# The effect of industrial policy on corporate performance:  Evidence from panel data

**Chiara Criscuolo**

*Centre for Economic Performance, London School of Economics*

**Ralf Martin**

*Centre for Economic Performance, London School of Economics*

**Henry Overman**

*Centre for Economic Performance, London School of Economics, CEPR*

**John Van Reenen**

*Centre for Economic Performance, London School of Economics, CEPR*

*This draft: July 16$^{th}$ 2007*

Abstract

Industrial or business support policies designed to raise productivity and employment are a common feature of the policy landscape. Despite this ubiquity, rigorous micro-econometric evaluation of their causal effect is rare primarily because of the difficulty of achieving credible identification. In this paper we tackle this problem by exploiting multiple changes in the area-specific eligibility criteria for a major business support scheme in the UK ("Regional Selective Assistance"). These changes arose because of the need to comply with the European Commission revisions of the eligibility criteria and coverage rules. We match over twenty years of administrative panel data on program participation and firm performance from the Census Bureau to investigate the causal impact of the policy on employment, investment, productivity and entry/exit. Using an instrumental variable approach we find that the program has had a positive effect on both employment and investment, which naïve estimators underestimate. There is no statistically significant effect on total factor productivity, however. There is also some evidence that the program, by supporting less efficient enterprises, may slow down reallocation from less efficient plants, negatively affecting aggregate productivity growth.

## I. INTRODUCTION

Business support policies are ubiquitous. Most governments have industrial policies that claim to foster productivity and employment, particularly in disadvantaged areas. Despite the ubiquity of such schemes, rigorous micro-econometric evaluation of their causal effect is rare (e.g. Griliches et al, 2000). This is somewhat surprising as the sums allocated are not trivial,[1] while the methods to analyse their "treatment effects" have advanced significantly in the last ten years or so. Labor economists have used these techniques most intensively, for evaluating a variety of government programs (see, for example, the survey by Heckman, LaLonde and Smith, 1999).

The basic concern these techniques try to address it that government programmes might simply finance activities the recipients – individuals or households in the labour literature and businesses in our case – would have undertaken anyways in absence of the programme. If this is the case, large amounts of taxpayer dollars could simply be wasted, even before we take into account the deadweight costs of taxation and other distortions induced by the program's design.

To address this concern requires the construction of a counter-factual: what would have happened in the absence of the support programme? Comparison to non-treated firms is one possibility but, of course, the problem is that in the absence of experimental data, those who participate in such programs are not random firms but are heavily selected, and thus participation is endogenous. Matching is one strategy used to deal with this problem, but it relies on the strong assumption of conditional independence which is unlikely to hold unless we have access to an unusually rich set of covariates to control for all the unobservables. Another solution, and the one we adopt in this paper, is to look for instrumental variables correlated with the likelihood of treatment, but not with individual firm performance.

One factor holding back the industrial policy evaluation literature has been the absence of obvious instruments. In this paper we are able to solve this problem by exploiting a useful quasi-experiment that induced exogenous changes in the eligibility

---

[1] For example, in 2005 countries in the EU-25 spent, on average 0.59% of GDP on state aid to industries. See http://ec.europa.eu/comm/competition/state_aid/studies_reports/key_indicators.xls

criteria governing the receipt of Regional Selective Assistance (RSA) in the UK. RSA is an investment subsidy program, administered by the Department for Trade and Industry (DTI) that gives grants to firms for investment in selected (economically disadvantaged) areas of Britain.[2] Grants totalling £110.6m (about $220m) were disbursed in the UK in 1998-1999 alone (DTI, 2001). Crucially for our identification strategy, new European Union guidelines caused a change in the areas eligible to receive these grants in 1993 and in 2000. We exploit the change in these "maps of assistance" to generate instrumental variables for the receipt of investment grants. This enables us (under certain assumptions discussed in detail below) to estimate the causal effect of the program on employment, investment, productivity, exit, entry and other factors.

Our data set is constructed by linking observations from three very rich administrative data sources. First, from the DTI, we have data on the population of all firms who applied to the RSA program since 1972 and the amount of assistance they received if they were successful. Second, from the Annual Business Inquiry (ABI) we have panel data on the population of all UK manufacturing plants. Finally, from the Interdepartmental Business Register (IDBR) we have location, entry and exit of all manufacturing plants in the UK. Further details are provided below.

Our substantive conclusion is that there is a large and significant "average effect of treatment on the treated" for employment, investment and the probability of exit. These effects are seriously underestimated if endogeneity is ignored, as the participants in the scheme appear to be weaker firms who would otherwise perform badly given their observable and unobservable characteristics. There appear to be no additional effects on productivity after controlling for the investment effects. We do find evidence, however, that the proportion of employment in entrants as a whole falls in an affected area, raising the possibility of negative aggregate productivity effects from lower reallocation through protection of inefficient incumbents.

The paper is structured as follows: the next section describes the policy in more detail and outlines how eligibility changes over time. Section III describes the econometric modelling strategy. In Section IV we describe the data we use and the characteristics of treated and non-treated firms in our sample. In Section V we report our results and a number of robustness checks. Finally we provide some conclusions

---

[2] The DTI is now the Department for Business, Enterprise and Regulatory Reform.

and a discussion on how we intend to forward the research. In Appendices we report more details on the data matching procedure and issues involved; the description of the variables used and we discuss some of the existing literature on evaluations of RSA and similar policies.

# II. INSTITUTIONAL FRAMEWORK: DESCRIPTION OF THE REGIONAL SELECTIVE ASSISTANCE POLICY

## IIA. Overview

During the period of our study, (1988-2003) Regional Selective Assistance (RSA) was the main regional business support scheme in the UK.[3] From the early 1970s it provided discretionary grants to companies in Assisted Areas. These are disadvantaged regions typically characterised by relatively high levels of unemployment and deprivation. It was designed to "create and safeguard employment". Emphasis was given to internationally mobile investments, new products and processes and the manufacturing sector. Assistance could be provided to establish a new business; to expand, modernise or rationalise an existing business; to set up research and development facilities or enable businesses to take the next step from development to production.

Because RSA has the potential to distort competition and trade between European countries it must comply with European Union legislation concerning state aid. In general, this type of assistance is prohibited by European law except in certain cases. In particular, Article 87 of the Treaty of Amsterdam allows for some state aid in support of the European Union's regional development policies. The guidelines designate very deprived "Tier 1 Areas" (previously called "Development Areas") in which higher rates of grant can be offered and slightly less deprived "Tier 2 Areas" (previously called "Intermediate Areas").[4] There is an upper threshold of support that

---

[3] We discuss our choice of study period below. According to Harris and Robinson (2004), in 1998-9 RSA represented 19% of the UK's industrial policy spending. In April 2004, the RSA scheme was replaced by the Selective Finance for Investment (SFI) scheme. Productivity became an official objective in April 2004, when RSA was replaced by SFI which explicitly requires that projects yield productivity improvements. We discuss the difference between the two schemes below.

[4] Article 87(3) of the Treaty of Amsterdam defines conditions where State aid may be compatible with EU laws. Article 87(3) (a) allows for "aid to promote the economic development of areas where the

is allowed, referred to as Net Grant Equivalent (NGE),[5] which essentially sets a maximum proportion of the firm's investment that can be subsidised by the member state government.

Since the main formulae which determine eligibility are decided at the European level at fixed seven yearly intervals and not at the UK level, this mitigates concern of endogeneity of policy decisions. And although the UK finance ministry has latitude to decide the overall amount of the annual budget for RSA they are not able to change the rules over which areas are eligible to receive some RSA. Thus, area-level eligibility is the key form of identification in our paper.

## IIB. Changes in eligibility over time

The map of the areas eligible for RSA changed twice during our study period: first in 1993 and then again in 2000. There were also changes in 1986 before our sample period begins and in 2006, after our sample period ends. These changes happen every seven years in conjunction with the periodic revision of the Structural Funds, the European Union's main policy for supporting economic development in less prosperous regions.

The map of the eligible areas is proposed by the UK but needs to be approved by the EU in accordance with the EU regional guidelines and in respect of Article 87 of the Amsterdam Treaty. The main criteria are (i) that only areas with underemployment and a low standard of living are eligible (ii) any map or change in the map must satisfy the threshold imposed by the EU defining the proportion of UK population that may be covered by Assisted Areas Status.

The eligibility criteria are outlined in the regional guidelines which are published two years before the implementation of the map (in our case 1991 and 1998). The UK government will then gather quantitative information on indicators of employment

---

standard of living is abnormally low or where there is serious underemployment" [Tier1/Development Areas] and Article 87(3) (c) allows for: "aid to facilitate the development of economic activities or of certain economic areas, where such aid does not adversely affect trading conditions to an extent contrary to the common interest." [Tier 2 or intermediate Areas] Additional restrictions apply to sectors with over-capacity: motor vehicles, synthetic fibres and yarns, iron and steel, coal, fishery and agricultural products.

[5] The Net Grant Equivalent (NGE) of aid is the benefit accruing to the recipient from the grant after payment of taxes on company profits. RSA grants must be entered in the accounts as income and are made subject to tax. Details for calculations of NGEs are available in OJ C74/19 10.03.1998.

level and deprivation at the relevant regional level based on the previous three years data where possible and will propose a new map.

Figures 1 through 4 show that eligibility changes in the maps for all areas in the UK.

Below we discuss each of the changes in turn.

*(a) The 1993 change*

The assisted area map for RSA was redrawn in 1993 on the basis of the new 1991 guidelines using "Travel to Work Areas" as the underlying spatial units.[6] The selection of Assisted Areas was based on several factors using a quantitative formula. The first set of factors used indicators of bad labor market conditions, such as persistently high unemployment, the proportion of long-term unemployed, participation rates and the likely future demand for jobs (based on growth/decline in local industries, demographic changes and expected major firm closures). The second set related to geographic features such as distance from major markets, low population density and urban problems.

The Assisted Areas fell into two categories: (a) Development Areas where aid could be granted up to a maximum of 30% NGE (Net Grant Equivalent - see above) and (b) Intermediate Areas where aid was limited to 20% NGE. The new 1993 maps implied a net reduction in the number of assisted areas with Development Areas covering 17%, and Intermediate Areas covering 19%, of the total UK population.

*(b) The change in 2000*

The EU Commission introduced new guidelines for State Aid in 1998 and the UK responded to that with the introduction of a new Assisted Area map in 2000. The maximum investment subsidy allowed for in these areas is 35% NGE for the most deprived (Tier 1) Areas.[7] These areas are the four eligible for funding under Objective 1 of the EU Structural Funds: Cornwall & the Isles of Scilly, Merseyside, South Yorkshire and West Wales & the Valleys.

---

[6] Travels to Work Areas are defined by the UK Census Bureau (Office for National Statistics). The fundamental criterion is that, of the resident economically active population, at least 75% actually work in the area, and also, that of everyone working in the area; at least 75% actually live in the area. Thus, in terms of definition, they are similar to the US Metropolitan Statistical Areas.

[7] Special status and a higher allowance are awarded to Northern Ireland, which is not included in our analysis.

The Tier 2 areas are more scattered. These 65 zones are constructed on the basis of groups of electoral wards.[8] Each grouping must have a population of at least 100,000 and the wards were selected according to four statistical indicators. Although the precise indicators differed from 1993, the main criteria were still labour market performance and the share of manufacturing.[9]

Within Tier 2 Areas the map identified four sub-tier areas eligible for different level of maximum NGE. The level of aid intensities proposed for these areas vary according to the seriousness and intensity of the problems in each region relative to the Community context, in particular as regards neighbouring EU countries.

For the most disadvantaged sub-tier areas, that were geographically distant and sparsely populated, a maximum subsidy rate of 30% NGE was allowed[10]. The maximum NGE level for relatively less deprived areas was 10%.[11] However, if those (less deprived) areas are adjoining to Tier 1 areas they have a 20% ceiling. The rest of the eligible areas aid ceilings are either an NGE of 20% or 15% (with the decision as to which applies made by referring to current conditions as well as the NGE in the 1993 map)

Finally, note that in assisted areas, a higher rate may be paid to small and medium-sized enterprises.[12] We plan to exploit this size difference in future work.

## IIC. Formal criteria for receipt of RSA

During our study period (1988-2003), RSA traditionally targeted manufacturing sectors, although support was also available for services sector firms that served national or foreign markets (i.e. not just the local market). The grants were

---

[8] The data used for the zone boundaries come from the 1991 Census of Population. A detailed list of the assisted areas by local authority within regions and the NGEs to which they are eligible is available from the authors upon request.

[9] The indicators used are residential employment and unemployment rates; workforce unemployment rates and manufacturing share of employment. The first two are used as indicators for potential labour market exclusion; the third as indication of low availability of any jobs and the last as indication of vulnerability to a continued contraction in manufacturing employment.

[10] These areas have a population density of less than 12.5 inhabitants per square kilometre and are mainly the Highlands in Scotland (1.2% of assisted areas population were in these areas).

[11] These are areas with a higher GDP per capita and lower unemployment rate than the Community average (covering 4.2% of assisted areas population).

