Department for Business Innovation & Skills

ANALYSIS

BIS ANALYSIS PAPER NUMBER 246

The Manufacturing Advisory Service (MAS) - Impact Analysis Methodology Study

FEBRUARY 2016

BIS Expert Peer Review for Evaluation

The BIS Expert Peer Review Group for Evaluation reviews all BIS impact evaluation publications and publications that make claims about the impact of policies, and provides an independent assessment of the methodological quality of the evaluation.

This publication was peer reviewed by Dr Maren Duvendack and Professor Henry Overman.

The peer reviewers' assessment can be found here:
www.gov.uk/government/publications/manufacturing-advisory-service-mas-impact-analysis

Contents

Executive Summary

The Manufacturing Advisory Service (MAS) was established in 2002 to provide support and advice locally to Small and Medium-sized Enterprises (SMEs). It was funded as a national scheme between 2012 and early 2016 and was managed independently by Grant Thornton. The main aims of the scheme were to support improvements in areas such as efficiency, strategy or innovation, and in some cases, following an independent review, awarded a small grant to applicants.

During 2014-15, BIS undertook a methodological study to assess the potential of using quasi-experimental data-matching methods to assess the economic benefits from the scheme. This involved comparing the growth of MAS clients to a matched comparison group of businesses that had not received MAS support. This analytical paper presents the findings of this internal BIS analysis and illustrates a potentially more robust and credible methodology for evaluating impact of business support programmes as it does not make use of self-reported assessment by grant recipients.

The methodological findings from this study will assist in helping to decide how a final economic evaluation of the scheme will be evaluated in the longer term. While the study found the methods were generally successful, following a peer review, some improvements to the method were recommended. Also, a full economic evaluation needs to be delayed and undertaken in a few years, when there has been sufficient time for the benefits from the support to materialise.

Aims of this paper

The aims of this analytical paper are:

- Establish a robust methodology for evaluating the average economic impact of MAS on its client businesses, making use of the best data sources available.

- Identify issues and unanswered questions and suggest avenues for future work.

Methodology

Previous impact evaluations have focused on self-reported assessments. This study examines quasi-experimental methods to find a robust counterfactual. The techniques employed are:

- *Data linking* – the database of MAS clients is linked to a comprehensive ONS database, covering variables such as turnover, employees and sector over the period 2010 to 2013[1]. This has allows a comparator group of businesses to be found that

[1] Note that there are timing issues with the update of the ONS IDBR data – which may be lagged for a few years, particularly for smaller businesses – this issue needs further investigation should the analysis be repeated in the future.

have not received MAS support and estimate a measure of each business's Gross Value Added (GVA) both **before** and **after** they have received MAS support.

- *Matching Methods* – matching methods are used to identify a group of businesses not in receipt of support (the "comparison group"), matched to a group of MAS clients (the "treatment group") on key parameters, such as sector, initial turnover and business birthdate. By assuming these parameters are important in determining economic performance and likelihood of receiving support, this matched group can be used as a counterfactual to assess the net economic impact.

- *Difference-in-difference (DiD)* – The impact of MAS support can then be estimated by calculating the difference in GVA growth over the period of interest between the treatment group and the control group. An adjustment also needs to be made for selection biases.

Conclusions and recommendations

The study presents a methodological framework for evaluating MAS that is a significant improvement on that used in previous evaluations, by identifying a counterfactual, estimating the scale of selection biases and avoiding using self-forecast or self-reported growth.

It is noted that the timing of undertaking this study for the reference year following the 2012 treatment period was relatively early – normally an evaluation will be undertaken some 3 to 5 years following an intervention.

There are also some uncertainties and methodological issues that potentially affect the conclusions of the analysis and need addressing in future evaluations. The main recommendations for improvements are:

- Validate or improve the methodology for estimating GVA using turnover data, particularly in terms of timing and availability of data for the required intervention period and the extensive use of imputation in the IDBR for smaller businesses;

- Repeat the quasi-experimental analysis annually to lengthen the time series and check for persistent impact;

- Consider and investigate self-selection bias further, where firms which are growing may be more likely to apply for MAS support;

- Investigate the implementation of a Randomised Controlled Trial for the allocation of grant funding to assess the scale of advisor-selection bias.

1. Introduction

1.1 Background

Originally set-up in 2002, MAS provided expert support to manufacturing SMEs. Manufacturers could benefit from a free review by a MAS advisor to identify key priority areas or access funding for improvement projects to increase efficiency, develop new products and boost sales.

MAS was administered by a private sector consortium of Grant Thornton, PERA, West Midlands Manufacturing Consortium and South West MAS (before 2012 this was done at a regional level through the RDAs). It was delivered on the ground by expert advisors, with relevant industrial experience, who diagnosed business needs, allocated direct funding for business improvement measures and connected firms to wider opportunities.

This analysis is focused on the time period after the scheme moved to a national delivery model, for 2012. This is mainly due to the fact that the process for collecting monitoring data changed at this point so there is no usable pre-2012 data available.

1.2 Eligibility

MAS offered funded support to SME manufacturers in England. An SME is a business with the following characteristics (defined by the EU[2]).

- Less than 250 employees.
- Turnover less than €50m or a Balance Sheet total of less than €43m.
- Not part of a group which in itself exceeds one of the criteria above.

The MAS guidelines gave a broad definition of the activities an applicant must undertake to be considered a manufacturer. All activities in the production cycle can be included, from research, design and development, production, logistics and service provision, to end of life management.

1.3 Levels of support

MAS services were categorised into five different levels (L1-L5), described briefly below:

L1 – Telephone Helpline and Email / Website contacts - undertaking a level of ‘triage’ to identify the most appropriate assistance for each business.

L2 – Manufacturing Reviews - identifying the interventions a business will undertake. This could be one-to-one (L4), one-to-many (L3) or brokered (L5) support and self-

[2] http://ec.europa.eu/enterprise/policies/sme/facts-figures-analysis/sme-definition/index_en.htm

implemented strategies. Some businesses undertaking a L2 review will not receive further MAS support.

L3 – Events - MAS organised a small number of events to deliver advice to multiple businesses simultaneously.

L4 – In-depth interventions - providing funding for in depth consultancy services, at three levels:

- The MAS Foundation Service - A contribution of up to £1,000 (or a maximum of 50%) to an improvement project. This will be targeted at businesses that need basic low level help.

- The MAS Step Change Service - Funding of up to £3,000 (or a maximum of 50%) towards a more significant improvement programme.

- The MAS Transformation Service - Funding of up to £10,000 for a strategic change to the business, available only in exceptional circumstances.

Foundation and Step Change projects could be committed by an advisor after completing a L2 review, subject to providing evidence that the target level of return of investment and jobs saved or created is achievable.

L5 – Active Referrals - brokering support from other Government schemes. This could happen at any stage in the MAS process, but particularly following an L2 review or the closing review of an L4 project. Where wider support was needed, the advisor referred the client to the most appropriate scheme.

Table 1 shows the percentage of businesses undertaking each level of MAS support.

Table 1 – Percentage of businesses receiving different levels of MAS support

	L1	L2 review	L2 diagnostic	L3	L4 project	L5 activity
% of businesses contacting MAS	89%	66%	2%	9%	22%	49%

Source: MAS monitoring data – Business Activity Tracker 2014

Although there are a number of different combinations of service a business could undergo, a very common pathway through the MAS system was a L2 review, followed by one or more L4 projects (often interspersed with additional L2 reviews). In this impact evaluation, we focus on this combination of L2 reviews and L4 projects, as these are the levels of support accounting for the majority of the MAS budget and also the levels thought to have the most significant impact on business performance. Focusing on these services also ensures that our sample sizes are large enough to allow us to draw statistically significant conclusions.

1.4 Logic model for MAS

Table 2 shows the proposed logic model for the MAS interventions described above. This provides a hypothesis for how inputs and intervention activities are translated into outputs, outcomes and impacts.

Table 2 – Logic model for MAS

Inputs	Activities	Outputs	Intermediate Outcomes	Impacts
		Narrative		
Government funding for MAS.	Business support interventions to manufacturing businesses.	Implementation of improvement projects.	Improvement projects increase competitiveness / access to growth opportunities.	Increased growth of UK manufacturing sector.
		Indicators		
Total BIS resources provided.	Reported number of MAS reviews and improvement projects.	Efficiency improvements, new strategies, access to finance, new supply chain links, etc.	New jobs, new capital expenditure, access to new markets, etc.	Increased employment and GVA in UK manufacturing industry.
		Data		
BIS finance data	MAS monitoring data.	MAS monitoring data. Qualitative evidence	MAS monitoring data. Qualitative evidence	National statistics. Inter-Departmental Business Register.

The aim of this analysis is to assess the scale of the observed impacts, relative to the known inputs, and use complementary qualitative evidence to judge the extent to which the observed impacts can be attributed directly to MAS interventions.

1.5 MAS monitoring data

The MAS monitoring database is managed by consultants, Grant Thornton. A data extract was provided to BIS analysts, covering key monitoring variables for all L2 and L4 support projects in 2012 and 2013. The dataset contains around 20,000 records and around 50 variables. It is populated with data recorded and supplied by MAS advisers during various stages of intervention. Summary statistics for the key variables in this dataset are provided below. The data used to carry out this analysis has been taken from this source unless otherwise stated.

Table 3 shows the numbers of L2 and L4 interventions and the amount of L4 grant funding provided in the four six-month periods between January 2012 and December 2013. The number of interventions has gradually increased as the national scheme has rolled out.

Table 3 – The numbers of L2 and L4 interventions and the amount of L4 grant funding over the period of interest

	Jan' 2012 - Jun' 2012	Jul' 2012 - Dec' 2012	Jan' 2013 - Jun' 2013	Jul' 2013 - Dec' 2013	Total
No. of L2 interventions	3,993	4,863	5,774	5,501	20,131
No. of L4 interventions	277	1,138	1,822	1,990	5,227
Total L4 grant	£410,741	£1,969,732	£3,025,905	£3,541,224	£8,947,602

Table 4 shows the regional distribution of L4 MAS projects. The distribution is reasonably uniform, with all regional shares except for two being between 10 to 13%. However, the South East has the highest share with 16% and North East the lowest share at only 4%.

