from SAMIS to the ARD and of making an incorrect link. When a plant is not linked from SAMIS to the ARD this means that a plant that belongs in the treatment group remains in the untreated group. Assuming a positive impact of treatment on a given outcome variables, this will lead to downwards biased estimates of treatment. If the wrong plant is linked, this not only leaves a plant that should be in the treatment group in the untreated group but also allocates a plant that belongs in the untreated group to the treatment group. This will generate estimates that are more biased than if the plant was not linked at all. As a result, considerable caution was taken to try to minimise the number of erroneous links.

Comparison of linked and non-linked plants

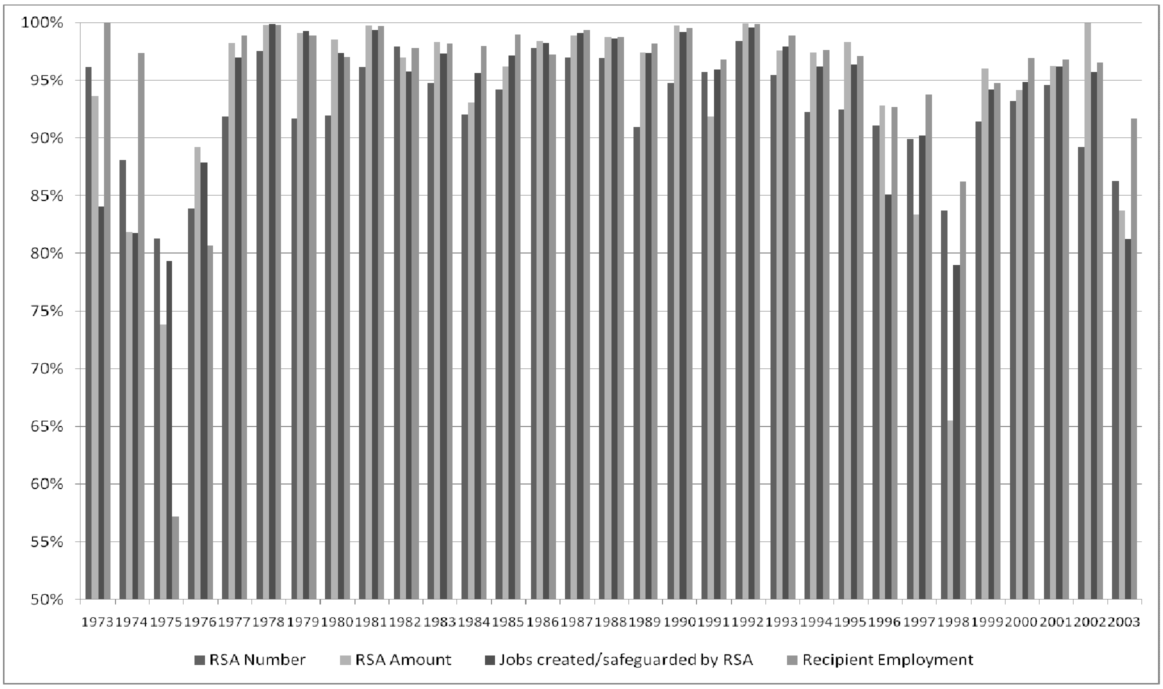
In the end, this process achieved a link from SAMIS to the ARD of 91.43% of manufacturing plants that received RSA in Scotland and 92.44% of manufacturing plants that received RSA in GB. This is higher than the level achieved in previous studies of the impact of receipt of an RSA grant (Harris and Robinson, 2004; Harris, 2005a; Criscuolo, Martin, Overman and Van Reenen, 2007) and therefore provides a firmer basis for empirical analysis. Nevertheless, it is important to probe whether the linking process may create any bias in the empirical results of later chapters. This would be the case if, for instance, the percentage of plants linked to the ARD rose over time and RSA grants became more successful in improving plant performance over time. The empirical results may then indicate that RSA had a large and positive impact on plant performance but these results would not be representative of the entire period.

To probe whether the linking process may have created any bias in the empirical results, a number of t-tests were performed. The mean of employment for the linked plants across all years is 206.66 while the corresponding figure for unlinked plants is 111.48. A t-test shows that the difference is significantly different from zero at the 99% level and therefore that the process has been more successful in linking larger plants. This is consistent with expectations as it was more difficult to find small plants that received assistance in the ARD when manually looking for a link. This is because the size distribution of plants in the economy, and hence the ARD, is skewed towards smaller plants and it is therefore difficult to identify the small plants that received grants as there can be many candidate plants in the ARD with the correct employment level, SIC code and year in which the application was made. This will create a bias in the estimation results if larger plants respond differently to receipt of an RSA grant than smaller plants.

Further t-tests were performed using the value of RSA grants and the number of jobs created or safeguarded within the recipient plant. The mean value of RSA grants across all years was £306,750 (2000 prices) for linked plants and £282,843 for plants that have not been linked. A t-test indicates that there is no statistically significant difference between these figures. On average, RSA created or safeguarded 100 jobs in the linked recipient plants and 81 in those that were not linked. This difference was also not statistically significant at the 90% level.

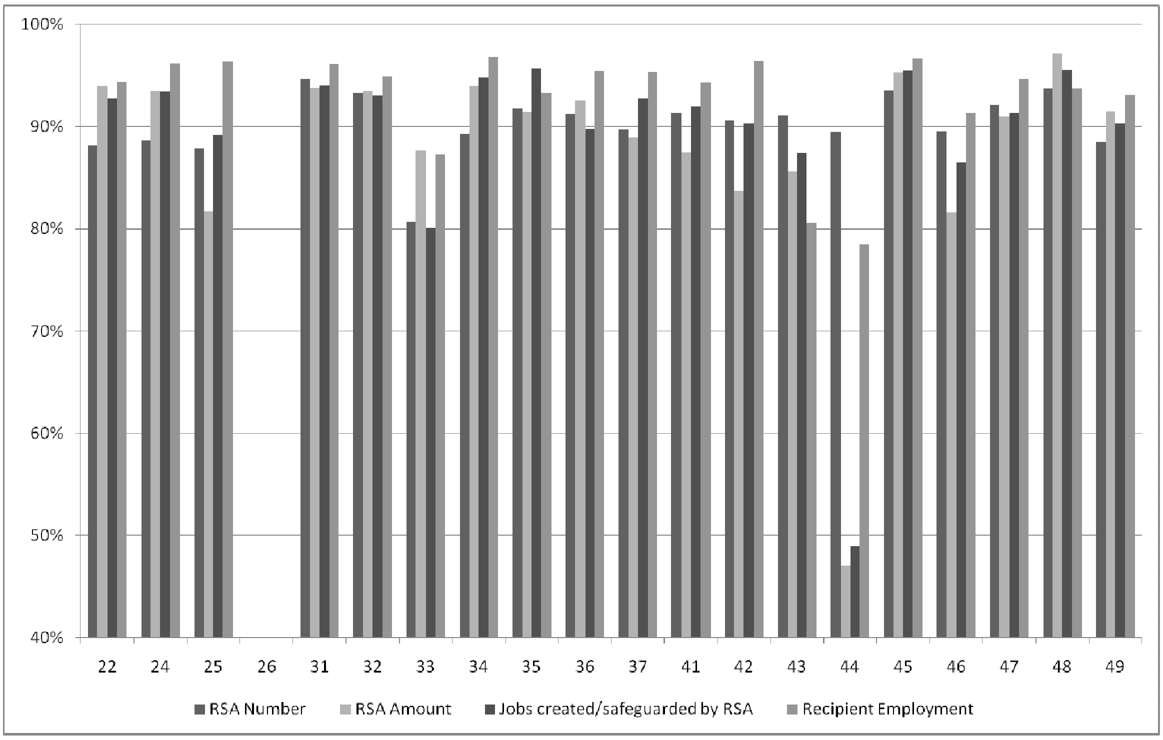
Figure 5.8 shows the extent of the link for plants located in Scotland achieved by year using several of the variables included within SAMIS: the first column shows the percentage of the number of RSA recipients linked to the ARD; the second gives the percentage of the value of grants awarded linked; the third shows the percentage of the jobs created or safeguarded in RSA recipients linked and the fourth shows the percentage linked of the employment of RSA recipients. Figure 5.8 gives the same information by 2-digit industry.

Figure 5.8: Percentage of SAMIS variables linked to the ARD by year in Scotland, 1972-2003



Source: SAMIS

Figure 5.9: Percentage of SAMIS variables linked to the ARD by 2-digit industry in Scotland, 1972-2003



Source: SAMIS

No figures are provided for SIC26 (production of man-made fibres) to avoid disclosure.

There are only two years which give cause for concern: 1975 and 1998. For these years, over 80% of the RSA recipients in SAMIS have been linked to the ARD. However, the proportion of the employment of RSA assisted plants in 1975 linked to the ARD is only 57% which indicates that the linked plants are smaller than those that have not been linked. In 1998, only 66% of the value of RSA grants provided have been linked which shows that plants that have not been linked in 1998 tended to receive larger grants than those that have been linked. With the exceptions of these years, the percentage linked of all the variables considered never falls below 80%. For the link achieved in 1975 or 1998 to create a greater bias in the empirical estimation than the link achieved in other years would require large plants that received a grant in 1975 or plants that received large grants in 1998 to respond differently to receipt of an RSA grant than other plants. As there is no obvious reason why this should be true, it is reasonable to conclude that differences in the percentage of plants linked across time are unlikely to create any bias in the empirical estimation of later chapters.

The percentage of RSA plants linked to the ARD is above 80% for all of the industries considered. For manufacture of leather and leather goods (SIC 44), the percentage of the value of RSA grants linked and the percentage of the jobs created and safeguarded linked is 47% and 49% respectively. As these figures are based on a sample size of only 13 grants and the plants in the industry are not used on their own for estimation in the empirical analyses, the poor quality of the link is unlikely to generate a large bias. With the exception of this industry, the percentage of the amount of RSA grants, jobs created and safeguarded by RSA and recipient plants' employment is above 80% for all the industries considered. This suggests that difference in the percentage of plants linked across industry will not bias the results of the empirical estimation.

In summary, the linking process undertaken here has managed to find a higher proportion of grant recipients in the ARD than any previous study. This allows greater confidence in the results as fewer plants mistakenly omitted from the treatment group will lead to less bias in estimates of the impact of assistance. Those plants that have not been linked tend to be smaller and this should be borne in mind when interpreting the results. The relatively constant extent of the link across time and industry in relation to number of grants, value of grants linked, jobs created or safeguarded and plant employment permits confidence that the results obtained will be representative of the impact of RSA across the entire period during which the scheme has been in operation.

5.4. Variables

In this section, a description of the variables that will be used in the empirical analyses is given. Explanations of the usefulness of these variables are given in the empirical chapters where they are used. The first variables described are intrinsic to the plant while the latter variables refer to the environment in which the plant is operating.

Gross Output

Gross output is the total value of sales made by the plant. It is available from the ABI at reporting unit level so was spread back using plant-level employment data from the IDBR to give estimates of gross output at the plant level.

GVA

GVA measures the value added by the plant to its intermediate inputs. It is calculated by subtracting intermediate inputs from gross output.

Employment

The employment variable measures the number of people employed at the plant and is taken at the plant level from the IDBR. In chapter 7, where the impact of receipt of an RSA grant on TFP is estimated, it was decided not to make an adjustment to employment to better reflect productive services from labour as the cost of retaining an unproductive worker encourages the plant to retain only those productive employees. The same argument cannot be made in relation to the capital stock variable and this explains the apparent asymmetry in the treatment of these variables (see below).

Intermediate Inputs

Intermediate inputs are materials and services consumed in the production process. This variable is available from the ABI at reporting unit level so was spread back using plant-level employment data to give estimates of intermediate input consumption at the plant level.

Capital

Investment data is available from the ABI at the reporting unit level. To give estimates of investment at the plant-level, it was spread back to the plant-level using plant-level employment data.

The major problem associated with transforming an investment series into a capital stock series is finding an appropriate rate of deterioration which reflects the loss of efficiency through time due to decay caused by use in production and obsolescence as a result to ageing. Following Harris and Drinkwater (2000), a net capital stock measure is calculated assuming straight-line deterioration. The ‘bought’ (as opposed to hired) capital stock variable is then calculated by summing a gross capital stock measure and the net capital stock measure using a weight of three to one as follows:

$$BK_{it} = \left(\frac{3}{4}\right)GK_{it} + \left(\frac{1}{4}\right)NK_{it}, \tag{5.1}$$

where BK_{it} is the bought capital stock in plant i at time t, GK_{it} is gross capital and NK_{it} is net capital. The attraction of this approach is that the pattern of deterioration is slow at first and then accelerates, reflecting the idea that firms will invest more in maintenance and repair to maintain the initial level of service from a piece of capital when it is relatively new. Denison (1972) provides a fuller discussion of the merits of this deterioration pattern.

Assuming that this method provides an accurate representation of deterioration, this measure will provide a good measure of the productive services available from bought capital. However, a plant may choose not to fully utilize the capital that it has at its disposal because of, for instance, a lack of demand for its outputs. For the TFP analysis that will be conducted in chapters 6 and 7, it is necessary to have a measure of the productive services from bought capital that were actually used rather than a measure of the productive services from capital that could have been used. Therefore, a further adjustment is required.

The first step in this adjustment required calculating for each plant and time period a measure of bought capital usage, Y_{it}/BK_{it}, where Y_{it} is gross output. The highest level of this measure, $Y_{it}*/BK_{it}*$, is then identified, which is, for now, assumed to be the maximum that the plant can produce from one unit of bought capital. If a plant has a lower level of Y_{it}/BK_{it}, it is regarded as not fully utilising its bought capital stock. The measure of bought capital utilisation is therefore:

$$KU_{it} = \left(\frac{Y_{it}}{BK_{it}} \bigg/ \frac{Y_{it}^{*}}{BK_{it}^{*}} \right). \tag{5.2}$$

One problem with this process is that the average level of bought capital utilisation across the entire sample calculated using (2) is rather low at around 60%. Ornaghi (2006), using data on Spanish manufacturing firms, finds 80% to be the average capital utilisation figure. To replicate this average within the sample, KU_{it} is scaled. This creates a bought capital utilisation figure which exceeds one for certain plants which requires a value of one to be redefined to be merely a high level of bought capital utilisation rather than the maximum level.

The adjusted estimate of the productive services from bought capital is then:

$$ABK_{it} = KU_{it} \times BK_{it}, \tag{5.3}$$

where ABK_{it} is the adjusted capital stock. There is also data on the hire of capital in the ABI at the reporting unit level. Once again, this can be spread back using employment shares. Under the assumption that hired capital is always fully utilised this can be simply added to the adjusted measure of bought capital to yield total capital as follows:

$$K_{it} = ABK_{it} + H_{it}, \tag{5.4}$$

where K_{it} represents the total capacity utilisation adjusted capital stock and H_{it} is the hire of capital.

Capital-to-Labour Ratio

The capital-to-labour ratio is calculated by dividing the unadjusted capital stock by the number of employees at the plant.

Labour Productivity

Labour productivity is a measure of plant productivity calculated by dividing gross output by employment.

Real Wage

A total labour costs variable is available in the ARD at the reporting unit level. To obtain a measure of the real wage, this variable is spread back to the plant using plant-level employment data. It is then divided by the number of employees and deflated by an output price index to obtain the real wage.

Age

The IDBR contains codes that allow the identification of a plant through time. As a result, for plants that began operation during or after 1970, it is straightforward to identify the year in which a plant first operated and calculate an age variable. However, because the IDBR was first collected in 1970, it is not possible to identify when a plant that began operation before 1970 started to produce. The age variable is therefore truncated.

FDI

The FDI variable is a dummy that takes the value of one if a plant is owned by a foreign firm. Information on the nationality of the plant's owner is available from the IDBR.

Single

The single variable takes the value of one if a plant does not belong to a wider enterprise group. This variable is created by identifying whether the plant possesses a unique enterprise group code. If this is the case, the plant is defined as a single plant.

Industry Dummies

Industry dummies take the value of one when the plant belongs to a particular industry. These dummies are derived from the SIC 80 codes contained in the IDBR.

Region Dummies

The regional dummies equal one when a plant is located within a particular region. These dummies are calculated from the government office region variable contained within the IDBR.

RSA dummy

The variable shows whether the plant has received an RSA grant. The standard specification of the dummy is one which takes the value of one in the period in which an RSA grant was received and from that period onwards. The way in which plants were identified as having received a grant was described at length in section 2.

Local Authority Industry Share

To capture specialisation or localisation externalities, the share of industry output within the local authority of the plant is calculated. This variable is calculated using gross output data from the ABI, SIC codes from the IDBR and local authority codes derived from the IDBR.

Diversity

To measure externalities that arise from being situated near a diverse range of plants, the number of SIC codes within a local authority is calculated.

Industry Growth Rate

The industry growth rate is the growth of aggregate gross output in the industry in which the plant is operating. The aggregate industry gross output data is calculated by weighting plant output, to avoid inaccurate estimates due to the sampling frame of the ARD, and then aggregating weighted plant gross output across industry.

Displacement

The displacement variable is the proportion of total industry gross output produced by plants in their first year of operation. The data used to calculate this variable is also weighted in order to avoid inaccuracy due to the sampling frame of the ARD.

Herfindahl Index

The Herfindahl Index is a measure of concentration within an industry (Herfindahl, 1950). It is calculated by summing the squared share of industry output of each establishment within an industry. Under the assumption that the elasticity of demand does not vary too greatly across industry, the Herfindahl Index can be used as a measure of market power (see, for example, Cabral, 2000). It can therefore also serve as a measure of the level of competition in an industry with larger values indicating greater market power and less competition within the industry. The share of industry output is calculated from the gross output data discussed above while the establishment's industry is identified from SIC codes in the IDBR.

Time

The time trend is calculated from the year variable in the ARD.

Weights

Because the probability of being surveyed in the ABI is determined by the size of the reporting unit to which the plant belongs, the sample of plants is not representative of the population of plants. To avoid obtaining results that are not representative of all plants, the data must be weighted. The weights are the inverse of the probability of being in the sample, given the size of the plant. This can be calculated because the IDBR contains the population of plants in GB and their employment level.

5.5. Comparison of Treated and Untreated Plants

This section will describe the differences in characteristics between RSA recipients and non-assisted plants. Descriptive statistics will be presented first and then results from a probit model will be given.

Descriptive Statistics

Table 5.2 presents means and standard deviations of the key variables for Scottish plants. Counts vary slightly due to the unavailability of certain variables for some observations. Plants classified as RSA-assisted are plants that have received an RSA grant in that period or at some period in the past. The data is weighted so as to be representative of the population of plants.

Table 5.2: Mean and standard deviation of variables by RSA status

	Non-Assisted Plants			RSA-Assisted Plants			Difference in Means
	Mean	Standard Deviation	Count	Mean	Standard Deviation	Count	
Gross output	5591.10	31262.48	47,590	20155.82	72293.57	5,105	-14564.72***
GVA	2186.22	13932.03	47,590	4781.97	27560.56	5,105	-2595.75***
Employment	55.50	128.62	47,716	175.54	226.56	5,107	-120.04***
Intermediate Inputs	3404.89	22191.55	47,590	15373.85	70286.22	5,105	-11968.96***
Unadjusted Capital Stock (1980 prices)	1101.49	7728.11	47,026	4909.69	14575.13	5,086	-3808.20***
Adjusted Capital Stock (1980 prices)	873.70	6833.99	44,580	3570.86	11811.29	4,995	-2697.16***
Labour Productivity	189.65	20723.07	47,447	70.36	102.77	5,105	119.29
Capital Per Worker (1980 prices)	0.02	0.24	47,233	0.02	0.25	5,090	-0.00
Real Wage	18.59	355.46	47,444	14.66	13.09	5,106	3.93
Age	13.61	10.66	47,507	16.98	9.13	5,107	-3.37***
Foreign Direct Investment (Dummy)	0.13	0.36	47,716	0.21	0.35	5,107	-0.08***
Single (Dummy)	0.23	0.44	47,716	0.32	0.44	5,107	-0.09***
Local Authority Industry Share	0.60	0.86	47,715	0.61	0.86	5,107	-0.01
Diversity	15.91	13.16	47,645	11.91	9.22	5,102	4.00***
Industry Employment Growth	0.04	0.76	45,118	-0.01	0.33	4,932	0.05***
Displacement	0.12	0.41	47,650	0.07	0.11	5,102	0.05***
Herfindahl Index	0.10	0.11	47,716	0.07	0.07	5,107	0.03***

* denotes statistical significance at the 90% level, ** denotes statistical significance at the 95% level, *** denotes statistical significance at the 99% level
Unless stated otherwise, financial variables are measured in £1,000 and 2000 prices
Source: SAMIS/ARD

Firstly, RSA-assisted plants are far larger on average than non-assisted plants: they tend to produce almost four times as much output, have over twice as much GVA and have over three times as many employees. Unsurprisingly, they also use a lot more intermediate inputs (over four times as many) and have far larger capital stocks (also four times as large). These differences are all statistically significant at the 99% level. Given that the main objective of RSA is to promote and safeguard employment, it is not surprising that RSA recipients tend to be relatively large.

RSA-assisted plants are less productive on average with labour productivity levels only 37% as large as non-assisted plants although this difference is not statistically significant due to the large variance of labour productivity among non-assisted plants. Non-assisted plants also give higher real wages on average than RSA-assisted plants although the difference is far smaller than the difference in labour productivity and not statistically significant at the 90% level. These statistics jointly imply that the unit labour costs of RSA recipients are higher than those of plants that do not receive assistance and are suggestive of why they may need assistance to safeguard employment.

There is no statistically significant difference in the mean of the capital per worker variable. This may be a consequence of RSA being a scheme that aims to boost employment through supporting investment rather than a scheme that attempts to increase one or the other.

As would be expected given that one of the uses of RSA is to attract FDI, a significantly larger proportion of RSA plants are owned by foreigners. RSA-assisted plants are also less likely to be part of a larger enterprise than non-assisted plants. As would be expected given their greater size, RSA recipients are also older by a statistically significant amount.

Turning to variables that relate to the environment in which the plant operates, there is no statistically significant difference between the proportion of industry output located in the areas in which RSA-assisted and non-assisted plants are situated. These areas are however less diverse, by a statistically significant amount, in terms of the number of plants from different industries operating in that area. The average growth rate for the industry in which RSA-assisted plants operate is minus 1% while the corresponding figure for non-assisted plants is 4%. This difference is statistically significant and suggests that RSA recipients tend to operate in struggling industries. A large and significant difference also exists with regard to displacement with RSA-assisted plants tending to operate in industries with less displacement than non-assisted plants. Finally, RSA-assisted plants tend to be found in more competitive industries as indicated by the lower value of the Herfindahl index.

Probit Model of Determinants of Receipt of an RSA Grant

Table 5.3 presents results from estimating a probit model of the determinants of receiving an RSA grant in Scotland. The advantage of the probit model over the simple comparison of means provided in table 5.2 is that it permits the investigation of the relationship between the explanatory variables and the RSA dummy, holding other variables constant. The explanatory variables are employment, the capital to labour ratio, plant age, a single plant dummy, an FDI dummy, the Herfindahl Index, the local authority's share of industry output, a measure of the diversity of the local authority's industrial base and a time trend (these variables are similar to those that are used in Harris and Robinson, 2004). The dependent variable is a dummy that takes the value of one in the year in which the plant receives treatment but reverts to zero in the following years. Continuous variables are logged and the data is weighted so that the results are representative of the population of plants.

Table 5.3: Probit Model of Determinants of Receipt of an RSA Grant

	Coefficient	Robust Standard Error
Ln(Empment)	0.22***	0.03
Ln(Capital per worker)	0.01	0.01
Ln(Herf)	-0.11***	0.04
Ln(Industry share)	0.09***	0.03
Ln(Diversity)	-0.08*	0.04
Ln(Age)	0.06	0.05
Ln(Time)	0.04	0.03
Single	0.29***	0.05
FDI	0.08	0.07
SIC 22	0.44	0.28
SIC 24	0.00	0.14
SIC 25	0.44***	0.16
SIC 31	0.34***	0.12
SIC 33	0.37***	0.14
SIC 34	0.34***	0.11
SIC 35	0.02	0.18
SIC 36	-0.01	0.17
SIC 37	0.21	0.18
SIC 41	-0.01	0.14
SIC 43	0.13	0.12
SIC 44	0.37*	0.21
SIC 45	0.52***	0.15
SIC 46	0.10	0.14
SIC 47	0.02	0.11
SIC 48	0.32***	0.12
SIC 49	-0.04	0.19
Constant	-3.78***	0.33
Pseudo R-Squared	0.1582	
Observations	50,914	

Source: SAMIS/ARD

The results mostly confirm what would be expected from table 5.2:[33] larger plants are more likely to receive an RSA grant; plants operating in more competitive industries (which therefore have a lower Herfindahl index) have a higher probability of receiving an RSA grant; plants operating in areas with a higher share of their industry's output are more likely to receive assistance and single plants are more likely to receive an RSA grant, ceteris paribus. The coefficients of the capital per worker, age, time and FDI variables are not statistically significant at the 90% level. The lack of significance of the FDI dummy is the only major surprise as RSA is frequently used to attract FDI.

[33] Note that the dependent variable used for the probit model differs from the variable used to denote treatment status in table 5.2.