[12] The additional support for smaller enterprises consists of fifteen percentage points gross in Tier 1 areas, and ten percentage points gross in Tier 2 areas.

discretionary and firms could only apply if the supported project satisfied the following criteria. (a) **Location**: The project had to be undertaken within one of the Assisted Areas. (b) **Investment:** The project had to involve capital expenditure on property, plant or machinery; (c) **Jobs:** The project should normally have been expected to lead to the creation of new employment or directly protect some or all of the jobs of existing workers which, without the project, would otherwise have been lost; (d) **Viability:** The project should have had good prospects of viability and should have been expected to help the business become more competitive; (e) **Need:** The applicant had to demonstrate that assistance was necessary to enable the project to proceed as envisaged in terms of nature, scale, timing or location;[13] (f) **Prior Commitments:** As RSA could only be offered where it would make the difference between the project going ahead and not proceeding, there should have been no prior commitment to the project, i.e. the DTI must have completed its appraisal of the project and issued a formal offer of assistance before the applicant entered into a commitment to proceed with the project; (g) **Other Funding:** The greater part of the funding for the project should have been expected to be met by the applicant or come from other sources in the private sector.[14]

Location, which also forms the basis for our instrumental variables, is objective, clearly defined and enforceable.

The process for application was as follows. Firms needed to fill in an application form, in which they needed to prove additionality, to provide business plans, accounts and reasons for wanting the grant. They then submitted this to the local DTI regional agency. The lag between the day the application was submitted and the decision depended on the amount that the firm was applying for. During the period analysed, the lag was normally between 35 and 60 days, and 100 days or more for grants above £2 million (about $4 million). The lag also depended on the time needed to appraise the application to ensure that all of the criteria were met and on negotiations between

---

[13] This may be to meet a funding gap, to reduce the risks associated with the project, or to influence the choice of location of a mobile project. It might also be to obtain parent company approval by meeting established investment criteria; or for some other acceptable reason – each case is considered on its own merits.

[14] These may include bank borrowings, hire purchase or lease finance, equity and loan finance from existing or new shareholders and loans from other organisations or institutions. Additional public sector assistance may however, be available towards the project. Any additional assistance must be cumulated with the RSA support and must not breach the European Union State Aid limits.

the government agency and the firm on the terms of assistance. If the application was accepted, the firm was paid the minimum necessary to get the project going. Additional payments started only after the jobs were created/safeguarded and the capital expenditure defrayed and were based on agreed fixed capital expenditure and job targets. The payments were given in installments – between two and seven and in the vast majority of cases these were spread across more than one financial year. The government agency monitored the project with visits (normally one per year, but more frequently for risky projects).

# III. ECONOMETRIC MODELLING STRATEGY

We next consider the econometric modeling strategy, starting with our basic approach and then considering extensions.

## IIIA. Basic Approach

Consider the outcome equation:

$$y_{it} = \alpha D_{it} + \beta X_{it} + u_{it} \qquad (1)$$

where $y_{it}$ is the outcome of interest for plant ("local unit") $i$ at time $t$. Note that a plant is uniquely located in an area, $r$, and in a firm ("reporting unit"), $j$. We suppress the sub-scripts on area and firm for notational simplicity unless needed. $D_{it}$ is the participation indicator which we will initially assume to be binary (see below for an extension to the continuous treatment intensity setting). Thus $D_{it}=1$ if a plant receives RSA in year $t$ and zero otherwise. Initially we assume a homogeneous treatment effect, $\alpha$, but we relax this assumption below. $X_{it}$ are other covariates used as controls such as age, industry, area, whether the plant belongs to a larger firm (and if so whether this firm is domestic or foreign). Outcomes can include employment; investment, productivity, and the precise set of $X_{it}$ will depend on the outcome of interest. The $u_{it}$ is an error term whose properties we discuss below.

Estimation of equation (1) by OLS will be fraught with difficulties because $D_{it}$ is likely to be correlated with $u_{it}$. For example, a firm might apply for RSA when they are already shedding jobs. As one objective of the scheme is to safeguard existing

employment, firms who have a low draw of $u_{it}$ are more likely to receive grants generating a possible downward bias on our estimate of $\alpha$. Working in the opposite direction is the fact that a second objective of RSA is to create new jobs, which may increase the likelihood of receiving a grant for firms who have experienced some positive shock $u_{it}$. The direction of the bias will depend on the balance of these two effects, but clearly the endogeneity of treatment needs to be controlled for if we are to consistently estimate $\alpha$.

Previous attempts to deal with this problem have had to rely on a set of observable controls, time dummies and fixed effects. Since we have panel data we can follow this practice by decomposing the error term into a correlated fixed effect, $\eta_i$, a set of time dummies, $\tau_t$, and $v_{it}$, a plant-specific time varying error term:

$$y_{it} = \alpha D_{it} + \beta X_{it} + \eta_i + \tau_t + v_{it} \tag{2}$$

Although equation (2) will purge our estimate of $\alpha$ of bias due to common macro-economic shocks (through $\tau_t$) and permanent correlated unobserved heterogeneity (through $\eta_i$), it will still be inconsistent if there are unobserved transitory shocks $v_{it}$ correlated with $D_{it}$.

Consequently we consider instrumental variables, $Z_{it}$, for program participation, $D_{it}$. The instrument we will construct exploits the fact that only plants located in certain areas of the UK are eligible for RSA (see Figures 1 to 4). Although these areas are fixed at a given point in time, the UK government changed the map of eligible areas twice (in 1993 and 2000) to comply with European Commission State Aid legislation. As a result, some areas ceased to be eligible and other areas that were ineligible became eligible. For example, Darlington a small city in the North-East of England was eligible pre-1993, but became ineligible in 1993. The changes were driven by a new European-wide formula for calculating which regions were eligible to receive subsidies under state aid rules (see Section III for more details). In addition to discrete indicators constructed on the basis of eligibility, the maximum investment subsidy also differs across eligible areas in a way that has changed over time. This allows us to construct measures for the intensity of treatment. Finally, there are size and industry-specific components of the subsidy which we are currently not exploiting, but which we plan to use in subsequent versions of the paper to improve efficiency.

With the instruments we can estimate equation (2) by instrumental variables. As reported below, we look carefully at the first stage to check for weak instruments issues. We also consider, in detail, the reduced form:

$$y_{it} = \pi_1 Z_{it} + \pi_2 X_{it} + \tilde{\eta}_i + \tilde{\tau}_t + \tilde{v}_{it} \tag{3}$$

Under the covariance assumption $E(Z_{it} \, \tilde{v}_{it}) = 0$ the estimate of $\pi_1$ by OLS is the "intent to treat" effect, which is of interest in its own right.

When moving from theory to implementation, one complication arises because of the unit of observation in the available data. We have written the analysis at the plant level, however the main data used for the analysis (ABI) is collected at the firm (reporting unit) level rather than at the plant (local unit) level. Although for most firms in the ABI the two levels of aggregation coincide (on average 80% of reporting units sampled are single plant firms), for the other firms measures of investment, output and materials are only available at the "reporting unit" level which combines several plants.[15] Employment and location are always available at the local unit level, even for multi-plant firms.

To deal with this issue we simply aggregate the relevant equation across all plants in the same firm. Equation (2), then becomes:

$$y_{jt} = \alpha D_{jt} + \beta X_{jt} + \eta_j + \tau_t + v_{jt} \tag{4}$$

For example, when $y_{it}$ is total employment in the plant, $y_{jt}$ is simply employment in the firm, summing across all plants $i$ in firm $j$, i.e. $y_{jt} = \sum_{i, i \in j} y_{it}$. All other variables are defined similarly, although there are some issues around the definition of $D_{jt}$ and the instruments, $Z_{jt}$. For the participation dummy we mainly continue to use a simple binary indicator if the firm received any treatment. But we also present checks on alternatives such as the amount of money received expressed as the fraction of total project costs covered by the grant. For the firm-level instruments we have $Z_{jt} = \sum_{i, i \in j} w_{it}^j Z_{it}$. We consider several different weighting factors, $w_{it}^j$, but one important issue is that choice of the weights could induce an endogeneity bias. For example, the current distribution of firm employment across plants across areas could

---

[15] We call this the firm level, $j$, but there could be many reporting units in one large firm.

be affected by the eligibility to RSA. Consequently we only ever use lagged data to construct weights (see sub-section IIIC below for a more detailed discussion).

Note that the interpretation of $\alpha$ subtly changes in the aggregated regression. Consider employment outcomes and assume that the number of plants is fixed. If a firm has two plants in two areas and then one area becomes ineligible for RSA, the firm could substitute employees from the plant in the ineligible area to the plant in the eligible area without changing total employment. Analysis at the plant level in equation (2) would find a positive program effect. Analysis at the firm level in equation (4) would find zero effect. In theory, program rules are meant to stop firms engaging in such switching, but in practice this is hard to enforce as the firm has more private information on the true counterfactual than the government agency. Given our data, and the fact that Equation (4) is arguably of more direct policy interest, we focus on firm level results in what follows.[16]

## IIIB. Extensions

*(a) Heterogeneous Treatment Effects*

If we relax the assumption that the response to participating is the same across firms we can re-write the plant-level equation of interest as:

$$y_{it} = \alpha_i D_{it} + \beta X_{it} + \eta_i + \tau_t + v_{it} \tag{5}$$

where $\alpha_i$ is now the plant specific effect of treatment.

There is much discussion in the evaluation literature over the interpretation of 2SLS estimation when the true model is equation (5) rather than equation (2).[17] The essential problem is that using observations from the whole population may give a selection of non-traded plants that does not actually provide a valid comparison group for those who participate and prevent us from consistently estimating the average effect of treatment on the treated (ATT). To address this we apply matching

---

[16] Comparison between the two estimates would be informative as regards such intra-firm switching behavior. In the paper we report our analysis for all of the firms in the sample and then separately for single plant firms. A comparison of the results for these two samples could give some indication of how this behaviour affects our results. However, we might be introducing additional selection bias as the single plant firms are a subset of the sample. Ideally, we would want to use the information on employment reported in the business register, (IDBR). We intend to exploit this information as a robustness check but we are worried about measurement error issues when using the IDBR employment information (see Data section for a description of the IDBR and related issues).

[17] For some examples see Angrist (2004), Imbens and Angrist (1994) or Heckman et al (1997, 1999).

techniques to link our set of participants with a set of control observations using the propensity score, trimming the sample of participants and controls so we have a common support. The combination of eligibility as an instrument and matching[18] is similar to Blundell, Costa-Dias, Meghir and Van Reenen (2004).

*(b) Aggregation to the area level*

We also examine the impact of treatment at the small regional level of a travel to work area (TTWA) rather than firms or plants.

$$y_{rt} = \alpha D_{rt} + \beta X_{rt} + \eta_r + \tau_t + v_{rt} \tag{6}$$

where $D_{rt} = \sum_{i, i \in r} w_{it}^R D_{rt}$ is the weighted total number of treatments in an area $r$ at time $t$. We look at several alternatives for $w_{it}^R$.

Analogously with equation (4) we are able to see what the area-wide effect is of the policy taking into account any possible substitution effects between participating and non-participating plants. For example, although the UK DTI performs a market and competition assessment prior to giving RSA, it is still possible that, ex-post, plants that receive RSA grow at the expense of other non-receivers in the same area.

More generally, the area-level analysis enables us to look at several aspects of the effect of RSA that we cannot identify at the plant or firm level. For example, at the firm level, we can only capture entry of new plants for incumbent multi-plant firms while at the area level we can examine plant entry (single or multi) by non-incumbents. In addition, the aggregate growth in employment in the area can be decomposed into the components coming from incumbent growth, exit and entry.

*(c) Indirect Effects*

Estimation of equation (2) by IV makes the usual Stable Unit Treatment Value Assumption (SUTVA), i.e. the treatment of unit $i$ only affects the outcome of unit $i$. A violation of SUTVA would be when there are "indirect" effects of the program on

---

[18] Using the propensity score function we restrict the sample to those firms that have a predicted propensity score larger than the 10th percentile of the propensity score distribution of treated firms and lower than the 90th percentile of non treated firms. We check the robustness of these results to more conservative thresholds. There are other ways to match including matching by area (looking at ineligible areas that are closer in observed characteristics to eligible areas) and matching within area by plant and firm observables.

non-treated plants. In the previous sections we have already considered some violations of this assumption: within a firm, there may be substitution away from some plants in non-eligible areas towards other eligible plants; within an area there may be substitution away from non-participants to participants. As discussed above, these indirect effects can be tested for and their extent estimated by comparing estimation results for the same specification, but at different levels of aggregation (cf Griliches, 1979 on R&D spillovers).

A more general indirect effect, that we have not yet discussed, may arise from substitution of activity between *different* areas for *different* firms (contrast equation (4) picks up within-firm substitution and equation (6) which picks up the within area). For example, when area A becomes eligible and area B remains ineligible, does employment fall in area B? A way to test this is to examine the impact of switches in eligibility in A on outcomes in neighbouring areas to A. Effectively we intend to use border effects to see if there are strong indirect effects.