Table 4 – Regional distribution of L4 MAS support

MAS Region	% of L4 projects by MAS Region
South East	16%
London	12%
Yorkshire & Humber	10%
South West	13%
East Midlands	10%
North West	11%
East of England	11%
West Midlands	13%
North East	4%

Table 5 shows the distribution of L4 MAS support by project type, as categorised on the MAS database. The distribution is highly skewed towards Strategy, Operational Improvement and Innovation projects. These categories are described below.

Table 5 – Distribution of L4 MAS support by project type

L4 support category	% of L4 projects by support category
Strategy	33%
Operational Improvement	39%
Supply Chain	4%
European Regional Development Funding (ERDF)	2%
Innovation (New Products/Processes)	20%
Expert Innovation	2%
Six Sigma	0%
Expert Finance	0%

Strategy – The advisor works with the business to develop a medium to long-term strategic manufacturing plan (usually between 3 and 7 years)

Operational Improvement / Six Sigma – The advisor suggests a package of measures to improve efficiency, often based on "lean" principles. "Six sigma" is a particular lean technique that uses statistical methods to improve consistency and efficiency. Other examples include, "value stream and process mapping", "improving layouts and space utilisation" and "improving quality and delivery".

Supply Chain – The advisor helps a business work effectively within a supply chain or supply base. For example, this could be by helping them adhere to some official set of standards or by mentoring a group of businesses within a specific supply chain.

European Regional Development Funding – Some projects can be supported by ERDF[3] and guided by MAS advisors.

Innovation / Expert Innovation – The advisor provides help to develop and introduce new products. For example, this could include advice on applying for intellectual property rights.

Expert Finance – The advisor helps the business overcome issues preventing them from getting access to finance.

Table 6 shows the distribution of L4 MAS clients by the number of repeat interventions. The majority of businesses undergo a single intervention with around a fifth receiving two interventions and a small number receiving more than that.

[3] https://www.gov.uk/government/policies/supporting-economic-growth-through-managing-the-european-regional-development-fund

Table 6 – Distribution of L4 MAS clients by the number of repeat interventions

No. of L4 interventions	Percentage of businesses contacting MAS
1	72.5%
2	20.1%
3	6.0%
4	1.2%
5	0.2%
6	0.1%

The subgroups outlined in these tables are used later in an attempt to assess with MAS clients benefit more from certain types or patterns of intervention. However, in general these subgroups contain too few individuals to draw statistically significant conclusions.

1.6 Supplementary qualitative evidence

Additional qualitative evidence has been gathered to assist with the interpretation of the econometric analysis described in this report. In particular, it is important to understand the role of the MAS advisor in selecting businesses for L4 support as it is likely that biases are introduced at this stage. We have reviewed the “MAS programme delivery manual”, which provides MAS advisors with a framework for the delivery of MAS services, and conducted a series of interviews with MAS advisors to gain a more detailed understanding of their decision making process. The findings are summarised below and discussed again in relation to the econometric analysis in section 4.

MAS programme delivery manual

The MAS programme delivery manual provides advisors with an overview of the scheme and services available to businesses. The document describes the different levels of support and covers some of the requirements businesses must satisfy to make the transition from a L2 review to a L4 project, and to receive repeated L4 support, but does not give a clear description of the decision making process undertaken by MAS advisors. However, advisors are offered training and induction to help them understand these processes.

The document gives guidance on how to record decisions on allocating MAS services and measure business progress following the MAS intervention. The manual implies that a range of factors must be considered in deciding to allocated L4 support to a business, usually relating to the business’s “potential for growth”, or “potential to benefit” from the support. The primary quantitative measure of success used by MAS advisors is the Gross Value Added (GVA) of the business, a measure of the business’s profits, which we define in more detail in section 3. The MAS assessment of GVA growth in the decision making process is broadly as follows:

1. As part of the L2 review, the client is asked provide its current GVA and to forecast its GVA the year after the intervention. The MAS advisors may provide some

assistance in estimating this forecast based on the L2 review. The ratio of the forecast GVA growth to the proposed L4 grant is then defined as the forecast return-on-investment (ROI). Note that this ROI uses the absolute change in GVA for an individual business and does not involve any assessment of a counterfactual scenario (i.e. what would have happened to the business in the absence of support).

2. To receive grant funding the forecast ROI must be above some target value. There is no explicit value given in the manual as the target appears to be context-specific (although the advisor interviews, covered in the next section, provided some typical values).

3. To receive repeat grant funding, the business must report back on its actual GVA growth following the initial L4 intervention. If this growth is consistent with the target ROI then further L4 grant funding may be provided.

The policy imperative is that MAS advisors do not pick businesses at random but pick businesses with high growth potential and that repeat-funding is contingent on businesses demonstrating actual high-growth potential and delivering the best ROI. However there is a risk that some of these businesses would be more likely to achieve growth in the absence of support.

In order to understand better the decision making process on allocating MAS funding we conducted telephone interviews with a small number of (seven) MAS advisors out of a total population of 85 nationally. In order to obtain a representative range of views, the advisors interviewed had a range of experience and came from a mixture of regions. The findings of these interviews are summarised in Annex 1.

1.7 Potential for bias

Decisions made by the business and advisors at different stages of the MAS process can introduce biases that could affect the results of this analysis.

- By choosing to call MAS, the business is subject to **self-selection bias**, which could mean that it is more proactive and therefore more likely to undertake activities to achieve growth than a similar business that did not seek Government support. This is covered in more detail in section 2.

- As described in section 1.6, part of the policy design is that MAS advisors choose which businesses will go on to receive a grant funded project, based on the business’s potential to benefit from the intervention and achieve high growth. Findings from a small number of advisor interviews suggest that there may be some deadweight associated with supporting businesses that report high-growth potential, engage most and are most willing to help. This could be termed as ‘advisor selection bias’ although there is no substantial empirical evidence yet to understand the scale or extent of this bias.

These biases are summarised in Table 7 below.

Table 7 –: Selection biases introduced at different stages of the business's interaction with MAS

Activity	Initial call	Decision on whether to provide grant funding	Decision on whether to provide follow-up grant funding
Bias	**Self-selection bias** - Business is proactive enough to seek Government support and may therefore be higher growth.	**Advisor-selection bias** - Business selected by MAS advisor as having potential for high growth.	**Further advisor-selection bias** - Business has to provide evidence of high growth for advisor to provide further grant funding.

.

2. Methodology

The dataset collected through the administration of MAS provides useful information for monitoring and evaluation. However, for an effective impact evaluation we require data on outcome variables from before and after the intervention. We also need data on a control group of businesses that have not received MAS services. This section explains how we have used data linking and quasi-experimental techniques to satisfy these requirements and provides information on the key variables used in this analysis.

2.1 Previous evaluations

The impacts of MAS have been evaluated on a number of occasions, most recently in 2007[4] and 2010[5] by consultants DTZ. In general, these evaluations have relied on *self-reported, self-forecasted* GVA and employment growth collected by the MAS advisors (see section 1.6). Using this data to assess impact may have limitations such as:

1. The GVA and employment growth recorded is *overall (gross)* GVA and employment growth, rather than the *additional* growth that can be attributed to MAS. In impact evaluation, it is this additional component of growth we seek to estimate.

2. The forecasted nature of the data means that it is both uncertain and likely to be subject to optimism bias, with businesses overestimating the extent of their future success.

3. The self-reported nature of the data means that it could be unreliable. Businesses may use a non-standard definition of GVA or be using accounts that have not been checked or audited[6]

For these reasons we have tried a complementary approach based on a quasi-experimental design applied to a longitudinal dataset of national statistics. This approach does go some way to addressing the first two problems described above, but there are a number of methodological issues that could affect still the results, which we describe later in the chapter.

2.2 The evaluation problem

In order to assess impact of a policy we need to establish a counterfactual – the outcome that would have occurred in the absence of the intervention. This is typically achieved by

---

[4] DTZ Consulting & Research, 2007, "Evaluation of the Manufacturing Advisory Service"

[5] DTZ Consulting & Research, 2010, "Review of the Manufacturing Advisory Service and Research to Support the Business Case for Continuing and Developing the Manufacturing Advisory Service"

[6] All of these limitations have now been addressed in the latest surveys for Business Growth Service

identifying a control group to compare with the group of individuals (businesses) in receipt of the policy treatment (MAS support). The Government publication, “The Magenta Book – Guidance for evaluation” states that a well-designed randomised, controlled trial (RCT) is the strongest research design available, as it is most likely to control for unobserved variables that could introduce bias[7]. However, time, resource and policy-delivery constraints can often make RCTs unattractive or difficult to implement effectively. In such cases, a well-designed quasi-experimental analysis can be more effective than a weak RCT.

Here we present a quasi-experimental methodology, based on matching methods and difference-in-difference (DiD) analysis. This general approach and its advantages and pitfalls have been widely documented in the academic literature.[8,9] In sections 4 and 5 we discuss the pros and cons of using a quasi-experimental design as opposed to an RCT and make recommendations for future evaluations.

Constructing a counterfactual by matching

MAS client businesses have different characteristics from the business population. This partly manifests itself in observable characteristics, such as sector, turnover and the number of employees (the MAS eligibility criteria dictate this to some extent). Furthermore, the MAS scheme is subject to selection-bias, i.e. MAS client businesses are also different from the business population in that they actively contact MAS themselves for support, i.e. they are “self-selecting”. Selection-bias is of concern for impact evaluation when there are factors that both affect the likelihood of receiving support and outcome variables.

In this analysis the main outcome variables of interest is the business GVA (we also look briefly at employment growth but do not cover this in depth). Many of the observable differences (sector, turnover, etc.) are likely to influence these outcome variables and should be controlled for. This is usually achieved using matching methods, such as direct nearest-neighbour matching, propensity score matching or coarsened exact matching (described in section 2.5). Using one of these statistical methods, a group of businesses, who we know have not received the intervention, is identified based a series of covariates, thought to influence both the likelihood of receiving support and the outcome variables.

This will control for observable differences between the treated businesses and the business population, but it does not guarantee **internal validity** as there is a chance that unobserved differences are influencing any observed impact.