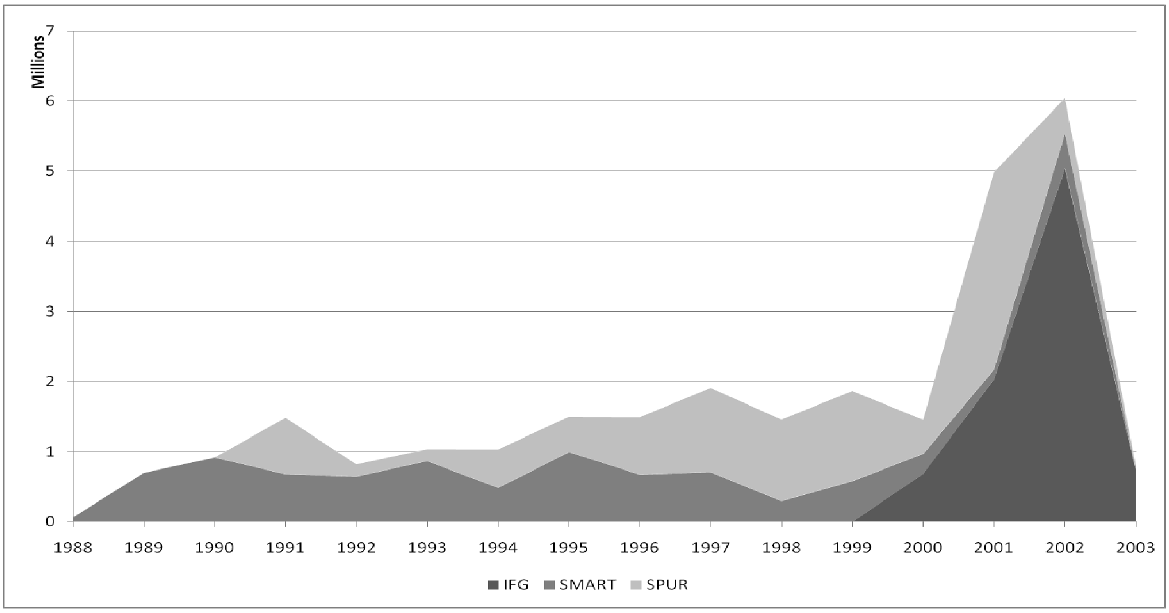
Plants operating in the following industries have a greater probability, at the 90% significance level, of receiving an RSA grant relative to the mechanical engineering industry (SIC 32) which serves as the baseline: chemical industry (SIC 25), metal goods not elsewhere specified (SIC 31), office machinery & data processing equipment (SIC33), electrical & electronic engineering (SIC 34), leather & leather goods (SIC 44), footwear & clothing industries (SIC 45) and processing of rubber & plastics (SIC 48).

5.6. Conclusion

This chapter has described the data that will be employed in the empirical analyses of chapters 6, 7 and 8. It began by providing descriptive statistics from the SAMIS database. These showed that RSA offers the largest number and the most generous grants of all the grant schemes in Scotland. The next section described how the SAMIS database was linked into the ARD. This process managed to link a higher proportion of RSA recipients into the ARD than previous researchers have managed which allows greater confidence in the results from the empirical analyses. A description of the variables that will be used in the empirical analysis was then given. The fifth section provided a comparison of the characteristics of RSA recipients with untreated plants. This showed that RSA recipients tended to be larger, older and were more likely to be foreign owned. They tended to operate in areas with a less diverse industrial base. The industries to which they belong tended to have lower employment growth, lower levels of displacement and greater competition. The probit model showed that, ceteris paribus, plant employment, the competitiveness of the industry in which the plant operates, the proportion of industry output produced in the plant's local authority and whether the plant is a single plant were all positive and statistically significant determinants of the probability of receiving an RSA grant.

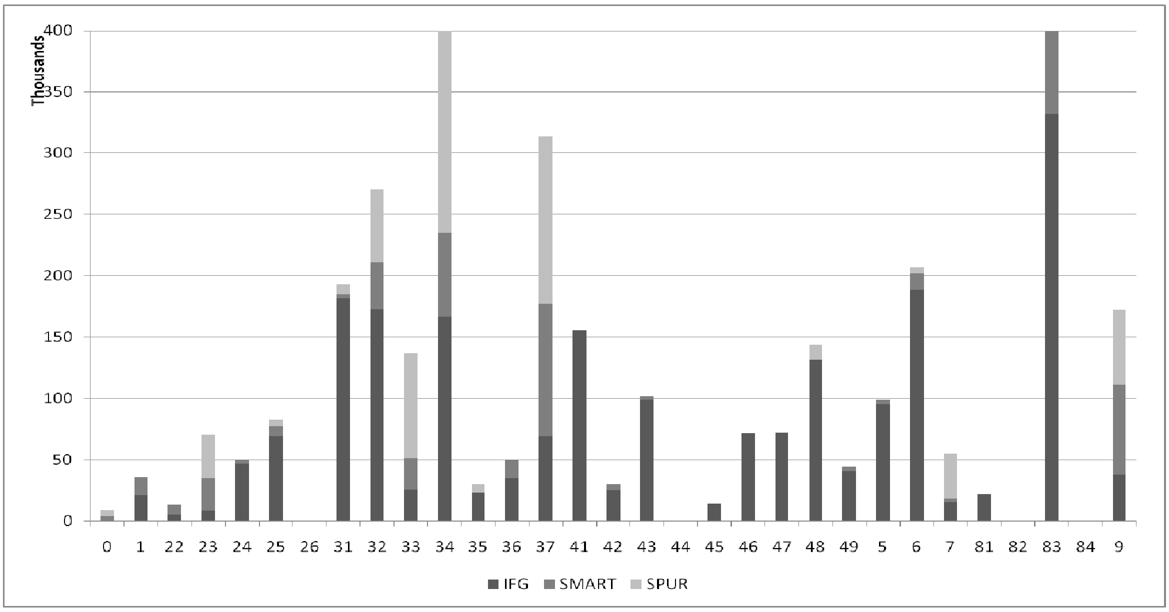
A5.1. Figures Showing the Value of Grants Awarded by Scheme excluding RSA by Year and Industry

Figure A5.1: Value of grants excluding RSA received in Scotland by year, 1988-2003 (2003 prices)



Source: SAMIS

Figure A5.2: Average value of grants excluding RSA received per year in Scotland by 2-digit industry, 1988-2003 (2003 prices)



Source: SAMIS

No figures are provided for SIC26 (production of man-made fibres) to avoid disclosure.

A5.2. SIC 80 codes

The following are the 1-digit SIC 80 codes used above to disaggregate by industry (Office for National Statistics, 1998):

0	Agriculture, forestry & fishing
1	Energy & water supplies
2	Extraction of minerals & ores other than fuels; manufacture of metals, mineral products & chemicals
3	Metal goods, engineering & vehicles industries
4	Other manufacturing industries
5	Construction
6	Distribution, hotels & catering (repairs)
7	Transport & communication
8	Banking, finance, insurance, business services & leasing
9	Other services

The following are the 2-digit SIC 80 codes used:

01	Agriculture & horticulture
02	Forestry
03	Fishing
11	Coal extraction & manufacture of solid fuels
12	Coke ovens
13	Extraction of mineral oil & natural gas
14	Mineral oil processing
15	Nuclear fuel production
16	Production & distribution of electricity, gas & other forms of energy
17	Water supply industry
21	Extraction & preparation of metalliferous ores
22	Metal manufacturing
23	Extraction of minerals not elsewhere specified
24	Manufacture of non-metallic mineral products
25	Chemical industry
26	Production of man-made fibres
31	Manufacture of metal goods not elsewhere specified

32	Mechanical engineering
33	Manufacture of office machinery & data processing equipment
34	Electrical & electronic engineering
35	Manufacture of motor vehicles & parts thereof
36	Manufacture of other transport equipment
37	Instrument engineering
41/42	Food, drink & tobacco manufacturing industries
43	Textile industry
44	Manufacture of leather & leather goods
45	Footwear & clothing industries
46	Timber & wooden furniture industries
47	Manufacture of paper & paper products; printing & publishing
48	Processing of rubber & plastics
49	Other manufacturing industries
50	Construction
61	Wholesale distribution (except dealing in scrap & waste materials)
62	Dealing in scrap & waste materials
63	Commission agents
64/65	Retail distribution
66	Hotels & catering
67	Repair of consumer goods & vehicles
71	Railways
72	Other inland transport
74	Sea transport
75	Air transport
76	Supporting services to transport
77	Miscellaneous transport services & storage not elsewhere specified
79	Postal services & telecommunications
81	Banking & finance
82	Insurance, except for compulsory social security
83	Business services
84	Renting of movables
85	Owning & dealing in real estate
91	Public administration, national defence & compulsory social security

92	Sanitary services
93	Education
94	Research & development
95	Medical & other health services; veterinary services
96	Other services provided to the general public
97	Recreational services & other cultural services
98	Personal services
99	Domestic services
00	Diplomatic representation, international organisations, allied armed forces

6. Productivity Growth Decomposition

6.1. Introduction

This chapter will decompose the growth of aggregate labour productivity and aggregate TFP of Scottish manufacturing between 1994 and 2004. This will reveal the contribution to productivity growth of five sources: productivity growth within continuing plants; redistributions of market share between continuing plants; productivity growth coinciding with increasing market share within continuing plants; entering plants and exiting plants. The share attributable to each source of productivity growth will also be split into shares accounted for by RSA grant recipients and plants that did not receive an RSA grant to give an indication of whether and how RSA grant recipients have contributed to or hindered aggregate productivity growth.

The relationship between the decomposition of aggregate productivity growth that will be performed in this chapter and the analysis that will be undertaken in chapter 7 requires clarification. The decomposition will show the contribution made by plants that received an RSA grant to aggregate productivity growth. The results will not be informative as to whether receipt of an RSA grant has a causal impact on plant productivity because the productivity performance of RSA-assisted plants is a reflection of numerous factors, many of which are unrelated to whether the plant received an RSA grant. The establishment of a causal relationship between receipt of an RSA grant and productivity is the subject of chapter 7.

The only paper that has examined the contribution of RSA recipients to aggregate productivity growth is by Harris and Robinson (2005). Also using a dataset created by merging SAMIS into the ARD,[34] they decompose the growth of labour productivity and TFP in UK manufacturing plants between 1990 and 1998. Their results show that RSA assisted plants made a proportionately large positive contribution to the growth of labour productivity but a negative contribution to the growth of TFP. The former effect is

[34] As discussed in chapter 5.3, their dataset is built upon to create the dataset used here and therefore contains more plants erroneously classed as not receiving an RSA grant.

primarily due to RSA grant recipients that improve their productivity between 1990 and 1998 also improving their market share while the latter is mostly due to RSA grant recipients with low TFP in 1990 increasing their market share. These results will be discussed in more detail below when they are compared with the results obtained here.

The next section sets out the method used to decompose the growth of aggregate productivity and how the measures of productivity are calculated; the third section will present the results from the decomposition and the final section concludes.

6.2. Decomposition Methodology

The index of aggregate productivity is calculated as follows:

$$\ln P_t = \sum_i \theta_{it} \ln P_{it}, \tag{6.1}$$

where θ_{it} is the share of aggregate output belonging to plant i at time t and P_{it} is the productivity of plant i at time t.

The geometric growth rate of aggregate productivity between 1994 and 2004 is:

$$\Delta \ln P_{2004} = \ln P_{2004} - \ln P_{1994}. \tag{6.2}$$

Following Haltiwanger (1997), this can be decomposed as follows:[35]

$$\begin{aligned} \Delta \ln P_{2004} = &\sum \theta_{i1994} \Delta \ln P_{i2004} + \\ &\sum \Delta\theta_{i2004} \left(\ln P_{i1994} - \ln P_{1994}\right) + \\ &\sum \Delta\theta_{i2004} \Delta \ln P_{i2004} + \\ &\sum \theta_{i2004} \left(\ln P_{i2004} - \ln P_{1994}\right) - \\ &\sum \theta_{i1994} \left(\ln P_{i1994} - \ln P_{1994}\right). \end{aligned} \tag{6.3}$$

The first term of the decomposition is the sum of the growth in plant-level productivity between 1994 and 2004 of plants that existed in both years, weighted by the plant's market share in 1994. It measures what aggregate productivity growth would have been, allowing for plant-level productivity growth within plants but holding market shares constant, and is termed the within component. The second term, which is the between component, captures what aggregate productivity growth would have been without plant-level productivity

[35] A justification for the choice of the Haltiwanger decomposition above other decompositions is provided in the appendix.

growth, entry or exit but allowing for reallocations in market share between continuing plants. It is the sum of the growth of continuing plants' market shares between 1994 and 2004, weighted by the deviation of the plant's 1994 productivity level from the average. The third term is the covariance component and is the sum of the change in continuing plants' productivity between 1994 and 2004 multiplied by the plant's change in market share between 1994 and 2004. It is a covariance effect measuring the contribution to the growth in aggregate productivity of the coincidence of increases in market share with improvements in productivity.

The fourth term gives the contribution to aggregate productivity growth of entrants. It is calculated as the sum of the deviation of each entrant's productivity from average productivity in 1994, weighted by its output share in 2004. A priori expectations are that this term will be positive as new plants will employ the latest technologies (see, for example, Samaniego, 2006). The final term enters negatively and is the sum of the deviation of exiting plants' productivity from average productivity in 1994, weighted by the plant's market share in 1994. Assuming that lower productivity plants are more likely to close, this term will be negative which implies a positive contribution of plants that exit to aggregate productivity growth.

In the decomposition presented below, each of the five components is decomposed further into two parts: one representing the component for plants that received an RSA grant between 1994 and 2004 and the other for plants that did not receive an RSA grant. It should be noted that this does not mean that the sample was first split into plants that received an RSA grant and plants that did not before performing the decomposition as this would not allow for reallocations in market share across plants that received a grant and those that did not and would therefore not give an accurate representation of the between and covariance components. Instead, the decomposition was performed using the full sample and then the components for plants that received a grant and those that did not were calculated.

Two measures of productivity will be considered here: labour productivity and TFP. Performing the analysis using these two measures of productivity allows a better understanding of the behaviour of RSA recipients. The natural logarithm of labour productivity is calculated as follows:

$$lp_{it} = y_{it} - e_{it}, \tag{6.4}$$

where y_{it} is GVA in plant i at time t and e_{it} is employment. Both variables are in logarithmic form.

TFP is calculated using the following log-linear Cobb-Douglas production function which is almost identical to that set out in chapter 7.2:

$$y_{it} = \beta_E e_{it} + \beta_K k_{it} + \beta_X x_{it} + (\eta_i + \upsilon_{it} + m_{it}), \tag{6.5}$$

where k_{it} represents the capacity utilisation adjusted capital stock (henceforth referred to simply as the capital stock) and x_{it} is a vector of variables thought likely to influence TFP.[36] All continuous variables are in logarithmic form. The only difference between equation (6.5) above and equation (7.5) is that the RSA dummy is excluded. This is done because this variable was not statistically significant in any of the models estimated in chapter 7.

The error term, ε_{it}, can be written as:

$$\varepsilon_{it} = \eta_i + \upsilon_{it} + m_{it}, \tag{6.6}$$

where η_i is an unobservable, plant-specific, time-invariant effect, v_{it} is a TFP shock which may be autoregressive, and m_{it} is a measurement error which is assumed to be serially uncorrelated.

Once the coefficients have been estimated, the logarithm of TFP can be calculated as follows:

$$\hat{tfp} = y_{it} - \hat{\beta}_E e_{it} - \hat{\beta}_K k_{it} = \hat{\beta}_X x_{it} + \hat{\varepsilon}_{it}. \tag{6.7}$$

Equation (6.5) is estimated in the same way as in chapter 7.3 where a detailed description of the estimation process is provided. Here only a brief summary is given. Because of the endogeneity of certain explanatory variables in equation (6.5), the model was estimated using the system GMM estimator. To allow for an autoregressive productivity shock which would otherwise invalidate the instruments used in the system GMM estimator, the following dynamic version of equation (6.5) was estimated:

$$y_{it} = \pi_1 y_{i,t-1} + \pi_2 e_{it} + \pi_3 e_{i,t-1} + \pi_4 k_{it} + \pi_5 k_{i,t-1} + \pi_6 x_{it} + \eta_i^* + \varepsilon_{it}, \tag{6.8}$$

[36] Information on these variables is provided in chapter 5.4.

where $\eta_i^* = (1-\alpha)\eta_i$ and $\varepsilon_{it} = \upsilon_{it} + m_{it} - \alpha m_{i,t-1}$. The long-run coefficients used to calculate TFP are calculated from the short-run coefficients in equation (6.8) as shown in chapter 7.3.

As discussed in chapter 7.4, it is undesirable to impose a common technology across disparate industries. As a result, equation (6.5) is estimated by 2-digit industry so that the coefficients are allowed to vary by industry.

6.3. Decomposition Results

In order to place the results in context, it is helpful to state that between 1994 and 2004, the GVA of Scottish manufacturing grew by an average of only 0.9% (Office for National Statistics, 2008). However, during this period, employment in Scottish manufacturing fell by a little over a quarter from almost 320,000 in 1994 (Office for National Statistics, 2009b). This implies a large increase in labour productivity to be decomposed below.

Table 6.1 gives the output shares, θ_{it}, attributable to each group. These are calculated using weighted data so that the output shares are representative of the population of plants rather than the stratified sample of plants selected for surveying in the ABI (see chapter 5.3 for more detail on the sampling frame of the ABI). Care was also taken to ensure that plants that were genuinely continuers but were only observed in either 1994 or 2004 were not incorrectly classified as exiters or enterers. It should be noted that the output shares attributable to each group are a function of the time that elapses between the chosen base and end years because the number of enterers and exiters is cumulative over time. This means that it is not possible to make direct comparisons between the output shares presented below and those presented in Harris and Robinson (2005) because their analysis decomposes the growth of productivity over eight years whereas the analysis below is concerned with productivity growth over ten years. The total number of plants in the sample was 2,703 in 1994 and 2,191 in 2004.

Table 6.1: Output Shares of Continuers, Exiters and Enterers in Scottish Manufacturing, 1994 and 2004 (%)

	Continuers (1994)	Exiters (1994)	Continuers (2004)	Enterers (2004)
Non-assisted	24.85 (20.94)	58.73 (75.58)	30.38 (25.83)	50.80 (70.10)
RSA-assisted	10.73 (2.18)	5.69 (1.29)	15.01 (2.69)	3.81 (1.37)
All	35.58 (23.12)	64.42 (76.88)	45.39 (28.53)	54.61 (71.47)

Figures in parentheses are the proportion of the total number of plants in that year in each group.
Source: SAMIS/ARD

Table 6.1 shows that, in 1994, almost two-thirds of output was produced by plants that would close before 2004, while in 2004, over half of output was produced by plants that had opened since 1994. This demonstrates the importance of entry and exit. The share of output attributable to continuers rose from slightly over a third to over 45% between 1994 and 2004. Comparing the output shares with the share of plants in each group shows that continuers tended to be larger than both exiters and enterers. Distinguishing now by RSA status, continuing plants that received an RSA grant produced a significant proportion of output (10.7% and 15.0% in 1994 and 2004 respectively). Exiters that received an RSA grant produced 5.7% of output in 1994 while the proportion of output produced by enterers that received a grant is small by comparison at just less than 4%. RSA-assisted plants are larger, on average, than non-assisted plants which accords with what was shown in the descriptive statistics of chapter 5.5.

Table 6.2 gives an indication of the differences in labour productivity and TFP across groups. The productivity indices are calculated as in equation (6.1) and are therefore in logarithmic form. Rather than present them in this form, the antilog is taken to provide a more easily interpretable measure of productivity.

Table 6.2: Productivity Indices of Continuers, Enterers and Exiters in Scottish Manufacturing Plants, 1994 and 2004 (£ thousands)

	Continuers (1994)	Exiters (1994)	Continuers (2004)	Enterers (2004)
	Labour Productivity			
Non-assisted	42.52	51.42	49.40	85.63
RSA-assisted	40.04	66.02	60.34	42.10
	TFP			
Non-assisted	22.42	27.11	26.84	34.81
RSA-assisted	17.46	46.99	24.29	45.60

Source: SAMIS/ARD

The table shows that for continuers in 1994, the output-weighted labour productivity of RSA-assisted plants was 6% lower than the corresponding measure for non-assisted plants. However, for the exiters, the index for RSA-assisted exiters is over 25% larger than the index for non-assisted plants. In fact, the productivity index for RSA-assisted exiters is only surpassed by the index for enterers in 2004 which is surprising as it would be expected that exiters have the lowest productivity of all the groups. This is a consequence of weighting by output shares as the unweighted mean of labour productivity for the exiters is only £15,640 per worker - the lowest unweighted mean of all the groups. This implies that a particularly strong positive relationship exists between output shares and labour productivity in this group. It is worth noting that Harris and Robinson (2005) also find that RSA-assisted exiters have the highest labour productivity index of all the groups in the base year.

By 2004, the labour productivity indices of non-assisted and RSA-assisted continuers have grown from their 1994 levels by 16% and 51% respectively so that the index for RSA-assisted continuers is now over 20% larger than the index for non-assisted continuers. The largest difference between non-assisted and RSA-assisted plants is found in the enterers. The index for non-assisted enterers is over twice as large as the corresponding index for RSA-assisted enterers. Very similar results are also found by Harris and Robinson (2005) for the continuers and enterers.

The situation is rather different from that found using labour productivity when TFP is used to measure productivity. In 1994, the index for continuers that received an RSA grant is less than 80% as large as the index for continuers that did not receive a grant. By contrast, exiters that received an RSA grant have an output-weighted productivity index which is almost 75% greater than non-assisted exiters and this is the largest index of all the groups considered. Again, this is found to be the product of weighting productivity by output shares as the unweighted mean of TFP would give this group the lowest TFP index. Harris and Robinson (2005) also find that RSA-assisted exiters have a higher index of TFP than non-assisted exiters but, unlike in table 6.2, the index is lower than that for both assisted and non-assisted continuers in the base year and for entering plants in the end year which accords better with a priori expectations.

By 2004, the index of TFP for RSA-assisted continuers remains lower than the index for non-assisted continuers although it is now only 10% smaller. This is the result of non-assisted continuers improving their TFP by 20% while RSA-assisted continuers increased their TFP by 39%. Harris and Robinson (2005) do not replicate this finding as their index of TFP is lower for continuers in 1990 than in 1998 although, as in table 6.2, they do find a slight TFP advantage towards non-assisted plants in both years. This is the first instance where the ranking of non-assisted and RSA-assisted plants differs between the two indices of productivity and this is explained by the large difference between the two groups in the amount of capital employed (see Appendix A6.2 for employment and capital indices that explain the differences in ranking across the two measures of productivity).

Finally, the output-weighted TFP index is 31% larger for RSA-assisted enterers than exiters. Again, this is the reverse of what is seen using labour productivity and is explained by the fact that enterers that receive an RSA grant have far larger capital stocks than enterers that do not receive a grant. This contradicts the finding of Harris and Robinson (2005) that non-assisted enterers had a higher index of TFP than RSA-assisted enterers.

Table 6.3 gives the results from the decomposition of labour productivity growth between 1994 and 2004. As shown in equation (6.3), the size of the within, entering and exiting components and therefore the continuers and total for both plants that received an RSA grant and plants that did not are partly determined by their output shares. These output shares, given in table 6.1, should therefore be borne in mind when assessing what represents a strong performance from RSA-assisted plants.

Table 6.3: Decomposition of Labour Productivity Growth between 1994 and 2004 (%)

	Total (A+B+C+D-E)	Within (A)	Between (B)	Covariance (C)	Continuers (A+B+C)	Entrants (D)	Exiters (E)
Non-assisted	28.85	0.01	-2.47	6.26	3.80	28.58	3.53
RSA-assisted	3.02	2.65	-6.44	9.07	5.28	-0.53	1.74
All	31.87	2.66	-8.91	15.33	9.08	28.05	5.27

Source: SAMIS/ARD

Overall, plants that received an RSA grant made a positive contribution towards aggregate labour productivity growth. Without these plants, aggregate labour productivity would have grown by 28.9% instead of 31.9%. However, the contribution to aggregate labour productivity from RSA-assisted plants is smaller than what would be expected, given their share of output in 1994.

The within-plant contribution to aggregate productivity growth is 0.01% from non-assisted plants and 2.7% from RSA-assisted plants. This indicates that continuers that received an RSA grant substantially improved their labour productivity between 1994 and 2004 while non-assisted plants barely improved their labour productivity at all. Taking account of the relative shares of output in 1994, this represents a very strong performance by RSA-assisted continuers relative to non-assisted continuers. The between plant component is negative for both non-assisted and RSA-assisted plants, indicating that in both groups, the market shares of plants with high productivity in 1994 tended to fall compared to the market share of plants with low productivity. By contrast, the covariance effect is large and positive for both non-assisted and RSA-assisted continuers at 6.3% and 9.1% respectively, which shows that continuers that improved their productivity between 1994 and 2004 also tended to increase their output shares. The performance of plants that received an RSA grant is particularly strong in this regard. In sum, therefore, the relative performance of RSA-assisted continuers in terms of their contribution to aggregate labour productivity growth is impressive.

The largest contributor to aggregate labour productivity growth by far is entrants that did not receive assistance who provide 28.6% to the growth in labour productivity. In contrast, the contribution from entrants that received an RSA grant is negative, albeit very small. This is a major concern as it suggests that RSA grants are being provided to a poorly performing subset of entrants.

Finally, the exiters' component for both non-assisted and RSA-assisted plants is, contrary to expectations, positive and therefore, in accordance with equation (6.4), these plants make a negative contribution to aggregate productivity growth. Given their relatively small share of output, the proportion of the exiters' component accounted for by RSA-assisted plants is particularly large. This is consistent with the objective of supporting high productivity plants although the fact that they closed also suggests that their subsequent performance was poor. However, this finding is the result of the strong positive relationship between output shares and labour productivity within this group mentioned earlier and may therefore be considered to be somewhat of an anomaly.

Table 6.4 provides the results from the decomposition of TFP growth. It should be noted that some plants are lost when TFP is used as a measure of productivity instead of labour productivity due to the greater number of variables needed to produce estimates of TFP. Therefore, these results are produced using 2,530 plants in 1994 and 2,037 in 2004 which is slightly fewer than were used to produce the results in table 6.3.

Table 6.4: Decomposition of TFP Growth between 1994 and 2004 (%)

	Total (A+B+C+D-E)	Within (A)	Between (B)	Covariance (C)	Continuers (A+B+C)	Entrants (D)	Exiters (E)
Non-assisted	14.36	-2.88	1.86	6.82	5.8	12.58	4.03
RSA-assisted	2.27	1.31	-3.11	5.36	3.56	2.25	3.54
All	16.63	-1.57	-1.25	12.18	9.36	14.83	7.57

Source: SAMIS/ARD

The overall contribution of plants that received an RSA grant to the growth of TFP is 2.3% compared to a contribution of 14.4% from plants that did not receive a grant. It is therefore slightly smaller than what would be expected given their share of output in 1994.