*(d) Dynamics*

Changes in eligibility are unlikely to have an immediate impact on outcomes. Most important is the fact that the need to collect information and properly assess applications means that there can be considerable delay between the application for a grant, its approval and eventual receipt of the subsidy. In addition to this administrative delay there may be the normal lags between investing and observing changes in production and employment (delivery lags, time to build, etc.). Consequently, our main instruments use eligibility dates at t-2, but we also experiment with lagging the eligibility by one or three years. This should also reduce further any concern that the selection of areas is endogenous to shocks to individual firm outcomes.

*(e) Expectations*

We are assuming that, from the firm's perspective, the change in eligibility is unexpected. Although there was some uncertainty over the exact placing of re-drawn boundaries and revised levels of subsidy it is possible that many firms knew about changes in advance. Although grants are only paid in arrears, the increased probability of subsidy could alter firm behaviour in advance of the change. For example, some firms may have increased investment as they would expect a better chance of a future

subsidy. We can check for the magnitude of these effects by including *future Z* in the reduced form of equation (3) that captures the intent to treat. The significance of these (in the presence of our preferred current or lagged *Z*'s) would signal the possibility of future looking behaviour by firms.

*(f) Continuous Treatment Intensity*

So far, we have focused on the discrete treatment case, but we can also exploit more information by using a continuous measure of treatment intensity. Our main continuous measure is simply to calculate the proportion of investment that is paid for by the program. If we denote the amount of grant received as *R* then in this case the participation variable is: $\left(\dfrac{R}{I}\right)_{it}$ where *I* is the total investment cost of the project. This investment subsidy can be directly calculated from available data.[19]

## IIIC. More detail on the Instrumental Variables

In the dataset we have information at the postcode level on whether plants located there are eligible and if so the Net Grant Equivalent.[20] As already discussed the map of eligibility and RSA rates changed in 1993 and in 2000 so the data contains variation in both the cross-section and the time-series dimension. Consider first the discrete variation in eligibility: when we use the cross-sectional variation, identification comes from firms located in eligible areas who did not get treated. When we use time series variation we use information from firms whose eligibility changes as a result of changes to the RSA map. Second, we exploit variation in the RSA rate since the higher the RSA rate the higher the participating in returns to applying for RSA. Identification is similar to that for eligibility except we now use differences in RSA rates or changes in RSA rates. Table 3 reports the number of such changes in eligibility and RSA rates both at the plant and the more aggregated firm level in our sample over both the 1993 and the 2000 eligibility map changes.

---

[19] An alternative is to include two variables, one a dummy for participation as before ( $D_{it}$ ) and the second $D_{it} * \ln(R_{it} / \overline{R})$ where $\overline{R}$ is the sample mean of *R* conditional on receiving any grant (to enable the average marginal effect to be read off more easily for the participants. Since the logarithmic transformation is undefined at zero we set this to an arbitrary value (0) which will be picked up by the participation dummy.

[20] In the UK, postcodes typically refer to one property or a very small group of dwellings.

*(a) Endogenous Eligibility?*

One concern is that areas that lose eligibility are also those who have improving economic conditions, thus generating a bias on our instrument. Consider the first differenced equivalent of the reduced form, equation (2), and ignoring time dummies for simplicity:

$$\Delta y_{it} = \pi_1 \Delta Z_{it} + \pi_2 \Delta X_{it} + \Delta w_{it} + \Delta v_{it} \qquad (7)$$

We have decomposed the error term into two components, $\Delta w_{it}$ which is correlated with the eligibility changes and a truly idiosyncratic error, $\Delta v_{it}$ which is not. The first thing to note is that since areas who are doing better are more likely to be made ineligible for RSA, i.e. E($\Delta Z_{it} \Delta v_{it}$) <0, this will lead to a *downwards* bias on the coefficient of interest, $\pi_1$, and make it harder to identify a policy effect.

Recall from the discussion in Section II, however, that the determination of area eligibility status depends on the European Commission's Regional Guidelines which are published two years prior to the map changes. The implementation of the guidelines, in turn, depend on data that available, at most, three to five years before the map changes (for example, in the 2000 change most of the indicators were actually based on the 1991 Census – nine years previous). So the magnitude of this possible bias will depend upon the correlation between variables like unemployment rates three years ago and current unobserved area-specific shocks. Note that variation in $Z_{it}$ is also driven by changes in the EU wide average GDP per capita and unemployment which change dramatically as new countries have entered the EU.

Although we think the size and direction of such biases are likely to be second order, we consider some checks of this. First, we include area-specific trends to proxy $\Delta v_{it}$ which are likely to pick up any longer run decline in an area that are not reflected in the covariates. Second, since we know the rules that are used to designate areas we can condition directly on the area-specific indicators used to determine eligibility, such as lagged area GDP per capita and lagged area unemployment rates. Since eligibility is now only driven by the exogenous decision of the EU commission, $Z_{it}$ is identified.

(b) *Weighting used in constructing the instrument*

As discussed above, information on the ARD is recorded at the reporting unit level rather than at the local unit level (i.e. for multi-plant firms information is aggregated at the firm level), so we are also faced with the additional issue of how to use the information on eligibility and rates for reporting units that have several local units some of which are eligible and others of which are not (and similarly for firms whose local units face different RSA rates).

As noted above an obvious concern is that using the current distribution of employment within the firm could create an endogeneity bias as this distribution could be affected by RSA eligibility. Even the location of plants within the firm could be affected by RSA eligibility (although this is less likely to be a problem due to sunk costs of plant entry and exit). Consequently we used only lagged information on the location of plants. To further reduce the risk that forward-looking firms take into account future changes in eligibility in deciding where to locate their plants our main results use the location of the oldest plant in the firm (i.e. the local unit owned by the firm for the longest amount of time) to calculate eligibility for RSA. The past geographical location of this plant is least likely to be affected by current changes in the eligibility map.

We also investigate the robustness of our results to less conservative alternatives constructions of the weight for the instrument such as using all plants of whatever age and employment distributions across regions. Generally the results are even stronger, which we can't exclude results from potential endogeneity of the instruments.

*(c) Functional form of the instrument*

Our instrument is the level of the maximum investment subsidy, the Net Grant Equivalent, available in the area. This variable takes on a number of discrete values ranging from zero in ineligible areas to 35% in the most deprived areas after 2000. Our baseline results for the analysis use mutually exclusive dummies for each of the different rates (with zero being the baseline), in order not to impose too much structure on the non-linear schedule. We also show the robustness to more parametric definitions of the instrument using a continuous measure.

The definitions of the instrument are therefore *NGE = x* where x ranges from zero to 0.35. For example, *NGE = 0.2* indicates that the firm's oldest plant is in an area where the maximum investment subsidy (Neg Grant Equivalent) is 20%.

## IV. DATA

We combine administrative data on support scheme participants[21] with independent business performance data. This involves matching Selective Assistance Management Information System (SAMIS) database of participants, the Interdepartmental Business Register (IDBR) and the Annual Respondents Database (ARD) which we describe in more detail below. We also describe the characteristics of the firms in our sample while we address in more detail the issues related to the matching procedure to the Appendix.

There are two main advantages of such an approach relative to evaluations based on industrial surveys. First, we can compare the firm before and after its exposure to the program – frequently data on program participants is only available after they have joined the program. Secondly and more importantly, we can compare the change in the participating firms' performance to a "control group" of firms who did not participate or were not eligible to participate in the program. Finally, independent performance data is less likely to be affected by strategic reporting by surveyed firms.

### *IVA. Administrative Data on program participants*

The Selective Assistance Management Information System (SAMIS) was used to monitor RSA projects. It contains information on more than 50,000 applications from 1972 to 2003. It includes for all applications information on the name, date and address of the applicant, a project description; the amount applied for, aims and date of application. For successfully completed applications it provides the date in which and the amount of the grant offered and paid (since 1988 additional payment information is available containing date and amount of first and last instalments). For those that were not completed it contains information on why; i.e. whether the project was withdrawn; was accepted but then the firm did not proceed; was not accepted by the firm; or was rejected by the DTI and if so for which reason.

---

[21] As described in more detail below we also have information on applicants to the scheme that were rejected for various reasons or had withdrawn their application.

Since the payment information from the SAMIS database is not always accurate we prefer to use additional information with more detailed payment information from the Payment RSA database available from 1988.

For reasons due to the quality of the match with the other data sources used, as described in more detail in the Appendix, and in order to have accurate information on payments and some information on post-treatment performance we use the applications made between 1st January 1988 and 31st December 2003.

## IVB. Interdepartmental Business Register (IDBR)

In order to be able to match the administrative information with production data the records from the SAMIS database needed to be matched with the Interdepartmental Business Register (IDBR), which contains both the names of the businesses and the identification numbers used by the Office for National Statistics to conduct the Annual Business Inquiry. The Interdepartmental Business Register (IDBR)[22] is essentially a list of all businesses in the UK, their addresses, type of activity and ownership/control structure compiled using a combination of tax records on VAT and PAYE, information lodged at Companies House, Dun and Bradstreet data, and data from other surveys using three aggregation categories: "local units" (plants), "enterprises" and "enterprise groups".[23] A plant or "local unit" is defined as "an enterprise or part thereof (e.g. a workshop, factory, warehouse, office, mine or depot) situated in a geographically identified place" and is identified by a unique identifier. A major advantage of the IDBR is that information is available at many disaggregated levels. For our analysis this is particularly useful since we also look at the effect of the policy at the regional level on employment and entry/exit. We therefore need employment and entry/exit information at the local unit level rather than at the enterprise level since enterprises can consist of local units in different regions.[24]

---

[22] The IDBR was introduced between 1994 and 1995. Previously, that sampling was on the basis of a Business Register maintained by the Office of National Statistics (the UK Census).

[23] Criscuolo et al (2003) report that in the 1998 IDBR the vast majority of enterprise groups and Reporting Units consist of just one local unit (92%, 149,326 out of 162,477 and 93%, 158,727 out of 171,271 respectively).

[24] Employment information on the IDBR comes from PAYE data if that is the source of the original inclusion and the enterprises operate a PAYE scheme, which in turn if operated at the local unit level, provides independent local unit employment data. Also the IDBR gathers and updates information on employment from the Annual Register Inquiry (see Criscuolo et al., 2003 for details) and the Annual

A stratified[25] random sample of enterprises is drawn every year from the IDBR to form the sampling frame for the Annual Business Inquiry (ABI), which provides information on employment, investment, materials, etc. and is described next.

## IVC. Annual Respondents Database (ARD)

The Annual Respondents Database (ARD)[26] is the UK equivalent of the US Longitudinal Respondents Database and is made available by the Office for National Statistics (ONS) based on information from the Annual Business Inquiry (ABI),[27] the mandatory annual survey of UK businesses. The ARD unit of observation is defined by the ONS as an autonomous business unit (also referred to as "reporting units").

Some of these business units are spread across several sites but in about eighty percent of all cases a business unit is located entirely at a single mailing address. We call this unit a "firm".

It is important to note that the ARD does not consist of the complete population of all UK businesses, since the sample is stratified with smaller businesses sampled randomly. It contains the population of larger businesses however (those over than 100 or 250 employees depending on the exact year). Each year the sampled firms account for around 90% of total UK manufacturing employment. For our analysis we use the unbalanced panel between 1985 and 2004.

The ARD contains a wealth of information, but most importantly for our study it contains information on employment, investment, intermediate inputs and gross output. We are particularly interested in the effect of RSA for employment, investment and productivity.

---

Business Inquiry (ABI). However, employment data is required to construct sampling frames and hence it will be interpolated from turnover data. The IDBR turnover information comes from VAT records if the original source of business information was VAT data; however this information is quite limited as it is only available for single-local unit enterprises that are large enough to pay VAT (the threshold was £52,000 in 2000–01) at both the enterprise and local unit level. For multi-local unit enterprises, no turnover information will be available for local units, since most multi-local unit enterprises do not pay VAT at the local unit level.

25 Stratification is broadly based on industry affiliation; regional location and size. For details see Criscuolo et al. (2003).

[26] More extensive description of the ARD can be found in Criscuolo, Haskel and Martin (2003), Griffith (1999) and Oulton (1997).

[27] Called the Annual Census of Production until 1998.

## IVD. The samples

The details of the linking procedure are described in more detail in the Appendix. Here we describe briefly the difficulties encountered and the characteristics of the sample focusing on the differences between firms that receive RSA and those who do not receive RSA in our sample.

### (a) Plant Sample

For the whole IDBR we have essentially all businesses and are able to match all the plants to firms who participated in RSA. The exception to this is where there are issues such as variations in spelling of names, changes in postcodes or typos in either of the databases. From manual examination of the merge results this does not seem to be a major issue.

A second problem is that for a large firm that consist of several plants it is not always clear which plant has been participating in the programme and which has not. This is not a problem for the firm-level analysis of course, but means that there is measurement error in classifying plants as "participating" or not. Our instrumentation strategy should deal with this problem, however.

We call this the plant sample and we can use this sample to conduct regional level analysis on exit rates and employment since when we conduct regressions at the area level as with equation (6), we aggregate the plant level data within an area to form an area panel.