2.2 Datasets

Two separate datasets are used in the impact analysis described below, the MAS monitoring data extract described in the introduction section and a second, longitudinal

---

[7] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/190984/Magenta_Book_quality_in_policy_impact_evaluation__QPIE_.pdf

[8] J. Payne, C. Payne, S. Lissenburgh and M. Range, 1999 - “Work-Based Training and Job Prospects for the Unemployed: An Evaluation of Training for Work” - Policy Studies Institute

[9] J. Månsson and B. Widerstedt, 2012 - "The Swedish Business Development Program - Evaluation and some methodological and practical notes" - ERSA conference papers ersa12 p858, European Regional Science Association

dataset which we have created by combining four years of data from the Inter-Departmental Business Register (IDBR)[10]. The IDBR is a comprehensive list of UK businesses used by government for national statistics, covering all sectors of the UK economy, other than some very small businesses and some non-profit making organisations. Comparable official datasets, such as the Annual Business Survey (ABS)[11] or the Business Register and Employment Survey (BRES)[12], are useful in that they contain a richer set of variables, but are surveys and are therefore not comprehensive and so not well suited to data linking exercises.

Key variables

We split variables into three different categories depending on their use in the analysis:

Linking variables – These variables are identifiers, such as name or postcode, which can be used to join records in the MAS monitoring data extract and the IDBR.

Treatment variables – These variables give details of MAS funding received by each business, such as the date and level of funding for each project.

Matching variables – These variables are business characteristics, thought to determine both a business's propensity to receive MAS support and its economic performance.

Outcome variables – These variables are used to assess the change in a business's performance over the period of interest.

Other variables – These variables are used to investigate whether observed impacts are different for different subgroups.

Table 8 highlights the key variables within the MAS monitoring data extract used in the analysis.

Table 8 –Key variables in the MAS monitoring dataset

Variable use	Variable names
Linking variables	• Business Name • Address • Postcode • Business registration number (where available)

[10] http://www.ons.gov.uk/ons/about-ons/products-and-services/idbr/index.html

[11] http://www.ons.gov.uk/ons/guide-method/method-quality/specific/business-and-energy/annual-business-survey/index.html

[12] http://www.ons.gov.uk/ons/guide-method/method-quality/specific/labour-market/business-register-and-employment-survey--bres-/index.html

Treatment variables	• Any MAS contact (L1, L2, L5, calls, emails, networking contacts) • Number of L2 activities in each of the four six month periods between January 2012 and December 2013. • Number of L4 activities in each of the four six month periods between January 2012 and December 2013. • L4 grant funding paid in each of the four six month periods between January 2012 and December 2013.
Other variables	• MAS Region – Geographical region within which MAS Support delivered. • L4 Project Type - The nature of the MAS project carried out with the client, selected from the following list: Operational Improvement (Productivity), Strategy (Business Strategy), Innovation (New Products/Processes), Expert Finance, Expert Innovation, Supply Chain • L5 activity

The longitudinal IDBR dataset contains over 2 million records and 26 key variables. The variable lists for the IDBR dataset is shown in Table 9.

Table 9 –Key variables in the IDBR extract

Variable use	Variable names
Linking variables	Enterprise reference number Business name Address Postcode Business registration number (where available)
Matching variables	• Birth date – Business start-up date. • Standard Industrial Classification (UK SIC 2007, 5 digit) • Employees, December 2011 • Turnover, December 2011
Outcome variables	• Difference in number of employees between December 2011 and December 2013. • Difference in turnover between December 2011 and December 2013.

For large and medium-sized enterprises, which have been selected to take part in the Annual Business Survey, turnover and employment data are collected once a year, and then inputted into the IDBR in September. For enterprises which have not been selected in the ABS, but are large enough to be VAT registered, the data is sourced from HMRC VAT data, which is again inputted into the IDBR in September. If an enterprise is only PAYE registered then turnover is imputed on the basis of employment multiplied by an average turnover per-head for the sector, based on enterprises with a valid turnover source and if possible of similar size. This is calculated and inputted in November.

This has implications when considering the length of time it takes for a MAS intervention to have an impact on a business's performance. For example, the turnover in the December 2013 IDBR will typically have been reported in September 2013 for 2012 - most likely for the calendar year or in some cases the 2012/13 financial year. So an intervention in 2013 will not have IDBR turnover data available for the 2013 reference year until at least the September 2014 IDBR.

For many smaller businesses, the turnover data may have never been reported and so is imputed, or could have been reported some years earlier and never been updated. Where turnover is imputed, this is from the sector's average turnover per employee derived from the (potentially) larger companies that report turnover. However, these companies are not necessarily representative of the activity of smaller companies.

At the time of undertaking this analysis, the timing and imputation issues of the IDBR turnover data was not fully understood and BIS has been working with ONS to gain a better understanding of these issues. These will need to be investigated further and more fully taken into account in any future evaluation studies.

In this analysis we mainly examine the impact of interventions between January 2012 and July 2012 on the outcome variables in December 2013. That means there will be contributions to the observed outcome variable from anywhere between 6 and 24 months after the intervention. So, we are assuming that an intervention has had, on average, roughly 15 months to have an impact.

Business Sector

The Standard Industrial Classification[13] (UK 2007 version) is commonly used to define the sector in which a business operates. SIC codes are arranged in a hierarchical structure, with the first two digits defining a broad sector Division and subsequent digits giving increasing levels of granularity, referring to more specific sets of products and processes. Table 10 demonstrates this hierarchy for one particular 5 digit SIC code - "Manufacture of corrugated paper and paperboard; manufacture of sacks and bags of paper", showing the parent "Class", "Group" and "Division" codes that sit above it.

Table 10 – Hierarchy of SIC codes.

	SIC code	Description
Division (2 digit)	17	Manufacture of paper and paper products
Group (3 digit)	17.2	Manufacture of articles of paper and paperboard
Class (4 digit)	17.21	Manufacture of corrugated paper and paperboard and of containers of paper and paperboard
Subclass (5 digit)	17.21/1	Manufacture of corrugated paper and paperboard; manufacture of sacks and bags of paper

The IDBR includes 5 digit SIC code information for each business, but in this analysis we primarily use the 3 digit, "Group" layer of the SIC hierarchy. This is for methodological reasons, discussed in more detail in section 2.5.

[13] http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/standard-industrial-classification/index.html

Gross Value Added

When considering benefits to the economy due to public expenditure, the Government's preferred outcome variable is the Gross Value Added (GVA). The basic definition of GVA is:

GVA = net profit before tax + depreciation + amortisation + payroll costs

In its analysis of the Annual Business Survey data, the Office for National Statistic (ONS) calculates an approximate value of GVA using the following general formula[14]:

GVA = output at basic prices – intermediate consumption

As the IDBR does not include GVA we propose a simple method for estimating GVA from turnover using the Annual Business Survey (ABS). Using the ONS's formula above, we assume that within a given sector, each business's output and therefore its turnover is roughly proportional its GVA.

From the ABS, we take the average ratio of GVA to turnover for each **4 digit sector** over the period 2010-2012 (more recent data is not currently available). We then estimate the GVA for each business in each year by multiplying the turnover in by the sector-specific GVA to turnover ratio.

This calculation is based on the assumption that selling similar amounts of similar products (i.e. within a sector) will lead to a similar level of value added to the economy. Although this assumption might not hold for individual businesses, on-aggregate it should give a reasonable estimate of the average GVA of a group of several hundred or thousand businesses (potential problems with this method are discussed in section 4).

As a check on this method, a GVA validation exercise is currently being carried out for a sample of MAS clients. Unfortunately, this audited GVA data was not collected in time to inform this impact analysis but should be revisited the next time MAS is evaluated.

2.3 Data linking

We have linked the MAS dataset to the IDBR using Company Reference Numbers (CRN), business name, address and postcode. Manual checking was carried out to validate partial matches. The final linking rates for different groups of businesses, receiving different levels of MAS support, are shown in Table 11.

[14] http://www.ons.gov.uk/ons/guide-method/method-quality/specific/business-and-energy/annual-business-survey/quality-and-methods/abs-technical-report---june-2014.pdf

Table 11 – Linking rate for businesses receiving different levels of MAS support

	Any MAS contact	L1	L2 Review	L2 Diagnostic	L3	L4 activity
No. of businesses	21228	18953	14101	400	2005	4659
No. of linked businesses	15178	14073	11009	317	1329	3878
Linking rate	71.5%	74.3%	78.1%	79.3%	66.3%	83.2%

As expected the linking rate is low for businesses that have contacted MAS but not received MAS support, mainly because the records are often incomplete for these businesses. The linking rate improves for L2 and L4 support. This may be due to improved data validation for businesses that have a longer term relationship with MAS.

Table 12 and Table 13 show the distribution of employees and turnover for linked businesses and the overall population of MAS clients. In general there is a low level of bias, except for very small businesses where a smaller percentage is linked. This is to be expected, given that very small businesses a less likely to have a CRN or may not have a record in the IDBR.

Table 12 – Distribution of employees for linked businesses and the overall population

Employees	% Linked	% Overall
0-4	27%	33%
5-9	18%	17%
10-24	27%	24%
25-249	28%	25%
250+	1%	1%

Table 13 – Distribution of turnover for linked businesses and the overall population

Turnover (£'000s)	% Linked	% Overall
0-249	26%	32%
250-999	27%	25%
1,000-2,499	22%	20%
2,500-24,999	24%	21%
25,000+	2%	1%

This bias should not affect the impact analysis results significantly but may mean that we should be cautious in drawing any specific conclusions on the impacts on very small businesses.

2.4 Characteristics of MAS clients

As discussed above, the characteristics of MAS clients are typically different from the population of businesses in the UK – partly because the MAS eligibility criteria require clients to be SME manufacturers and partly because self-selection bias may skew the characteristics of MAS clients.

Here we compare observable characteristics – turnover, employees, sector and birthdate – for the population of MAS clients under consideration (all successfully linked businesses that have contacted MAS between January 2012 and December 2013) and the population of businesses covered by the IDBR. As the differences are large, we have opted to present this visually, rather than detailing the quantitative differences between the populations. We use Stata "Kernel density"[15] plots to display the distributions for the different populations.

Figure 1 shows the distribution of turnover for the two groups.