The within component is negative for continuers that did not receive a grant and positive for continuers that did receive a grant. This indicates that RSA-assisted continuers increased their TFP while non-assisted continuers experienced a fall in TFP. This represents a strong performance from RSA-assisted plants and is similar to the labour productivity case in which the within component was far larger for RSA-assisted than non-assisted plants. The situation is reversed when considering the between component with a contribution to aggregate TFP growth of 1.9% from non-assisted continuers and a negative contribution of 3.1% from RSA-assisted continuers. This shows that those non-assisted continuers with initially high TFP in 1994 increased their output share while RSA-assisted continuers with high TFP in 1994 experienced a fall in their share of output. The covariance component is positive for both non-assisted and RSA-assisted continuers at 6.8% and 5.4% respectively, showing that, for both groups, improvements in TFP coincided with increases in output share. Overall, the contribution from continuers that received an RSA grant is 3.6% which, considering output shares in 1994, is a relatively large part of the total contribution from continuing plants to the growth of aggregate TFP.

The largest contribution to aggregate TFP growth comes from entering plants that did not receive an RSA grant which contributed 12.6% to the growth of aggregate TFP. However, unlike with labour productivity, the contribution from RSA-assisted entering plants is positive at 2.3% which is a larger part of the total enterers' component than would be expected given their share of output in 2004. As with labour productivity, the contribution to aggregate TFP growth of both non-assisted and RSA-assisted exiting plants is negative at 4.0% and 3.5% respectively because the weighted average of TFP of both these groups of plants is higher than the average in 1994. Taking account of the relative output shares of these groups, this represents a very large part of the total negative contribution of exiters to aggregate TFP growth from plants that received an RSA grant but this is once more the consequence of the relatively strong relationship between TFP and output in this group.

Overall, these results show that plants that received an RSA grant did not make as large a contribution to either labour productivity or TFP growth as their output shares in 1994 suggest they ought to have made although the size of their contribution is slightly more impressive when the measure of productivity is TFP rather than labour productivity.

The proportion of the total contribution that comes from RSA-assisted plants that were in operation in both 1994 and 2004 is larger than what would be expected, given their share of output, and most of this contribution is made through the process by which plants that improve their productivity also improve their output shares. This implies that continuers that receive an RSA grant tend to be more dynamic than plants that do not receive support. However, entrants that received an RSA grant made a negative contribution of 0.5% to the growth in aggregate labour productivity but a positive contribution of 2.3% to the growth in aggregate TFP. This is the result of RSA-assisted entrants having higher TFP than average but lower labour productivity than average which, unsurprisingly given the main objective of the RSA scheme, suggests that RSA recipient are more labour intensive than the average. Given the very large contribution from non-assisted entrants, this suggests that the entrants need to be better targeted for receipt of RSA grants if the scheme is to make a larger contribution to the growth of productivity. Turning to the exiters, there is a large and negative contribution to aggregate productivity growth from exiters that received a grant due to their high productivity indices. However, this is slightly misleading as RSA-assisted entrants are, on the basis of an unweighted average, the least productive using either measure of productivity. It is only that there is a particularly strong positive relationship between the share of output and productivity within this group.

The results are in many ways similar to those obtained by Harris and Robinson (2005) despite differences in the area, length and period of time analyzed and the quality of the dataset as well as slight differences in the methodology. Overall, they also found that RSA-assisted plants made a positive contribution to productivity growth although the contribution is over twice as large as that identified here. In terms of the sources of this contribution, in accordance with the results presented in table 6.3, they found that RSA-assisted plants made a small contribution to aggregate labour productivity growth through the within component, a large positive contribution through the covariance component (although less than half the size of that identified here), a negative contribution through the entering plants' component and a negative contribution through the exiting plants' components. The results differ in that they find that RSA-assisted continuers made a positive contribution through the between effect which indicates that RSA-assisted plants with high levels of labour productivity tended to increase their market share between 1990 and 1998.

Turning to the decomposition of TFP growth, Harris and Robinson (2005) find that RSA-assisted plants contributed negatively whereas table 6.4 shows that RSA-assisted plants contributed positively to aggregate TFP growth. The within and between components are similar in both sign and magnitude but, unlike the results in table 6.4, they find that enterers made a small negative contribution to aggregate TFP growth which is the result of RSA-assisted enterers being less productive than average while exiters made a small positive contribution because they were less productive than the average. The largest difference is in the covariance effect as Harris and Robinson find that RSA-assisted plants make a contribution of only 0.8% to aggregate productivity growth whereas table 6.4 shows a contribution of 5.4%. This means that UK manufacturing plants that improved their TFP between 1990 and 1998 were less successful at improving their market share than Scottish manufacturing plants that improved their TFP between 1994 and 2004.

6.4. Conclusion

This chapter has decomposed the growth of aggregate labour productivity and aggregate TFP in Scotland between 1994 and 2004 in order to identify the extent to and the channels through which plants assisted by RSA contributed to productivity growth. The decomposition showed that RSA-assisted plants made a positive but small contribution, relative to the output share of these plants in 1994, to the growth in aggregate labour productivity and TFP. For both labour productivity and TFP, the bulk of this contribution came from plants that operated in both 1994 and 2004 and, more specifically, through the process by which plants that improve their productivity also improve their market share. This suggests that continuers that receive an RSA grant are more dynamic than continuers that do not receive support.

The contribution of entrants that received an RSA grant is negative to aggregate labour productivity growth and positive and small, relative to output shares in 2004, to aggregate TFP growth which implies, in accordance with the objectives of the scheme, that RSA-assisted entrants are relatively labour intensive. The small contribution to productivity growth from RSA-assisted entrants is a major concern as the largest contribution to the growth of both measures of productivity comes from non-assisted entrants. This suggests that RSA grants should be better targeted at entering plants.

A6.1. Decomposition Methods

There are three competing methods that can be employed to decompose the growth of productivity (Disney, Haskel & Heden, 2003a). A description of the Haltiwanger (1997) approach has been provided above. This appendix will describe the other two approaches and explain why the Haltiwanger approach was chosen instead of them.

Bailey, Hulten and Campbell (1992) suggest a decomposition of the form:

$$\begin{aligned}\Delta \ln P_t = {} & \sum \theta_{it} \Delta \ln P_{it} + \\ & \sum \Delta\theta_{it} \ln P_{it} + \\ & \sum \theta_{it} \ln P_{it} - \\ & \sum \theta_{it-k} \ln P_{it-k}.\end{aligned} \tag{A6.1}$$

The first term is identical to the within component in the Haltiwanger decomposition. It shows what the contribution to productivity growth of continuing plants would have been, had output shares been held constant. However, while the Haltiwanger decomposition has two terms – the between and covariance components - which jointly capture the contribution to aggregate productivity of changes in the share of continuing plants' output between the base and end years, there is only one such term in equation (A6.1). This term is calculated as the sum of the change in output shares multiplied by the productivity of the continuers. It therefore captures both the contribution to productivity growth of plants with initially higher productivity increasing their market share and the contribution of plants that improve their productivity between the base and end year and also increase their share of output. The ability of the Haltiwanger decomposition to provide more information on the nature of the contribution from change in output shares is a clear advantage.

The contribution to productivity growth from exiters in equation (A6.1) is simply the sum of the productivity of entrants in the end year weighted by their output shares while the contribution from exiters is the sum of the productivity of exiters in the base year weighted by their output shares. As shown by Haltiwanger (1997), even if entrants are more productive than exiters, this does not guarantee that the net contribution from entry and exit, calculated using equation (A6.1) is positive if the output share of entrants is smaller in the end year than the output share of exiters in the base year. This is an undesirable property and one which the Haltiwanger decomposition and the decomposition discussed next does not share.

The following decomposition is proposed by Griliches and Regev (1992):

$$\begin{aligned}\Delta \ln P_t = &\sum \bar{\theta}_i \Delta \ln P_{it} + \\ &\sum \Delta \theta_{it} \left(\ln \bar{P}_i - \ln \bar{P}\right) + \\ &\sum \theta_{it} \left(\ln P_{it} - \ln \bar{P}\right) - \\ &\sum \theta_{it-k} \left(\ln P_{it-k} - \ln \bar{P}\right).\end{aligned} \tag{A6.2}$$

where a bar denotes the average calculated over the end and base year. The main difference between this approach and the Haltiwanger approach is that the latter uses deviations of productivity from the average in the base year whereas this approach employs deviations of productivity from the average over both base and end year. This has the advantage of making equation (A6.2) less sensitive to measurement error than the Haltiwanger approach (Disney, Haskel & Heden, 2003a).

The first term is the sum of the growth in the productivity of continuers, weighted by their average market share over the base and end year. The second component is the change in market shares of continuers multiplied by the deviation in average plant productivity over base and end year from average aggregate productivity over the base and end year. The entrants' term is the output share of plants in the end year multiplied by the deviation in their productivity levels from aggregate productivity over base and end year and the exiters' contribution is calculated as the output share of exiters in the base year multiplied by the deviation of their productivity from aggregate productivity over base and end year.

The major problem with equation (A6.2) is that the first component does not represent a recognisable component of productivity growth. It cannot be regarded as the as the contribution from continuers that improve their productivity, holding output shares constant because $\bar{\theta}_i$ is determined by the growth of market shares between the base and the end year. For this reason, the Haltiwanger decomposition is judged to be a better way of decomposing the growth of productivity and was used above.

A6.2. Employment and Capital Indices

To assist in the understanding of differences between the two measures of productivity, table A6.1 provides the weighted average of employment and capital for each group. Assuming the coefficients on employment and capital in the production function are greater than zero but less than one, equations (6.4) and (6.7) show that plants with a relatively high index of employment would prefer to be compared on the basis of TFP, ceteris paribus. On the other hand, these equations also show that plants with a relatively high index of capital would prefer to be compared on the basis of labour productivity, ceteris paribus. When one group have both more employment and capital, it is not immediately clear which measure of productivity provides a more favourable comparison. This will depend upon the coefficients on employment and capital in the production function.

Table A6.1: Employment and Capital Indices, 1994 and 2004

	Continuers (1994)	Exiters (1994)	Continuers (2004)	Enterers (2004)
	Employment			
Non-assisted	652.94	355.70	472.61	399.20
RSA-assisted	996.04	1039.74	694.12	483.59
	Capital			
Non-assisted	12713.69	10184.64	9446.06	20311.78
RSA-assisted	17534.60	8327.98	23994.05	5641.72

Capital is measured in measured in £1,000 (1980 prices).
Source: SAMIS/ARD

The employment index for RSA-assisted continuers in 1994 is over 50% larger than the corresponding index for non-assisted continuers. The capital index for RSA-assisted plants is almost 40% larger than the index for plants that did not receive an RSA grant. Table 6.2

shows that RSA-assisted continuers are at less of a productivity disadvantage in a labour productivity comparison than a TFP comparison. In accordance with equations (6.4) and (6.7), this implies that the disadvantage to this group that arises in a TFP comparison from having a larger capital stock is outweighed by the lower weight that is placed on its larger level of employment.

For exiters in 1994, the index of employment is almost three times larger for RSA-assisted than non-assisted plants but the capital index is over 20% larger for non-assisted plants. This implies that RSA-assisted plants would prefer to be compared on the basis of TFP and this is indeed the case as table 6.2 shows a larger productivity advantage for RSA-assisted exiters than non-assisted plants using TFP as the measure of productivity.

Continuers in 2004 are similar to continuers in 1994 in that RSA-assisted plants have a higher weighted average for both employment and capital. The former is larger by almost 50% while the latter is larger by over 150%. The TFP index is again more favourable to RSA-assisted plants as it ranks them above non-assisted plants while non-assisted plants have a higher index of labour productivity. The explanation for this is identical to that set out for continuers in 1994.

Finally, the employment index for RSA-assisted enterers in 2004 is over 20% larger than the corresponding index for non-assisted enterers while the capital index for non-assisted enterers is almost four times larger for non-assisted enterers. This suggests that a comparison based on TFP would favour RSA-assisted plants and this is precisely what is shown in table 6.2 where RSA-assisted entrants have a far lower index of labour productivity but a larger index of TFP.

7. The Causal Impact of RSA on Productivity

7.1. Introduction

This chapter will examine whether receipt of an RSA grant has a causal impact on plant TFP. To tackle the problem of self-selection into the treatment group outlined in chapter 3.2, propensity score matching and instrumental variables will be employed. In order to control for the endogeneity of other variables in the model, all estimations will be performed using the system GMM estimator.

There are two measures of productivity upon which the impact of receipt of an RSA grant could be measured: labour productivity and TFP. Labour productivity is calculated by dividing output or GVA by employees. TFP measures the contribution to output not attributable to factor inputs and, as such, captures technology and efficiency. TFP is chosen instead of labour productivity as the measure of productivity because the latter, unlike the former, is determined by factor input levels in addition to levels of efficiency and technology (Harris, 2005b). To be precise, more capital or higher TFP leads to higher labour productivity while higher employment is associated with lower labour productivity under the assumption of diminishing returns to labour. As RSA is a capital grant scheme which has the promotion and safeguarding of employment as its main aim, any estimated impact of RSA on labour productivity will be the sum of the impact of an RSA grant on employment, capital and TFP. It will therefore be more difficult to interpret than when TFP is used as the measure of productivity.

As discussed in chapter 4.2, there are two main channels through which an RSA grant may improve TFP. The first is by allowing the acquisition of modern capital which demands the reorganisation of the plant along more efficient lines. The second is by allowing the plant to create a new product that can be produced with greater efficiency than older product lines.

As shown in chapter 4.4, previous studies have generally failed to find a statistically significant impact of receipt of an RSA grant on productivity. One exception is Harris and Robinson (2004) who, using a control group consisting of all untreated plants in GB, find a positive and statistically significant impact on TFP. However, this result is not replicated when the control group consists of untreated plants from the assisted areas, which implies that untreated plants in the assisted areas performed better, in terms of productivity, during the period under investigation. Criscuolo, Martin, Overman and Van Reenen (2007), using an instrumental variable approach which is partially replicated below, also do not find a significant effect of RSA on either labour productivity or TFP. Finally, Hart, Driffield, Roper and Mole (2008a), using the control functions approach on data taken from a telephone survey of Scottish firms, do not find a positive impact of receipt of an RSA grant on labour productivity. While the latter two papers use appropriate methods to tackle the consequences of self-selection into the treatment group, they fail to deal with the endogeneity of other explanatory variables. Harris and Robinson (2004), by contrast, do not employ a sufficiently sophisticated method to control for self-selection but do control for the endogeneity of control variables. The analysis below tackles both sources of bias.

The next section will set out the econometric model that will be estimated. The third section will describe how the propensity score matching and instrumental variables estimators, described in chapter 3.3, are implemented and discusses the system GMM estimator which is used to handle the endogeneity of control variables in the model. The fourth section presents the results and the final section concludes.

7.2. Econometric Model

Consider the following Cobb-Douglas production function (Cobb and Douglas, 1928):

$$Y_{it} = A_{it} E_{it}^{\beta_E} K_{it}^{\beta_K}, \tag{7.1}$$

where Y_{it} is GVA in plant i at time t, E_{it} represents employment, K_{it} represents the capacity utilisation adjusted capital stock and A_{it} is TFP. Taking natural logs of equation (7.1) gives:

$$y_{it} = \beta_E e_{it} + \beta_K k_{it} + a_{it}, \tag{7.2}$$

where the lower case is used to denote the natural logarithm of a variable.

It is now postulated that the natural logarithm of TFP can be modelled as follows:

$$a_{it} = \beta_X x_{it} + \beta_{ATT} D_{it} + (\eta_i + \upsilon_{it} + m_{it}), \tag{7.3}$$

where x_{it} is a vector of variables thought to influence TFP (in which continuous variables are logged) and D_{it} is a dummy taking the value of one if a plant receives an RSA grant in that period or has done so in the past. The error term is composed of η_i, an unobservable, plant-specific, time-invariant effect, v_{it}, a TFP shock, and m_{it}, a measurement error which is assumed to be serially uncorrelated. The RSA dummy is the key variable in the model as its coefficient, β_{ATT}, will provide the estimate of the impact of receiving an RSA grant on TFP.[37]

The TFP shock takes the following form:

$$\upsilon_{it} = \alpha \upsilon_{i,t-1} + e_{it} \qquad |\alpha < 1|. \tag{7.4}$$

It is autoregressive if $\alpha \neq 0$.

The model is therefore:

$$y_{it} = \beta_E e_{it} + \beta_K k_{it} + \beta_X x_{it} + \beta_{ATT} D_{it} + (\eta_i + \upsilon_{it} + m_{it}). \tag{7.5}$$

The x_{it} variables are included in equation (7.5) to avoid a biased estimate of the ATT caused by observed variables that are correlated with both the treatment variable and the error term (see chapter 5.4 for descriptions of all the variables used in the model). The first of these variables is the Herfindahl Index (Herfindahl, 1950), calculated at the four-digit industry level. The Herfindahl Index is a measure of the concentration of output and hence competition within an industry. Intuitively, one would expect that greater competition (which implies a lower Herfindahl index) demands that plants operate more efficiently. However, it is arguable that the level of competition may be inversely related to productivity if monopoly rents are required for management to invest in R&D which in turn leads to improvements in TFP (Aghion, Harris, Howitt and Vickers, 2001).

The next two variables included in x_{it} represent attempts to measure two types of agglomeration externalities. Agglomeration externalities are reductions in costs or improvements in productivity which accrue to plants located in the vicinity of other plants. The first type of such externalities is localisation externalities. These arise due to the concentration of plants from a particular industry in a given area and are termed Marshallian externalities (Marshall, 1890). By contrast, Jacobian externalities arise as a

[37] It is clear that receiving an RSA grant will have an impact upon employment and capital as well as on TFP. Therefore, the coefficient on the RSA dummy cannot be interpreted as the full impact of receiving an RSA grant on output. As interest here lies entirely in its impact on TFP, this presents no problems.

result of diversity in the activities of plants in a particular area (Jacobs, 1969). The different types of externality are a reflection of difference views of information spillovers. Marshall believed that information spillovers arose primarily within industries while Jacobs believed that they arose primarily between industries (Van Der Panne, 2004). In what follows, Marshallian externalities are captured by a variable measuring the proportion of industry output located within the local authority. Jacobian externalities are measured by a variable calculated as the number of different SIC codes within the local authority.

An age variable is also included in x_{it}. A priori, it is not immediately clear whether older plants will have higher TFP. On one hand, old plants are less likely to be employing the most modern technologies. On the other hand, they may have higher TFP as their survival indicates that they are the best of their cohort of plants.

A time trend is also included to control for common improvements in TFP through time. The model was also run using year dummies but this made no substantive difference to the estimate of the ATT.

A single plant dummy, equal to one if that plant is the only plant owned by the firm, is also included in x_{it}. If technology is shared within multi-plant enterprises, this may confer a TFP benefit to being part of a multi-plant enterprise which would imply that the coefficient on the single plant dummy should be negative.

A foreign ownership dummy was also included in the model. This is justified by the observation that multinationals must possess advantages that allow them to overcome the costs of operating in a foreign country (Hymer, 1976). If one of these advantages is superior technology, foreign owned plants would have higher TFP (Harris and Li, 2007). It may also be hypothesised that foreign owned plants are a self-selected group of the population of plants as multinationals tend to acquire plants that have high levels of TFP. On the other hand, domestic plants may experience difficulties in adjusting to the technologies of the multinational owner which implies that FDI plants may have low levels of TFP (Kronborg and Thomsen, 2008). It is therefore not obvious whether FDI plants should have higher or lower TFP than other plants in the population.

7.3. Estimation Strategy

As discussed at length in chapter 3.2, the major econometric issue that must be tackled when estimating the impact of an RSA grant is that of self-selection into the treatment group. If all variables that determine both treatment status and the outcome variables are observed, the only problem with estimating equation (7.5) using OLS is that of misspecification. To deal with this, propensity score matching will be used. On the other hand, if there are unobservable variables that determine both treatment status and the outcome variable, the matching estimator will not provide unbiased estimates of the treatment effect. To overcome the problems posed by the potential existence of such variables, an instrumental variables approach will be adopted. Regardless of the approach used to tackle the consequences of self-selection into the treatment group, it is necessary to control for correlation between the covariates and the error term in equation (7.5). This will be done using the system GMM estimator.

Propensity Score Matching

The major difficulty in creating a matched sample when the dataset is a panel is to avoid matching on variables that are themselves affected by treatment status (chapter 3.3 provides a more detailed discussion of why this is important). If this is not done, treated and untreated plants are matched on the basis of characteristics that are similar only because of differences in treatment status. Had they had the same treatment status, they would have different characteristics and therefore different values of the outcome variables. This is therefore a violation of the conditional independence assumption which underpins the matching estimator. When the dataset is cross-sectional, the researcher can either ignore this issue or only match on variables that are not affected by treatment status. When the dataset is a panel, treated observations can be matched to untreated observations using data from the period prior to treatment. This is the approach taken here for those variables that are affected by treatment.

The first stage in performing propensity score matching is to estimate the propensity score. The probit model that is used to estimate the propensity score is:

$$ID_{it} = \gamma_1 end_{it-1} + \gamma_3 exo_{it} + \varepsilon_{it}, \qquad (7.6)$$

where end_{it-1} is a once lagged vector of variables that may be affected by treatment and exo_{it} is a vector of variables that are not affected by treatment. Included in end_{it} are employment, the capital stock, the Herfindahl Index and the Marshallian specialisation externalities variables. exo_{it} consists of Jacobian diversification externalities; the single plant dummy, the foreign ownership dummy and government office region dummies. ID_{it} is a dummy taking the value of one in the year in which a plant receives an RSA grant for the first time. This is different from D_{it} which takes the value of one if a plant receives an RSA grant in that period or has done so in the past.

The matching proceeds by year on the basis of the predicted values from equation (7.6) so that treated plants are only matched to untreated plants from the same year. This is done because matched plants are supposed to act as matches for each other throughout the sample time period. The same treated plant cannot be matched to different plants in different years because this would involve matching on variables that are affected by treatment after the treatment has occurred. It is however impossible to avoid 'broken' matches with this approach as plants (treated or untreated) may become unobserved due to closure or being removed from the selected group of plants in the ARD. This is one reason why a multivariate regression approach is used on the matched sample as this will control for remaining differences in the distributions of the observables across the treatment and control groups.

The precise form of matching was nearest neighbour matching where treated plants were matched to 50 untreated plants.[38] Such a large number of treated plants were chosen because of the relatively few plants in the sample that received RSA and the need for many observations to facilitate the use of the system GMM estimator which requires a large number of observations. The cost of so many neighbours is that the treated and control group plants are not as well matched as they would be with fewer matches. Again, this provides further justification for estimating a multivariate regression rather than simply a comparison of means across treatment and control groups because differences in the distributions of the observed covariates are likely to remain in the matched sample.

[38] Propensity score matching is performed in STATA 9.2 using the 'psmatch2' command developed by Leuven and Sianesi (2003).

As our focus is on the impact of RSA on Scottish manufacturing plants, the treatment group always consists exclusively of Scottish plants. However, the ARD contains information on plants from throughout GB and these are allowed to form part of the control group in the matched sample. The main advantage of this approach is that, in the matched sample, plants in the treatment group should be more closely matched to those in the control group in terms of the observed covariates in equation (7.5) than would be the case had only Scottish plants been available to form the control group. The disadvantage of this approach is that if non-Scottish plants have unobserved characteristics that differ from those in the Scottish treatment group, this will generate a bias in the estimate of the ATT. It should however be noted that the inclusion of region dummies in (7.6) reduces the probability that non-Scottish untreated plants are in the matched sample so that this problem, if it exists at all, should not be too severe.

Instrumental Variables

Finding a genuine IV is the major obstacle to implementing the IV estimator. Here, the instrument that is used is the location of the plant, AA_{it}, defined as a dummy variable which takes the value of one if the plant is situated in an assisted area and zero otherwise. This approach is based on that taken by Criscuolo, Martin, Overman and Van Reenen (2007).

The difference between the approach taken by Criscuolo, Martin, Overman and Van Reenen (2007) and that taken here is that they use a series of dummies that equal one when the plant is located in different tiers of the assisted areas in which different proportions of project costs can be covered by RSA grants. This is not done here because some of these instruments will not satisfy the monotonicity assumption, discussed in chapter 3.3, which must be satisfied to estimate the LATE. This assumption requires that if some plants are induced to apply and receive a grant because the dummy equals one, no plants can be induced not to apply and receive a grant because the dummy equals one. But some plants may be induced to apply for an RSA grant when, for example, 20% of project costs can be covered by the grant while others may only apply for an RSA grant if a higher proportion of project costs can be covered by the grant as they regard 20% of project costs as insufficient for it to be worthwhile receiving a grant. A dummy that equals one when a plant is situated in an area where RSA can provide grants up to a level of 20% of project costs would then not satisfy the monotonicity condition as some plants could be induced to apply for an RSA grant when the dummy takes the value of one but some would be induced not to apply when the dummy equals one. No such problem exists when a single assisted areas dummy is used.[39]

As discussed in chapter 3.3, an instrumental variable must be correlated with the RSA dummy and uncorrelated with all components of the error term, ($\eta_i + v_{it} + m_{it}$), in equation (7.5). With regard to the former criterion, the RSA dummy must be a function of the assisted areas dummy as a plant has to be situated in an assisted area if it is to receive an RSA grant. With regard to the latter, there is no reason to suppose that it would be correlated with the measurement error component of the error term, m_{it}. The productivity shock, v_{it}, is arguably not correlated with the assisted areas variable because the map of assisted areas is drawn using data on economic deprivation which is at least three years obsolete by the time the map comes into force. The assisted areas dummy should therefore not capture contemporaneous factors that would determine the productivity shock, v_{it}. However, if economic conditions are slow to change in an area and are not captured by the time-invariant effects, there may still be correlation between the productivity shock and the assisted areas dummy.

It is even more difficult to argue that there will be no correlation between the assisted areas dummy and the time-invariant effects, η_i, if plants derive TFP advantages and disadvantages from being situated in different locations.[40] If no variables were included in x_{it} which captured the influence of locational factors on output, the assisted areas variable will be correlated with the time-invariant effects, η_i, which would be a function of unobserved locational factors from all periods. As assisted areas are poorly performing areas, it is expected that this correlation would be negative. However, x_{it} includes the Marshallian and Jacobian externalities variables discussed earlier. If these are sufficient to measure the influence of locational influences on plant performance, they will purge from the time-invariant effects, η_i, (and also from the productivity shock, v_{it}) those locational

---

[39] Another difference between the approach of Criscuolo, Martin, Overman and Van Reenen (2007) and that taken here is their use of data at the reporting unit level. As discussed in chapter 5.3, because RSA grants are awarded to individual plants, the local unit is the appropriate unit at which to conduct analysis of RSA.