### (b) Firm Sample

For analysis of investment and productivity we need to use the match between the SAMIS and ARD sample. This is a (large) subset of the total population of all businesses in Britain so we miss some businesses that participate in a programme but might not be sampled in the ARD in a particular year. Moreover since for our analysis we need to have observations both pre- and post-treatment this issue might be more severe for smaller plants. However, since sampling of the ARD is random this is not a problem for consistency of the estimates, although it does mean we miss out on some of the smaller firms.

## IVE. Some Data Description: Characteristics of RSA recipients

Table 1 reports descriptive statistics distinguishing between firms which never participate in RSA and firms which at some point in their lives participate in RSA. For the participants we report characteristics both before and after they joined the program.

Rows 1 and 2 shows that both in terms of labour productivity, measured as value added per employee (VA/L), and Total Factor Productivity (TFP) participating firms are significantly less productive when they enter the program. After receiving RSA the gap tends to decrease in terms of labour productivity and treated firms become more labour productive than non-treated firms. This is true along the whole distribution as shown in columns (4) to (8) reporting the $25^{th}$, $50^{th}$ and $75^{th}$ percentile of the distribution. In terms of TFP firms that participate in RSA have significantly lower productivity than non-participants firms both in the period before and after the program.

The Table also reports figures for employment and gross output. According to both measures participating firms are on average larger than non-participants before treatment. For treated firms the difference between pre- and post-program periods for both of these measures is not significant.

Looking at investment (and capital intensity) shows that firms which receive RSA observe a statistically significant post-program increase in the level of investment. One point to note here is that even though at the mean non-participants are more capital intensive than participating firms, in the pre-program periods this result is driven by a few outliers at the top of the non-treated distribution. Looking at the medians shows that on average participating firms had larger investments and were (slightly) more capital intensive.

One other interesting feature is that growth rates, both of labour productivity and employment, do not seem to differ between the pre- and post-RSA receipt. This is a first indication that RSA seems to affect levels rather than growth rates.[28] In summary, in the pre-treatment period firms who receive RSA tend to be larger and less

---

[28] This initial finding is confirmed in unreported regression analysis where we control for industry and region effects, which is available from the authors upon request.

productive. After RSA, participating firms seem to be more capital intensive and more productive and to remain larger. These features may not be causal, of course, and the econometric framework seeks to control for confounding influences that could make these initial findings spurious.


## V. RESULTS

We report firm level results for employment, investment and productivity. We then present some robustness tests before moving on to the area level analysis.

### *VA. Firm Level Analysis – Basic Results*

We start by reporting the results from estimating equation (5) using simple OLS regressions for each of the following outcome variables: employment; investment; labor productivity. Our key variable, RSA is a dummy which is equal to unity for all the periods in and after a firm has participated in the RSA program and zero otherwise. In the vector of explanatory variables we include a dummy for whether the firm is part of a domestic group or of a foreign group, a quadratic polynomial in firm age, a dummy for firms that entered before 1980 to control for left censoring of the age variable, a full set of four-digit industry dummies, regional dummies and time dummies. Note that in all the regressions involving our instruments we allow for clustering at the area level to take into account that the instruments variation is at this level.

We first turn to analyzing employment in Table 4. The first four columns do not include fixed effects whereas the last four columns do include fixed effects. The first column simply reports the basic OLS regression results. The RSA program participation dummy is positive and significant with a coefficient that indicates that RSA participation is associated with about a 37% increase in employment. Column (2) reports the reduced form where we regress firm employment on our policy instruments – dummies for the level of investment subsidy open to the firm. The omitted category is if the firm does not have a plant that is eligible for an investment subsidy. The policy dummies are all positive and jointly and individually significant (except for the least generous subsidy level, below 10% NGE). They show a sensible

pattern rising almost monotonically with the generosity of the subsidy. Column (3) reports the first stage of the 2SLS estimates where we regress the RSA dummy on all the exogenous covariates. The policy variables are again jointly significant with the largest effect from the area with the most generous subsidy.[29] In column (4) we present the instrumental variable results. The coefficient on the RSA dummy is much larger than in column (1) suggesting substantial downwards bias in OLS.

To control for permanent unobserved heterogeneity we include a full set of firm dummies and repeat the same specifications. Column (5) reports OLS estimates where the coefficient on the RSA dummy is under half the size that of column (1), but remains positive and significant at the 5% level. Columns (6) and (7) report the reduced form and first stage respectively. The policy instruments are less significant than before, but continue to have significant explanatory power, especially for the highest categories of RSA subsidies. Finally in our preferred specification of column (8) we see the IV results where the RSA coefficient is positive, significant and again larger than the OLS column (by a factor of three). This confirms the evidence of downward bias which is consistent with the view that RSA is being awarded to those firms who face negative shocks, exactly what one would expect from the policy aimed at under-performing firms in deprived areas.

Table 5 repeats the same order of specifications as Table 4 but uses investment as the dependent variable. The broad pattern of results is similar to that for employment. First, the magnitude of all RSA effects is reduced when we control for fixed effects (the OLS coefficient falls by half from 0.6 in column (1) to 0.3 in column (5)). Second, the policy instruments are informative in both the reduced firm and first stage both with and without firm effects. Third, and most importantly, the IV results are positive and significant being larger in magnitude than the OLS results (0.871 in the final column).

Table 6 reports some estimates for production functions. To save space we report fixed effects results only as similar patterns are revealed under OLS. The dependent variable in all columns is labor productivity as measured by the ratio of real gross output to employment. The coefficient on the RSA dummy is small and insignificant in the OLS regressions of column (1). When we instrument this with the policy

---

[29] The F-statistics for the excluded instruments is 16.90. This is significant at the 5% level using Table 1 and 3 of Stock and Yogo (2002) with a b=.10 (desired maximal bias of the IV estimator relative to OLS).

variables, however, the coefficient rises and is significant at the 10% level in column (4). The next four columns control for capital and materials. Once we do this, the RSA variable is insignificant even after using instrumental variables. The main reason for this is conditioning on capital. We saw from Table 5 that RSA increases investment, once we control for this effect there is no impact of RSA on labor productivity.

We confirmed the absence of an RSA effect on total factor productivity (TFP) in a number of ways such as calculating TFP as a residual and regressing this on (instrumented) RSA. We also estimated a value added per worker regression instead of a gross output regression.

In summary, our basic results suggest that RSA has a causal effect in increasing the employment and investment of participating firms. After controlling for the increase in these factor inputs, however, the policy has no impact on productivity.


## VB. Firm Level Analysis – Robustness Tests

### (a) Heterogeneous Treatment Effects

With heterogeneous treatment effects, it is important to examine whether our estimates of the average effect of treatment on the treated (ATT) could be biased due to a non-overlapping support of the distributions of participants and non-participants. Table 7 shows the results when we restrict ourselves to having a common support through using propensity score matching based on pre-treatment characteristics. We restrict the sample to those firms that have a predicted propensity score larger than the tenth percentile of the propensity score distribution of participating firms and lower than the ninetieth percentile of non-participating firms. We are left with about half of the initial sample by imposing this constraint.

Across all the dependent variables in Table 7 the qualitative results are similar to those presented earlier. There is an increase in the coefficient between OLS and IV, the instruments have power and we obtain a significant and positive treatment effect. For employment, however, we obtain a lower estimate of the treatment effect using the common support restriction than in earlier tables (0.496 rather than 0.646). This does suggest some evidence of heterogeneous treatment effects and we regard the lower number as our current preferred estimates.

*(b) Single plant firms*

Since twenty percent of firms in our data are aggregations over multiple plants the instrument is also aggregated. To investigate biases associated with this we restrict the sample to single plant firms and re-run all our results. Table 8 presents some of these, showing slightly higher effects than for the main sample. This suggests that some of the RSA effect at the plant level comes from intra-firm switching of employment from non-RSA eligible areas to RSA eligible areas.

*(c) Continuous Treatment Intensity*

The discrete dummy for program participation does not take into account the differential amounts of money to different firms. We simply estimate an average effect on the treated. Obviously this discards some useful information on the intensity of the treatment. In Table 9 we report the estimates obtained when we use the actual RSA grant intensity awarded to the firms, calculated using the ratio between the grant amount and the total cots of the investment project.

Our preferred IV specifications suggest a semi elasticity of 5.171 for employment and 6.576 for investment. Given that the average treated firm is covered by the RSA grant for about 12% of the investment; for the average treated firm the treatment effect will be 0.62 (=5.2*0.12) for employment and 0.79 (=6.6*0.12) for investment. These estimates are very similar to the ones obtained using the binary dummy variable for treatment.

## VC. Area-level analysis

The firm level regressions reported earlier suggest that the causal effect of RSA is to increase employment and investment (but not total factor productivity). A problem with this identification strategy is that it relies entirely on within incumbent firm variation and therefore cannot capture the effect of RSA on firm entry.[30] Another issue is that we only capture the direct effect of RSA on participating firms in the eligible areas but not the indirect impact of RSA on the non-participating firms in the

---

[30] We do however capture plant entry within existing reporting units but we cannot distinguish and therefore implicitly assume that there are no differential effects between RSA driven increases in employment within existing plants and creation of new plants within the firm.

eligible areas. RSA support might give supported firms a competitive edge so that their market share in the eligible areas increases at the expense of other firms.[31]

We make a first attempt to estimate the net impact of both of these effects at the level of the travel to work areas (henceforth "area"). We therefore estimate equation (6) where the outcome of interest that we focus on is aggregate employment and the number of plants at the TTWA level (investment is not available at the local unit level). Our treatment variables become the number of treatments an area receives at a given point in time and the amount of support money that flows in the area. Table 2 reports descriptive statistics at the Travel to Work Area level in terms of number of treatments and amount of financial support received expressed also in average Net Grant Equivalent terms for all eligible areas in our sample over the period analysed 1988 to 2003.

Table 10 reports reduced form regressions for this type of analysis; i.e. the explanatory variable of interest becomes the support intensity (Net Grant Equivalent rate, NGE) of an area. In column (1) we find a positive and significant effect on area employment. The coefficient implies that increasing the maximum support level (NGE) by ten percentage points is associated with a 2.5 percent increase in employment. The second column of Table 10 uses the total number of plants as the outcome – RSA appears to also, unsurprisingly increase the number of plants in an area.

The results as a whole tell a very straightforward story. RSA increases employment and investment[32] for the firms who receive it. This is disguised in the standard OLS results because the firms tend to be subject to negative unobservable (to the econometrician) shocks causing a downwards bias on the treatment effect. While we find some evidence that RSA increased the number of plants in an area, our results suggest that the employment effects on incumbents are stronger than any positive effect on entry. This area level growth is more from incumbents than new entry, as RSA tends to dampen reallocation effects (a smaller share of entrant and exit

---

[31] A further type of indirect effect is the impact of RSA on non-eligible areas. Part of this is captured when we look at the firm level, but part may be missed due to (e.g.) lower entry in non-eligible areas due to RSA. We are examining these effects by looking at the effect of RSA changes to plants in *neighbouring* areas.

[32] In unreported analysis we also look at the impact of RSA on survival of firms. However, we did not report the results of the analysis here as we could not control for unobserved heterogeneity and therefore the results might be biased. However the results obtained without controlling for unobserved heterogeneity suggest that RSA treatment is positively correlated with survival of treated firms.

employment). Consequently, since there is no productivity effect from receiving RSA grants on treated firms, the fact that they are relatively large with low productivity prior to treatment implies that RSA dampens reallocation effects from more productive to less productive plants. This is likely to dampen aggregate productivity growth both in the affected areas and in the economy as a whole.

## VI. CONCLUSIONS

There are surprisingly few micro-econometric analyses of the causal effects of industrial policy, despite their ubiquity in policy making. In this paper we have examined one business support policy – Regional Selective Assistance. We use exogenous changes in the eligibility of businesses to receive support driven by policy changes at the European level. These changes are based on areas and are exogenous to firm characteristics and the party in power in the UK.

We find that OLS and matching techniques that fail to account for the endogenous selection of firms contain a large downward bias. When we correct for this we find evidence for a positive effect of treatment on the treated in terms of employment and investment. We find no effects on (total factor) productivity, however. Since participants tend to have below average productivity and RSA helps them to expand in size this might yield a negative effect on aggregate productivity. In terms of further work we aim to investigate more carefully the indirect effects on other plants outside the eligible areas. Secondly, we could use the policy to obtain credible identification of structural parameters in the production function. The coefficient on capital in firm or plant-level production functions is difficult to estimate due to selection and endogeneity problems[33]. Investment subsidies can be an external instrument that shifts the capital stock exogenously under the assumption that RSA does not have a direct effect on TFP (consistent with what we are finding here). Finally, we have not attempted a full cost-benefit analysis, due to concerns that we have not incorporated general equilibrium effects. Nevertheless we plan some simple policy simulations in future work.

---

[33] See inter alia Marschak and Andrews (1944), Griliches and Mairesse (1998); Olley and Pakes (1996) and Ackerberg et al (2007).

## REFERENCES

Ackerberg, Daniel, A.; Kevin Caves and Garth Frazer (2007) "Structural Identification of Production Functions", available from http://www.econ.ucla.edu/ackerber/

Angrist, Joshua D. (2004) "Treatment effect heterogeneity in theory and practice" *Economic Journal*, vol. 114(494), pp. C52-C83.