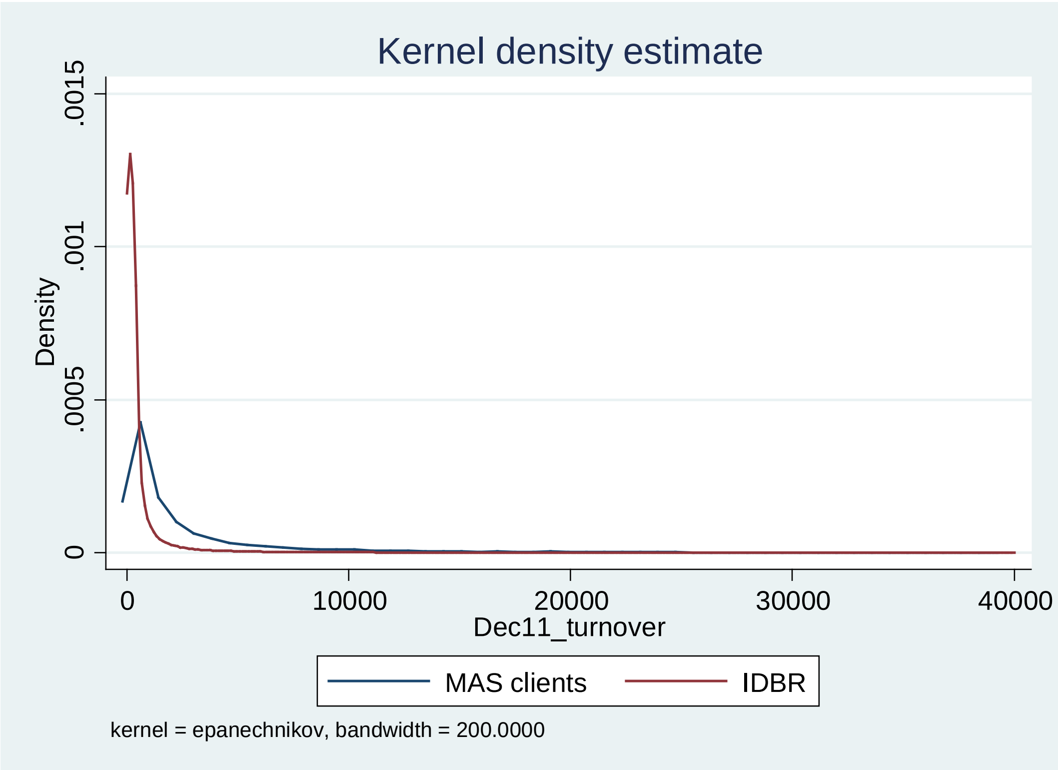

Figure 1 – Turnover distribution for MAS clients and the IDBR population

Consistent with previous BIS analysis[16], the turnover distribution is highly skewed towards business with turnover less than £2m, with a long tail of larger businesses. On the other hand, the population of MAS clients has far fewer very low-turnover businesses and a

[15] http://www.stata.com/manuals13/rkdensity.pdf

[16] https://www.gov.uk/government/publications/analysis-of-the-distribution-of-private-sector-enterprises-by-turnover

more significant proportion of medium-sized businesses with turnover between £2m and £10m. Only SMEs are eligible for MAS grant funding but there are a small number of higher turnover MAS clients in the dataset who self-fund MAS consultancy services.

Figure 2 shows the distribution of the number of employees for the two groups.

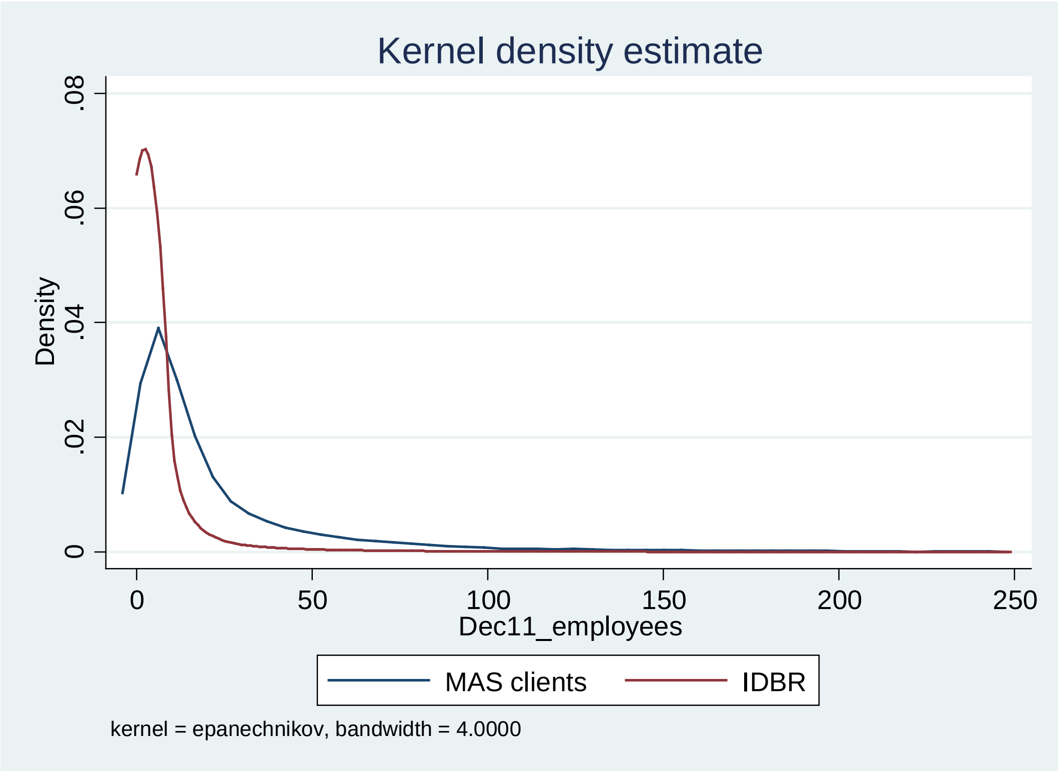

Figure 2 - Employee distribution for MAS clients and the IDBR population

The employee distributions are similar in shape to the turnover distributions, with the MAS client population containing a smaller proportion of very small businesses and a larger proportion of medium-sized businesses, with between 10 and 100 employees.

Figure 3 shows the distribution of sector for the two groups, using 3 digit SIC codes.

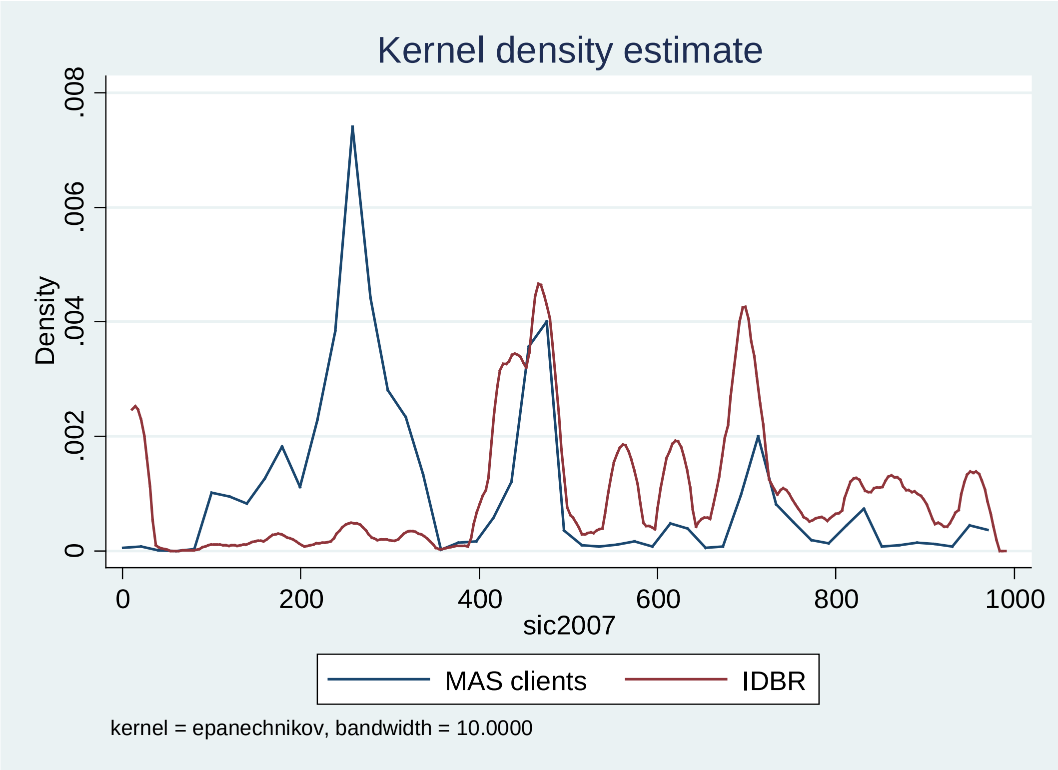

Figure 3 - Sector distribution for MAS clients and the IDBR population

Table 14 defines broad business sectors relative to specific 3 digit SIC code ranges, so that the peaks in figure 3 can be identified.

Table 14 – Broad business sectors by 3 digit SIC code

3 digit SIC code range	Broad sector
011-099	Agriculture and mining
101-332	Manufacturing
351-390	Utilities
411-439	Construction
451-532	Vehicles and transport
551-990	Various service sectors

The most obvious difference is that there are significantly more manufacturing businesses in the MAS client population, as expected. The construction, vehicles and transport sectors are reasonably well represented in both populations. There are fewer MAS client businesses in the service sector, although there is a peak around 700 under the “Professional, Scientific and Technical Activities” categories, which relates to services in direct support of manufacturing activities.

Figure 4 shows the distribution of birthdate for the two groups. The date variable has been converted to a numerical variable, with zero being 01/01/1960 and 1 unit being a day.