[40] Criscuolo, Martin, Overman and Van Reenen (2007) employ the fixed effects estimator, which eliminates the time-invariant effects, so do not need to concern themselves with correlation between the treatment dummy and the time-invariant effects.

influences on output which, if ignored, would create a correlation between the assisted areas dummy and the error term. They would therefore justify the exclusion of the assisted areas dummy from the outcome equation and hence its use as an instrumental variable. However, it is doubtful whether these variables will be sufficient to remove from the error term all locational influences on plant performance. If these doubts are well founded, the estimate of the ATT obtained using this strategy will be biased downwards because of the negative correlation between the assisted areas dummy and the error term.

System GMM

In addition to correlation between the error term and the RSA dummy resulting from self-selection into the treatment group, there will also be correlation between the factor inputs and the productivity shock in the error term. This is the product of simultaneity and attrition bias.[41] Simultaneity bias arises because plants may have some knowledge about the value of the productivity shock in equation (7.5) and use this knowledge to choose the level of inputs in the production function (Marschak and Andrews, 1944). Attrition bias is present if plants base their exit decisions on their productivity level. As plants with a larger capital stock will be able to withstand lower productivity levels, this will generate a negative correlation between the productivity shock and the capital stock variable. More generally, capital and labour are endogenous if the demand curve for the output of the plant is downward sloping and firms maximise profits. Although the main variable of interest is the RSA dummy, it is essential to deal with the endogeneity of other explanatory variables if an unbiased estimate of the ATT is to be obtained. This point is made forcefully by Frölich (2008) who shows that the asymptotic bias of the estimate of the treatment effect can be large if the endogeneity of other variables in the model is ignored.

Therefore, the coefficients in equation (7.5) will be estimated using the system GMM estimator developed initially by Arellano and Bond (1991), augmented by Arellano and Bover (1995) and further improved by Blundell and Bond (1998) (see Bond, 2002 for an introduction). This estimates the equation as a system, using lagged levels and lagged first differences of the endogenous variables as instruments for the equations in first differences

[41] Van Beveren (2007) lists other potential sources of correlation between factor inputs and the error term. Here, consideration is given only to the two that are most frequently discussed in the literature.

and levels respectively.[42] The endogenous variables in our model that will be dealt with in this way are employment, capital, the Herfindahl index and the Marshallian externalities variable. It is assumed that the other variables in x_{it} are exogenous. A more detailed description of the system GMM estimator is provided in the appendix along with an explanation of why the system GMM estimator was chosen ahead of a two stage approach in which the popular Levinsohn and Petrin (2003) approach is used to estimate TFP.

However, if the productivity shock, v_{it}, is autoregressive so that $\alpha \neq 0$ in equation (7.4), it is not possible to estimate equation (7.5) using the system GMM estimator (Blundell and Bond, 2000). This is because the instruments for the endogenous variables will be correlated with the error term regardless of the number of times they are lagged. It is therefore necessary to transform equation (7.5) so that the following dynamic equation is estimated:

$$\begin{gathered} y_{it} = \pi_1 y_{i,t-1} + \pi_2 e_{it} + \pi_3 e_{i,t-1} + \\ \pi_4 k_{it} + \pi_5 k_{i,t-1} + \pi_6 x_{it} + \pi_7 D_{it} + \eta_i^* + \varepsilon_{it}, \end{gathered} \tag{7.7}$$

where $\eta_i^* = (1-\alpha)\eta_i$ and $\varepsilon_{it} = \upsilon_{it} + m_{it} - \alpha m_{i,t-1}$. This error term is now serially uncorrelated if there is no measurement error and it is a first-order moving average if there is measurement error. In either case, the system GMM estimator using suitably lagged instruments will provide unbiased estimates of the parameters of the model. The long run coefficients in equation (7.5) on employment, capital, x_{it} and the RSA dummy are given by the following:

$$\beta_E = \frac{\pi_2 + \pi_3}{1-\pi_1},\ \beta_K = \frac{\pi_4 + \pi_5}{1-\pi_1},\ \beta_X = \frac{\pi_6}{1-\pi_1},\ \beta_{ATT} = \frac{\pi_7}{1-\pi_1}. \tag{7.8}$$

7.4. Results

Rather than performing the analysis using the entire sample, it is done separately for four industries. This is done to avoid the imposition of common coefficients across disparate industries. In particular, it is undesirable to impose common coefficients for labour and capital as different industries operate with different technologies. If the imposition was not valid, it would not be possible to argue that the coefficients on the other variables

---

[42] This estimator is can be implemented in STATA 9.2 using the ‘xtabond2’ command developed by Roodman (2005).

accurately measure their impact on TFP. The industries that will be used are the food (SIC 41), textiles (SIC 43), footwear and clothing (SIC 45) and paper, printing and publishing (SIC 47) industries. These are chosen because they receive a relatively large number of RSA grants and the estimated parameters were stable across instrument sets.

All the results presented below are obtained using weighted data. Weighting is required to make the results representative of the population of plants because of the stratified sampling frame of the ABI. The sampling frame of the ABI and the construction of the weights are discussed in more detail in chapter 5 (Harris, 2005b gives a more general discussion of the need for weighting).

The long-run coefficients obtained from estimation of equation (7.7) using no mechanism to control for self-selection into the treatment group are displayed in table 7.1. These results are useful in establishing a baseline set of estimates that allow comparison with the estimates obtained using propensity score matching and instrumental variables. To allow for measurement error, instruments are lagged at least three times for textiles, footwear and clothing and paper, printing and publishing. This proved to be a sufficient number of lags to avoid rejection of the null of valid instruments in the Hansen test. For the food industry, however, the instruments had to be lagged at least five times to avoid rejection of the null of valid instruments at the 95% level in the Hansen test.

Table 7.1: Estimates of Augmented Production Function obtained using no control for self-selection

	Food	Textiles	Footwear & Clothing	Paper, Printing & Publishing
ln(Employment)	0.612*** (0.132)	0.732*** (0.159)	0.647*** (0.166)	0.665*** (0.099)
ln(Capital)	0.379*** (0.131)	0.327** (0.143)	0.376** (0.150)	0.383*** (0.082)
ln(Herfindahl Index)	0.269** (0.108)	-0.065 (0.070)	0.225** (0.094)	0.017 (0.024)
ln(Marshallian Externalities)	0.314 (0.298)	-0.048 (0.120)	0.092 (0.128)	0.143 (0.131)
ln(Jacobian Externalities)	-0.424 (0.322)	-0.345*** (0.093)	-0.146 (0.158)	-0.292 (0.196)
ln(Age)	-0.382 (0.255)	-0.832*** (0.261)	-0.437*** (0.166)	0.464*** (0.116)
ln(Time)	0.151 (0.113)	-0.179*** (0.044)	-0.092 (0.089)	-0.038 (0.035)
Single	-0.120 (0.113)	0.032 (0.101)	0.021 (0.096)	-0.344*** (0.070)
FDI	0.114 (0.083)	0.193 (0.130)	0.205 (0.174)	0.088 (0.064)
RSA	0.124 (0.167)	0.214 (0.130)	-0.050 (0.121)	-0.155 (0.092)
Lags	5	3	3	3
Hansen statistic	100.33*	92.48	80.33	92.22
Observations	4,095	1,897	1,198	3,262

* denotes significance at the 90% level, ** denotes significance at the 95% level, *** denotes significance at the 99% level
Standard errors are in parentheses

The estimates of the coefficients on employment and capital are all statistically significant at the 95% level and of a reasonable order of magnitude. This is important as it gives confidence that the coefficients on the other variables are genuinely estimates of their impact on TFP. With the exception of the textiles industry, which has a larger employment coefficient and a smaller capital coefficient than the other industries, the estimates of the coefficients are similar across industry. The estimates of the coefficient on capital are particularly close across the other three industries.

The coefficient on the Herfindahl Index is positive and statistically significant at the 95% level for the food and the footwear and clothing industry but not significant at the 90% level for the textiles and the paper, printing and publishing industry. Given the earlier discussion, this suggests that monopoly rents may be required to encourage plants to invest in R&D in the food and the footwear and clothing industries. The shortcomings of the

Herfindahl Index as a measure of competition should also be noted as the Herfindahl Index does not take account of either potential or international competition and is dependent on the definition of the industry (Okada, 2005).

The Marshallian specialisation externalities variable is not statistically significant at the 90% level for any of the industries considered which suggests that productivity benefits of being situated near plants from the same industry are negligible in these industries. By contrast, the Jacobian externalities variable is negative and statistically significant for the textiles industry and negative but not significant for the other three industries. This may suggest that the benefits from being in areas with plants from many different industries are outweighed by disadvantages that arise from congestion.

For the textiles, footwear and clothing and the paper, printing and publishing industries, older plants are found to have lower TFP, ceteris paribus. This is unsurprising as older plants are less likely to be employing the latest technologies. The time variable is negative and statistically significant for the textiles industry but not significant for the other three industries. This is difficult to explain as TFP would be expected to improve over time.

The coefficient on the single plant dummy is negative and statistically significant for the paper, printing and publishing industry but not for any of the other industries considered. A negative coefficient is precisely what would be expected if plants derive benefits from being part of a multi-plant enterprise. While always positive, the coefficient on the FDI dummy is not statistically significant for any of the industries considered.

The most important coefficient is, of course, that associated with the RSA dummy. It is positive for food and textiles but negative for footwear and clothing and paper, printing and publishing. It is not significant at the 90% level for any of the industries considered. Whether this lack of statistical significance is the product of misspecification of the observed variables or RSA recipients possessing unobserved characteristics that make them prone to poor performance is the issue that will now be investigated.

The results from estimating equation (7.7) using the matched sample are given in table 7.2. It should be borne in mind that all the observed covariates have been used to calculate the propensity score. As a result, they have less variance in the matched sample and their coefficients will therefore be less reliable. Appendix A2 shows the extent to which differences in the distribution of these variables across treated and untreated groups are reduced by moving from the full to the matched sample.

Table 7.2: Estimates of Augmented Production Function obtained using Propensity Score Matching

	Food	Textiles	Footwear & Clothing	Paper, Printing & Publishing
ln(Employment)	0.581*** (0.105)	0.243 (0.162)	0.174 (0.186)	0.841*** (0.127)
ln(Capital)	0.405*** (0.131)	0.796*** (0.141)	0.668*** (0.161)	0.176 (0.138)
ln(Herfindahl Index)	0.116 (0.079)	0.053 (0.077)	0.019 (0.089)	0.063 (0.042)
ln(Marshallian Externalities)	-0.124 (0.162)	0.126 (0.096)	0.887*** (0.237)	0.149 (0.177)
ln(Jacobian Externalities)	0.038 (0.152)	-0.411*** (0.128)	-0.645*** (0.233)	-0.120 (0.175)
ln(Age)	-0.517** (0.250)	-1.825*** (0.338)	-0.924** (0.422)	-0.334* (0.183)
ln(Time)	0.076** (0.067)	-0.174*** (0.061)	-0.072 (0.062)	0.077 (0.069)
Single	-0.248*** (0.065)	0.041 (0.109)	-0.023 (0.115)	-0.112 (0.076)
FDI	0.076 (0.104)	-0.212 (0.145)	-0.015 (0.220)	0.123** (0.058)
RSA	0.032 (0.087)	0.151 (0.155)	-0.126 (0.194)	-0.099 (0.114)
Lags	3	3	3	3
Hansen statistic	103.14	102.11	103.45	102.38
Observations	2,115	1,748	1,391	2,213

* denotes significance at the 90% level, ** denotes significance at the 95% level, *** denotes significance at the 99% level
Standard errors are in parentheses

Generally, the coefficients on the control variables are of the same sign as those in table 7.1 although there are notable variations in significance. The coefficient on the employment variable is positive but not statistically significant for the textiles and footwear and clothing industries while the coefficient on the capital stock is also positive for all industries but not significant for the paper, printing and publishing industry. Unlike in table 7.1, the Marshallian Externalities variable is positive and statistically significant for the footwear and clothing industry while the Jacobian Externalities variable is negative and statistically significant for this and the textiles industry. The age variable is now negative and statistically significant for all of the industries instead of only for some as in table 7.1. The coefficient on time is positive and significant at the 95% level for the food industry and negative and significant at the 99% level for the textiles industry. The latter is the same as in the results using no control for self-selection. The single plant enterprise variable is negative and statistically significant for the food industry.

Most importantly, for all four industries, the RSA dummy is once more not statistically significant. The point estimate is positive for the food and the textiles industry but negative for the footwear and clothing industry and the paper, printing and publishing industry.

Table 7.3 gives estimates of equation (7.7) using instrumental variables. The only difference between these estimates and those presented in table 7.1 is that the RSA dummy has been replaced in the instrument set by the assisted areas dummy. The Hansen test is of particular relevance in this set of estimates as it provides some evidence as to whether or not the instrument is valid. The null of valid instruments is not rejected at the 95% level for any of the industries considered although it is rejected at the 90% level for the food industry. This was also the case for the estimates presented in table 7.1 so this does not cast additional doubt on the validity of the instrument. However, it is important to note that the Hansen test is weak when there are many instruments (Roodman, 2009). Given that there were 96, 104, 101 and 106 instruments for the food, textiles, footwear and clothing and paper, printing and publishing industries respectively, this suggests that the Hansen test may not be very useful as a test of the validity of the assisted areas dummy as an instrument.

Table 7.3: Estimates of Augmented Production obtained using Instrumental Variables

	Food	Textiles	Footwear & Clothing	Paper, Printing & Publishing
ln(Employment)	0.551*** (0.151)	0.755*** (0.164)	0.684*** (0.194)	0.682*** (0.105)
ln(Capital)	0.413*** (0.148)	0.236* (0.142)	0.327* (0.171)	0.353*** (0.087)
ln(Herfindahl Index)	0.290** (0.117)	-0.106 (0.064)	0.203** (0.090)	0.017 (0.025)
ln(Marshallian Externalities)	0.363 (0.231)	0.055 (0.126)	0.058 (0.097)	0.191 (0.140)
ln(Jacobian Externalities)	-0.487** (0.239)	-0.266*** (0.090)	-0.162 (0.169)	-0.309 (0.210)
ln(Age)	-0.429 (0.266)	-0.567** (0.243)	-0.329 (0.205)	-0.375*** (0.121)
ln(Time)	0.127 (0.122)	-0.123*** (0.047)	-0.031 (0.086)	0.019 (0.049)
Single	-0.214 (0.168)	0.070 (0.108)	0.104 (0.094)	0.259** (0.113)
FDI	0.047 (0.109)	0.243* (0.137)	0.366* (0.201)	0.122* (0.066)
RSA	0.842 (1.235)	-0.354 (0.482)	-0.633 (0.586)	-0.572 (0.510)
Lags	5	3	3	3
Hansen statistic	99.03*	90.02	87.63	96.30
Observations	4,095	1,897	1,198	3,262

* denotes significance at the 90% level, ** denotes significance at the 95% level, *** denotes significance at the 99% level
Standard errors are in parentheses

As with the results in table 7.1, the coefficients on the factor inputs are all positive and statistically significant at the 90% level. With the exception of the coefficient on the RSA dummy, there are few differences with table 7.1 in terms of the size or significance of the coefficients on the explanatory variables. This is unsurprising given that the only difference relates to the replacement of one variable in the instrument set. In terms of coefficients that are statistically significant, the coefficient on the Jacobian externalities variable remains negative and is still statistically significant; the age coefficient is no longer statistically significant for the footwear and clothing industry; strangely, the coefficient on the single plant dummy in the paper, printing and publishing industry has changed signs but retained its significance while, in accordance with expectations, the FDI dummy is now significant at the 90% level for all industries apart from the food industry.

More importantly, comparing these results with those in table 7.1 that are obtained using no control for self-selection, the coefficient on the RSA dummy has fallen for all industries with the exception of the food industry. That it is not statistically significant for any of the industries considered implies that RSA had no impact on TFP. This is the same result as that obtained by Criscuolo, Martin, Overman and Van Reenen (2007) using a similar instrumental variables approach.

Comparing the magnitudes of the coefficients obtained using different estimators can reveal the direction of bias from using different estimators. For all of the industries considered, the estimate of the ATT using the matched sample is very close to the estimate obtained using no control for self-selection bias. This suggests that the latter estimates are not greatly contaminated by misspecification. It should be recalled that neither method will provide unbiased estimates of the treatment effect if there is correlation between the treatment variable and the error term due to the existence of unobserved variables that determine treatment status and the outcome variable as both sets of estimates depend on the conditional independence assumption holding to obtain unbiased estimates of the treatment effect.

The instrumental variables estimate of the ATT is lower than the estimate obtained using no control for self-selection and using the matched sample for all industries with the exception of the food industry. Assuming that the instrumental variable used is valid, this suggests that, for these three industries, there are unobserved covariates that are positively correlated with treatment status that are generating an upwards biased estimate of the ATT in tables 7.1 and 7.2. An alternative interpretation is that the instrumental variable is not valid and that the negative coefficients can be explained by the instrument being negatively correlated with the error term. Although it is not possible to prove conclusively which is the correct interpretation, it is the view of the author that the latter is more likely for two reasons: firstly, the direction of the movement of the estimate of the ATT between the matched sample and instrumental variable estimates conforms with what would be expected if the instrument was not valid; secondly, while it is quite conceivable that receipt of an RSA grant would have no impact on TFP, a negative impact of the magnitude suggested for the textiles, footwear and clothing and paper, printing and publishing industries is difficult to understand. Therefore, the preferred results are those obtained from the matched sample.

In order to probe the robustness of the results in tables 7.2 to 7.3, a number of variations on equation (7.7) were estimated. Firstly, interactions between the employment, capital, x_{it} variables and the RSA dummy were included to allow for different technologies across treated and untreated groups. These interactions were generally not statistically significant so this route was not pursued any further.

Secondly, as SAMIS provides information on the value of grants received, the RSA dummy was substituted for a variable measuring the size of grant received. This specification of the treatment variable was statistically significant for one of the industries considered above for the matched sample at the 95% level but was not significant at the 90% level either for any of the other industries using either the matched sample or instrumental variables.

Thirdly, SAMIS also states whether an RSA grant was provided to safeguard or promote employment. In order to check whether the impact on productivity differed for grants provided for different purposes, equation (7.7) was estimated using two treatment dummies: the first took the value of one when a plant received a grant to increase employment and the second equalled one if a plant received a grant to safeguard employment. Neither dummy was statistically significant at the 90% level using either the matched sample or instrumental variables.

Fourthly, the results above are obtained by treating plants that received multiple grants in exactly the same way as plants that received a single grant. Therefore, equation (7.7) was estimated using dummies that took the value of one from the time at which a plant received a second grant, a third grant and so on. Although these dummies were occasionally statistically significant, there was no obvious pattern indicating that receipt of more grants led to greater improvements or deteriorations in TFP. Taken together, these robustness tests provide further evidence that the receipt of an RSA grant tends to have no statistically significant impact on TFP.

7.5. Conclusion

This chapter has sought to establish the existence of a causal impact of receipt of an RSA grant on plant TFP. In confirmation of results from other analyses of the RSA scheme, no statistically significant effect was identified for any of the industries considered. This is a worrying finding as an important objective of the RSA scheme is to improve productivity. It also calls into question whether jobs safeguarded or created by RSA grants will endure.

The use of different estimators to estimate the ATT allows the researcher to see if the estimated treatment effect is robust to the choice of estimation strategy. When these estimates differ, the differences can be explained with reference to the assumptions underpinning the different estimators. The small difference between the estimates obtained using no control for self-selection and the estimates obtained using a matched sample suggests that there is little problem of misspecification from using no control for self-selection. However, for three of the four industries considered, the instrumental variables estimate of the treatment effect was considerably lower than the estimate obtained using the matched sample. Under the assumption that the instrument is valid, this suggests that plants that receive a grant have unobserved characteristics that lead to higher TFP which, if not controlled for, generate a misleadingly high estimate of the impact of an RSA grant. On the other hand, if the instrumental variable is correlated with the error term and therefore not a valid instrument, this is what will be driving the difference in results.

Although it is impossible to state conclusively which conclusion ought to be drawn, the latter is more likely because the very large negative coefficient on the estimate of the ATT obtained using instrumental variables is consistent with the existence of a negative correlation between the assisted areas dummy and treatment variable in spite of attempts to remove this. Therefore, the preferred results are those obtained using the matched sample although, as no statistically significant coefficient on the treatment variable is found using any of the estimators employed, the conclusion that receipt of an RSA grant has no causal impact on TFP is not put in doubt by the issue of which set of estimates are preferable.

A7.1. Production Function Estimation

This appendix will provide a description, based upon that provided by Bond (2002), of the system GMM estimator that will be used to estimate equation (7.7). It will then set out the Levinsohn and Petrin (2003) approach to estimating the parameters of production functions and discuss why their approach will not be used here.

The first part will describe the difference GMM estimator of which the system GMM estimator is an extension. To simplify the exposition, assume that we wish to estimate a simple AR(1) model:

$$y_{it} = \alpha y_{it-1} + \eta_i + \upsilon_{it}, \tag{A7.1}$$

where y_{it} is an observation of some variable pertaining to plant i at time t, η_i is an unobserved, plant-specific, time-invariant effect and υ_{it} is an error term which is assumed to be independently distributed across plants. It is assumed that the number of plants for which data is available is large while the number of time periods for which data is available is small.

The presence of y_{it-1} amongst the explanatory variables means that equation (A7.1) cannot be estimated using OLS because y_{it-1} is correlated with the η_i. Similarly, the fixed effects estimator cannot be used because, although the within-transformation removes η_i, a correlation exists between the transformed lagged dependent variable, $y_{it-1} - \frac{1}{T-1}(y_{i1} + \ldots + y_{it} + \ldots + y_{iT-1})$, and the transformed error term, $\upsilon_{it} - \frac{1}{T-1}(\upsilon_{i2} + \ldots + \upsilon_{it-1} + \ldots + \upsilon_{iT})$. This is because the transformed error term includes every realisation of the error term.

Another means of removing the time-invariant effects is first-differencing. First-differencing equation (A7.1) gives:

$$\Delta y_{it} = \alpha \Delta y_{it-1} + \Delta \upsilon_{it}. \tag{A7.2}$$

This cannot be estimated by OLS because of the correlation between y_{it-1} in the Δy_{it-1} term and υ_{it-1} in the $\Delta \upsilon_{it}$ term. However, the crucial difference between the within-transformed error term and the first differenced error is that the latter does not include every realisation of the error term. If y_{it} is assumed to be predetermined – that is, y_{i1} is assumed to be uncorrelated with the error term in future periods – and υ_{it} is assumed to be serially uncorrelated, the vector (y_{i1}, y_{i1}, ..., y_{iT-2}) can be used as instrumental variables.[43] Therefore, consistent estimates of the parameters in equation (A7.2) can be obtained by 2SLS.

[43] If υ_{it} is MA(1), $\Delta\upsilon_{it}$ is MA(2) and y_{iT-2} is no longer a valid instrument. In this case, the instrument set becomes (y_{i1}, y_{i1},..., y_{iT-3}).

However, 2SLS is not efficient because the complete set of moment conditions is not exploited. The difference GMM estimator does fully exploit all available moment conditions. It uses the following instrument matrix:

$$Z_i = \begin{bmatrix} y_{i1} & 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & y_{i1} & y_{i2} & \dots & 0 & \dots & 0 \\ . & . & . & \dots & . & \dots & . \\ 0 & 0 & 0 & \dots & y_{i1} & \dots & y_{iT-2} \end{bmatrix}, \tag{A7.3}$$

where the first row contains the instrument set for period 3, the second row the instrument set for period 4 and the final row the instrument set for period T. The moment conditions are given by:

$$Cov(Z'_i \,, \Delta\upsilon_{it}) = 0, \tag{A7.4}$$

where $\Delta\upsilon_i = (\Delta\upsilon_{i3}, \Delta\upsilon_{i4}, \dots, \Delta\upsilon_{iT})'$.

The difference GMM estimator minimises the following criterion:

$$J_N = \left(\frac{1}{N}\sum_{i=1}^{N}\Delta\upsilon'_i Z_i\right) W_N \left(\frac{1}{N}\sum_{i=1}^{N} Z'_i \,\Delta\upsilon_i\right), \tag{A7.5}$$

where the weighting matrix, W_N, is given by:

$$W_N = \left[\frac{1}{N}\sum_{i=1}^{N} Z'_i \,\Delta\hat{\upsilon}_i \Delta\hat{\upsilon}'_i \, Z_i\right]^{-1}, \tag{A7.6}$$

where $\Delta\hat{\upsilon}_i$ is a consistent estimate of the first difference error taken from a preliminary consistent estimator.

When data is available for more than 3 time periods, the validity of the moment conditions can be tested using the Sargan test of overidentifying restrictions. Under the null that the moment conditions are valid, NJ_N has an asymptotic χ^2 distribution.

Suppose now that, instead of estimating the simple AR(1) model in equation (A7.1), we wish to estimate the following model:

$$y_{it} = \alpha y_{it-1} + \beta x_{it} + \eta_i + \upsilon_{it}, \tag{A7.7}$$

where x_{it} is an additional vector of explanatory variables that are assumed to be correlated with η_i.

Once again, first-differencing removes the η_i so equation (A7.7) becomes:

$$\Delta y_{it} = \alpha \Delta y_{it-1} + \beta \Delta x_{it} + \Delta \upsilon_{it}, \tag{A7.8}$$

If it is assumed that x_{it} is correlated with υ_{it}, it is treated in the same way as y_{it-1}, so x_{it-2}, x_{it-3} and longer lags can be used as instruments.

So far, the difference GMM estimator has been described. Consider now estimation of equation (A7.7) if we are willing to assume that Δx_{it} is uncorrelated with the η_i:

$$Cov(\Delta x_{it}, \eta_{it}) = 0, \tag{A7.9}$$

but we continue to assume that x_{it} is correlated with the error term. In this situations, Δx_{it-2}, Δx_{it-3} and longer lags are available as instruments for the estimation of equation (A7.6).