Armstrong, Harvey W. (2001) "Regional Selective Assistance: Is the Spend Enough and Is It Targeting the Right Places?" *Regional Studies*, vol. 35(3) May, pp. 247-57.

Blundell, R.; Monica Costa Dias; Costas Meghir and John Van Reenen (2004) "Evaluating the Employment Impact of a Mandatory Job Search Program" *Journal of the European Economic Association*, vol. 2(4), pp. 569-606.

Criscuolo Chiara; Jonathan Haskel and Ralf Martin, (2003) "Building the evidence base for productivity policy using business data linking", *Economic Trends* vol. 600, November, pp. 39-51, www.statistics.gov.uk/articles/economic_trends/ETNov03Haskel.pdf.

Department of Trade and Industry (2001) "DEPARTMENT OF TRADE AND INDUSTRY (2001). The government's expenditure plans, 1999–2000, the regions, p.41

Devereux, Griffith and Simpson (2007)

Foster, Lucia; John Haltiwanger and C.J. Krizan (1998), "Aggregate productivity growth: lessons from microeconomic evidence", NBER Working Paper (6803).

Griffith Rachel S. (1999) "Using the ARD establishment level data to look at foreign ownership and productivity in the UK." *Economic Journal* vol. 109, pp F416-F442.

Griliches, Zvi (1979) "Issues in Assessing the Contribution of Research and Development to Productivity Growth" *Bell Journal of Economics*, The RAND Corporation, vol. 10(1), pp. 92-116, Spring.

Griliches, Zvi, and Jacques Mairesse (1998) "Production Functions: The Search for Identification," in Z. Griliches, *Practicing Econometrics: Essays in Method and Application*, Cheltenham, UK: Elgar.

Klette Tor Jakob; Jarle Møen and Zvi Griliches (2000) "Do subsidies to commercial R&D reduce market failures?" *Research Policy* vol. 29, 4-5, pp. 471-496.

Heckman, James J.; Hidehiko Ichimura, and Petra Todd (1997), "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program", *Review of Economic Studies*, vol. 64, pp. 605-654.

Heckman, James J., Hidehiko Ichimura, Jeffrey A. Smith, and Petra Todd (1998) "Characterizing selection bias using experimental data" *Econometrica* vol. 66, pp.1017–1098.

Heckman, James J.; Robert, J. LaLonde and Jeffrey A. Smith (1999) "The Economics and Econometrics of Active Labor Market Programs" in Handbook of Labor Economics, Volume 3 O. Ashenfelter and D. Card eds., pp. 1865-2097.

Heckman, James J., Jeffrey A. Smith, Nancy Clements (1997). "Making the most out of social experiments: accounting for heterogeneity in programme impacts" *Review of Economic Studies* vol. 64, pp. 487–535.

Imbens, Guido W and Angrist, Joshua D. (1994) "Identification and Estimation of Local Average Treatment Effects" *Econometrica*, Econometric Society, vol. 62(2), pp. 467-75, March.

Jones, Jonathan and Colin Wren (2004) "Do Inward Investors Achieve their Job Targets?" *Oxford Bulletin of Economics and Statistics*, vol. 66(4), pp. 483-514.

Jones, Jonathan and Colin Wren (2004) "Inward Foreign Direct Investment and Employment: A Project-Based Analysis in North-East England" *Journal of Economic Geography*, vol. 5(4), pp. 517-44.

Marschak, Jacob and Andrews, William H. (1944) "Random Simultaneous Equations and the Theory of Production" *Econometrica,* July-October, vol. 12(3-4): pp. 143-205

Olley, Steve and Ariel Pakes (1996) "The dynamics of Productivity in the Telecommunications equipment industry", *Econometrica* vol. 64 (6), pp. 1263-1297

Oulton Nicholas (1997) The ABI Respondents Database: a new resource for industrial economics research. *Economic Trends* vol. 528, pp. 46–57.

Roome, Marjorie (2005),'Linking DTI Business Support Schemes data to the IDBR and ARD with implications for the future use of the administrative data in scheme assessment', DTI report.

Stock, James H. and Motohiro Yogo (2002) "Testing for weak instruments in linear regressions", NBER Technical Working Paper 284 http://www.nber.org/papers/T0284

Wren, Colin (2005) "Regional Grants: Are They Worth It?" *Fiscal Studies*, vol 26(2), pp. 245-75.

Wren, Colin (2003) "Investment Scale as a Signal in Industrial Assistance Schemes with Employment Objectives" *Economica*, vol. 70, pp. 331-52.

Wren, Colin and David Storey (2002). "Evaluating the Effect of 'Soft' Business Support upon Small Firm Performance" *Oxford Economic Papers*, vol. 54, pp. 334-65.

Wren, Colin and J. Taylor. (1999) "Industrial Restructuring and Regional Policy" *Oxford Economic Papers*, vol. 51, pp. 487-516.

# Table 1: Summary statistics for RSA participating and non-participant firms

| Variable | sample | | (1) mean | (2) sig | (3) Sd | (4) p25 | (5) p50 | (6) p75 | (7) Obs. |
|---|---|---|---|---|---|---|---|---|---|
| **Real Value added per worker (VA/L)** | non treated | | 31.05 | | 162.51 | 17.66 | 24.27 | 34.09 | 136524 |
| | | before | 26.32 | ** | 23.51 | 16.45 | 22.38 | 30.59 | 7247 |
| | Treated | after | 31.78 | ** | 36.51 | 19.14 | 26.02 | 35.55 | 10647 |
| **Total Factor Productivity (TFP)** | non treated | | 0.02 | | 0.33 | -0.16 | 0.01 | 0.17 | 134755 |
| | | before | -0.03 | *** | 0.29 | -0.19 | -0.03 | 0.12 | 7925 |
| | Treated | after | -0.05 | *** | 0.31 | -0.20 | -0.04 | 0.11 | 9946 |
| **Employment (L)** | non treated | | 253 | | 737 | 48 | 111 | 246 | 145389 |
| | | before | 417 | *** | 957 | 83 | 171 | 383 | 8209 |
| | Treated | after | 350 | *** | 846 | 63 | 144 | 336 | 10750 |
| **Gross output (GO)** | non treated | | 26774 | | 136448 | 2422 | 6622 | 18329 | 136524 |
| | | before | 39401 | *** | 151614 | 4034 | 10256 | 28990 | 7247 |
| | Treated | after | 42774 | *** | 177942 | 3583 | 9440 | 28015 | 10647 |
| **Investment (I)** | non treated | | 1082.76 | | 8471.20 | 32.49 | 147.70 | 562.03 | 145382 |
| | | before | 1624.35 | *** | 7204.89 | 90.99 | 310.03 | 1055.98 | 8209 |
| | Treated | after | 1941.11 | *** | 8975.45 | 73.63 | 317.74 | 1217.20 | 10750 |
| **Growth in Labor Productivity Δln(VA/L)** | non treated | | 0.02 | | 0.30 | -0.12 | 0.02 | 0.16 | 85693 |
| | | before | 0.02 | *** | 0.27 | -0.11 | 0.02 | 0.16 | 5108 |
| | Treated | after | 0.03 | ** | 0.30 | -0.12 | 0.03 | 0.18 | 6782 |
| **Growth in Employment Δln(L)** | non treated | | 0.00 | | 0.19 | -0.07 | 0.00 | 0.06 | 93021 |
| | | before | 0.02 | *** | 0.17 | -0.05 | 0.01 | 0.08 | 5872 |
| | Treated | after | 0.01 | *** | 0.19 | -0.07 | 0.00 | 0.08 | 6926 |
| **Intermediate Inputs per worker (M/L)** | non treated | | 62.18 | | 1090.92 | 19.18 | 32.69 | 58.02 | 136524 |
| | | before | 48.35 | *** | 61.61 | 20.37 | 32.54 | 55.60 | 7247 |
| | Treated | after | 57.79 | *** | 74.60 | 23.17 | 38.71 | 66.04 | 10647 |
| **Capital Stock per worker (K/L)** | non treated | | 179.32 | | 6880.82 | 24.02 | 42.76 | 80.77 | 42951 |
| | | before | 69.17 | *** | 99.35 | 25.57 | 43.11 | 75.25 | 2446 |
| | Treated | after | 69.04 | *** | 89.83 | 27.03 | 45.11 | 77.02 | 4331 |
| **Age** | non treated | | 15 | | 8 | 9 | 15 | 20 | 145389 |
| | | before | 14 | *** | 6 | 9 | 15 | 18 | 8209 |
| | Treated | after | 17 | *** | 8 | 10 | 18 | 24 | 10750 |

**Notes:** Column 1, mean, reports the mean of the variables of interest separately for non treated firms and for the group of treated for the period before and after treatment. Column 2, sig, reports the significance of a t-test of equality between the values for treated firms, both before and after treatment, relative to the group of non treated firms. Column 3, "sd" reports standard deviations, while columns (4) to (6) describe the distribution – 25th; median and 75th percentile respectively, of the variables for non-treated; treated in the pre-treatment period and treated firms post treatment. Finally, column 7 reports the number of observations for each cell.

**Source:** Authors' calculation using the ARD SAMIS matched data.

## Table 2: Descriptive statistics across TTWA

| year | Treatments (1) Mean | Treatments (2) Sd | Annualized payments (3) mean | Annualized payments (4) sd | Net grant equivalent (NGE) (5) mean | Net grant equivalent (NGE) (6) sd |
|------|------|-------|-----------|-----------|------|------|
| 1988 | 3.42 | 3.92 | 903,552 | 2,033,946 | 0.23 | 0.09 |
| 1989 | 7.82 | 10.50 | 1,271,698 | 5,205,792 | 0.21 | 0.10 |
| 1990 | 11.05 | 14.19 | 936,613 | 1,513,561 | 0.21 | 0.10 |
| 1991 | 11.27 | 13.61 | 1,247,630 | 2,192,776 | 0.21 | 0.10 |
| 1992 | 12.99 | 15.61 | 1,789,744 | 3,119,778 | 0.19 | 0.12 |
| 1993 | 12.91 | 18.30 | 1,701,525 | 3,253,045 | 0.19 | 0.11 |
| 1994 | 12.33 | 16.55 | 1,372,372 | 2,934,404 | 0.19 | 0.11 |
| 1995 | 13.46 | 19.95 | 1,732,582 | 3,084,355 | 0.20 | 0.10 |
| 1996 | 12.88 | 18.83 | 1,482,756 | 2,859,229 | 0.19 | 0.11 |
| 1997 | 12.03 | 16.30 | 1,195,726 | 2,055,161 | 0.20 | 0.10 |
| 1998 | 11.21 | 15.55 | 1,190,799 | 2,779,583 | 0.20 | 0.10 |
| 1999 | 9.79 | 12.82 | 1,126,135 | 1,947,954 | 0.15 | 0.14 |
| 2000 | 7.02 | 8.08 | 969,325 | 1,626,163 | 0.17 | 0.14 |
| 2001 | 5.17 | 5.80 | 837,742 | 1,357,481 | 0.17 | 0.14 |
| 2002 | 3.85 | 4.84 | 630,586 | 1,319,403 | 0.18 | 0.14 |
| 2003 | 4.63 | 4.35 | 992,399 | 1,487,250 | 0.15 | 0.12 |

**Notes:** Column (1), mean, reports the mean of the number of RSA grants given within an eligible area. Column (2), "sd" is standard deviation. Columns (3) and (4) report mean and standard deviation of annualised deflated RSA grants values given within an eligible area. Finally, column (5) and (6) reports the average NGE rate and its standard deviation. Note that there are 297 areas, but the table reports statistics only over those areas that are eligible for RSA subsidies (about 100 areas in each year).

**Source:** Authors' calculation using the ARD SAMIS matched data.

## Table 3: Source of identification

| Year | Changes in support status local unit level | Changes in support status reporting unit level | Changes in support intensity local unit level | Changes in support intensity reporting unit level |
|------|------|------|------|------|
| 1993 | 3929 | 3062 | 1478 | 1251 |
| 2000 | 9076 | 5088 | 6411 | 3808 |

**Notes:** Column (1) reports the number of plants ("local units") in the regression sample for which a change in eligibility is observed. The top panel is for the 1993 map change and the bottom panel for the 2000 change. Column (2) reports similar figures for firms ("reporting units"). Column 3 and 4 report only number of changes in eligibility level (i.e. from NGE=x% to NGE=z% where x≠z) for plants and firms respectively. For example, in 1993 there were 3062 changes in support status at the Reporting Unit level. 41% of these changes (1251/3062) were in eligibility intensity.