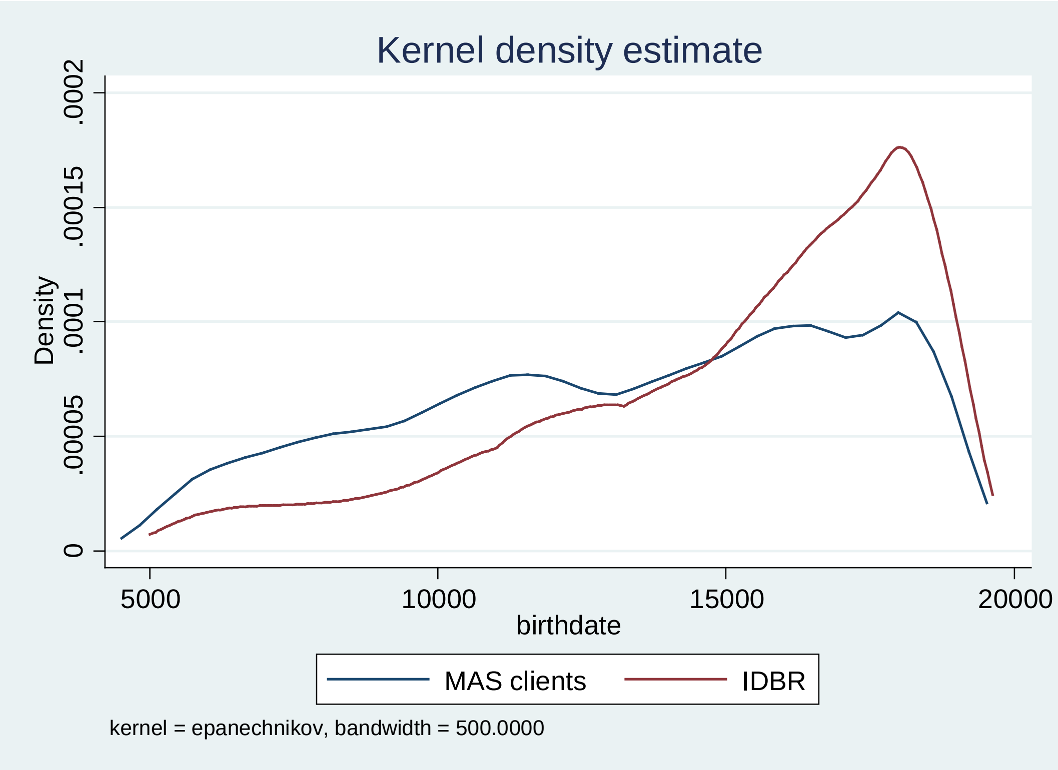

Figure 4 - Birthdate distribution for MAS clients and the IDBR population

The IDBR business population contains more very young businesses, with a large peak in around 2009 (roughly 18,000 on the x-axis of figure 4). This suggests that MAS clients are more likely to have been established for a longer period of time. Recent recessions can be observed as dips in the number of businesses in the early 1990s and around 2008 (roughly 13,000 and 17,500 on the x-axis of figure 4).

These large differences between the MAS client population and the IDBR business population demonstrate that these characteristics play a role in determining the likelihood of businesses contacting MAS for support, and help to make the case for using matching methods to find a comparator group.

2.5 Matching methods

Matching methods are typically applied to a dataset in which a subset of individuals have received some “treatment” in order to identify a second “control” group of individuals who have not received the treatment but are matched on a series of characteristics, known as covariates. There are three widely used matching methods that can be considered for this type of analysis:

- *Nearest-neighbour matching (NNM)* – This method is the simplest approach. Each business is compared to every other business simultaneously across each covariate. A distance metric (e.g. “Euclidean”) is used to find the closest untreated individual to each treated individual. The result is a group of paired businesses, closely matched on all covariates simultaneously, which can then be directly compared. One key disadvantage of this approach is that it is very computationally intensive for more than two or three covariates.

- *Propensity-score matching (PSM)* – This method was developed in order to match using a large number of covariates. The technique traditionally uses a simple probit or logistic regression model to estimate the propensity score - the probability of receiving the treatment – for each individual. Matches are then identified by finding pairs of treated and untreated individuals with similar propensity scores. This does not guarantee that they are well matched on each covariate. The extent to which PSM provides closely matched pairs is known as “balance”. Such regression models are more likely to provide a balanced set of matches when there is a clear, linear separation of the treated individuals in the “covariate space”. When this is not the case, the model may have to be modified to include nonlinear terms. In the extreme case that treated individuals are “clumped” into separate groups, literature suggests using “machine learning” techniques to arrive at a more sophisticated treatment model[17]. Alternatively, coarsened-exact matching, described below, offers a simpler solution to such complications.

- *Coarsened-exact matching (CEM)* – This method works by coarsening the covariate space into a limited number of bins and then assigning individuals to bins. Each treated individual is then matched at random to an untreated individual within the same bin. There are a number of advantages to this method, such as avoiding the need to choose an approximate model and improving the chance of achieving balanced matches – i.e. closer to full NNM[18]. CEM can also be used in conjunction with NNM by acting as an initial filter to restrict the population to the “region of common support”, i.e. ignoring any areas of the covariate space in which there are no treated individuals[19].

We conducted a series of initial analyses trialling these different methods on the linked MAS-IDBR dataset. The findings of this analysis and reasons for selecting the chosen approach are outlined below:

- Taking advantage of the small number of covariates available, we devised a way of implementing NNM that led to well-balanced matches but did not take an excessive amount of computational time. This involved splitting the dataset into a series of subsets according to 3 digit SIC code sector and carrying out NNM on turnover, employees and birthdate within that subset. Using 3 digit, as opposed to 4 or 5 digit, was found to be a compromise between sufficient sector detail and having subsets that were large enough to sample from. **This is the preferred method**, as NNM typically gives much closer matches than PSM or CEM. However, as PSM and CEM are widely used in the literature and might prove a useful sensitivity check, we decided to investigate whether these techniques could be used on the dataset.

[17] B. K. Lee, J. Lessler, and E. A. Stuart, 2010, “Improving propensity score weighting using machine learning” - Statist. Med., 29: 337–346. doi: 10.1002/sim.3782

[18] G. King, 2011, “Comparative Effectiveness of Matching Methods for Causal Inference” - http://gking.harvard.edu/publications/comparative-effectiveness-matching-methods-causal-inference

[19] M. Blackwell, S. Iacus, G. King, 2009, "cem: Coarsened exact matching in Stata" - The Stata Journal 9, Number 4, pp. 524–546

- Analysis of the matching covariates demonstrated that they are not well suited to the traditional probit or logistic regression model used in PSM. For example, the SIC code sector covariate is particularly “clumpy”, with several distinct sector groupings (see figure 3). Other covariates are not as clumpy but do not appear to be well-described by simple logistic regression models. For this reason, we had no success in running a PSM matching analysis to produce balanced matches and the technique was abandoned.

- We found that CEM could be applied relatively easily, but the matches were generally not as close as matches obtained with NNM. We also experimented with using CEM to restrict the analysis to the region of common support before running a NNM analysis. This was found to speed up computation time considerably (by filtering out the very large businesses in the IDBR) and was therefore implemented alongside NNM.

The specific methodology used for each separate analysis is broken down in steps in section 2.8.

2.6 Difference-in-difference

Difference-in-difference (DiD) analysis is a simple technique that can be applied to two distinct time series datasets. The technique is widely used to control for changes in the outcome variable that would have occurred equally for both treated and control individuals. In essence we find the difference in the value of the outcome variable **before** the intervention for treatment and control groups and subtract it from the difference the value of the outcome variable **after** the intervention for treatment and control groups. For businesses, this technique is intended to control for changes in the background economy. This is particularly powerful when combined with matching methods, as we expect well-matched groups of businesses to respond to background changes in very similar ways.

2.7 Software tools

The methodology described above has been implemented in a number of Stata routines. We use a series of routines built-in to Stata 12, but also two user-written routines to implement the matching methods, which have been well documented and peer reviewed. These routines are references and described below.

cem[20] - Implements CEM to restrict to the region of common support.

nnmatch[21] **-** Implements NNM to find well-matched control group and estimate average treatment effects. This routine matches with replacement so that the same control business can be picked as a match for multiple treated businesses.

[20] http://gking.harvard.edu/cem/

[21] http://ideas.repec.org/c/boc/bocode/s439701.html

2.8 Experiments

We have split the analysis into four separate experiments (or quasi-experiments) involving different treatment and control groups. Experiments A and B both involve matching to businesses that have never contacted MAS, which means that the entire IDBR is used as the pool from which to match. For these experiments we split the dataset into sector subsets to reduce the amount of computational time required (as described in section 2.5). Experiments C and D involve matching to businesses that have contact MAS, which reduces the size of the pool considerably and means that it is not necessary to split the dataset.

Experiments A and B have larger populations and very low numbers of duplicate matches in the control group compared to Experiments C and D, which means the results generally have a much lower variance and conclusions tend to be statistically significant. For this reason we focus our interpretation on the results of Experiments A and B and include Experiments C and D mainly for illustrative purposes. Their results are also used to provide a sense-check for the results of Experiments A and B.

In each case, the outcome variables and matching covariates are as follows:

Outcome variables

- Change in GVA between December 2013 and December 2011 (as discussed in section 2.2 this gives around 15 months for the average intervention in the first 6 months of 2012 to take effect)

- Change in employees between December 2013 and December 2011 (again this gives around 15 months for an effect to be observed)

Matching covariates

- Initial turnover (in December 2011)

- Initial employees (in December 2011)

- Sector (3 digit SIC code)

- Birthdate

The full details of each experiment and the methodologies for the two pairs of experiments are provided in Annex 2 and summarised below.

Experiment A – L4 vs No MAS (886 pairs of businesses)

In Experiment A we compare businesses that receive a MAS L4 grant funded project and matched businesses who have never contacted MAS. We therefore expect the results to be influenced by self-selection bias, advisor-selection bias (selection for a L4 project) and the actual impact of the MAS support.

Treatment group – Businesses receiving MAS L4 support between 01/01/12 and 31/12/12.

Control group – Matched businesses who have never contacted MAS.

Experiment B – L2 vs No MAS (2922 pairs of businesses)

In Experiment B we compare businesses that receive a MAS L2 review, but no L4 grant, and matched businesses who have never contacted MAS. Self-selection bias will affect the results of this experiment but advisor-selection bias will not, as the treatment group are not selected for a L4 project. We assume that the L2 review does not have any direct impact on GVA (although this assumption is somewhat uncertain). For this reason we propose to treat the entire GVA DiD in this experiment as self-selection bias, which allows us to isolate the effect of this bias and use the results to analyse the other experiments.