Whether lags of Δy_{it} can be used as instruments for estimation of equation (A7.6) depends upon whether or not the following condition holds:

$$Cov\left[\left(y_{i1} - \left(\frac{\eta_i}{1-\alpha}\right)\right), \eta_i\right] = 0. \tag{A7.10}$$

Equation (A7.10) states that the initial value of the y_{it} does not differ systematically from the value, $\left(\frac{\eta_i}{1-\alpha}\right)$, towards which the series converges. Equation (A7.10) then implies:

$$Cov(\Delta y_{i2}, \eta_i) = 0. \tag{A7.11}$$

If it is also assumed that that the time-invariant effects, η_i, are uncorrelated with υ_{it-1}, the following moment conditions are available:

$$Cov(\Delta y_{it-1}, (\eta_i + \upsilon_{it})) = 0. \tag{A7.12}$$

The main benefit of the additional moment conditions is that estimation of the parameters in equation (A7.7) no longer depends entirely on the use of lagged levels of variables as instruments for their first-differences as in the difference GMM estimator. This can be problematic because, as α approaches one or as $\mathrm{var}(\eta_i)/\mathrm{var}(\upsilon_{it})$ grows, lagged levels are only weakly correlated with subsequent first differences and the difference GMM estimator is hampered by the problem of weak instruments (Blundell, Bond and Windmeijer, 2000). This problem is diminished when use is made of the additional moment conditions which allow estimation of the equation in levels.

Another method that is often used in this type of analysis is a two-stage approach in which estimates of TFP are obtained in the first stage and then used as the dependent variable in a second stage regression in which the treatment effect is estimated. Usually, the method developed by Olley and Pakes (1996) and modified by Levinsohn and Petrin (2003) is used in the first stage to calculate the measure of TFP. However, the validity of this approach is subject to question. Their method will now be set out and an explanation provided of why it will not be used here.

The Levinsohn and Petrin approach will be discussed here as this is now the more commonly used of the two. Consider the following model:

$$y_{it} = \beta_0 + \beta_E e_{it} + \beta_N n_{it} + \beta_K k_{it} + \omega_{it} + \upsilon_{it}, \tag{A7.13}$$

where ω_{it} represents productivity shocks observed by the plant but not by the researcher and n_{it} is intermediate inputs. It is accepted that input choices will be based, at least on part, on the realisation of these productivity shocks. Assuming that demand for intermediate inputs is a function of capital and the productivity shock, intermediate inputs can be written as follows:

$$n_{it} = n_{it}(\omega_{it}, k_{it}). \tag{A7.14}$$

This can be inverted to obtain the following function:

$$\omega_{it} = \omega_{it}(n_{it}, k_{it}). \tag{A7.15}$$

Substituting equation (A7.15) into equation (A7.13) gives:

$$\begin{aligned} y_{it} &= \beta_E e_{it} + \phi_{it}(n_{it}, k_{it}) + \upsilon_{it}, \\ \phi_{it} &= \beta_0 + \beta_N n_{it} + \beta_K k_{it} + \omega_{it}(n_{it}, k_{it}), \end{aligned} \tag{A7.16}$$

which can be estimated using a high-order polynomial in n_{it} and k_{it} to approximate ω_{it}. This is said to provide an estimate of the coefficient on employment although the coefficients on capital and intermediate inputs are not identified since these variables enter equation (A7.16) more than once.

In the second stage, the dependent variable is output net of the contribution from labour, y^*, so that the following is estimated:

$$y^*_{it} = y_{it} - \beta_E e_{it} = \beta_0 + \beta_N n_{it} + \beta_K k_{it} + \omega_{it} + \upsilon_{it}. \tag{A7.17}$$

Levinsohn and Petrin, following Olley and Pakes, assume that the productivity shock follows a first-order process and that capital does not respond immediately to differences between the expected value and the realised value of the productivity shock. This difference is given by:

$$\varepsilon_{it} = \omega_{it} - E[\omega_{it} \mid \omega_{i,t-1}]. \tag{A7.18}$$

An estimate of $E[\omega_{it} \mid \omega_{i,t-1}]$ can be taken from the estimates from equation (A7.16).

Equation (A7.17) can then be rewritten as:

$$y_{it}^{*} = \beta_0 + \beta_N n_{it} + \beta_K k_{it} + E[\omega_{it} \mid \omega_{i,t-1}] + \varepsilon_{it} + \upsilon_{it}. \qquad \text{(A7.19)}$$

By assumption, capital is uncorrelated with the error term. This is not true of intermediate inputs which may respond to the innovation in productivity. To estimate the coefficient on intermediate inputs, Levinsohn and Petrin use the moment condition implied by the fact that the lagged value of intermediate inputs will be uncorrelated with the productivity innovation. All the coefficients required to calculate TFP are therefore identified.

This approach, which has become popular in recent years, suffers from a number of drawbacks. It has been pointed out by Ackerberg, Caves and Frazer (2006) that, if the ideas underpinning identification in the model are applied consistently, the coefficient on employment is not, in fact, identified unless some unappealing assumptions are made. To be specific, as employment and intermediate inputs are chosen at the same time and are assumed to be perfectly-variable, non-dynamic variables, it is natural to assume that they are both functions of the same variables, k_{it} and ω_{it}, so that employment can be written as:

$$e_{it} = e_{it}(\omega_{it}, k_{it}). \qquad \text{(A7.20)}$$

Substituting equation (A7.15) into equation (A7.20) gives:

$$e_{it} = e_{it}(\omega_{it}(n_{it}, k_{it}), k_{it}) = h(n_{it}, k_{it}). \qquad \text{(A7.21)}$$

Employment is therefore also a function of intermediate inputs and the capital stock which implies perfect collinearity between employment and $\omega_{it}(n_{it}, k_{it})$ in the first-stage of the model and, therefore, that the coefficient on employment is not identified in the first-stage.

Another drawback of this approach is that, unlike the system GMM estimator, it does not allow for unobserved, time-invariant effects in the error term that are correlated with the factor inputs. Instead, the productivity shock is the only component of the error term which is allowed to be correlated with the factor inputs. As time-invariant effects are likely to be important due to the existence of constant unobserved variables (such as managerial ability) that determine output, this is a major shortcoming of the Levinsohn and Petrin approach.

Another problem noted by Ackerberg, Caves and Frazer (2006) is that there must be strict monotonicity between intermediate input demand and the productivity shock. If the latter does not hold, then the intermediate input function cannot be inverted and used as a proxy for the productivity shock.[44] Because of these three problems, the system GMM method outlined above is preferred.

A7.2. Covariate Balance in the Full and Matched Sample for Productivity Analysis

This appendix gives information on the distribution of each variable across treated and untreated groups in the full and the matched sample for each industry. The mean of each variable is presented and t-tests are employed to establish if this difference is statistically significant. Kolmogorov-Smirnov tests are also performed which test the null that the variable in the treated and untreated groups are drawn from the same distribution. This information is valuable as it provides an indication of the extent to which the balance of the observed covariates across the treated and untreated groups is improved by moving from the full to the matched sample and therefore the extent to which problems of misspecification caused by self-selection are alleviated.

Table A7.1 shows the results for the food industry.

[44] The major difference between the Levinsohn & Petrin method and the Olley & Pakes method is that, in the latter, investment performs the function of intermediate inputs. Levinsohn and Petrin choose to use intermediate inputs because they argue that the monotonicity condition is more likely to be satisfied using intermediate inputs because of the 'lumpiness' of investment.

Table A7.1: Distribution of Observed Covariates across the Full and Matched Sample for the Food Industry for Productivity Analysis

	Full Sample				Matched Sample			
	Non-RSA	RSA	Difference	Combined K-S	Non-RSA	RSA	Difference	Combined K-S
ln(Employment)	2.52	3.48	-0.96***	0.40***	3.29	3.63	-0.34***	0.16***
ln(Capital)	3.29	4.62	-1.33***	0.34***	4.43	4.91	-0.48***	0.14***
ln(Herfindahl Index)	-2.11	-1.77	-0.35***	0.24***	-1.64	-1.58	-0.06	0.08**
ln(Marshallian Externalities)	-1.06	-0.73	-0.33***	0.18***	-0.68	-0.65	-0.03	0.06
ln(Jacobian Externalities)	2.26	1.86	0.40***	0.21***	1.76	1.75	0.01	0.05
ln(Age)	1.98	2.06	-0.07	0.11***	2.17	2.18	-0.01	0.07*
Single	0.19	0.36	-0.17***	0.23***	0.13	0.16	-0.04**	0.06
FDI	0.03	0.07	-0.04***	0.06*	0.09	0.13	-0.04**	0.06

* denotes significance at the 90% level, ** denotes significance at the 95% level, *** denotes significance at the 99% level
Source: SAMIS/ARD

With the exception of the age variable, the differences in the mean of the explanatory variables across treated and untreated groups are all statistically significant at the 99% level in the full sample. In the matched sample, the differences in the mean have fallen for all variables. While the differences remains statistically significant for the employment, capital, age and single plant enterprise variable, the differences in the mean for the other variables are no longer statistically significant. In the full sample, the Kolmogorov-Smirnov test statistic indicates a rejection of the null that the treated and untreated groups are drawn from the same distribution for each variable at the 90% level. In the matched sample, the null is only rejected at this level for the employment, capital, Herfindahl Index and the age variable.

Table A7.2 provides the same information for the textiles industry.

Table A7.2: Distribution of Observed Covariates across the Full and Matched Sample for the Textiles Industry for Productivity Analysis

	Full Sample				Matched Sample			
	Non-RSA	RSA	Difference	Combined K-S	Non-RSA	RSA	Difference	Combined K-S
ln(Employment)	3.31	3.46	-0.15***	0.10***	3.64	3.63	0.01	0.07*
ln(Capital)	4.51	4.54	-0.03	0.08**	5.05	5.07	-0.02	0.07*
ln(Herfindahl Index)	-2.55	-2.07	-0.48***	0.20***	-1.99	-2.07	0.08*	0.09**
ln(Marshallian Externalities)	-0.43	-0.30	-0.19**	0.17***	-0.15	-0.28	0.13**	0.18***
ln(Jacobian Externalities)	1.92	1.61	0.31***	0.14***	1.75	1.70	0.05	0.06
ln(Age)	2.44	2.29	0.15**	0.10***	2.48	2.47	0.01	0.03
Single	0.24	0.23	0.01	0.03	0.26	0.24	0.03	0.04
FDI	0.05	0.07	-0.02**	0.04	0.09	0.11	-0.02	0.03

* denotes significance at the 90% level, ** denotes significance at the 95% level, *** denotes significance at the 99% level
Source: SAMIS/ARD

In the full sample, the difference in mean across the treated and untreated groups is statistically significant at the 95% level for all variables apart from the capital and the single plant enterprise variable. In the matched sample, the difference in mean has fallen for all variables apart from the single plant variable and remains statistically significant at the 90% level for the Herfindahl index and Marshallian Externalities variables. The Kolmogorov-Smirnov test indicates a difference in the distribution of all variables, with the exception of the single plant and FDI variable, at the 95% level in the full sample. In the matched sample, the difference is only statistically significant at the 90% level for the employment, capital, Herfindahl Index and Marshallian Externalities variable.

The results for the footwear and clothing industry are given below in table A7.3.

Table A7.3: Distribution of Observed Covariates across the Full and Matched Sample for the Footwear and Clothing Industry for Productivity Analysis

	Full Sample				Matched Sample			
	Non-RSA	RSA	Difference	Combined K-S	Non-RSA	RSA	Difference	Combined K-S
ln(Employment)	3.08	3.58	-0.49***	0.21***	3.48	3.67	-0.19***	0.11***
ln(Capital)	3.31	4.05	-0.74***	0.21***	3.99	4.24	-0.25***	0.14***
ln(Herfindahl Index)	-2.56	-2.01	-0.55***	0.27***	-1.95	-1.84	-0.11**	0.10**
ln(Marshallian Externalities)	-0.84	-0.44	-0.40***	0.17***	-0.48	-0.50	0.02	0.18***
ln(Jacobian Externalities)	2.57	1.80	0.77***	0.26***	1.62	1.45	0.17***	0.19***
ln(Age)	2.18	2.15	0.03	0.15***	2.24	2.15	0.08**	0.08*
Single	0.30	0.26	0.05*	0.08**	0.27	0.21	0.06***	0.11***
FDI	0.06	0.14	-0.08	0.16***	0.01	0.09	-0.08***	0.11***

* denotes significance at the 90% level, ** denotes significance at the 95% level, *** denotes significance at the 99% level
Source: SAMIS/ARD

In the full sample, there is a statistically significant difference in the mean of all the variables, apart from the age and FDI variables across treated and untreated groups. In the matched sample, the magnitude of the difference falls for all variables apart from the age and the single plant enterprise variable. However, the differences are statistically significant at the 90% level for all variables except the single plant variable. The null that the treated and untreated groups are drawn from the same distribution is rejected for all variables at the 95% level in the full sample. Although the test statistic decreases for all but the Marshallian Externalities and the age variable, the null is still rejected at the 90% level for all variables in the matched sample.

Table A7.4 provides the equivalent statistics for the paper, printing and publishing industries.

Table A7.4: Distribution of Observed Covariates across the Full and Matched Sample for the Paper, Printing and Publishing Industry for Productivity Analysis

	Full Sample				Matched Sample			
	Non-RSA	RSA	Difference	Combined K-S	Non-RSA	RSA	Difference	Combined K-S
ln(Employment)	2.58	3.88	-1.30***	0.51***	3.46	3.65	-0.19***	0.17***
ln(Capital)	3.89	5.60	-1.71***	0.37***	5.08	5.69	-0.61***	0.22***
ln(Herfindahl Index)	-3.01	-2.88	-0.13	0.13***	-2.40	-2.84	0.44***	0.17***
ln(Marshallian Externalities)	-1.03	-0.79	-0.24***	0.16***	-0.78	-0.91	0.12**	0.15***
ln(Jacobian Externalities)	2.49	2.22	0.27***	0.14***	1.94	2.17	-0.23***	0.16***
ln(Age)	2.33	2.36	-0.03	0.14***	2.30	2.63	-0.33***	0.23***
Single	0.22	0.37	-0.16***	0.19***	0.23	0.40	-0.17***	0.19***
FDI	0.12	0.17	-0.05***	0.11***	0.14	0.18	-0.04*	0.09*

* denotes significance at the 90% level, ** denotes significance at the 95% level, *** denotes significance at the 99% level
Source: SAMIS/ARD

In the paper, printing and publishing industry, the difference in mean across the treated and untreated groups is statistically significant at the 99% level for all but the Herfindahl Index and the age variables. The difference in mean decreases for all but the Herfindahl Index, age and the single plant enterprise variables in the matched sample but is still statistically significant at the 95% level for all variables. The null of the treated and untreated groups being drawn from the same distribution is rejected at the 99% level for each variable in the full sample. While the distributions are better matched across treatment and control groups in the matched sample, the null is still rejected for all variables at the 90% level.

Taken together, these tables show that the matching process has brought into closer alignment the distribution of the observed covariates in equation (7.7) across the treated and untreated groups. This is unsurprising as this is precisely what matching is designed to do. The extent to which the difference in the distribution of variables across treatment and untreated groups between the full and matched sample is reduced is a reflection of the size of the coefficient on the variable in the probit model used to generate the propensity scores. For instance, employment is an important determinant of treatment so has a large coefficient in a probit model of treatment status. As a result, for all industries, there is a large reduction in the difference of the mean of employment across treated and untreated group when the matched sample is compared to the full sample. On the other hand, for those variables which are not such important determinants of treatment status, the difference across treated and untreated groups may remain large or even increase in the matched sample. This suggests that the approach adopted of including these variables in the outcome regression, rather than simply comparing the mean of the outcome variable across treatment and control groups, is the correct one. Failure to do so would not control for remaining differences in the distribution of the observed covariates across treatment and control group and would therefore lead to biased estimates of the treatment effect.

8. The Causal Impact of RSA on Survival

8.1. Introduction

This chapter will examine whether receipt of an RSA grant has a causal impact on plant survival. It will do so by estimating a Cox proportional hazards model. In order to control for the consequences of self-selection into the treatment group, the model will be estimated using a sample created by propensity score matching.

As discussed in chapter 3.3, the instrumental variables estimator provides a way of obtaining unbiased estimates of the treatment effect when the treatment group is a self-selected group of the population of plants with unobservable characteristics that differ from those of the untreated group. However, implicit in that earlier discussion is the assumption that the relationship between the explanatory variables and the dependent variable is linear. When the relationship is non-linear, as in proportional hazards models, the instrumental variable estimator as described in chapter 3.3 cannot be applied. Instrumental variables estimators for hazards models are now being developed (see, for example, Bijwaard, 2008) but these are not yet available in statistical packages. This is unfortunate as it means that it is not possible to estimate a hazard model that allows for unobservable variables that are correlated with both the treatment variable and the hazard rate. Therefore, the conditional independence assumption discussed in chapter 3.3 must be assumed to hold if the estimates below are to be unbiased estimates of the treatment effect.

The only paper that has evaluated the impact of receipt of an RSA grant on the probability of survival is by Harris and Robinson (2005). Using a similar dataset and methodology to that which will be used below, but with no control for the consequences of self-selection, they find that receipt of an RSA grant lowered the probability of closure by 32.1% for plants aged one or less and that this rises to nearly 57.1% for plants aged over ten years. Harris and Trainor (2007) also employ a Cox proportional hazards model to examine the impact of SFA - the Northern Irish equivalent of RSA - on the probability of closure and find that SFA recipients had on average a 24.1% lower probability of closure than non-recipients. Noting the potential existence of self-selection bias in these estimates, they then re-estimate the model using the predicted values from a Tobit regression of value of SFA received upon relevant variables instead of the actual values. This method suggests that receipt of SFA reduced the probability of closure by 15%. However, it should be borne in mind that this approach, which amounts to an instrumental variables approach, is invalid because of the non-linear relationship between the explanatory variables and the dependent variable in the Cox proportional hazards model. Finally, as is done below, Girma, Görg and Strobl (2007b) combine propensity score matching with the Cox proportional hazard model to evaluate the impact of government grants on closure probability in Irish manufacturing plants. Their results also indicate that receiving a government grant lowered the probability of closure by a statistically significant amount (these papers are critically reviewed in chapter 4.4).

The next section will describe the model that is used to estimate the impact of receiving an RSA grant on closure probability; the third section will describe how the matched sample was created and how the proportional hazards model was estimated; the fourth section gives results and the fifth section concludes.

8.2. Econometric Model

The hazard function is defined as the probability of closure in period *t*, having survived until period *t*:

$$h(t; X(t)) = P[T = t \mid T \geq t, X(t)], \tag{8.1}$$

where *T* is the year in which the plant closes and *X*(*t*) is a vector of time varying covariates.

A proportional hazards model takes the following form:

$$h(t) = h_0(t)\exp(x(t)\beta), \tag{8.2}$$

where $h_0(t)$ is a non-parametric baseline hazard function – the hazard rate for a plant with all covariates set to zero – that is shared by all plants and $\exp(x(t)\beta)$ is a parametric function of plant characteristics. The proportional hazards model is therefore a semi-parametric model. This specification of the hazards function implies that covariates multiplicatively shift the baseline hazard as follows:

$$\frac{h(t \mid x_i)}{h(t \mid x_j)} = \frac{\exp(x_i)\beta}{\exp(x_j)\beta} = \frac{\exp(x_i)}{\exp(x_j)}, \tag{8.3}$$

where $i \neq j$. This implies that for two plants with different covariate values, the ratio of their hazard functions is not a function of time.

The specification of the explanatory variables in equation (8.2) is:

$$x(t)\beta = \beta_X x_{it} + \beta_{ATT} D_{it}, \tag{8.4}$$

where D_{it} is a dummy that takes the value of one from the time at which the plant received a grant. x_{it} are a vector of variables included in the hazard model to avoid a biased estimate of the ATT due to correlation between observable covariates and the treatment dummy. All continuous variables are entered in logarithmic form. The following will describe the variables included in x_{it} and explain why they are thought likely to determine the probability of survival. Chapter 5.4 provides more detail on how the variables were constructed.

The decision of an enterprise to close a plant depends ultimately on the contribution of the plant to the profits of the enterprise. When the difference in the discounted expected profits over future periods of the enterprise with the plant and without the plant exceeds the liquidation value of the plant, then the enterprise will choose not to close the plant. Given its role as a key determinant of the contribution of the plant to the profits of the enterprise, efficiency is of great importance in this decision (Harris and Li, 2007). As discussed in chapter 7.1, TFP is a better measure of efficiency than labour productivity because it is not determined by factor input levels. However, because TFP must be estimated in a first stage regression, its inclusion in x_{it} would introduce problems of inference with the estimate of the hazard rate on the treatment variable (see Wooldridge, 2007 for an introduction to the issues that arise when generated regressors are used). This is because the standard errors of all the estimated hazard rates are incorrect when one of the regressors is generated. To avoid these problems, labour productivity is used to proxy for efficiency.

The theoretical model of learning and market selection by Jovanovich (1982) motivates the inclusion in the model of plant size and age as further proxies for efficiency levels. In this model, efficiency levels are random, unobserved, time-invariant and differ across enterprises which begin with identical prior beliefs about their efficiency, believing themselves to be a random draw from the distribution of efficiency. On the basis of their estimate of their own efficiency, enterprises choose their output levels. Ceteris paribus, a higher estimate of the efficiency leads to a higher level of output. Actual production costs are determined by both efficiency and a stochastic error. Over time, each enterprise's estimate of its efficiency becomes more precise as they update their estimate on the basis of observed actual production costs. Enterprises close when its estimate of its efficiency falls below a threshold. As enterprises with such expectations will not have grown as fast as plants with greater efficiency, plant size is negatively associated with closure (Colombo and Delmastro, 2001). Moreover, as the probability of a change in the estimate of efficiency that leads to closure falls over time due to the increasing precision with which 'true efficiency' is estimated, enterprise age is negatively associated with closure probability. The employment and age variables are therefore included, alongside labour productivity, in x_{it} to proxy for efficiency levels.

Also included in x_{it} is the capital to labour ratio. As plants with a large capital to labour ratio will have a large fixed to variable costs ratio, such plants are more likely to be able to avoid closure when costs exceed revenues because they are better able to cover their fixed costs (Doms, Dunne and Roberts, 1995). The inclusion of the capital to labour ratio can also be justified on the grounds that it acts as a proxy for sunk costs. As such assets cannot be profitably employed in other industries, they deter exit (Siegfried and Evans, 1994). From a different perspective, the model by Hopenhayn (1992) provides another explanation for why the existence of sunk costs may reduce the probability of exit. Since sunk costs deter entry, they insulate incumbents, who do not have to pay the sunk cost, from market selection based on productivity levels. Therefore, plants with low productivity levels that would otherwise close do not in industries where entrants have to incur sunk costs.

Also included among the covariates is the Herfindahl Index (Herfindahl, 1950) calculated at the four-digit industry level. The Herfindahl Index is a measure of the degree of concentration and hence competition within an industry. Lower values indicate greater competition. A priori expectations are that large plants within a highly concentrated industry with little competition will have a lower probability of closure while small plants within the same industry will have a higher probability of closure because large plants can behave in a retaliatory manner in a bid to preserve their position (Audretsch, 1994).

The proportion of newly opened plants' output in total industry output is included in x_{it} to capture displacement. Displacement occurs when incumbent plants are driven out of the market by more aggressive and efficient entrants (see Siegfried and Evans, 1994; Caves and Porter, 1976). Gabszewicz and Thisse (1980) provide a theoretical model in which prices are chosen noncooperatively and find that the number of firms in the market is subject to an upper bound. Having reached this upper bound, the entry of new firms necessitates the exit of others.

Another variable included in x_{it} is the growth of industry employment. This represents an attempt to capture perceptions of future profits. As employment will generally rise as perceptions of profits improve, it is expected that the growth of industry employment will be negatively related to closure probability. An advantage of using employment rather than output is that, given the difficulties of dismissing employees, the growth of employment is more likely to indicate an improvement in perceptions of future profits than rises in output which do not reflect such a commitment.

A foreign ownership dummy is also included among the covariates. It is difficult to predict a priori whether being owned by a multinational will be positively or negatively related to survival. If foreign-owned plants have access to superior foreign technologies by virtue of its link to the home country of the multinational, this would imply that FDI plants would be less likely to close (Harris and Li, 2007). Furthermore, if the multinational itself has proprietary assets which it shares with its subsidiaries, this should also reduce the probability of closure. It may also be hypothesised that multinationals tend to acquire high quality plants and that FDI plants are therefore a selected group of the population of plants with characteristics that make them less likely to close. On the other hand, a lack of knowledge of operating in the foreign market may impede plants owned by multinationals. The acquired plant may also experience difficulties in adjusting to the technologies of the multinational. Furthermore, multinationals may be more inclined to close foreign subsidiaries than plants at home because home country stakeholders have more influence over company policy than host country stakeholders (Kronborg and Thomsen, 2008). These latter considerations suggest that FDI plants may have a higher probability of closure.

A change in ownership dummy taking the value of one from the point at which ownership changed is also included among x_{it}. This is because the acquiring enterprise may be unable to purchase only those plants that it wishes to purchase but rather has to buy all the plants belonging to the enterprise that is being bought. If some of these plants are not actually wanted by the enterprise they will be closed if they cannot be resold (McGuckin and

Nguyen, 1995). Another related motivation for the inclusion of this dummy is that restructuring within the acquiring enterprise following the acquisition of a new group of plants may require the closure of some of the newly acquired plants (McGuckin and Nguyen, 2001).

Finally, a dummy that equals one when the plant is not part of a larger enterprise is included in x_{it}. Owners of single plant enterprises are expected to have a lower opportunity cost of closure than owners of multi-plant establishments and are therefore expected to be willing to accept lower rates of return (Audretsch, 1994). This is because closure of a plant within a multi-plant enterprise does not mean the closure of the entire enterprise or the exiting of the enterprise from the market if production is then transferred to other plants within the enterprise (Colombo and Delmastro, 2000). As a result, if the multi-plant enterprise decided to re-open the plant, it would not incur the same re-entry costs as would the single plant enterprise.