**Source:** Authors' calculation using the ARD SAMIS matched data.

## Table 4: ln(Employment) Regressions

| Dependent Variable | (1) Ln(EMP) | (2) ln(EMP) | (3) RSA | (4) ln(EMP) | (5) ln(EMP) | (6) ln(EMP) | (7) RSA | (8) ln(EMP) |
|---|---|---|---|---|---|---|---|---|
| | OLS | Reduced Form | First Stage | IV | FE | Reduced Form | First Stage | IV |
| RSA | 0.370*** | | | 4.658*** | 0.166*** | | | 0.646*** |
| | (0.020) | | | (0.530) | (0.013) | | | (0.154) |
| NGE = 10% | | 0.197 | 0.063 | | | -0.062* | 0.015 | |
| | | (0.127) | (0.041) | | | (0.032) | (0.020) | |
| NGE = 15% | | 0.368*** | 0.122*** | | | -0.005 | 0.068*** | |
| | | (0.049) | (0.023) | | | (0.019) | (0.014) | |
| NGE = 20% | | 0.382*** | 0.050*** | | | 0.027*** | 0.015*** | |
| | | (0.017) | (0.008) | | | (0.003) | (0.004) | |
| NGE = 30% | | 0.406*** | 0.091*** | | | 0.033*** | 0.027*** | |
| | | (0.031) | (0.009) | | | (0.006) | (0.005) | |
| NGE = 35% | | 0.236*** | 0.197*** | | | 0.033 | 0.098*** | |
| | | (0.043) | (0.024) | | | (0.023) | (0.014) | |
| Observations | | | | 157771 | | | | |
| Number of firms | | | | 28882 | | | | |
| F-stats for excluded instruments | | | 29.52 | | | | 16.90 | |
| Fixed effects | NO | NO | NO | NO | YES | YES | YES | YES |

**Notes:** *RSA* equals unity for all the periods in and after a firm has participated in the program and zero otherwise. *NGE* is Net Grant Equivalent (maximum investment subsidy) at the area-level. Eligibility for investment subsidies used as an instrumental variable in columns (4) and (8). *NGE = x%* indicates that the firm's has a reference plant in an area that is eligible for up to x% in investment subsidy. All columns include controls for whether a firm is foreign owned, whether it is part of a domestic multi-firm group, a quadratic in age; an age censoring dummy for firms born before 1980 and a full set of four digit industry dummies, regional and time dummies. Standard errors below coefficients are robust to heteroscedacity and arbitrary serial correlation (they are clustered by firm in columns (1) and (5) and by area in columns (2), (3) (4), (6), (7) and (8)). The last four columns also include a full set of firm dummies. Time period is 1985-2004.

## Table 5: ln(Real Investment) Regressions

| Dependent Variable | (1) ln(INV) | (2) ln(INV) | (3) RSA | (4) ln(INV) | (5) ln(INV) | (6) ln(INV) | (7) RSA | (8) ln(INV) |
|---|---|---|---|---|---|---|---|---|
| | OLS | Reduced Form | First Stage | IV | FE | Reduced Form | First Stage | IV |
| RSA | 0.623*** | | | 5.606*** | 0.306*** | | | 0.871** |
| | (0.031) | | | (0.705) | (0.024) | | | (0.358) |
| NGE = 10% | | 0.174 | 0.084** | | | -0.065 | 0.021 | |
| | | (0.216) | (0.043) | | | (0.106) | (0.022) | |
| NGE = 15% | | 0.377*** | 0.116*** | | | -0.004 | 0.069*** | |
| | | (0.077) | (0.023) | | | (0.050) | (0.015) | |
| NGE = 20% | | 0.433*** | 0.051*** | | | 0.037*** | 0.016*** | |
| | | (0.028) | (0.008) | | | (0.010) | (0.004) | |
| NGE = 30% | | 0.492*** | 0.091*** | | | 0.035** | 0.028*** | |
| | | (0.048) | (0.009) | | | (0.014) | (0.005) | |
| NGE = 35% | | 0.266*** | 0.198*** | | | 0.063 | 0.099*** | |
| | | (0.082) | (0.025) | | | (0.059) | (0.015) | |
| Observations | | | | 143334 | | | | |
| Number of Firms | | | | 28022 | | | | |
| F-stats for excluded instruments | | | 30.16 | | | | 16.32 | |
| Fixed effects | NO | NO | NO | NO | YES | YES | YES | YES |

**Notes:** *RSA* equals unity for all the periods in and after a firm has participated in the program and zero otherwise. *NGE* is Net Grant Equivalent (maximum investment subsidy) at the area-level. Eligibility for investment subsidies used as an instrumental variable in columns (4) and (8). *NGE = x%* indicates that the firm's has a reference plant in an area that is eligible for up to x% in investment subsidy. All columns include controls for whether a firm is foreign owned, whether it is part of a domestic multi-firm group, a quadratic in age; an age censoring dummy for firms born before 1980 and a full set of four digit industry dummies, regional and time dummies. Standard errors below coefficients are robust to heteroskedacity and arbitrary serial correlation (they are clustered by firm in columns (1) and (5) and by area in columns (2), (3) (4), (6), (7) and (8)). The last four columns include a full set of firm dummies. Time period is 1985-2004.

34

## Table 6: Labor Productivity Regressions
## ln(Real Gross Output/employment)

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| **Dependent Variable** | ln(GO/EMP) | ln(GO/EMP) | RSA | ln(GO/EMP) | ln(GO/EMP) | ln(GO/EMP) | RSA | ln(GO/EMP) |
| | FE | Red. form | First Stage | IV | FE | Red. form | First Stage | IV |
| RSA | 0.002 | | | 0.248* | 0.007 | | | 0.106 |
| | (0.008) | | | (0.133) | (0.006) | | | (0.088) |
| NGE = 10% | | 0.030 | 0.017 | | | 0.013 | 0.020 | |
| | | (0.032) | (0.020) | | | (0.023) | (0.020) | |
| NGE = 15% | | 0.032** | 0.068*** | | | 0.014 | 0.069*** | |
| | | (0.013) | (0.014) | | | (0.009) | (0.014) | |
| NGE = 20% | | 0.005 | 0.012*** | | | 0.005** | 0.012*** | |
| | | (0.003) | (0.004) | | | (0.002) | (0.004) | |
| NGE = 30% | | 0.013*** | 0.022*** | | | 0.006** | 0.020*** | |
| | | (0.004) | (0.005) | | | (0.003) | (0.005) | |
| NGE = 35% | | -0.014 | 0.090*** | | | -0.012 | 0.091*** | |
| | | (0.017) | (0.014) | | | (0.011) | (0.014) | |
| ln(Materials/ Employment) | | | | | 0.494*** | 0.494*** | 0.011*** | 0.493*** |
| | | | | | (0.005) | (0.007) | (0.003) | (0.007) |
| ln(Capital/ Employment) | | | | | 0.012*** | 0.012*** | 0.004*** | 0.012*** |
| | | | | | (0.001) | (0.001) | (0.001) | (0.001) |
| ln(Employment) | | | | | -0.128*** | -0.127*** | 0.044*** | -0.132*** |
| | | | | | (0.004) | (0.004) | (0.005) | (0.005) |
| Observations | | | 148102 | | | | 142937 | |
| Number of Firms | | | 28865 | | | | 28082 | |
| F-stats for excluded instruments | | | 15.31 | | | | 14.74 | |
| Fixed effects | YES | YES | YES | YES | YES | YES | YES | YES |

**Notes:** RSA equals 1 for all the periods in and after a firm has participated in the program and zero otherwise. NGE is Net Grant Equivalent at the area-level; eligibility for investment subsidies used as an instrumental variable in columns (4) and (8). NGE = x% indicates that the firm's has a plant that is eligible for up to x% in investment subsidy. All columns control for a quadratic in age; regional dummies; an age censoring dummy for firms born before 1980 and a full set of four digit industry dummies, time dummies and include a full set of firm dummies. Standard errors below coefficients are robust to heteroskedacity and arbitrary serial correlation (they are clustered by firm in columns (1) and (5) and by area in columns (2), (3) (4), (6), (7) and (8)).

# Table 7: ln employment –
## Using Propensity Score Matching to obtain Common Support

| Dependent Variable | (1) ln(EMP) | (2) RSA | (3) ln(EMP) | (4) ln(INV) | (5) RSA | (6) ln(INV) | (7) ln(GO) | (8) RSA | (9) ln(GO) |
|---|---|---|---|---|---|---|---|---|---|
| | FE | First Stage | IV | FE | First Stage | IV | FE | First Stage | IV |
| RSA | 0.141*** | | 0.496*** | 0.346*** | | 0.871** | 0.002 | | 0.083 |
| | (0.015) | | (0.164) | (0.028) | | (0.390) | (0.006) | | (0.089) |
| NGE = 10% | 0.000 | 0.000 | | | 0.008 | | | 0.008 | |
| | (0.000) | (0.000) | | | (0.023) | | | (0.020) | |
| NGE = 15% | | 0.005 | | | 0.085*** | | | 0.083*** | |
| | | (0.020) | | | (0.019) | | | (0.018) | |
| NGE = 20% | | 0.082*** | | | 0.018*** | | | 0.014*** | |
| | | (0.018) | | | (0.004) | | | (0.004) | |
| NGE = 30% | | 0.016*** | | | 0.026*** | | | 0.021*** | |
| | | (0.004) | | | (0.005) | | | (0.005) | |
| NGE = 35% | | 0.025*** | | | 0.107*** | | | 0.100*** | |
| | | (0.006) | | | (0.019) | | | (0.017) | |
| ln(materials/employment) | | | | | | | 0.000 | -0.006 | 0.001 |
| | | | | | | | (0.005) | (0.007) | (0.004) |
| ln(capital/employment) | | | | | | | 0.003 | -0.005 | 0.003 |
| | | | | | | | (0.003) | (0.004) | (0.003) |
| ln(Employment) | | | | | | | 0.520*** | 0.009** | 0.519*** |
| | | | | | | | (0.006) | (0.004) | (0.008) |
| | | | | | | | 0.012*** | 0.006*** | 0.011*** |
| | | | | | | | | | |
| Observations | | 93701 | | | 87575 | | | 87354 | |
| Number of Firms | | 13402 | | | 13401 | | | 13396 | |
| F-stats for excluded instruments | | 14.63 | | | 15.00 | | | 14.03 | |
| Fixed effects | YES | YES | YES | YES | YES | YES | YES | YES | YES |

**Notes:** *RSA* equals unity for all the periods in and after a firm has participated in the program and zero otherwise. *NGE* is Net Grant Equivalent (maximum investment subsidy) at the area-level. Eligibility for investment subsidies used as an instrumental variable in columns (4) and (8). *NGE = x%* indicates that the firm's has a reference plant in an area that is eligible for up to x% in investment subsidy. All columns include controls for whether a firm is foreign owned, whether it is part of a domestic multi-firm group, a quadratic in age; an age censoring dummy for firms born before 1980 and a full set of four digit industry dummies, regional and time dummies. Standard errors below coefficients are robust to heteroskedacity and arbitrary serial correlation (they are clustered by firm in columns (1) and (5) and by area in columns (2), (3)  (4),  (6), (7) and (8)).  The last four columns include a full set of firm dummies. Time period is 1985-2004.

# Table 8: ln(Employment) – single plants

| Dependent Variable | (1) ln(EMP) | (2) ln(EMP) | (3) RSA | (4) ln(EMP) | (5) ln(EMP) | (6) ln(EMP) | (7) RSA | (8) ln(EMP) |
|---|---|---|---|---|---|---|---|---|
| | OLS | Reduced Form | First Stage | IV | FE | Reduced Form | First Stage | IV |
| RSA | 0.343*** | | | 4.988*** | 0.168*** | | | 0.801*** |
| | (0.022) | | | (0.712) | (0.014) | | | (0.160) |
| NGE = 10% | | 0.272* | 0.042 | | | -0.032 | 0.022 | |
| | | (0.156) | (0.057) | | | (0.034) | (0.020) | |
| NGE = 15% | | 0.374*** | 0.123*** | | | -0.015 | 0.057*** | |
| | | (0.059) | (0.028) | | | (0.020) | (0.016) | |
| NGE = 20% | | 0.398*** | 0.048*** | | | 0.030*** | 0.014*** | |
| | | (0.020) | (0.009) | | | (0.004) | (0.005) | |
| NGE = 30% | | 0.415*** | 0.087*** | | | 0.045*** | 0.036*** | |
| | | (0.030) | (0.012) | | | (0.009) | (0.006) | |
| NGE = 35% | | 0.257*** | 0.190*** | | | 0.028 | 0.090*** | |
| | | (0.052) | (0.031) | | | (0.021) | (0.017) | |
| Observations | | | | | 105346 | | | |
| Number of Firms | | | | | 24549 | | | |
| F-stats for excluded instruments | | | 20.11 | | | | 14.12 | |
| Fixed effects | NO | NO | NO | NO | YES | YES | YES | YES |

**Notes:** *RSA* equals unity for all the periods in and after a firm has participated in the program and zero otherwise. *NGE* is Net Grant Equivalent (maximum investment subsidy) at the area-level. Eligibility for investment subsidies used as an instrumental variable in columns (4) and (8). *NGE = x%* indicates that the firm's has a reference plant in an area that is eligible for up to x% in investment subsidy. All columns include controls for whether a firm is foreign owned, whether it is part of a domestic multi-firm group, a quadratic in age; an age censoring dummy for firms born before 1980 and a full set of four digit industry dummies, regional and time dummies. Standard errors below coefficients are robust to heteroskedacity and arbitrary serial correlation (they are clustered by firm in columns (1) and (5) and by area in columns (2), (3), (4), (6), (7) and (8)). The last four columns include a full set of firm dummies. Time period is 1985-2004.