Treatment group – Businesses receiving MAS L2 support between 01/01/12 and 31/12/12, but no L4 support in 2012 or 2013

Control group – Matched businesses who have never contacted MAS

Experiment C – L4 vs L2 (597 pairs of businesses)

In Experiment C we compare businesses that receive a MAS L4 grant funded project and matched businesses who receive a MAS L2 review, but no L4 grant. As both treatment and control groups are subject to self-selection bias, this bias should not influence the GVA DiD. We expect the results to be influenced by advisor-selection and the actual impact of the MAS support. By looking at the results of this experiment alongside those of Experiment A, we can carry out a check on the impact of self-selection bias estimated in Experiment B.

Treatment group – Businesses with MAS L4 support between 01/01/12 and 31/12/12, but not afterwards

Control group – Businesses with MAS L2 service between 01/01/13 and 31/12/14, but no L4 service in 2012 or 2013

Experiment D – L4 vs later L4 (597 pairs of businesses)

In Experiment D we compare businesses that receive a MAS L4 grant funded project in 2012 and matched businesses who receive a MAS L4 grant funded project in 2013. This is a "phased" approach, where both treatment and control groups receive the same intervention but at different times. In theory this experiment is very powerful as it controls for both self-selection bias and advisor-selection bias, leaving only the actual impact of the MAS support. However, additional bias could be introduced due to the fact that businesses have chosen to seek MAS support at a particular time. For example, businesses may contact MAS following a change of leadership or direction of strategy or at a time when they are "ready to grow" – i.e. have other measures/investments in place to coincide with MAS funding. This type of timing bias means that we might expect to see higher growth for the earlier treatment group in this experiment.

Treatment group – Businesses with MAS L4 support with MAS L4 service between 01/01/12 and 31/12/12 but not afterwards

Control group – Businesses with MAS L4 support with MAS L4 service between 01/01/13 and 31/12/13 but not before

Matching with replacement – variance adjustment

The nnmatch routine carries out matching with replacement. If the population is small, it is more likely that a significant number of duplicates will be matched into the control group, and appropriate adjustments to the variance need to be made.

We found that, for Experiments A and B, where the matching population was all of the businesses in the IDBR, there was very little replacement (less than 1% of the control group), so we have assumed that we can use a standard paired T-test to compare distributions without any variance adjustment. The advantage of this is it is much easier to implement when splitting and recombining the sector datasets.

For Experiments C and D, where the matching population was much smaller, there was more significant replacement (around 10-15% of the control group), so we have used the “average treatment effect” tools in the nnmatch routine[19], which automatically makes the appropriate variance adjustment.

2.9 Sensitivity analysis

Sections 2.5 - 2.8 outlined the primary methodology used in this impact evaluation, but alternative analyses have been carried out in parallel as a sensitivity or sense check:

- *Removing covariates from matching* – We have run repeated Experiments A - D leaving out different covariates, such as initial turnover, from the matching routine. We do not present the results of this analysis here, but the impacts were found to be broadly similar to those found using the full set of covariates.

- *Regression analysis* – We have also carried out equivalent regression analyses for Experiments A - D. We use GVA growth as the dependent variable and our independent variables are initial turnover, initial employees, sector (1000 0/1 dummy variables), birthdate and treatment (0/1 dummy variable). The regression coefficient of the treatment variable then gives an estimate of the GVA DiD between the treatment and control groups, which can be compared to the estimate obtained through matching as a sense check of the results. The main results of these analyses are presented alongside the matching results in section 3.

3. Results

This section covers the results of the four experiments described in section 2. We examine time series of the outcome variables for treatment and control groups in each case and highlight the key differences between these groups of businesses. We carry out distributional analysis for the grant funded businesses in Experiment A, to investigate whether different groups of businesses benefit more than others. This section also includes a discussion of statistical significance and some sensitivity analysis. The results are interpreted in section 4.

3.1 Experiment A – L4 vs No MAS

As outlined in section 2.8, the aim of Experiment A is to find the GVA difference during the period between December 2013 and December 2011 for businesses that had received a MAS L4 grant funded project and matched businesses who had not contacted MAS.

This experiment provides the main difference-in-difference results for businesses receiving MAS grant funding (recall that we focus on this experiment due to the larger sample size and small number of duplicate matches in the control group, leading to results that have a relatively low variance and are generally statistically significant). For this reason, we also include a distributional analysis for this experiment, looking at the GVA DiD results for different subgroups of businesses. This additional analysis is not provided for the other experiments as they are largely included in order to estimate the impact of selection biases.

Matching

The first stage of the analysis is to identify a control group or a matched comparison using the matching methods described in section 2.5. In order to validate the results of any comparisons between treatment and comparison groups we must first check the quality of our matches on the different covariates. We have done this by carrying out two-sample Kolmogorov-Smirnov (KS) tests, and plotting distributions for treated and untreated matches as a visual check.

Figure 5 shows distributions of the initial turnover covariate for the treated and untreated, matched groups. The KS test for similarity of distributions gives a combined KS P value of 94.7%.

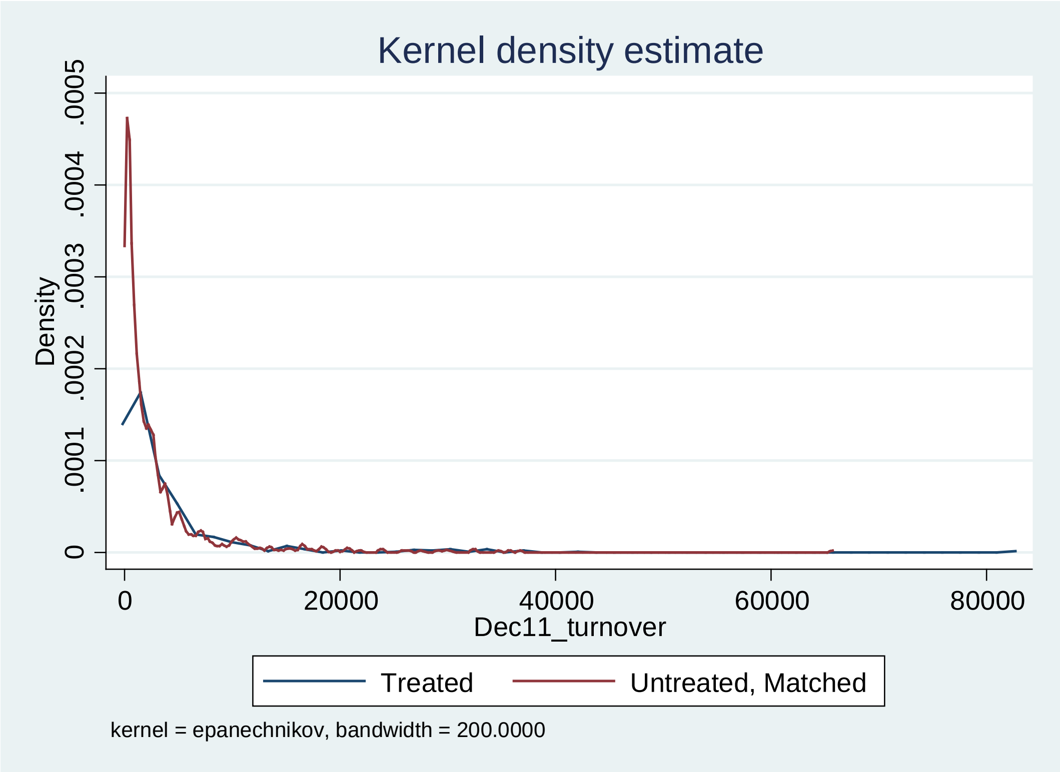

Figure 5 – Turnover distributions for treatment and control groups

Figure 6 shows distributions of the initial number of employees for the treated and untreated, matched groups. The KS test for similarity of distributions gives a combined KS P value of 94.8%.

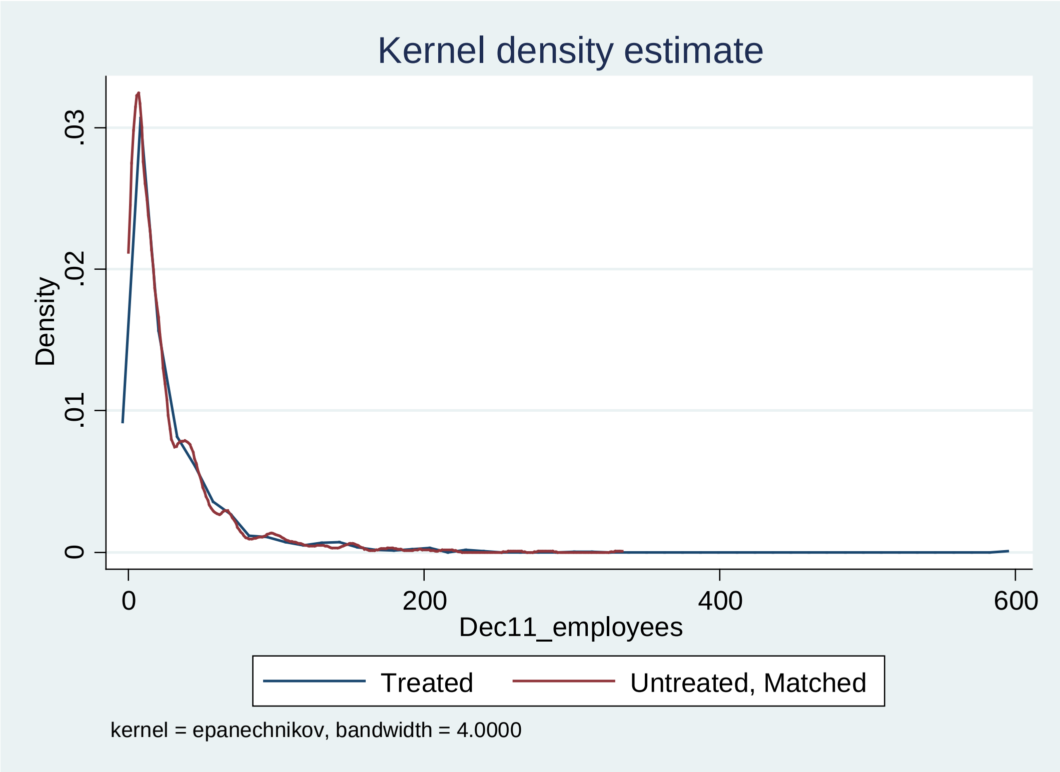

Figure 6 - Employee distributions for treatment and control groups

Figure 7 shows distributions of the initial birthdate covariate for the treated and untreated, matched groups. The KS test for similarity of distributions gives a combined KS P value of 99.9%.