The inclusion of variables in x_{it} that may be determined by treatment status begs the question of whether their coefficients must be taken account of to generate an estimate of the impact of the total impact of receipt of an RSA grant on closure probability. Two such variables are labour productivity and employment. Consideration of the rationale for including these variables in x_{it} shows that it is not necessary to consider the impact of these variables on closure probability to come to an estimate of the total impact of RSA on survival probability. This is because both these variables are only included in x_{it} as proxies for efficiency and chapter 7 has shown that receipt of an RSA grant has no statistically significant impact on TFP, which is a better measure of efficiency. Another variable in x_{it} that will be determined by treatment status is the capital to labour ratio which is included to proxy for sunk costs. The fact that RSA grant may be spent on sunk costs, which are expected to be negatively associated with closure probability, is undeniable and the coefficient on the capital to labour ratio should therefore be taken into account when estimating the total impact of receipt of an RSA grant.

The inclusion of variables to proxy for efficiency in x_{it} deserves further attention as it casts doubt on whether a negative causal relationship between RSA and closure probability is desirable. A negative relationship would imply that RSA grants could allow plants with low productivity to survive and therefore, in accordance with Schumpeterian notions of ‘creative destruction’, hinder the process by which resources shift from low to high productivity plants (Schumpeter, 1943). In this situation, the causal impact of RSA on aggregate productivity, and therefore income per capita, would be negative.

8.3. Estimation Strategy

This section will describe how equation (8.2) will be estimated. The section is divided into two parts: the first describes how the matched sample was created and the second explains how the proportional hazards model is estimated.

Propensity Score Matching

The way in which the matched sample was created is directly analogous to the way in which the matched sample was created in chapter 7 where the impact of receipt of an RSA grant on TFP was examined. Therefore, to avoid unnecessary repetition, the following will only set out the way in which the approaches differ which is in the estimation of the probit model which contains different variables due to the different variables contained in the outcome equations.

The probit model used to estimate the propensity score is:

$$ID_{it} = \gamma_1 end_{it-1} + \gamma_3 exo_{it} + \varepsilon_{it}, \qquad (8.5)$$

where end_{it} is a vector of variables that may be affected by treatment and exo_{it} is a vector of variables that are not affected by treatment. Included in end_{it-1} are the lag of labour productivity, employment, the capital to labour ratio and the Herfindahl Index. Included in exo_{it} is plant age, displacement, the growth of industry employment, the foreign ownership dummy, the ownership change dummy and the single plant dummy.

Maximum Likelihood Estimation

This section will describe how proportional hazards models of the form of equation (8.2) can be estimated. It will then describe how the assumption of proportional hazards can be tested.

To estimate equation (8.2), the shape of the baseline hazard can be assumed to take a particular form. For instance, the Weibull distribution could be assumed:

$$h_0(t) = pt^{p-1}, \tag{8.6}$$

where p is a shape parameter. If $h_0(t)$ was assumed to take that form, the proportional hazards model becomes a parametric model. However, it is difficult to say what form $h_0(t)$ will take and specifying an incorrect baseline hazard will lead to biased estimates of β. Fortunately, Cox (1972) showed that it is not necessary to specify a functional form for $h_0(t)$ so this problem can be avoided. The cost of not assuming a functional form for the baseline hazard is a loss in efficiency but, in situations where the shape of the baseline hazard is unknown, this is a cost that must be accepted.

The parameter estimates in the Cox proportional hazards model are obtained by maximising the following partial likelihood function:[45]

$$L(\beta) = \prod_{j=1}^{k} P_j, \tag{8.7}$$

where P_j is the conditional probability that plant j closes in the period in which it closes and k is the number of periods in which plants are observed to close. Equation (8.7) is a partial rather than a full likelihood assumption because only those periods in which plants are observed to fail are used to calculate β. Given equation (8.2), the conditional probability that plant j closes in the period in which it closes is given by:

$$P_j = \left(\frac{h_0(t_j)\exp(x_j\beta)}{\sum_{i \in R_j} h_0(t_j)\exp(x_i\beta)} \right), \tag{8.8}$$

where t_j is the year in which plant j closes and R_j is the number of plants in the risk pool at time t_j. As can be seen from equation (8.8), the $h_j(t_j)$ terms cancel which means that it is not necessary to specify a baseline hazard function.

Substituting equation (8.8) into equation (8.7) gives:

$$L(\beta) = \prod_{j=1}^{k} \left(\frac{\exp(x_j\beta)}{\sum_{i \in R_j} \exp(x_i\beta)} \right). \tag{8.9}$$

---

[45] The Cox proportional hazards model is performed in STATA 9.2 using the ‘stcox’ command (see Cleves, Gould and Gutierrez, 2002 for an introduction to this command).

When more than one plant is observed to fail during a particular time period, the issue of how to handle tied failures arises. If different plants failed at different times but are observed to have failed at the same time because of limitations in the precision with which failure times are measured, the marginal method should be used. To simplify the exposition, define:

$$r_i = \exp(x_i \beta). \tag{8.10}$$

To simplify further, assume that there are only three plants at risk of failure in period 1. Suppose that plants 1 and 2 are observed to fail. P_{12} is the probability that plant 1 fails before plant 2 and P_{21} is the probability that plant 2 fails before plant 1. P_{12} is given by:

$$P_{12} = \frac{r_1}{r_1 + r_2 + r_3} \frac{r_2}{r_2 + r_3}, \tag{8.11}$$

$$P_{12} = \frac{r_2}{r_1 + r_2 + r_3} \frac{r_1}{r_1 + r_3}. \tag{8.12}$$

To obtain the marginal estimate of β, $P_{12} + P_{21}$ are substituted in equation (8.7) in place of P_1.

As the marginal method is computationally demanding, the Breslow (1974) method for handling tied failures is often used as an approximation. For this reason, and the more practical reason that Stata 9.2 does not allow the marginal method to be used when the data is weighted, the Breslow approximation is used below.

The difference between the Breslow approximation and the marginal method is that the Breslow approximation does not adjust the risk pool for plants that fail after the first failure when calculating the conditional probabilities. Therefore, instead of representing the probability that plant 1 fails before plant 2 and plant 2 fails before plant 1 as in equation (8.11) and (8.12), these conditional probabilities are given by:

$$\begin{aligned} P_{12} &= \frac{r_1}{r_1 + r_2 + r_3} \frac{r_2}{r_1 + r_2 + r_3} = \frac{r_1 r_2}{(r_1 + r_2 + r_3)^2}, \\ P_{21} &= \frac{r_2}{r_1 + r_2 + r_3} \frac{r_1}{r_1 + r_2 + r_3} = \frac{r_2 r_1}{(r_1 + r_2 + r_3)^2}. \end{aligned} \tag{8.13}$$

The contribution to the likelihood function for period 1 is then:

$$P_{12} + P_{21} = \frac{2 r_1 r_2}{(r_1 + r_2 + r_3)^2}. \tag{8.14}$$

Comparison of the contributions to the likelihood function from tied failures when the Breslow approximation is employed rather than the marginal method shows why the former is less computationally demanding.

As discussed above, proportional hazards models assume that all plants share the same baseline hazard. In many cases, the assumption of a single baseline hazard for all plants is not satisfied. It is then necessary to use a stratified proportional hazards model in which the parameters of the model are constrained to be the same for all plants, but different groups of plants are allowed to have different baseline hazards. Tests of the proportional hazards assumption are based on residuals developed by Schoenfeld (1982). These residuals are regressed against time with a statistically significant coefficient on time indicating that the proportional hazards assumption does not hold.[46]

In a typical regression, one set of residuals is obtained for the entire regression. From the Cox proportional hazards model, a different set of Schoenfeld residuals is obtained for each covariate. In the absence of tied failures times, the Schoenfeld residual for covariate x_u is calculated as follows:

$$\varepsilon_{uj} = x_{uj} - \frac{\sum_{i \in R_j} x_{ui} \exp\left(x_i \hat{\beta}\right)}{\sum_{i \in R_j} \exp\left(x_i \hat{\beta}\right)}. \tag{8.15}$$

Equation (8.15) states that ε_{uj} is the difference between the value of x_u for the plant that failed and the weighted average of x_u for those plants at risk of failure in that period. The weights are determined from the estimated hazard rates in the Cox model with plants that are more likely to fail having higher hazard rates. To perform the proportional hazards test, we postulate that β is a function of time as follows:

$$\beta_u(t) = \beta_u + q_j g(t), \tag{8.16}$$

where β_u is the coefficient associated with x_u and $g(t)$ is a function of time. If the proportional hazards assumptions holds, q_j should equal zero. The Schoenfeld residuals can be scaled so that the following holds (Grambsch and Therneau, 1994):

$$E\left(\varepsilon_{uj}^{*} + \beta_u\right) = \beta_u(t), \tag{8.17}$$

[46] This is done in STATA 9.2 using the stphtest command.

where ε^*_{uj} are the rescaled Schoenfeld residuals associated with x_u. The test of proportional hazards is then performed by regressing ε^*_{uj} on t_j or $g(t)$. If the coefficient on the time variable is statistically significantly different from zero, this indicates that the proportional hazards assumption is not appropriate.

8.4. Results

The analysis is performed using the entire sample but also for the same industries as were used in chapter 7 - the food, textiles, footwear and clothing and paper, printing and publishing industries. Unlike in chapter 7, the model is also estimated using the entire sample because there is not such an obvious case against doing so as there is when the model is a production function due to the undesirability of imposing a common technology across different industries. Nevertheless, it is preferable to avoid imposing restrictions across disparate industries and this is why the model is also estimated using data from the four industries.

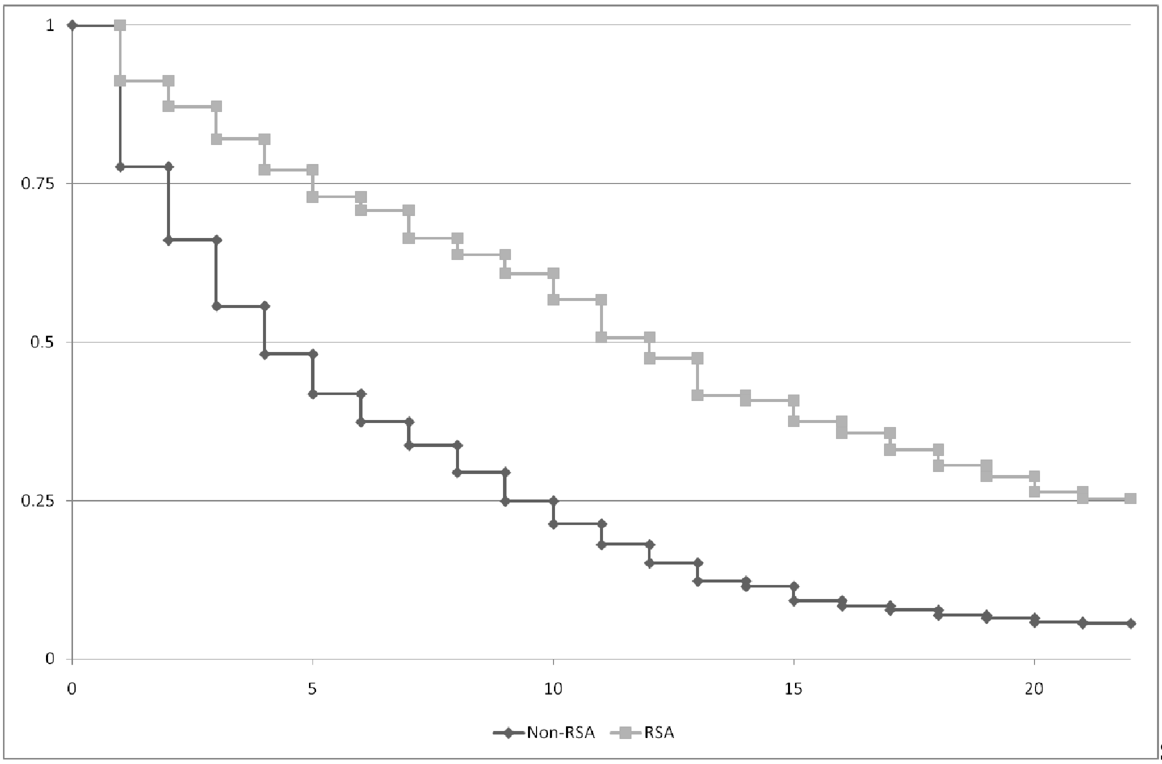
The survivor function, $S(t)$, is a useful means of examining differences in the probability of closure across treatment and untreated groups. It measures the probability of survival past time t. The Kaplan-Meier estimate of the survivor function is given by (Kaplan and Meier, 1958):

$$\hat{S}(t) = \prod_{j|t_j \leq t} \left(\frac{n_j - d_j}{n_j} \right), \tag{8.18}$$

where n_j is the number of plants in the risk pool at time t_j and d_j is the number of plants that close at time t_j.

Figure 8.1 gives Kaplan-Meier estimates of the survivor function for plants that received an RSA grant and for those that did not receive a grant for the full sample containing all industries.

Figure 8.1: Kaplan-Meier estimate of the survivor function for all industries



Source: SAMIS/ARD

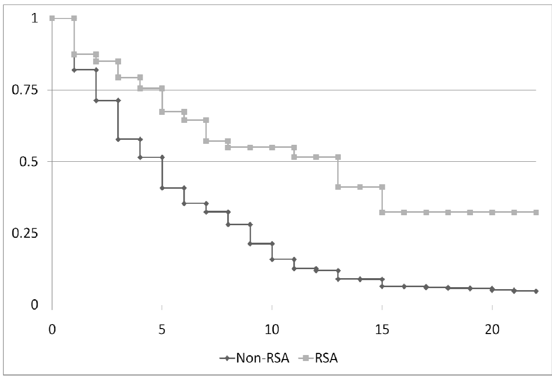
The estimate of the survivor function is higher for RSA recipients than for plants that did not receive assistance at all times which shows that RSA recipients always have a higher probability of survival after receiving assistance.[47]

Figure 8.2 provides Kaplan-Meier estimates of the survivor function for the industries that will be used for estimating equation (8.2).

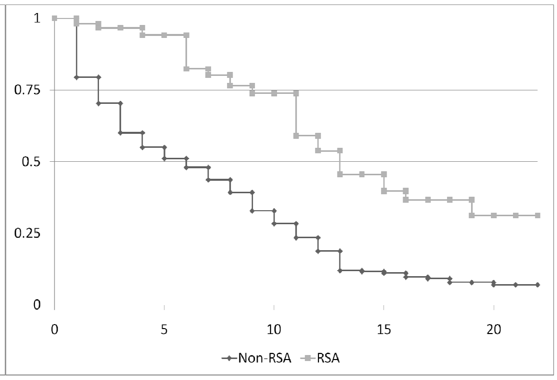
[47] The estimates of both survivor functions are equal to one in the first year because plants that fail in their first year are not observed.

Figure 8.2: Kaplan-Meier estimate of the survivor function for individual industries

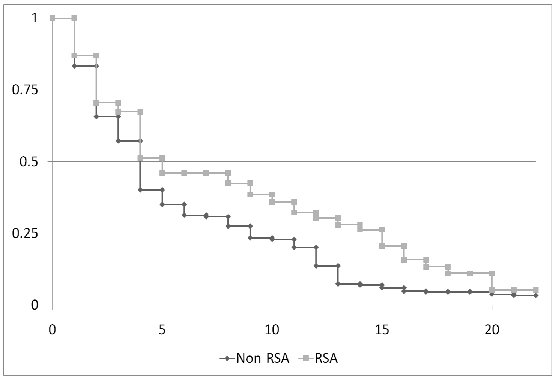
Food



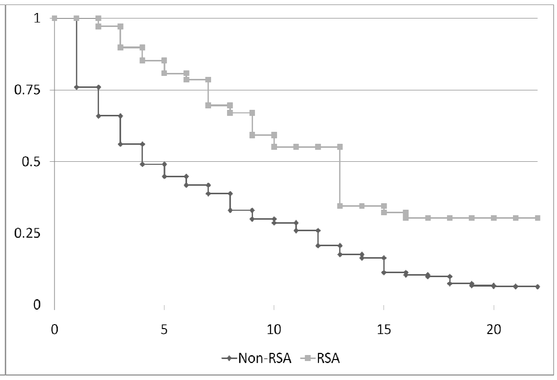
Textiles



Footwear and Clothing



Paper, Printing and Publishing



Source: SAMIS/ARD

For all four industries, plants that have received an RSA grant have a higher survivor function than plants that have not received a grant for each period of time. The gap in the probability of survival is particularly large for the textiles industry but quite small for the footwear and clothing industries. Nevertheless, it is clear that the probability of survival is larger for RSA recipients in all four industries. The purpose of the remainder of the chapter is to establish whether this difference in survival probabilities is caused by the difference in treatment status or whether it is the consequence of other differences in plant characteristics between plants that received an RSA grant and plants that did not.

The proportional hazards test revealed that employment, single plant status and age did not satisfy the proportional hazards assumption requiring covariates to multiplicatively shift the baseline hazard by the same amount through time. The model was therefore stratified by three employment size bands, age and the single plant dummy.[48] As the tables below reveal, the null of proportional hazards cannot now be rejected at the 10% level in any of the models estimated.

As in chapter 7, the models had to be estimated using weighted data to make the results representative of the population of plants because of the stratified sampling frame of the ABI. The sampling frame of the ABI and the construction of the weights are discussed in detail in chapter 5.

The estimated hazard rates of the Cox proportional hazards model using the full sample are displayed in table 8.1. These estimates are obtained using no control for self-selection bias and are presented to allow comparison with the results obtained using the matched sample. Hazard rates are reported so a value greater (less) than one should be interpreted as meaning that larger values of the variable are associated with a larger (smaller) probability of closure.

---

[48] Many studies (see, for example, Disney, Haskel and Heden, 2003b and Harris and Li, 2007) include interactions between each variable in x_{it} and the age variable to control for the influence of age. This is not necessary here because the influence of age is controlled for by the stratification of the model.

Table 8.1: Estimates of Cox Proportional Hazards Model obtained using no control for self-selection

	All Industries	Food	Textiles	Footwear and Clothing	Paper, Printing and Publishing
RSA	0.878** (0.054)	0.753 (0.181)	0.565*** (0.110)	0.713** (0.112)	1.071 (0.223)
ln(Labour Productivity)	0.901*** (0.011)	0.816*** (0.036)	0.820*** (0.058)	0.909 (0.068)	0.940** (0.025)
ln(Employment)	0.923*** (0.012)	0.976 (0.038)	0.937 (0.058)	0.771*** (0.049)	0.819*** (0.032)
ln(Capital to Labour Ratio)	1.251*** (0.013)	1.242*** (0.038)	1.196*** (0.051)	1.303*** (0.072)	1.290*** (0.049)
ln(Herfindahl Index)	1.026* (0.014)	0.967 (0.044)	1.027 (0.077)	0.874* (0.063)	1.047 (0.036)
ln(Displacement)	0.984 (0.014)	0.998 (0.039)	1.007 (0.055)	1.075 (0.112)	1.091 (0.076)
ln(Industry Employment Growth)	0.835** (0.073)	1.110 (0.403)	0.956 (0.351)	1.729 (0.681)	0.847 (0.388)
Ownership change	1.238*** (0.035)	1.181*** (0.070)	1.460*** (0.210)	1.306 (0.233)	1.504*** (0.163)
FDI	0.939** (0.026)	0.762*** (0.079)	0.768* (0.111)	0.731 (0.140)	0.989 (0.107)
Log Pseudolikelihood	-35220.57	-3380.24	-711.61	-724.32	-1603.46
Proportional Hazards Test	35.58	1.55	2.55	1.38	0.95
Observations	45,617	6,501	2,526	1,771	4,762

* denotes significance at the 90% level, ** denotes significance at the 95% level, *** denotes significance at the 99% level
Robust standard errors are in parentheses
A full set of 2-digit industry dummies is included but not reported for the ‘all industries’ model

The statistical significance and magnitude of the hazard rates on the x_{it} variables do not vary much across the industries considered. Therefore, it is sufficient to discuss the coefficients obtained using the full sample. These show that plants with higher labour productivity and plants with more employees are less likely to exit. As these variables are in the model to proxy for efficiency, this is what would be expected. However, as in Harris and Trainor (2007), the hazard rate associated with the capital per worker variable is greater than one. This is surprising as if this variable is a good proxy for sunk costs, it should be negatively related to closure probability. As discussed above, because plants may use an RSA grant to incur sunk costs, this implies that, through this channel, receipt of an RSA grant has a positive impact on closure probability. But, given that this result is likely to be the consequence of the capital to labour ratio being an unsatisfactory proxy for sunk costs, this inference is of dubious value. Therefore, we may henceforth focus on the hazard rate associated with the RSA dummy as measuring the full impact of receipt of an RSA grant on closure probability.

Another strange feature of the results is that the Herfindahl Index is positively associated with closure, meaning that plants operating in less competitive industries are more likely to close. As discussed in chapter 7.4, the deficiencies of the Herfindahl Index may be the cause of this result. Another surprise is that the displacement variable is negatively related to closure probability, although the impact is very small. A priori expectations are that greater displacement should be positively associated with closure probability. The hazard ratio associated with industry employment growth is less than one which reflects this variable's inclusion to capture improving or deteriorating market conditions. Also in line with expectations, the ownership change variable is positively related to closure probability. Finally, FDI plants are found to be less likely to close.

Of greatest interest here is the hazard ratio associated with the RSA dummy. For the full Scottish sample, this is less than one and statistically significant at the 95% level, indicating that being an RSA grant recipient is associated with a lower probability of closure, ceteris paribus. This is also the case for the textiles and the footwear and clothing industries. The estimated hazard rate on the RSA dummy is less than one for the food industry but not statistically significant. For the paper, printing and publishing industry, the estimated hazard rate actually suggests a positive relationship between receipt of an RSA grant and closure but is not significantly different from one.

Table 8.2 presents results obtained from estimating the proportional hazards model using a matched sample. As all the control variables were included in the probit model, from which the predicted values were taken and used for matching, their variance in the matched sample is reduced. As interest lies mainly in the hazard rate associated with the RSA dummy rather than the other variables included in the model, this does not present a problem.

Table 8.2: Estimates of Cox Proportional Hazards Model obtained using Propensity Score Matching

	All Industries	Food	Textiles	Footwear and Clothing	Paper, Printing and Publishing
RSA	0.945 (0.056)	0.609 (0.195)	1.042 (0.098)	0.903 (0.164)	0.907 (0.160)
ln(Labour Productivity)	0.871*** (0.017)	0.911* (0.050)	0.876** (0.055)	0.699 (0.161)	0.844 (0.134)
ln(Employment)	0.969 (0.026)	1.009 (0.103)	0.745* (0.115)	0.716* (0.130)	1.128 (0.122)
ln(Capital to Labour Ratio)	1.185*** (0.021)	1.081 (0.077)	1.477*** (0.135)	1.242 (0.217)	1.035 (0.111)
ln(Herfindahl Index)	1.018 (0.022)	1.076 (0.087)	1.112 (0.104)	0.451** (0.146)	1.003 (0.102)
ln(Displacement)	0.958 (0.026)	1.079 (0.067)	0.980 (0.064)	0.355** (0.163)	0.835 (0.158)
ln(Industry Employment Growth)	0.876 (0.097)	0.095** (0.090)	1.663 (0.605)	2.892 (3.775)	0.410 (0.508)
Ownership change	1.125*** (0.045)	1.171 (0.190)	1.146 (1.146)	1.218 (0.234)	1.146 (0.233)
FDI	0.883*** (0.035)	0.867 (0.217)	0.653** (0.111)	0.731 (0.266)	1.011 (0.182)
Log Pseudolikelihood	-5566.07	-195.60	-122.10	-59.35	-188.45
Proportional Hazards Test	8.86	1.26	0.54	1.03	1.93
Observations	28,501	2,667	2,552	1,551	2,811

* denotes significance at the 90% level, ** denotes significance at the 95% level, *** denotes significance at the 99% level
Robust standard errors are in parentheses
A full set of 2-digit industry dummies is included but not reported for the 'all industries' model

In the matched sample for all industries, the coefficients on labour productivity, the capital to labour ratio, the ownership change dummy and the FDI dummy all retain their sign and statistical significance from the full sample. The displacement variable remains not statistically significant at the 90% level. Employment, the Herfindahl Index and the industry employment growth variables all lose their statistical significance. Generally, the loss of statistical significance in the matched sample compared to the full sample is also seen in the individual industries considered. This is the effect of their reduced variance in the matched sample.

For the matched sample that includes all industries, the hazard rate on the RSA dummy is no longer statistically significant. This suggests that receipt of an RSA grant had no impact on the probability of survival. Compared to the results from using the full sample, the point estimate of the hazard ratio is closer to one which implies that misspecification leads to an overestimate of the impact of RSA on closure probability.

The same loss of significance is seen when the model is estimated by industry. When the matched sample is used, there is no longer any statistically significant reduction in closure probability associated with receipt of an RSA grant for the textiles and footwear and clothing industries. For the paper, printing and publishing industry, the hazard rate associated with the RSA dummy is not statistically significant in the full sample and this remains the case in the matched sample. The most curious result comes from the food industry in which the hazard rate associated with the RSA dummy is lower in the matched than the full sample, implying a larger reduction in closure probability from receiving an RSA grant. However, it remains not statistically significant. This shows that misspecification caused by self-selection can lead to both over and under estimates of the ATT but, mostly, misspecification leads to an overestimate of the reduction in closure probability caused by receipt of an RSA grant.

In order to test the robustness of these results, a number of variations on equation (8.2) were estimated. Firstly, the treatment dummy was substituted for a variable measuring the total value of RSA grants that the plant had received up to that point in time. This would show the impact on closure probability of receipt of an extra pound of RSA. However, in neither the matched sample including all industries or the matched samples for individual industries was this variable statistically significant.

The next approach was to replace the simple RSA dummy used in the specification above with a dummy that equalled one from the time at which a grant for safeguarding employment was provided and a dummy that equalled one from the time at which a grant for creating new employment was provided. Neither dummy tended to be statistically significant using the matched samples which indicates that the motivation for seeking the grant does not lead to differing impacts on survival.