# Table 9: ln(Employment) - Continuous Treatment Intensity

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| **Dependent Variable** | ln(EMP) | | NGE_RSA | ln(EMP) | ln(EMP) | | NGE_RSA | ln(EMP) |
| | OLS | Reduced Form | First Stage | IV | FE | Reduced Form | First Stage | IV |
| NGE_RSA | 2.954*** | | | 35.295*** | 0.876*** | | | 5.171*** |
| | (0.185) | | | (4.586) | (0.095) | | | (1.029) |
| NGE = 10% | | 0.197 | 0.007 | | | -0.062* | 0.002 | |
| | | (0.127) | (0.005) | | | (0.032) | (0.002) | |
| NGE = 15% | | 0.368*** | 0.007*** | | | -0.005 | 0.006*** | |
| | | (0.049) | (0.002) | | | (0.019) | (0.001) | |
| NGE = 20% | | 0.382*** | 0.006*** | | | 0.027*** | 0.003*** | |
| | | (0.017) | (0.001) | | | (0.003) | (0.001) | |
| NGE = 30% | | 0.406*** | 0.013*** | | | 0.033*** | 0.004*** | |
| | | (0.031) | (0.001) | | | (0.006) | (0.001) | |
| NGE = 35% | | 0.236*** | 0.022*** | | | 0.033 | 0.015*** | |
| | | (0.043) | (0.004) | | | (0.023) | (0.003) | |
| Observations | | | | 157771 | | | | |
| Number of Firms | | | | 28789 | | | | |
| F-stats for excluded instruments | | | 26.76 | | | | 18.21 | |
| Fixed effects | NO | NO | NO | NO | YES | YES | YES | YES |

**Notes:** *NGE_RSA* is the investment subsidy received by the firm (total payment divided by total investment cost). *NGE* is Net Grant Equivalent (maximum investment subsidy) at the area-level. Eligibility for investment subsidies used as an instrumental variable in columns (4) and (8). *NGE = x%* indicates that the firm's has a reference plant in an area that is eligible for up to x% in investment subsidy. All columns include controls for whether a firm is foreign owned, whether it is part of a domestic multi-firm group, a quadratic in age; an age censoring dummy for firms born before 1980 and a full set of four digit industry dummies, regional and time dummies. Standard errors below coefficients are robust to heteroskedacity and arbitrary serial correlation (they are clustered by firm in columns (1) and (5) and by area in columns (2), (3) (4), (6), (7) and (8)). The last four columns include a full set of firm dummies. Time period is 1985-2004.

## Table 10: Regressions at the travel to work area level

| Dependent Variable | (1) Ln(EMP) | (2) ln(Number of plants) | (3) Share of new entrants in total employment | (4) Share of exitors in total employment |
|---|---|---|---|---|
| NGE(t) | 0.250*** | 0.091*** | -0.034** | -0.027* |
|  | (0.051) | (0.029) | (0.013) | (0.015) |
| Year dummies | YES | YES | YES | YES |
| Area fixed effects | YES | YES | YES | YES |
| Observations | 5929 | 5929 | 5336 | 5336 |
| Number of TTWA |  | 297 |  |  |

**Notes:** NGE is Net Grant Equivalent level; i.e. the maximum percentage of support firms can apply for in an area. All columns include a full set of travel to work areas dummies and time dummies. Standard errors below coefficients are robust to heteroscedacity and arbitrary serial correlation. Time period is 1985-2004.

**Figure 1: Assisted Areas Map prior to August 1<sup>st</sup> 1993**



**Notes:** The shaded areas are those which are eligible for some Regional Selective Assistance. The dark shaded areas are the very deprived areas eligible for an investment subsidy of up to 30% NGE (Net Grant Equivalence). The light shaded areas are eligible for up to 20% NGE.
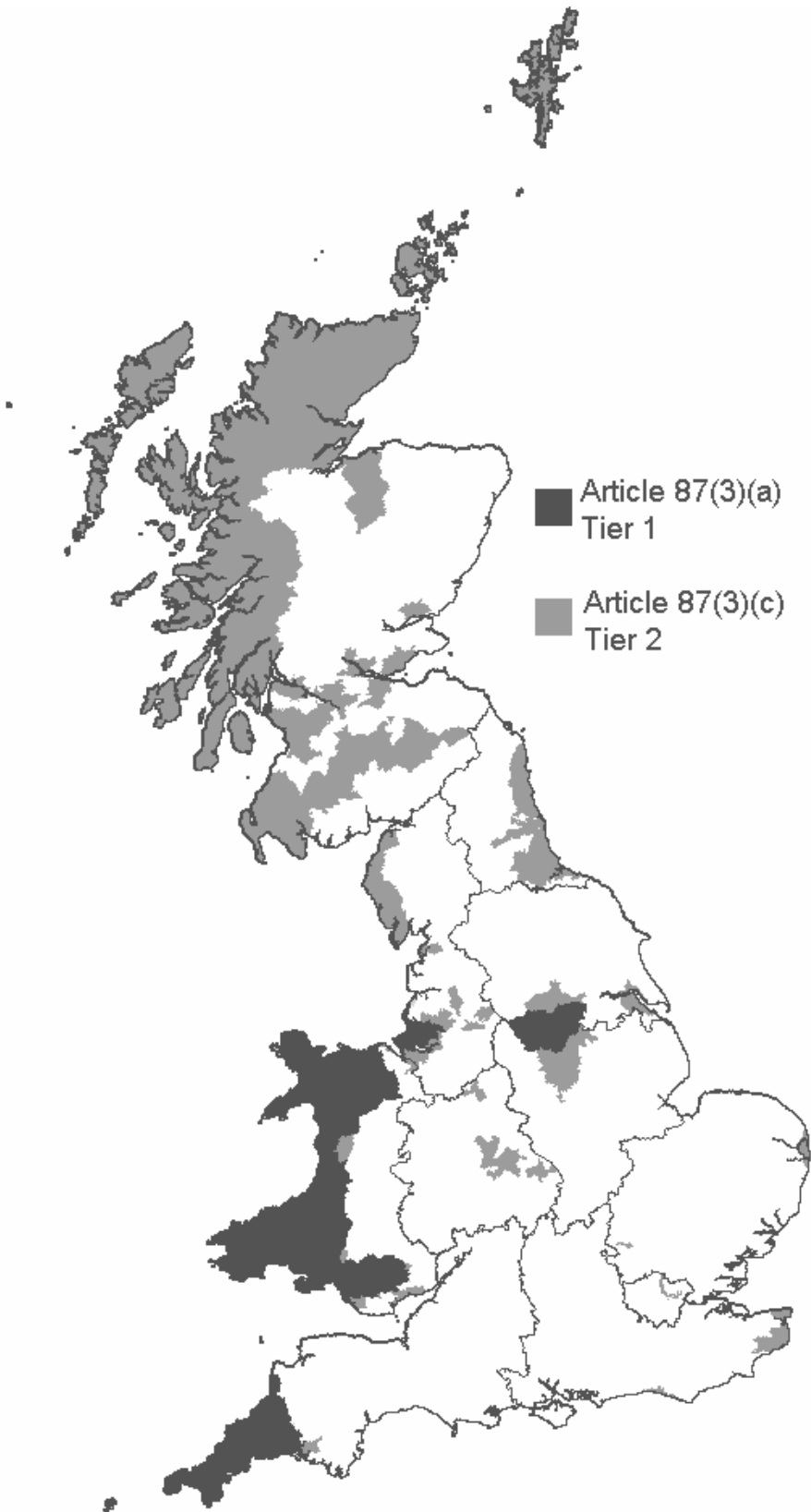**Source:** Department of Trade and Industry

**Figure 2: Assisted Areas Map after August 1<sup>st</sup> 1993 and prior to January 1<sup>st</sup> 2000**



**Notes:** The shaded areas are those which are eligible for some Regional Selective Assistance. The dark shaded areas are the very deprived areas eligible for an investment subsidy of up to 30% NGE (Net Grant Equivalence). The light shaded areas are eligible for up to 20% NGE.
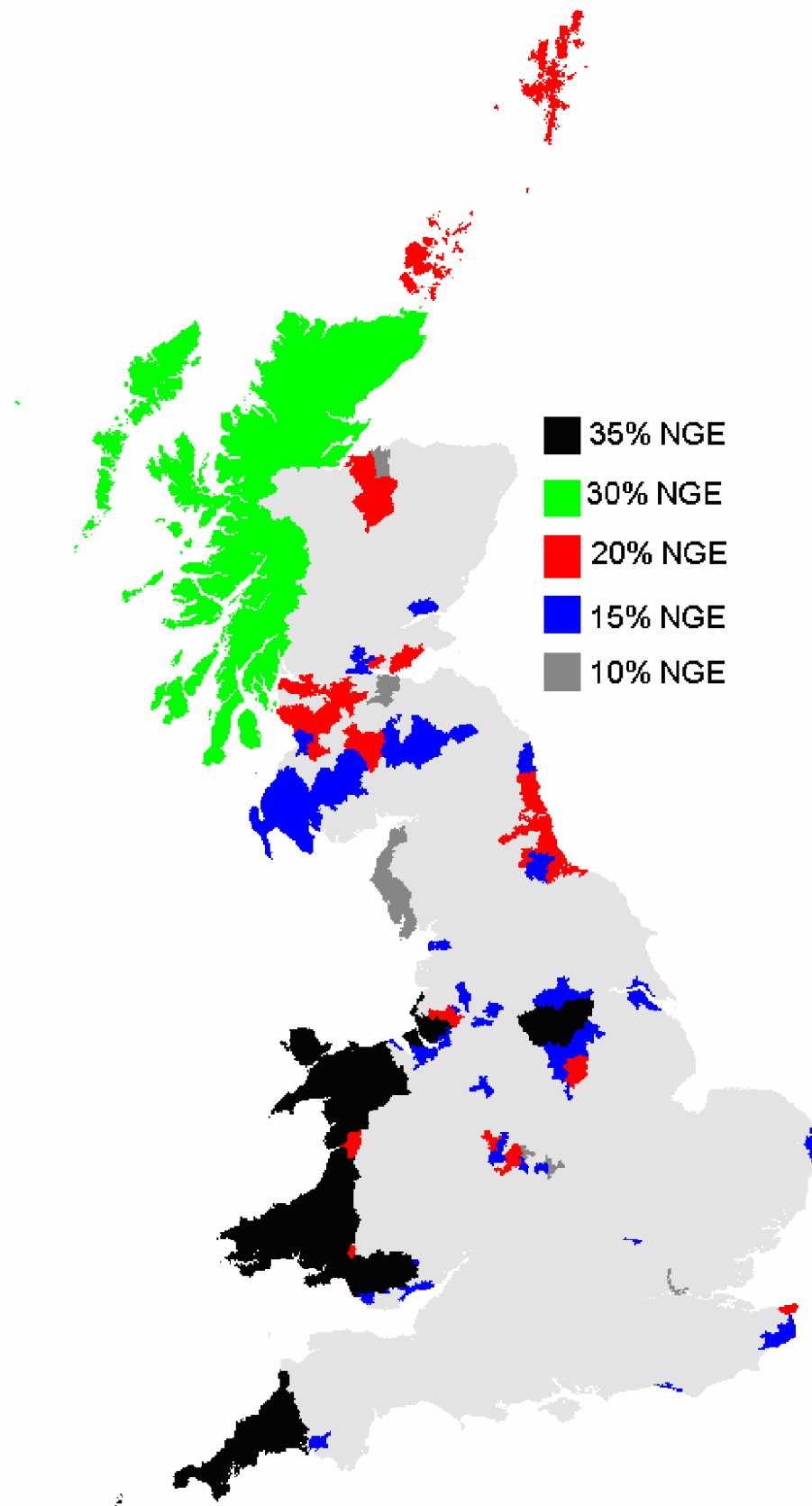**Source:** Department of Trade and Industry

**Figure 3: Assisted Areas Map after January 1st 2000**



Article 87(3)(a)
Tier 1

Article 87(3)(c)
Tier 2

**Notes:** The shaded areas are those which are eligible for some Regional Selective Assistance.
**Source:** Department of Trade and Industry

**Figure 4: Assisted Areas Map with detailed NGE rates  after January 1st 2000**



**Source:** Department of Trade and Industry
**Notes**: This shows all the different levels of NGE by area

## *Appendices*

### *A. Details of data matching*

Since the performance data comes from data sources unrelated to programme participation several problems arise in matching. Firstly, our dataset with productivity information, the ARD, is a survey with stratified random sampling (i.e. not a census) so programme participants might not be sampled in every year. Secondly, it may be that the company identifier in the program applicants' database does not give a unique match to the ARD.