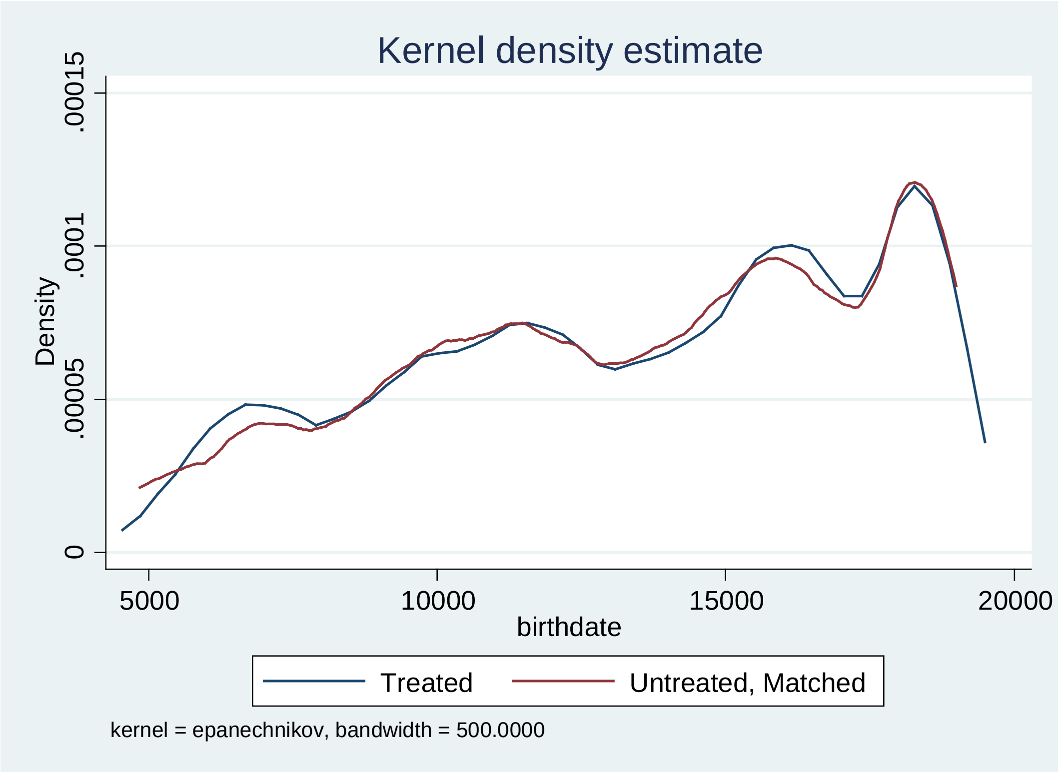

Figure 7- Birthdate distributions for treatment and control groups

For the sector covariate we have matched exactly on 3 digit SIC code. The SIC code variable is difficult to classify. It lies somewhere between a categorical and an ordinal variable because there are similarities between some, but not all, pairs of neighbouring codes. Also, the closeness of these similarities is somewhat subjective. Tables 15 and 16 give examples of 4 digit subsectors within a 3 digit group to provide a sense of the types of subsectors we are allowing to be matched up.

Table 15 – 4 digit SIC codes within "Manufacture of glass and glass products"

23.1	**Manufacture of glass and glass products**
23.11	Manufacture of flat glass
23.12	Shaping and processing of flat glass
23.13	Manufacture of hollow glass
23.14	Manufacture of glass fibres
23.19	Manufacture and processing of other glass, including technical glassware

Table 16 – 4 digit SIC codes within "Manufacture of other fabricated metal products"

25.9	**Manufacture of other fabricated metal products**
25.91	Manufacture of steel drums and similar containers
25.92	Manufacture of light metal packaging
25.93	Manufacture of wire products, chain and springs
25.94	Manufacture of fasteners and screw machine products
25.99	Manufacture of other fabricated metal products n.e.c.

These tables provide qualitative evidence for the quality of the sector matches, which must be judged somewhat subjectively, as opposed to using a quantitative KS test.

GVA analysis

After carrying out the matching analysis, we can examine the time series of the outcome variables for the treatment and control groups. Figure 8 shows time series of the mean GVA for the treatment and control groups between December 2010 and December 2013.

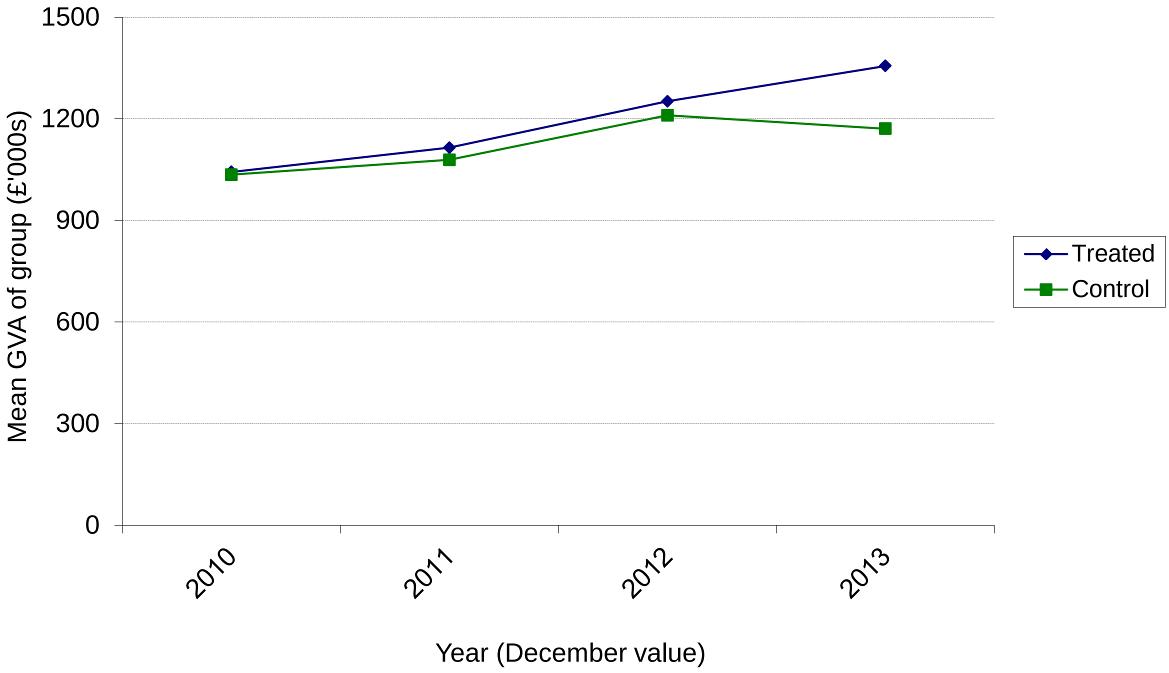

Figure 8 - GVA time series for treatment and control groups.

The two groups appear to be closely matched up until 2012, at which point the mean GVA values diverge, with a DiD between 2013 and 2011 of around £150,000. However, we cannot ascribe the whole difference to the MAS intervention as we need to take account of selection bias, covered in section 4.

Companies in the comparison group were more likely to go out of business. Figure 9 shows that by December 2013, nearly 14% of the original comparison group had gone out of business whereas only 4% of the treatment group had gone out of business over the same period. This factor could be correlated both with firms' growth potential and a potential positive impact of MAS support.

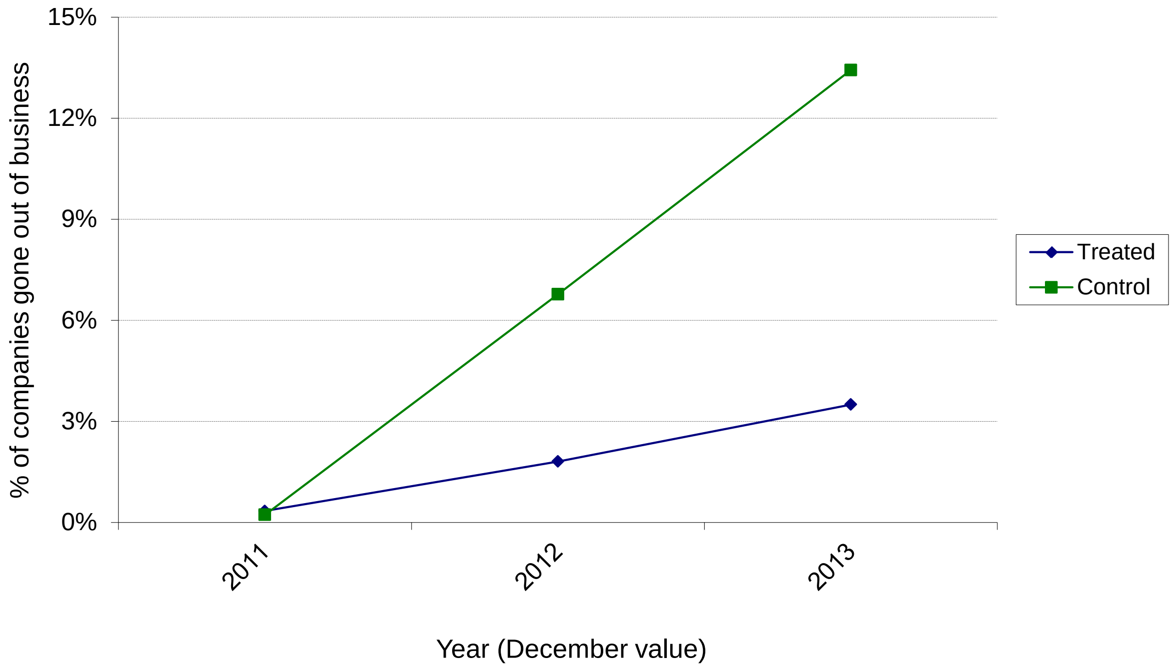

Figure 9 - Percentage of companies that ceased trading in treated and control groups

Figure 10 shows the time series of average GVA of treatment and control groups if companies who go out of business are excluded. There is still a GVA DiD of around £90,000.