Tests were also run to see if receiving different numbers of RSA grants has a differential impact on survival probability. This was done by replacing the RSA dummy employed above with a series of dummies that took the value of one if the plant had received a given numbers of grants. Although some of these dummies were statistically significant, there was no clear pattern indicating that more grants led to a greater reduction or increase in closure probability. The reliability of these results is also dubious given that, particularly for individual industries, there can be very few plants that received large numbers of grants.

These robustness tests therefore cast little doubt on the central finding of this chapter that receipt of an RSA grant has no statistically significant impact on plant survival when a matched sample is used to combat self-selection bias.

8.5. Conclusion

This chapter has sought to establish the existence of a causal impact of receipt of an RSA grant on the probability of survival. In line with the results obtained by Harris and Trainor (2007), receiving an RSA grant was found to reduce the probability of closure using an unmatched sample containing all industries. This was also the case for the textiles and the footwear and clothing industries. No statistically significant effect was found for the food and the paper, printing and publishing industries.

When the Cox model was estimated a matched sample of plants from all industries, no statistically significant impact of receipt of an RSA grant on closure probability was found. This differs from the results obtained by Girma, Görg and Strobl (2007b) in their analysis of the impact of government grants on survival in Ireland who found that receipt of a grant reduced the probability of closure. The impact of RSA on closure probability was also not statistically significant in any of the models estimated for the individual industries. This implies that, in general, self-selection, if ignored, leads to an overestimate of the impact of receipt of an RSA grant on closure probability.

A8.1 Covariate Balance in the Full and Matched Sample for Survival Analysis

This appendix gives information on the distribution of each variable across treated and untreated groups in the full and the matched sample for the datasets used to produce the results given in tables 8.1 and 8.2. The mean of each variable is presented and t-tests are employed to establish if this difference is statistically significant. Kolmogorov-Smirnov tests are also performed which test the null that the variable in the treated and untreated groups are drawn from the same distribution. This information is valuable as it provides an indication of the extent to which the balance of the observed covariates across the treated and untreated groups is improved and therefore the extent to which problems of misspecification caused by self-selection are alleviated by using the matched instead of the full sample.

Table A8.1 provides this information for the sample containing all industries.

Table A8.1: Distribution of Observed Covariates across the Full and Matched Sample for All Industries for Survival Analysis

	Full Sample				Matched Sample			
	Non-RSA	RSA	Difference	K-S	Non-RSA	RSA	Difference	K-S
Ln(Labour Productivity)	3.18	2.58	0.60***	0.24***	2.38	2.40	-0.02	0.05***
Ln(Employment)	2.73	3.67	-0.94***	0.38***	3.58	3.68	-0.10***	0.09***
Ln(Capital to Labour Ratio)	-5.27	-3.70	-1.57***	0.29***	-3.25	-3.18	-0.07***	0.03***
Ln(Herfindahl Index)	-2.59	-2.20	-0.40***	0.16***	-2.04	-2.07	0.03**	0.03***
Ln(Age)	2.26	2.27	-0.01	0.14***	2.30	2.31	-0.01	0.05***
Ln(Displacement)	-2.57	-2.12	-0.45***	0.17***	-2.03	-1.99	-0.03**	0.02
Industry Employment Growth	0.04	-0.01	0.04***	0.06***	-0.02	-0.02	.00*	0.02**
Ownership Change	0.32	0.44	-0.12***	0.25***	0.45	0.50	-0.06***	0.10***
FDI	0.13	0.21	-0.08***	0.15***	0.15	0.18	-0.03***	0.05***
Single	0.24	0.33	-0.09***	0.14***	0.20	0.23	-0.04***	0.05***

* denotes significance at the 90% level, ** denotes significance at the 95% level, *** denotes significance at the 99% level
Source: SAMIS/ARD

In the full sample, the difference in the mean of each variable, with the exception of the age variable, across treated and untreated groups is statistically significant at the 99% level. Moving from the full to the matched sample leads to a reduction in the difference for all variables. However, these differences remain statistically significant at the 90% level for all variables apart from labour productivity and age in the matched sample. The Kolmogorov-Smirnov statistics tell a similar story. The null of the treated and untreated groups being drawn from the same distribution is rejected at the 99% level for all variables in the full sample. In the matched sample, the null is rejected at this level for all variables apart from the displacement variable and the industry employment growth variable although the size of the test statistic is once more smaller for every variable. This shows that although propensity score matching improves the balance of the covariates across treated and untreated groups, large differences in distribution remain in the matched sample.

Table A8.2 shows the results for the food industry.

Table A8.2: Distribution of Observed Covariates across the Full and Matched Sample for the Food Industry for Survival Analysis

	Full Sample				Matched Sample			
	Non-RSA	RSA	Difference	K-S	Non-RSA	RSA	Difference	K-S
Ln(Labour Productivity)	2.67	2.38	0.29***	0.12***	2.22	2.28	-0.07*	0.12***
Ln(Employment)	2.60	3.54	-0.94***	0.43***	3.54	3.92	-0.38***	0.24***
Ln(Capital to Labour Ratio)	-5.43	-3.74	-1.70***	0.34***	-2.83	-3.10	0.27***	0.21***
Ln(Herfindahl Index)	-2.15	-1.80	-0.36***	0.25***	-1.53	-1.54	0.00	0.06
Ln(Age)	2.03	2.09	-0.06	0.13***	2.15	2.23	-0.08***	0.14***
Ln(Displacement)	-2.47	-2.17	-0.31***	0.10***	-1.87	-2.03	0.16***	0.15***
Industry Employment Growth	0.03	0.01	0.02	0.05	-0.01	-0.02	0.00	0.07*
Ownership	0.25	0.35	-0.10***	0.19***	0.49	0.43	0.06***	0.12***
FDI	0.03	0.07	-0.04***	0.06*	0.07	0.16	-0.09***	0.16***
Single	0.19	0.36	-0.17***	0.24***	0.12	0.17	-0.05***	0.09**

* denotes significance at the 90% level, ** denotes significance at the 95% level, *** denotes significance at the 99% level
Source: SAMIS/ARD

Only the difference in the mean of the age and the industry employment growth variables are not statistically significant at the 99% level. In the matched sample, this is true of the labour productivity, Herfindahl Index and industry employment growth variables. The size of the difference in mean falls between the full and the matched samples for all variables apart from the age and FDI variables. The Kolmogorov-Smirnov statistics are statistically significant for all variables at the 95% level with the exception of the industry employment growth and FDI variables in the full sample and the Herfindahl Index and industry employment variables in the matched sample. As is clear from above, it is possible that some variables are more poorly balanced in the matched than in the full sample. This occurs when well balanced variables in the full sample are correlated with poorly balanced variables. Because the former will have small coefficients in the probit as they have little explanatory power for treatment status and the latter will have large coefficients because they have large explanatory power for treatment status, matching on the propensity score leads to an improvement in the balance of those variables that were poorly balanced in the full sample at the expense of those that were well balanced.

Table A8.3 provides the same information for the textiles industry.

Table A8.3: Distribution of Observed Covariates across the Full and Matched Sample for the Textiles Industry for Survival Analysis

	Full Sample				Matched Sample			
	Non-RSA	RSA	Difference	K-S	Non-RSA	RSA	Difference	K-S
Ln(Labour Productivity)	2.66	2.23	0.43***	0.17***	2.30	2.38	-0.08***	0.10***
Ln(Employment)	3.41	3.56	-0.15***	0.13***	3.70	3.58	0.12***	0.14***
Ln(Capital to Labour Ratio)	-4.99	-3.79	-1.20***	0.29***	-3.38	-3.40	0.02	0.05
Ln(Herfindahl Index)	-2.60	-2.11	-0.49***	0.21	-2.04	-2.06	0.02	0.06
Ln(Age)	2.50	2.34	0.15***	0.19***	2.52	2.48	0.04	0.08**
Ln(Displacement)	-2.95	-2.37	-0.57***	0.20***	-2.41	-2.37	-0.04	0.02
Industry Employment Growth	-0.01	0.04	-0.06	0.07*	-0.05	-0.06	0.00	0.06
Ownership	0.30	0.31	-0.01	0.09***	0.37	0.47	-0.10***	0.16***
FDI	0.05	0.07	-0.03**	0.04	0.08	0.10	-0.02	0.04
Single	0.24	0.24	0.00	0.05	0.25	0.24	0.02	0.07*

* denotes significance at the 90% level, ** denotes significance at the 95% level, *** denotes significance at the 99% level
Source: SAMIS/ARD

The difference in mean across treated and untreated groups is statistically significant at the 95% level for all variables except the industry employment growth, ownership change and single enterprise plant variables in the full sample. In the matched sample, only the difference in mean for the labour productivity, employment and ownership change variables is statistically significant. The difference in mean falls for all variables apart from the ownership change and single plant variables. The Kolmogorov-Smirnov statistics also fall for most variables and are statistically significant at the 95% level for all variables but the Herfindahl Index, industry employment growth, FDI and single plant enterprise variables in the full sample and for the labour productivity, employment, age and ownership change variables in the matched sample. The covariates are therefore relatively well matched in the matched sample compared to the matched sample for all industries and the food industry.

Table A8.4 provides the same statistics for the footwear and clothing industry.

Table A8.4: Distribution of Observed Covariates across the Full and Matched Sample for the Footwear and Clothing Industry for Survival Analysis

	Full Sample				Matched Sample			
	Non-RSA	RSA	Difference	K-S	Non-RSA	RSA	Difference	K-S
Ln(Labour Productivity)	2.70	2.13	0.56***	0.32***	2.02	1.83	0.19***	0.20***
Ln(Employment)	3.18	3.64	-0.46***	0.23***	3.63	3.62	0.01	0.10**
Ln(Capital to Labour Ratio)	-6.54	-4.36	-2.18***	0.42***	-4.18	-3.66	-0.52***	0.32***
Ln(Herfindahl Index)	-2.63	-2.01	-0.62***	0.34***	-2.05	-1.79	-0.25***	0.19***
Ln(Age)	2.24	2.17	0.07	0.20***	2.35	2.16	0.19***	0.28***
Ln(Displacement)	-2.14	-1.68	-0.46***	0.25***	-1.70	-1.56	-0.15***	0.14***
Industry Employment Growth	-0.03	-0.05	0.02	0.05	-0.03	-0.04	0.01	0.07
Ownership	0.23	0.36	-0.13***	0.24***	0.33	0.35	-0.02	0.16***
FDI	0.06	0.14	-0.09***	0.16***	0.02	0.11	-0.10***	0.15***
Single	0.31	0.26	0.05*	0.11***	0.26	0.18	0.08***	0.12***

* denotes significance at the 90% level, ** denotes significance at the 95% level, *** denotes significance at the 99% level
Source: SAMIS/ARD

The age, industry employment growth and single plant enterprise variables are the only variables for which the difference in the mean across treated and untreated groups is not statistically significant in the full sample. The difference in the mean falls between the full and the matched sample for every variable apart from the age and the FDI variables. In the matched sample, the difference in mean across treated and untreated groups is statistically significant for all variables with the exceptions of the employment, industry employment growth and ownership change variables. The Kolmogorov-Smirnov test statistics are statistically significant for all variables but the industry employment growth variable in both the full and matched samples. The covariates are therefore similarly balanced across treated and untreated groups in the matched sample for the footwear and clothing industry as in the matched sample for all industries and the food industry but not as well balanced as in the matched sample for the textiles industry.

Table A8.5 provides the same information for the paper, printing and publishing industry.

Table A8.5: Distribution of Observed Covariates across the Full and Matched Sample for the Paper, Printing and Publishing Industry for Survival Analysis

	Full Sample				Matched Sample			
	Non-RSA	RSA	Difference	K-S	Non-RSA	RSA	Difference	K-S
Ln(Labour Productivity)	3.14	2.75	0.39***	0.20***	2.41	2.71	-0.30***	0.33***
Ln(Employment)	2.67	3.83	-1.16***	0.52***	3.49	3.70	-0.21***	0.19***
Ln(Capital to Labour Ratio)	-4.90	-3.75	-1.15***	0.27***	-2.92	-3.14	0.22***	0.22***
Ln(Herfindahl Index)	-3.08	-2.84	-0.25***	0.17***	-2.32	-2.70	0.38***	0.19***
Ln(Age)	2.39	2.32	0.07	0.20***	2.26	2.48	-0.22***	0.33***
Ln(Displacement)	-2.36	-2.29	-0.07	0.13***	-1.91	-2.31	0.39***	0.29***
Industry Employment Growth	0.00	-0.02	0.02**	0.20***	-0.01	-0.02	0.01*	0.23***
Ownership	0.35	0.45	-0.10***	0.25***	0.42	0.52	-0.10***	0.28***
FDI	0.12	0.17	-0.05***	0.11***	0.13	0.12	0.01	0.07
Single	0.22	0.37	-0.15***	0.20***	0.23	0.34	-0.10***	0.18***

* denotes significance at the 90% level, ** denotes significance at the 95% level, *** denotes significance at the 99% level
Source: SAMIS/ARD

In the full sample, the difference in mean across treated and untreated groups is statistically significant for all variables apart from the displacement and industry employment growth variables at the 95% level. In the matched sample, the difference is statistically significant at the 95% level for every variable but the ownership change and industry employment growth variables. The difference is larger in the matched than the full sample for the Herfindahl Index, age and displacement variables and smaller for all other variables. The Kolmogorov-Smirnov statistics indicate rejection of the null that the treated and untreated groups are drawn from the same distribution for all variables in the full sample and all variables apart from the FDI variables in the matched sample. This is therefore the worst balanced matched sample of all those created.

Overall, the extent to which the distribution of the observed covariates across treatment and untreated groups are balanced by propensity score matching is similar to that achieved in chapter 7 in which the dependent variable was output rather than the hazard rate. Although the balance is substantially improved in the matched sample, there remain large differences. This therefore implies that estimating a multivariate hazard model rather than simply comparing hazard rates across treatment and control groups in the matched sample is the correct approach because failure to control for differences in the distribution of covariates in the matched sample could generate biased estimates of the treatment effect.

9. Conclusion

9.1. Introduction

This thesis has investigated the impact of receipt of an RSA grant on plant performance. It has done so using a database created by linking a register of recipients of RSA grants into the longitudinal ARD which contains the financial information on factor inputs and outputs necessary for empirical analysis. The first empirical chapter consisted of a decomposition of the growth of labour productivity and TFP in Scottish manufacturing plants between 1994 and 2004. The next two empirical chapters contained analyses of the impact of receipt of an RSA grant on TFP and of the impact of receipt of an RSA grant on survival in Scottish manufacturing plants between 1984 and 2004.

The next section will describe the contribution to the literature that has been made by this thesis. The third section will set out the main findings from the empirical analyses of chapters 6, 7 and 8. The fourth section will make some policy recommendations on the basis of these findings. The fifth section will provide some suggestions for future work that could be done in this area. The final section concludes.

9.2. Contribution to the Literature

The contribution to the literature has been primarily twofold. Firstly, the dataset that was used for the empirical analyses is an improvement on previous datasets created by linking SAMIS with the ARD that have been used for estimating the impact of receipt of an RSA grant on plant performance. This is because over 91% of the plants that received an RSA grant in Scotland have been identified in the dataset as doing so. This is a higher proportion than has been achieved in previous studies. This is of crucial importance when estimating treatment effects because, assuming that there is some impact of receiving a grant, incorrectly classifying a plant that received treatment as not having received treatment will lead to downward biased estimates of the treatment effect.

The second contribution to the literature was methodological. The chapter that investigated the impact of receipt of an RSA grant on TFP used appropriate methods for dealing with both the endogeneity of factor inputs and the consequences of self-selection into the treatment group. Previous papers that have analysed the impact of receipt of an RSA grant on plant performance have only dealt with one of these sources of bias (see, for example, Criscuolo, Martin, Overman and Van Reenen, 2007; Harris and Robinson, 2004). The chapter in this thesis that examined the impact of receipt of an RSA grant on the probability of survival dealt with self-selection bias by using a sample created using propensity score matching. This is the first time that the impact of receipt of an RSA grant on the probability of survival has been analysed using any method that deals with the consequences of self-selection.

9.3. Main Findings

The first empirical chapter decomposed the growth of aggregate labour productivity and TFP between 1994 and 2004. This showed that plants that received an RSA grant in this period of time made a positive but small contribution to the growth of both measures of productivity. The bulk of this contribution came from plants that existed in both 1994 and 2004 and, more specifically, through the coincidence of improvements in productivity and increases in market shares. More concerning is the negative contribution to labour productivity growth and the small (but positive) contribution to TFP growth from RSA-assisted entrants. The contribution from non-assisted entrants to the growth of productivity is the largest component in both decompositions which suggests that RSA grants need to be better targeted at entrants.

Chapter 7 revealed that receipt of an RSA grant had no statistically significant impact on TFP for the food, textiles, footwear and clothing and paper, printing and publishing industries. Although there was considerable movement in the point estimate of the treatment effect across different estimators, this finding of no statistically significant impact was found for all four industries when no control for self-selection bias was employed, when an instrumental variables estimator was used and when the sample used for estimation was created using propensity score matching. It is therefore robust to choice of estimators.

Chapter 8 showed that receipt of an RSA grant did not have a statistically significant impact on the probability of closure when the Cox proportional hazards model was estimated using a sample created by propensity score matching. This was the case regardless of whether a matched sample containing plants from all industries was used or whether individual industries were used. When the full sample was used, a statistically significant reduction in closure probability was found for the sample containing plants from all industries and for the textiles and footwear and clothing industries. This implies that the finding of a statistically significant impact is driven by a misspecification of the model.

In sum, the results from the empirical analysis are not generally supportive of the RSA scheme in its current form. Although it must be recalled that previous studies have shown that receiving an RSA grant leads to increases in employment, investment and therefore output (see chapter 1.3) and therefore that the scheme makes a contribution to the growth of GDP per capita, Productivity enhancing effects of receiving an RSA grant would clearly lead to a larger contribution to the growth of living standards. An increase in TFP caused by receipt of an RSA grant would allow the owners of the firm to make larger profits or to provide higher wages, both of which would contribute to the growth of GDP per capita. An increase in TFP may also induce the firm to increase employment beyond the levels they are obliged to employ as a condition of receiving the grant if the demand curve is sufficiently elastic. That receipt of an RSA grant does not lead to such an improvement in TFP therefore implies that, assuming that they are able to stay open, RSA grant recipients are not contributing as much to the growth of living standards as they would if RSA grants had productivity enhancing effects.

However, the assumption that RSA grant recipients are able to continue operating is not a trivial one, especially if the grant was provided to safeguard employment. It is reasonable to assume that plants that require a grant to safeguard employment and are unable to obtain the necessary finance for capital investment from elsewhere are at higher risk of closure than the average. The finding that receipt of an RSA grant does not lead to a statistically significant reduction in closure probability is therefore worrying. The problem is not so great for plants that apply for a grant to increase their levels of employment as they are less likely to be at risk of closure, even though they must also, according to the rules of the scheme, be unable to obtain finance from the private sector. Given that the creation and safeguarding of employment is the main aim of the scheme, this finding is concerning as it casts doubt upon whether the jobs that are created and safeguarded by RSA grants will endure and suggests that the contribution to living standards from increases in employment, capital and output occasioned by receiving an RSA grant may be transitory.

9.4. Policy Recommendations

Overall, these results suggest that the RSA scheme should be geared more towards improving the productivity of recipient plants rather than merely safeguarding and promoting employment. This would allow plants that receive a grant to make a greater contribution to the growth of living standards. As discussed in chapter 2.6, an eligibility criterion to focus the English version of the RSA scheme more on improving productivity was introduced in 2004. This stipulates that the jobs created or safeguarded by projects must be relatively productive as determined by a comparison of GVA per employee with the sectoral and national averages. A similar criterion should also be applied to the RSA scheme in Scotland.

In order to ensure that projects satisfy this criterion, grants should only be provided to support investment in capital which will significantly improve the productivity of the plant. This will invariably mean capital which embodies the latest technologies. One example of such capital could be information and communications technology, the accumulation of which has been shown by some studies to lead to improvements in TFP (see, for example, Van Ark, 2001; O'Mahony and Vecchi, 2003).

If the RSA scheme cannot be changed so that grants have a strong productivity enhancing impact, it is arguable from a Schumpeterian perspective that grants should no longer be provided for projects that seek to safeguard employment. As firms that apply for grants to safeguard employment must be unable to maintain their current levels of employment without a grant, they are likely to have lower productivity levels than those that apply for a grant to increase employment. Successfully safeguarding low productivity jobs will impede the process of creative destruction and lead to a lower rate of productivity growth and living standards.[49]

[49] Note that the results from chapter 8 cast doubt upon whether grants provided to safeguard employment are successful in achieving this aim.

9.5. Suggestions for Future Research

The size of the theoretical literature review relative to the empirical literature review suggests that there is considerable scope for additional work on the impacts of schemes such as RSA using general equilibrium models of the economy. Such studies would be useful in providing a greater understanding of the impact of such schemes at the macroeconomic level. In particular, dynamic models in the mould of Samaniego (2006) that allow for the productivity levels of the firm to vary over time assist in providing an understanding of how such schemes enhance or impede growth in the economy. Of particular relevance to schemes that provide grants in specific areas of the country such as RSA would be economic geography models that take explicit account of the significance of geography to firm performance.

Although it is clear that the failure to identify every plant that received an RSA grant in the ARD will, assuming the treatment effect is positive, lead to underestimates of the impact of receipt of a grant, it would be helpful to have some guidance on the extent to which the estimated treatment effect will be biased. Such guidance could be provided by simulation. The impact of failing to find difference percentages of plants that receive grants in the dataset could be analysed using different sizes of treated and untreated groups and different magnitudes of treatment effects. This information could then be used to develop standard errors for treatment effects that take account of the fact that the sample contains some plants that are erroneously classified as untreated. Given the large number of studies that depend on matched datasets, such a study would be useful to many researchers (see Harris and Trainor, 2005; Girma, Görg and Strobl, 2007b for examples of papers that have used linked data).

Simulation work would also be useful in order to establish the magnitude of bias in the estimate of the treatment effect that arises due to misspecification of the observable covariates when observations in the treatment group are self-selected. In chapter 8, moving from the full to the matched sample removed the statistical significance of the estimate of the treatment effect for some industries. This shows that the balance of the observed covariates across treatment and untreated groups is an important issue. Simulation work that showed, for differences in the extent to which the covariates are balanced across treated and untreated groups, the size of the bias for different specifications of the ‘real’ outcome equation would be helpful in allowing the researcher to understand whether creating a matched sample is necessary.

Further empirical work could be done to analyse the impact of receipt of an RSA grant on other dimensions of plant performance. In particular, it would be possible to analyse the impact of receipt of an RSA grant on R&D expenditure and innovativeness using a database created by linking SAMIS into the Community Innovation Survey (see Harris and Robinson, 2001a for information on the Community Innovation Survey). Such work would be of considerable interest in Scotland as the poor innovation performance of Scottish firms is a source of major concern (see Scottish Government, 2009c for information on the state of innovation in Scotland). The impact of receipt of an RSA grant on exporting could also be analysed by linking SAMIS into the Global Connections Survey (see Harris and Reid, 2009 for information on the Global Connections Survey). This would be of interest because of the benefits to both individual firms and the host country associated with exporting (see Harris and Reid, 2009 for a summary of the benefits associated with exporting).

9.6. Conclusion

This chapter is a conclusion to the thesis. It began by setting out the contribution made by this thesis to the literature. This was in the use of a superior dataset to those which have been used previously to analyse the impact of receipt of an RSA grant on plant performance and in the use of a superior econometric methodology which deals with both self-selection into the treatment group and other sources of endogeneity.

It then described the main findings from the empirical chapters. The decomposition showed that plants that received a RSA grant made a small but positive contribution to the growth of both labour productivity and TFP. Chapter 7 and 8 showed that receipt of an RSA grant had no statistically significant impact on TFP or the probability of survival.

On the basis of these findings, some policy recommendations were made. These were mainly that the RSA scheme should become more focused on improving the productivity of recipient plants by offering support for the acquisition of only the most technologically advanced capital. This would help to ensure that the jobs created or safeguarded by receipt of an RSA grant endure.

Finally, some suggestions for future work were made. Firstly, more general equilibrium models that look at the impact of grant schemes such as RSA at the macroeconomic level are required as there are currently very few of such papers. Secondly, simulation work that quantifies the implications of incorrectly classifying plants that received a grant as not having received a grant would be beneficial to the applied researcher as many empirical studies that estimate treatment effects rely on linking two datasets. Simulation work that provides a guide to the magnitude of bias caused by misspecification of the observable covariates when they are not perfectly balanced across the treated and untreated group would also be useful to the applied researcher. Finally, it was suggested that the impact of receipt of an RSA grant on R&D spend, innovativeness and exporting could be analysed by linking SAMIS with other available databases.