For any dataset it is useful to keep in mind the unit of observation (i.e. the unit that defines a row in the dataset). The ARD unit of observation is referred to as a "reporting unit" (RU) and provides data on these units on an annual basis. Reporting units are composed of many local units (LUs) and we have population information (on employment and location) on these local units through the Inter-Departmental Business Register (IDBR) which is linked to the ARD. We refer to the RUs as firms and the LUs as plants. Some firms are part of larger "enterprise groups" and this is also recorded (and controlled for in the regressions). The DTI administrative data on the other hand reports on applicants and participants' in the RSA programme. The key difficulty is to identify which ARD reporting unit or IDBR local unit has potentially been affected by a particular incidence of programme participation.

DTI uses name and postcode from its administrative data to match a list of participants (and possibly applicants) to the Inter-Departmental Business Register. This matching may occur at the local unit, reporting unit, enterprise unit, and sometimes PAYE unit and Company's House Reference Numbers levels. We exploit the match at each level to get a pointer from the administrative record to the reporting units that potentially could have been affected by the RSA subsidies. There are three possible outcomes: (i) we cannot map the support to any reporting unit in the ARD; (ii) we map the support uniquely to one reporting unit in the ARD or (iii) we map the support to multiple ARD RUs. Figure A1 illustrates this graphically. For each ARD reporting unit that is matched to a record of DTI support, we then examine if we have sufficient pre- and post- programme participation performance data to perform econometric evaluation analysis[34].

This raises a number of issues. Consider a binary treatment (i.e. simply a one for getting an incident or a zero for not getting it)[35]. Firstly, as Figure A1 illustrates, even if we have a unique match between an observation in the DTI's dataset and an ARD reporting unit that unit might report for several local units and only a subset of them may actually be affected by the programme. The fact that it only affects a subset of local units may lead us to understate the programme impact if we do not control for this in the evaluation. Similarly, for records of DTI support matched with multiple ARD RUs we might have the problem that not all ARD RUs are affected by the programme.

---

[34] This is not an issue for the area-level analysis where there are many observations per area (see Figure 2).

[35] These problems are less severe if the treatment is continuous and measured in cash terms. For example, if a firm received an RSA grant of £100,000 then this is equivalent to two grants of £50,000. It is not so obvious how to aggregate a binary treatment (employment weights are the practical solution in many cases).

The fundamental problem with these issues is that we generally do not know which case applies, which makes it hard to control for it in an econometric model. RSA should by definition only apply to local units. However, if the head office applies for the RSA on behalf of one of its local units then the administrative dataset would hold the postcode of the head office. As a consequence we would associate the administrative record with the head office local unit instead of the programme participating local unit.

Our main response to this is to estimate at the reporting unit level where we effectively aggregate over all local units. Being aware of the problem in each case we consider many robustness tests. This can be done for example by looking at differences in evaluation results when using only a sample of RUs with unique match to a treatment unit or by aggregating over several RUs in the case of multiple matches.

In the SAMIS database, there is information on 54,322 program applications, whether or not the application has been successful from 1972 to 2004. Using name, postcode and CRN numbers, the information in the DTI files was matched to the IDBR for 68% of the cases. The improvement of matching rate over time, shown in Table A1, reflects the fact that the IDBR was introduced in 1994 and does not contain information for "units" that have closed down before 1993. Given the lower match rate in earlier years and fewer observations in the last few years we focus only on data between 1988 through 2003 in the econometric analysis.

Concerning the matching to the IDBR, there are three main issues. First, Name and address identifiers from the DTI SAMIS database for one investment project are sometimes attributed to several IDBR units. Secondly, there may be only one unique match but the company name and postcode does not necessarily refer to the local unit that is affected by the support scheme. Sometimes, especially if the grant is given to a group to open up a new plant in a deprived area, where the application is likely handled by the Headquarters the recorded address and postcode will not correspond to the plant which got the money. Thirdly, one "IDBR unit" might have applied for and received several grants.

We start by considering all matches as valid. In a few instances "IDBR units" might have received more than one grant. We proceed as follows in those cases: when the IDBR unit has applied several times to RSA we keep this information separately in the data; however if there are more than one application in the same year we consider the sum of these applications (i.e. the total sum awarded) for that year, since our production and employment data is yearly.

Finally, we conducted a detailed comparison of the characteristics of projects and project participants of firms that DTI matched with IDBR relative to all the projects in the database. We do not report these here (available from the authors) but the analysis shows that the set of "IDBR matches" do not significantly differ from the rest of the projects in the database and this is the case for both unsuccessful and successful applications. The variables we considered are: application amounts, headquarter location, a dichotomous variable which is one if the application was handled by the London office of the DTI, foreign owned, and a DTI code that seeks to identify "internationally mobile" jobs.

### A.II Matching with ARD Sample

When matching to the ARD sample, the aim is to investigate whether there is sufficient information in the matched RSA-ARD sample to undertake meaningful econometric evaluation. Sample sizes for number of successful firms matched are larger than those for unsuccessful despite the fact that both outcomes are roughly equally likely in the DTI data.

We find that the numbers appear sufficiently large to examine programme effects with at least a five year window and possibly beyond. From our analysis we conclude that a large fraction of units in our sample are unique matches and while there is quite a significant number of administrative records that are linked to several RUs in most cases only one of these RUs has sampled information in the ARD.

### B. Variable Definitions

#### Regression based TFP

There are numerous ways to obtain a TFP measure, a subject of ongoing debate in the economic literature. For the purpose of this study we experiment with a number of different TFP measures to see if our results are sensitive to the method chosen. The simplest measure we use is regression based TFP. This involves running a regression of gross output per employee deflated using 2-digit sectoral producer price indices (source EUKLEMS database) on capital stock per employee, deflated material inputs using 2-digit sectoral material price indices (source EUKLEMS database) per employee and employment:

$$\ln \frac{GO_{it}}{L_{it}} = \beta_K \ln \frac{K_{it}}{L_{it}} + \beta_M \ln \frac{M_{it}}{L_{it}} + \beta_L \ln L_{it} + TFP_{it} \tag{1}$$

where GO is gross output, K is capital, M is inputs and L is employment.

The TFP measure is then obtained by calculating the residual from that regression:

$$\widehat{TFP}_{it} = \ln \frac{GO_{it}}{L_{it}} - \beta_K \ln \frac{K_{it}}{L_{it}} - \beta_M \ln \frac{M_{it}}{L_{it}} - \beta_L \ln L_{it}$$

To analyse if treatment has any effects on TFP we could then run a second regression of estimated TFP on the treatment indicator; i.e.

$$\widehat{TFP}_{it} = \delta D_{it} + \varepsilon_{it}$$

It is more convenient, however, to run both regressions in one step as

$$\ln \frac{GO_{it}}{L_{it}} = \beta_K \ln \frac{K_{it}}{L_{it}} + \beta_M \ln \frac{M_{it}}{L_{it}} + \beta_L \ln L_{it} + \delta D_{it} + \varepsilon_{it} \tag{2}$$

These two procedures are equivalent if $D_{it}$ is not correlated with any of the production factor variables. It is very likely that this condition is not met, however, because treatment may both shift TFP and lead to adjustments of the factors mix. In this case using two steps is not only less convenient but may also give biased results. For this reason, we focus only on the one-step regressions below.

## C. Previous evaluations of the policy

Most of the previous evaluation studies of RSA are based on "industrial survey" techniques where senior personnel of a randomly drawn sample of assisted firms are asked to give their subjective assessment of what the counterfactual situation would have been had they not received the grant (see AEP NERA 2003, Cambridge Economics). Few other studies have used firm-level econometric techniques to evaluate the direct impact of RSA (Wren, 1994).

Devereux, Griffith and Simpson (2007) look at the role of RSA in affecting location decision of greenfield investments by foreign-owned multinationals and UK-owned multi-plant groups across different counties using information from matching the Annual Respondents Database (which we describe below) with publicly available information published by DTI on the subset of the largest RSA grants' offers of above £75,000 over the period 1986-1992.[36] Using econometric methods[37] to solve the problem of endogeneity of firm and industry characteristics in the location choice equation their analysis suggests that grants are a very poor predictors of firms' location choices relative to agglomeration effects, such as locating near other foreign-owned plants in the same industry and natural advantages. In fact they find that an increase in the expected grant of £100,000 raises the probability of a Greenfield locating in an assisted area from 1% to 1.01%; this rises to 1.03% when taking into account that the location incentive of grant offers increases as the economic activity in the entrant's industry increases in the assisted area.

Harris and Robinson use the ARD matched the SAMIS database (which we describe in more detail below) over the period 1990-1998 to look at differences in survival rates for RSA recipients vs. non-recipients using a hazard model and sources of productivity growth in assisted vs. non-assisted areas using the productivity decomposition techniques taken by Haltiwanger (1997). Their results show that RSA recipients have significantly higher survival rates and that while in terms of labour productivity growth RSA recipients make a significant contribution to aggregate growth; in terms of TFP growth RSA treated plants experienced negative growth, mainly because plants with low initial TFP increased their market share. The authors therefore conclude that "plants in receipt of RSA generally experience market share growth despite having relatively lower productivity" (*ibidem* p.763).

Harris and Robinson look at how RSA is related to aggregate productivity growth using a decomposition technique

Finally, Jones and Wren (2003) look at differences in survival between "treated plants", i.e. recipients of RSA grants and non-treated firms and find that treated firms have shorter survival durations.

---

[36] Not that their matching is coarser than our as they can only match the two datasets using postcode and industry affiliation of the plant, while we also use information on the name of the applicant.

[37] Thy estimate a conditional logit equation of location choices –where any variable that do not vary across region drop out - on the predicted value of grant. The latter is estimated in a first stage where non-plant specific explanatory variables are included and corrected for self-selection using a Heckman selection equation method. The selection equation is estimated using data on a set of potential applicants and including firm-specific variables that affect the probability of application but not the amount of grant offered.

Therefore the econometric evidence seems to suggest some positive effect of RSA on employment but very small effect on location and on productivity (growth) and survival.

## Table A1: the RSA database and the success rate of IDBR matching

| Year | Applications | Linkage % |
|---|---|---|
| 1972 | 357 | 34.73 |
| 1973 | 1,458 | 39.64 |
| 1974 | 1,119 | 42.45 |
| 1975 | 872 | 41.40 |
| 1976 | 1,024 | 47.07 |
| 1977 | 1,214 | 46.05 |
| 1978 | 1,199 | 51.79 |
| 1979 | 1,418 | 52.75 |
| 1980 | 584 | 47.60 |
| 1981 | 743 | 43.20 |
| 1982 | 1,305 | 47.51 |
| 1983 | 1,144 | 46.50 |
| 1984 | 1,363 | 54.22 |
| 1985 | 1,437 | 63.12 |
| 1986 | 2,034 | 62.73 |
| 1987 | 3,357 | 61.45 |
| 1988 | 3,119 | 64.67 |
| 1989 | 2,553 | 61.22 |
| 1990 | 2,782 | 63.66 |
| 1991 | 2,589 | 70.14 |
| 1992 | 2,336 | 75.30 |
| 1993 | 2,732 | 80.78 |
| 1994 | 2,710 | 80.89 |
| 1995 | 2,900 | 81.24 |
| 1996 | 2,516 | 84.70 |
| 1997 | 2,219 | 87.07 |
| 1998 | 1,775 | 86.99 |
| 1999 | 1,913 | 86.83 |
| 2000 | 1,003 | 89.13 |
| 2001 | 874 | 90.50 |
| 2002 | 675 | 92.00 |
| 2003 | 705 | 89.79 |
| 2004 | 222 | 91.89 |
| 2005 | 71 | 97.18 |
| Total | 54,322 | 67.83 |

**Notes:** This Table shows the absolute number of RSA applications and the proportion of those applications we were able to match into other datasets. In the regressions we only use data from 1988 to 2003.

**Figure A1: Matching unique and multiple matches**
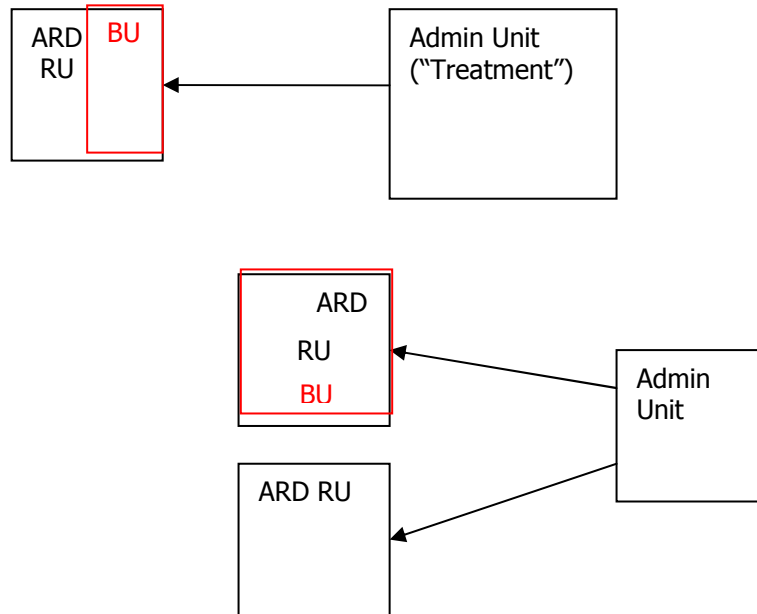
Notes: This shows heuristically the matching of the different datasets. See text for details