10. Bibliography

Ackerberg, D., Caves, K. and Frazer, G. (2006), 'Structural Identification of Production Functions', University of Toronto Working Paper, Available from: http://www.econ.ucla.edu/ackerber/ACF20withtables.pdf

Aghion, P., Harris, C., Howitt, P. and Vickers, J. (2001), 'Competition, imitation and growth with step-by-step innovation', Review of Economic Studies, 68, 3, 467–492

Altman, M. (2000), 'A behavioral model of path dependency: the economics of profitable inefficiency and market failure', The Journal of Socio-Economics, 29, 2, 127-145

Angrist, J. and Imbens, G. (1995), 'Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity', Journal of the American Statistical Association, 90, 430, 431-442

Angrist, J. and Krueger, A. (2001), 'Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments', Journal of Economic Perspectives, 15, 4, 69-85

Angrist, J. and Pischke, J. (2009), Mostly Harmless Econometrics: An Empiricist's Companion, Princeton and Oxford: Princeton University Press

Angrist, J., Imbens, G. and Krueger, A. (1999), 'Jackknife Instrumental Variables Estimation', Journal of Applied Econometrics, 14, 1, 57-67

Ankarhem, M., Daunfeldt, S., Quoreshi, S. and Rudholm, N. (2009), 'Do Regional Investment Grants Improve Firm Performance', Ratio Working Paper, 137

Arellano, M. and Bond, S. (1991), 'Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations', Review of Economic Studies, 58, 2, 277-97

Arellano, M. and Bover, O. (1995), 'Another Look at the Instrumental Variable Estimation of Error-Components Models', Journal of Econometrics, 68, 1, 29-51

Armstrong, H. and Taylor, J. (2000), Regional economics and policy, Oxford: Blackwell

Arthur, B. (1989), 'Competing technologies, increasing returns and lock-in by historical events', Economic Journal, 99, 394, 116-131Audretsch, D. (1994), 'Business Survival and the Decision to Exit', International Journal of the Economics of Business, 1, 1, 125-137

Bailey, M., Hulten, C. and Campbell, D. (1992), 'Productivity Dynamics in manufacturing plants', Brookings Papers on Economic Activity, Microeconomics, 187-249

Baltagi, B. (2005), Econometric Analysis of Panel Data, Chichester: Wiley

Baumol, W. (1984), 'On Productivity Growth in the Long Run', Atlantic Economic Journal, 12, 3, 4-10

Beason, R. and Weinstein, D. (1996), 'Growth, Economies of Scale, and Targeting in Japan (1955-1990)', Review of Economics and Statistics, 78, 2, 286-295

Bergström, F. (2000), 'Capital Subsidies and the Performance of Firms', Small Business Economics, 14, 3, 183-193

Bijwaard, G. (2008), 'Instrumental Variables Estimation for Duration Data', Tinbergen Institute Discussion Paper, 2008-032/4

Bjorklund, A. and Moffitt, R. (1987), 'The Estimation of Wage Gains and Welfare Gains in Self–Selection Models', Review of Economics and Statistics, 69, 1, 42–49

Blundell, R. and Bond, S. (1998), 'Initial conditions and moment restrictions in dynamic panel data models', Journal of Econometrics, 87, 1, 115-43

Blundell, R. and Bond, S. (2000), 'GMM estimation with persistent panel data: an application to production functions', Econometric Reviews, 19, 3, 321-340

Blundell, R. and Costa Dias, M. (2008), 'Alternative Approaches to Evaluation in Empirical Microeconomics', CETE Discussion Papers, 0805

Blundell, R. and MaCurdy, T. (1998), 'Labour Supply: A Review of Alternative Approaches,' Institute for Fiscal Studies Working Paper, W98/18

Blundell, R., Bond, S. and Windmeijer, F. (2000), 'Estimation in Dynamic Panel Data Models: Improving on the Performance of the Standard GMM Estimator', Institute for Fiscal Studies Working Paper, W00/12

Blundell, R., Dearden, L. and Sianesi, B. (2005), 'Evaluating the Effect of Education on Earnings: Models, Methods and Results from the National Child Development Survey', Journal of the Royal Statistical Society, 168, 3, 473-512

Bond, S. (2002), 'Dynamic panel data models: A guide to theory and practice', Cemmap Working Paper, CWP09/02

Bound, J., Jaeger, D. and Baker, R. (1995), 'Problems with Instrumental Variables Estimation when the Correlation between the Instruments and the Endogenous Explanatory Variable is Weak', Journal of the American Statistical Association, 90, 430, 443–50

Breslow, N. (1974), 'Covariance analysis of censored survival data', Biometrics, 30, 1, 89-99

Bronzini, R. and de Blasio, G. (2006), 'Evaluating the impact of investment incentives: The case of Italy's Law 488/1992', Journal of Urban Economics, 60, 2, 327-349

Cabral, L. (2000), Introduction to Industrial Organization, Cambridge, Massachusetts: MIT Press

Canning, D., Moore, B. and Rhodes, J. (1987), 'Economic Growth in Northern Ireland: Problems and prospects,' in Teague, P. (ed.), Beyond the Rhetoric, London: Lawrence and Wishart, 211-235

Card, D. and Robins, P. (1998), 'Do Financial Incentives Encourage Welfare Recipients To Work?' Research in Labor Economics, 17, 1, 1-56

Caves, R. and Porter, M. (1976), 'Barriers to Exit', in Qualls, P. and Masson, R. (eds.), Essays on Industrial Organisation in Honour of Joe S. Bain, Cambridge, MA: Ballinger, 36-69

Cleves, M., Gould, W. and Gutierrez, R. (2002), An introduction to survival analysis using Stata, College Station, Texas: Stata Press

Cobb, C. and Douglas, P. (1928), 'A Theory of Production', The American Economic Review, 18, 1, Supplement, Papers and Proceedings of the Fortieth Annual Meeting of the American Economic Association, 139-165

Cochran, W. and Rubin, D. (1973), 'Controlling bias in observational studies', Sankyha: The Indian Journal of Statistics, Series A, 35, 4, 417–46

Colombo, M. and Delmastro, M. (2000), 'A note on the relation between size, ownership status and plant's closure: sunk costs vs. strategic size liability', Economics Letters, 69, 3, 421-427

Colombo, M. and Delmastro, M. (2001), 'Technology use and plant closure', Research Policy, 30, 1, 21-34

Cox, D. (1972), 'Regression Models and Life-Tables', Journal of the Royal Statistical Society, 34, 2, 187-220

Criscuolo, C., Martin, R., Overman, H. and Van Reenen, J. (2007) 'The effect of industrial policy on corporate performance: Evidence from panel data', Centre for Economic Performance Working Paper, Available from: http://personal.lse.ac.uk/criscuol/papers/criscuolo.pdf

David, P. (1985), 'Clio and the economics of qwerty', American Economic Review, 75, 2, 332-337

De Long, J. and Summers L. (1991), 'Equipment Investment and Economic Growth', Quarterly Journal of Economics, 106, 2, 445–502

Dehejia, R. and Wahba, S. (2002), 'Propensity score matching methods for nonexperimental causal studies', Review of Economics and Statistics, 84, 1, 151–161

Denison, E. (1972), 'Final Comments', Survey of Current Business, 52, 5, 95-110
Diamond, A. (2004), 'Schumpeter's Central Message', Paper Presented to Conference of International Schumpeter Society in Milan, Available from: http://cba.unomaha.edu/faculty/adiamond/web/diamondpdfs/schumpcent.pdf
Disney, R., Haskel, J. and Heden, Y. (2003a), 'Restructuring and Productivity Growth in UK Manufacturing', The Economic Journal, 113, 489, 666-694
Disney, R., Haskel, J. and Heden, Y. (2003b), 'Entry, Exit and Establishment Survival in UK Manufacturing', The Journal of Industrial Economics, 51, 1, 91-112
Doms, M., Dunne, T. and Roberts, M. (1995), 'The role of technology use in the survival and growth of manufacturing plants', International Journal of Industrial Organisation, 13, 4, 523-542
EU (2010), State Aid Control, (online), Available from: http://ec.europa.eu/competition/state_aid/studies_reports/expenditure.html (Accessed March 2010)
European Commission (2009), Access to Finance, (online) Available from: http://ec.europa.eu/youreurope/business/access-to-finance/index_en.htm (Accessed March 2010)
Faulk, D. (2002), 'Do State Economic Development Incentives Create Jobs? An Analysis of State Employment Tax Credits', National Tax Journal, 55, 2, 263-280
Frölich, M. (2008), 'Parametric and Nonparametric Regression in the Presence of Endogenous Control Variables', International Statistical Review, 76, 2, 214-227
Fuest, C. and Huber, B. (2000), 'Why do governments subsidise investment and not employment', Journal of Public Economics, 78, 1-2, 171-192
Gabe, T. and Kraybill, D. (2002), 'The Effect of State Economic Development Incentives on Employment Growth of Establishments', Journal of Regional Science, 42, 4, 703-730
Gabszewicz, J. and Thisse, J. (1980), 'Entry (and Exit) in a Differentiated Industry', Journal of Economic Theory, 22, 2, 327-338
Gertler, M. (2003), 'Tacit Knowledge and the Economic Geography of Context, or the Undefinable Tacitness of Being (There)', Journal of Economic Geography, 3, 1, 75-99
Girma, S., Görg, H. and Strobl, E. (2007a), 'The effect of grants on plant level productivity', Economic Letters, 94, 3, 439-444
Girma, S., Görg, H. and Strobl, E. (2007b), 'The effects of government grants on plant survival: A micro-econometric analysis', International Journal of Industrial Organisation, 25, 4, 701-720

Gorter, C., Hassink, W., Nijkamp, P and Pels, E. (1997), ‘On the Endogeneity of Output in Dynamic Labour-Demand Models’, Empirical Economics, 22, 3, 393-408

Grambsch, P. and Therneau, T. (1994), ‘Proportional hazards tests and diagnostics based on weighted residuals’, Biometrika, 81, 3, 515-526

Griffith, R. (1999), ‘Using the ARD Establishment Level Data to look at Foreign Ownership and Productivity in the United Kingdom’, The Economic Journal, 109, 456, F416-F442

Griliches, Z. and Regev, H. (1992), ‘Productivity and firm turnover in Israeli industry 1979-88’, NBER Working Paper, 4059

Haltiwanger, J. (1997), ‘Measuring and analyzing aggregate fluctuations: the importance of building from microeconomic evidence’, Federal Reserve Bank of St Louis Review, May, 55–77

Harhoff, D. (2000), ‘R&D Spillovers, Technological Proximity, and Productivity Growth - Evidence from German Panel Data’, Schmalenbach Business Review, 52, 3, 238-260

Harris, R. (1991a), Regional Economic Policy in Northern Ireland, Aldershot: Avebury

Harris, R. (1991b), ‘The Employment Creation Effects of Factor Subsidies: Some Estimates for Northern Ireland Manufacturing Industry, 1955-1983’, Journal of Regional Science, 31, 1, 49-64

Harris, R. (2005a), ‘Longitudinal Microdata Study of Business Support Programmes’, Report to DTI

Harris, R. (2005b), ‘Economics of the Workplace: Special Issue Editorial’, Scottish Journal of Political Economy’, Scottish Journal of Political Economy, 52, 3, 323-343

Harris, R. (2005c), ‘Deriving measures of plant-level capital stock in UK manufacturing, 1973–2001’, Report to Department of Trade and Industry

Harris, R. (2010), ‘The Impact of SDI Support: Inward Investment 2001-2006’, Report to SQW

Harris, R. and Drinkwater, S. (2000), ‘UK Plant and Machinery Capital Stocks and Plant Closures’, Oxford Bulletin of Economics and Statistics, 62, 2, 243-265

Harris, R. and Li, C. (2007), ‘Firm Level Empirical Study of the Contribution of Exporting to Productivity Growth’, Report to UK Trade and Investment

Harris, R. and Reid, R. (2009), ‘The Role of Knowledge in Lowering Barriers to Internationalisation’, CPPR Working Paper, 21

Harris, R. and Robinson, C. (2001a), ‘Contract A: An Analysis of Current DTI Industry Support Patterns’, Report to Department of Trade and Industry

Harris, R. and Robinson, C. (2001b), 'Contract B: A critical review of empirical research on hindrances to business development and productivity growth and the relative importance of different constraints on UK business', Report to Department of Trade and Industry

Harris, R. and Robinson, C. (2004), 'Industrial Policy in Great Britain and its Effect on Total Factor Productivity in Manufacturing Plants, 1990-1998', Scottish Journal of Political Economy, 51, 4, 528-543

Harris, R. and Robinson, C. (2005), 'Impact of Regional Selective Assistance on Sources of Productivity Growth: Plant-level Evidence from UK Manufacturing, 1990-98', Regional Studies, 39, 6, 751-765

Harris, R. and Trainor, M. (2005), 'Capital Subsidies and their Impact on Total Factor Productivity: Firm-level Evidence from Northern Ireland', Journal of Regional Science, 45, 1, 49-74

Harris, R. and Trainor, M. (2007), 'The Impact of Government Intervention on Employment Change and Plant Closure in Northern Ireland, 1983-1997', Regional Studies, 41, 1, 51-63

Hart, M. and Scott, R. (1994), 'Measuring the Effectiveness of Small Firm Policy: Some Lessons from Northern Ireland', Regional Studies, 28, 8, 849-858

Hart, M., Driffield, N., Roper, S. and Mole, K. (2008a), 'Evaluation of Regional Selective Assistance (RSA) in Scotland', Report to Scottish Government

Hart, M., Driffield, N., Roper, S. and Mole, K. (2008b), 'Evaluation of Regional Selective Assistance (RSA) and its successor, Selective Finance for Investment in England (SFIE)', BERR Occasional Paper, 2

Hart, M., Driffield, N., Roper, S. and Mole, K. (2008c), 'Evaluation of Selective Financial Assistance (SFA) in Northern Ireland, 1998-2004', Report to DETINI

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

Heckman, J. and Navarro-Lozano, S. (2004), 'Using Matching, Instrumental Variables, and Control Functions to Estimate Economic Choice Models', Review of Economics and Statistics, 86, 1, 30-57

Heckman, J. and Vytlacil, E. (1999), 'Local Instrumental Variables and Latent Variable Models for Identifying the Bounding Treatment Effects', Proceedings of the National Academy of Sciences, 96, 8, 4730-4734

Heckman, J., Ichimura, H. and Todd, P. (1997), 'Matching as an econometric evaluation estimator: evidence from evaluating a job training programme', Review of Economic Studies, 64, 4, 605–54

Heckman, J., Ichimura, H., Smith, J. and Todd, P. (1998), 'Characterising Selection Bias Using Experimental Data', Econometrica, 66, 5, 1017-1098

Heckman, J., Smith, J. and Clements, N. (1997), 'Making The Most Out Of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts', Review of Economic Studies, 64, 4, 487-535

Herfindahl, O. (1950), 'Concentration in the US Steel Industry', PhD dissertation, Columbia University

Highlands and Islands Enterprise (2009), Regional Selective Assistance, (online) Available from: http://www.hie.co.uk/regional-selective-assistance.htm (Accessed September 2009)

HM Treasury (2003), Green Book: Appraisal and Evaluation in Central Government, London: The Stationary Office

Hopenhayn, H. (1992), 'Entry, Exit and Firm Dynamics in Long-Run Equilibrium', Econometrica, 60, 5, 1127-1150

House of Commons (2009), Industrial Development Act 1982 Annual Report, London: The Stationary Office

Hymer, S. (1976), The International Operations of National Firms: A Study of Direct Foreign Investment, Cambridge, MA: MIT Press

Imbens, G. and Angrist, J. (1994), 'Identification and Estimation of Local Average Treatment Effects', Econometrica, 62, 2, 467-475

Imbens, G. and Lemieux, T. (2008), 'Regression Discontinuity Designs: A Guide to Practice', Journal of Econometrics, 142, 2, 615-635

Imbens, G. and Wooldridge, J. (2009), 'Recent Developments in the Econometrics of Program Evaluation', Journal of Econometric Literature, 47, 1, 5-86

Jacobs, J. (1969), The economy of cities, New York: Random House

Jovanovich, B. (1982), 'Selection and the Evolution of Industry', Econometrica, 50, 3, 649-670

Kangasharju, A. and Venetoklis, T. (2002), 'Business Subsidies and Employment of Firms: Overall Evaluation and Regional Extensions', Government Institute for Economic Research Discussion Paper, 268

Kaplan, E. and Meier, P. (1958), 'Nonparametric estimation from incomplete observations', Journal of the American Statistical Association, 53, 282, 457–481

Keuschnigg, C. (1998), 'Investment Externalities and a Corrective Subsidy', International Tax and Public Finance, 5, 4, 449-469

Kronborg, D. and Thomsen, S. (2008), 'Foreign Ownership and Long-term survival', Strategic Management Journal, 30, 2, 207-219

Krugman, P. (1997), The Age of Diminished Expectations, Cambridge, MA: MIT Press

LaLonde, R. (1986), 'Evaluating the Econometric Evaluations of Training Programs with Experimental Data', American Economic Review, 76, 4, 604-620

Lechner, M. (2008), 'A Note on Endogenous Control Variables in Causal Studies', Statistics and Probability Letters, 78, 2, 190-195

Lee, J. (1996), 'Government Interventions and Productivity Growth', Journal of Economic Growth, 1, 3, 391-414

Leuven, E. and Sianesi, B. (2003), 'PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing' Available from: http://ideas.repec.org/c/boc/bocode/s432001.html. Version 1.2.3

Levinsohn, J. and Petrin, A. (2003, 'Estimating Production Functions Using Inputs to Control for Unobservables', Review of Economic Studies, 70, 2, 317-342

Little, R. and Rubin, D. (1987), Statistical Analysis with Missing Data, New York: John Wiley and Sons

Maddala, G. (1993), Limited Dependent and Qualitative Variables in Econometrics, New York: Cambridge University Press

Maddala, G. (2001), Introduction to Econometrics, New York: John Wiley and Sons

Mahalanobis, P. (1936), 'On the generalised distance in statistics', Proceedings of the National Institute of Sciences in India, 2, 1, 49-55

Marschak, J. and Andrews, W. (1944), 'Random Simultaneous Equations and the Theory of Production', Econometrica, 12, 3, 143-205

Marshall, A. (1890), Principles of Economics, London: MacMillan

McGuckin, R. and Nguyen, S. (1995), 'On Productivity and Plant Ownership Change: New Evidence from the Longitudinal Research', The RAND Journal of Economics, 26, 2, 257-276

McGuckin, R. and Nguyen, S. (2001), 'The impact of ownership changes: a view from labor markets', International Journal of Industrial Organisation, 19, 5, 739-762

Metcalfe, S. (2000), Evolutionary Economics and Creative Destruction, London: Routledge

Metcalfe, S. and Georghiou, L. (1997), 'Equilibrium and Evolutionary Foundations of Technology Policy', Science Technology Industry Review, 22, 75-100

Moore, B. and Rhodes, J. (1973), 'Evaluating the Effects of British Regional Economic Policy', Economic Journal, 83, 329, 87-110

Myles, G. (1995), Public Economics, Cambridge: Cambridge University Press

Neyman, J. and Scott, E. (1948), 'Consistent estimation from partially consistent observations', Econometrica, 16, 1, 1-32

O'Mahony, M. and Vecchi, M. (2003), 'Is there an ICT impact on TFP? A heterogeneous dynamic panel approach', NIESR Working Paper, 219

OECD (2003), The Sources of Economic Growth, Available from: http://www.oecd.org/dac/ictcd/docs/otherdocs/OtherOECD_eco_growth.pdf.

Office for National Statistics (1998), LFS User Guide, Available from: http://rlab.lse.ac.uk/data/reference/docs/sic80.pdf

Office for National Statistics (2008), Regional, sub-regional and local gross value added, Available from: http://www.statistics.gov.uk/pdfdir/gva1208.pdf

Office for National Statistics (2009a), Population Estimates: August 2009, Available from: http://www.statistics.gov.uk/pdfdir/pop0809.pdf

Office for National Statistics (2009b), Scotland: Employee Jobs by Industry, (online) Available from: http://www.scotland.gov.uk/Topics/Statistics/Browse/Labour-Market/DatasetsEmployment (Accessed March 2010)

Okada, Y. (2005), 'Competition and Productivity in Japanese Manufacturing Industries', Journal of the Japanese and International Economies, 19, 4, 586--616

Olley, S. and Pakes, A. (1996), 'The Dynamics of Productivity in the Telecommunications Equipment Industry', Econometrica, 64, 6, 1263-1295

Ornaghi, C. (2006), 'Assessing the Effects of Measurement Error on the Estimation of the Production Function', Journal of Applied Econometrics, 21, 6, 879-891

Oulton, N. (1997), 'The ABI respondents database: A new resource for industrial economics research', Economic Trends, 528, 46-57

Pellegrini, G. and Centra, M. (2006), 'Growth and efficiency in subsidised firms', Paper Presented to 'The Evaluation of Labour Market, Welfare and Firm Incentive Programmes' Workshop in Venice, Available from: http://valutazione2003.stat.unipd.it/pdf/VeneziaMag2006/Paper/Pellegrini_Centra.pdf

Puhani, P. (2000), 'The Heckman Correction for Sample Selection and Its Critique', Journal of Economic Surveys, 14, 1, 53-68

Restuccia, D. and Rogerson, R. (2008), 'Policy Distortions and Aggregate Productivity with Heterogeneous Plants', Review of Economic Dynamics, 11, 4, 707-720

Robjohns, J. (2006), 'ARD2: the new Annual Respondents Database', Economic Trends, 630, 43-51

Roodman, D. (2005), 'xtabond2: Stata module to extend xtabond dynamic panel data estimator' Available from: http://econpapers.repec.org/software/bocbocode/s435901.htm

Roodman, D. (2009), 'How to Do xtabond2: An Introduction to 'Difference' and 'System' GMM in Stata', The Stata Journal, 9, 1, 86-136

Roper, S. and Hart, M. (2005), 'Small Firms, Economic Growth and Public Policy in the UK: What exactly are the connections?' Aston Business School Working Paper, RP0504

Roper, S. and Hewitt-Dundas, N. (2001), 'Grant Assistance and Small Firm Development in Northern Ireland and the Republic of Ireland', Scottish Journal of Political Economy, 48, 1, 99-117

Rosenbaum, P. and Rubin, D. (1983), 'The Central Role of the Propensity Score in Observational Studies for Causal Effects', Biometrika, 70, 1, 41-55

Rubin, D. (1973), 'Matching to Remove Bias in Observational Studies', Biometrics, 29, 1, 159-183

Rubin, D. (1974), 'Estimating the Causal Effects of Treatments in Randomized and Non-Randomized Studies', Journal of Education Psychology, 66, 5, 688-701

Rubin, D. (1977), 'Assignment to a Treatment Group on the Basis of a Covariate', Journal of Educational Statistics, 2, 1, 1-26

Rubin, D. (1979), 'Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies', Journal of the American Statistical Association, 74, 366, 318-328

Samaniego, R. (2006), 'Industrial subsidies and technology adoption in general equilibrium', Journal of Economic Dynamics and Control, 30, 9-10, 1589-1614

Schoenfeld, D. (1982), 'Partial residuals for the proportional hazards regression model', Biometrika, 69, 1, 239-241

Schumpeter, J. (1943), Capitalism, Socialism and Democracy, London: Unwin

Scottish Enterprise, (2009), R&D Grant Evaluation, Available from: http://www.evaluationsonline.org.uk/evaluations/Search.do?ui=basic&action=show&id=348

Scottish Government (2003), Regional Selective Assistance: Annual Report 2002-03, Available from:
http://www.scottishbusinessgrants.gov.uk/rsa/files/Annual%20Report%202002-03.pdf
Scottish Government (2004), Regional Selective Assistance: Annual Report 2003-04, Available from:
http://www.scottishbusinessgrants.gov.uk/rsa/files/Annual%20Report%202003-04.pdf
Scottish Government (2005), Regional Selective Assistance: Annual Summary 2004-05, Available from:
http://www.scottishbusinessgrants.gov.uk/rsa/files/Annual%20Summary%202004-05.pdf
Scottish Government (2006), Regional Selective Assistance: Annual Summary 2005-06, Available from:
http://www.scottishbusinessgrants.gov.uk/rsa/files/Annual%20Summary%202005-06.pdf
Scottish Government (2007a), Regional Selective Assistance: Annual Summary 2006-07, Available from:
http://www.scottishbusinessgrants.gov.uk/rsa/files/Annual%20Summary%20200607.pdf
Scottish Government (2007b), The Government Economic Strategy, Available from:
http://www.scotland.gov.uk/Resource/Doc/202993/0054092.pdf
Scottish Government (2008), Regional Selective Assistance: Annual Summary, Available from: http://www.scottishbusinessgrants.gov.uk/rsa/files/Annual%20Summary%202007-08.pdf
Scottish Government (2009a), Regional Selective Assistance: Annual Summary, Available from:
http://www.scottishbusinessgrants.gov.uk/rsa/files/Annual%20Summary%20Doc.pdf
Scottish Government (2009b), Scottish Business Grants, (online) Available from:
http://www.scottishbusinessgrants.gov.uk (Accessed September 2009)
Scottish Government (2009c), Innovation for Scotland, Available from:
http://www.scotland.gov.uk/Publications/2009/06/26143501/0
Siegfried, J. and Evans, L. (1994), ‘Empirical Studies of Entry and Exit: A Survey of the Evidence’, Review of Industrial Organisation, 9, 2, 121-155
Snow, J. (1855), On the Mode of Communication of Cholera, London: John Churchill
Swales, J. (1997) ‘The Ex Post Evaluation of Regional Selective Assistance’, Regional Studies, 31, 9, 859-865
Taylor, J. and Wren, C. (1997), ‘UK Regional Policy: An Evaluation’, Regional Studies, 31, 9, 835-848

Teece, D. (1996), 'Firm organisation, industrial structure and technological innovation', Journal of Economic Behaviour and Organisation, 31, 2, 193-224

Teece, D. and Pisano, G. (1994), 'The Dynamic Capabilities of Firms: an Introduction', Industrial and Corporate Change, 3, 3, 537-556

Theil, H. (1953), Repeated Least Squares Applied to Complete Equation Systems, The Hague: Central Planning Bureau

Van Ark, B. (2001), 'The Renewal of the Old Economy: An International Comparative Perspective', OECD Science, Technology and Industry Working Paper, 2001/5

Van Beveren, I. (2007), 'Total factor productivity estimation: A practical review', LICOS Discussion Paper, 182/2007

Van Der Klaauw, W. (2008), 'Regression Discontinuity Analysis: A Survey of Recent Developments in Economics', Labour, 22, 2, 219-245

Van Der Panne, G. (2004), 'Agglomeration Externalities: Marshall versus Jacobs', Journal of Evolutionary Economics, 14, 5, 593-604

Wooldridge, J. (1995), 'Selection corrections for panel data models under conditional mean independence assumptions', Journal of Econometrics, 68, 1, 115-132

Wooldridge, J. (2007), Econometric Analysis of cross section and panel data, Cambridge, MA: MIT Press

Wooldridge, J. (2009), 'On estimating firm-level production functions using proxy variables to control for unobservables', Economics Letters, 104, 3, 112-114

Wren, C. (1996), Industrial subsidies: the UK experience, Basingstoke: MacMillan

Wren, C. (2005), 'Regional Grants: Are They worth It?' Fiscal Studies, 26, 2, 245-275

Wright, P. (1928), The Tariff on Animal and Vegetable Oils, New York: MacMillan

Zabel, J. (1992), 'Estimating fixed and random effects models with selectivity', Economics Letters, 40, 3, 269-272

Zhao, Z. (2004), 'Using Matching to Estimate Treatment Effects: Data Requirements, Matching Metrics, and Monte Carlo Evidence', Review of Economics and Statistics, 86, 1, 91-107