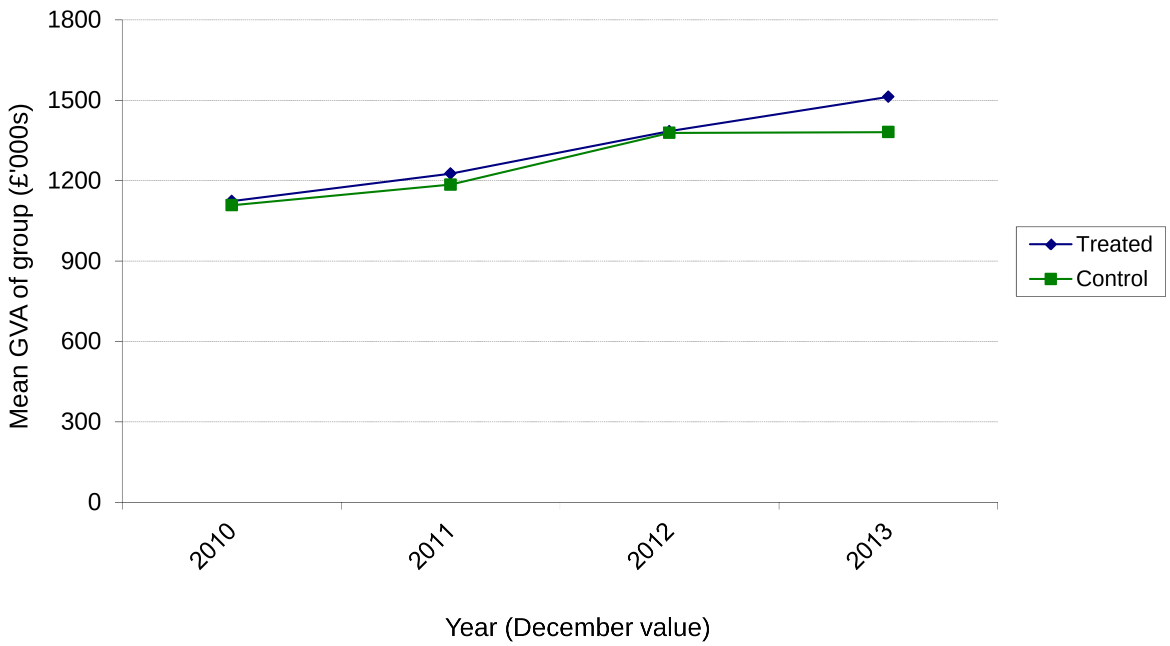

Figure 10- GVA time series for treatment and control groups.

Figure 11 and Table 17 look at the differences between the two groups (including companies who go out of business) in more detail. Figure 11 shows how the difference evolves over time (dashed lines show the upper and lower bound of the 95% range of the distribution) .

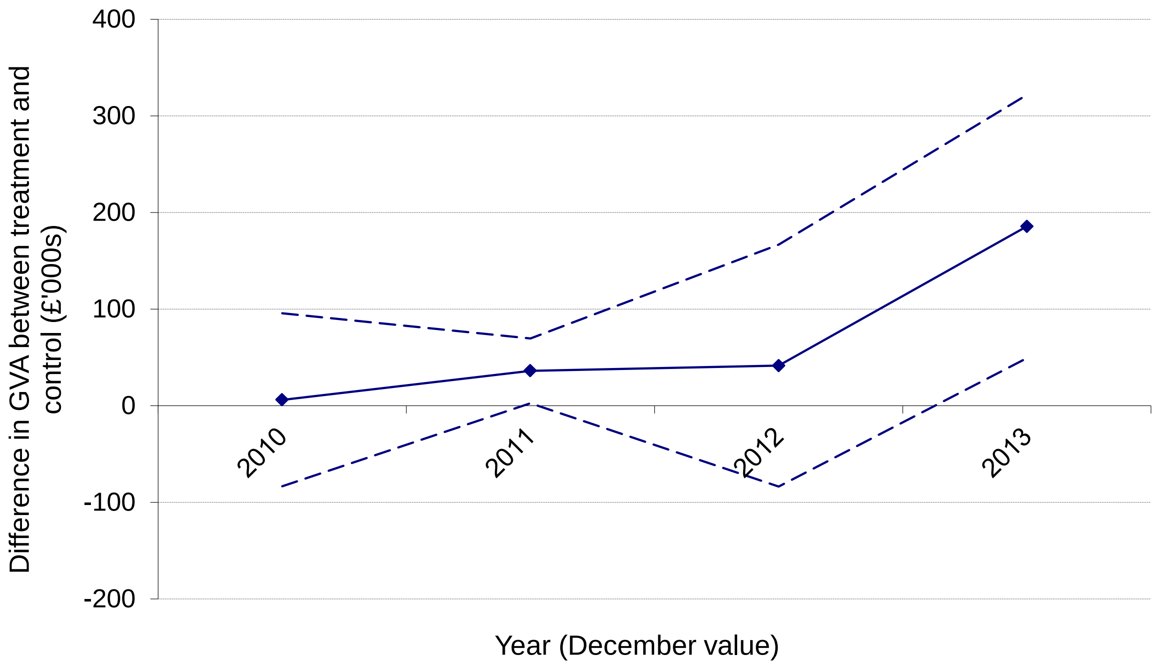

Figure 11 – Difference in GVA between treatment and control groups. Dashed lines show upper and lower bounds of the 95% range

Table 17 shows the results of a paired t-test on the difference in GVA between December 2011 and December 2013 for treatment and control groups (the difference-in-difference test).

Recall that, due to the methodological constraints of dealing with a large dataset, we have opted to use a paired T-test for the difference-in-dfference analysis in Experiments A and B. In Experiments C and D we use the Average Treatment effect on the Treated (ATT) algorithm included within the nnmatch Stata routine.

Table 17 – Results of paired T-test for Experiment A. Number of observations = 886

(£ '000s)	Mean GVA DiD	St. Err.	St. Dev.	95% range lower bound	95% range upper bound
Treatment GVA difference	241	52	1535	140	343
Control GVA difference	92	75	2238	-56	240
Difference-in-difference	149	74	2189	5	294

Note that the standard deviation is very large due to a small number of very large outliers. This is confirmed by the much more modest 95% range.

Table 18 shows the statistical significance of the T test and demonstrates than there is a significant difference at a 5% level.

Table 18 – Statistical significance in paired T-test for Experiment A

Statistical significance of paired T-test	
Pr(T < t)	0.97
Pr(|T| > |t|)	0.04
Pr(T > t)	0.02

As discussed in section 2.9, we have also run a regression analysis using the Experiment A treatment and control populations. We use GVA growth as the dependent variable. The matching covariates are then used as independent variables are initial turnover, initial employees, sector (1000 0/1 dummy variables), birthdate and treatment (0/1 dummy variable).

Table 19 shows the main results of this regression analysis. As this is simply a sensitivity check, we do not cover these results in more depth.

Table 19 – Results of regression analysis for Experiment A

(£ '000s)	Coef.	Std Err	T value	P>T	95% range lower bound	95% range upper bound
GVA difference-in-difference	130	45	2.91	0.00	43	218

Using this method the GVA DiD is estimated to be £130,000, which is broadly similar to the £150,000 estimated using the matching approach.

Employees analysis

Figures 12 and 13 show the difference in the number of employees between the treatment and control groups over the period December 2010 to December 2013.

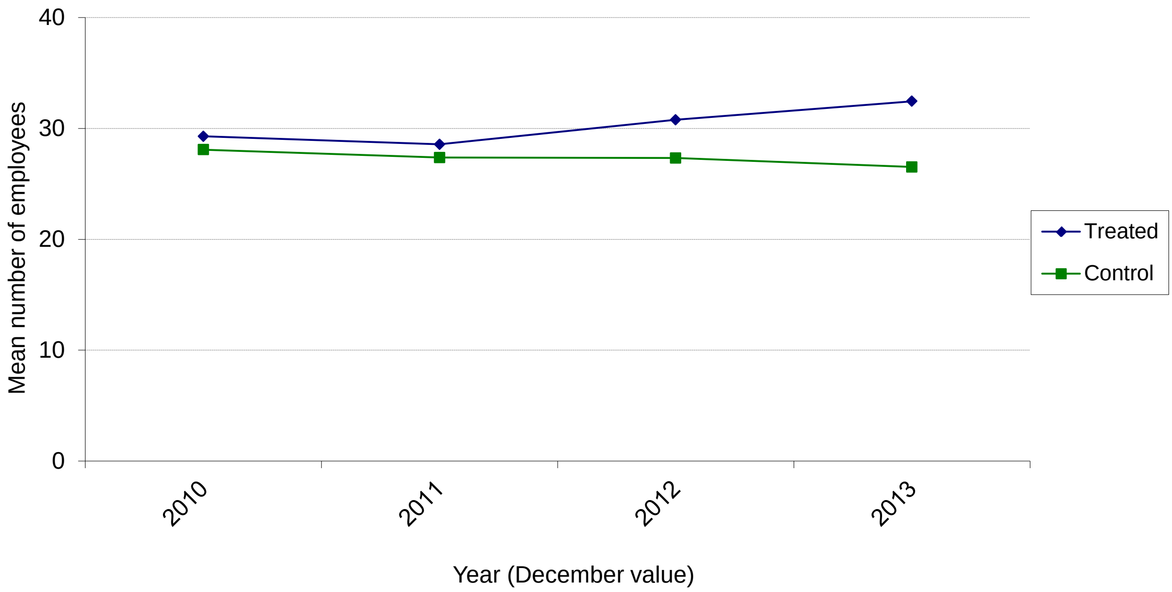

Figure 12 - Employee time series for treatment and control groups

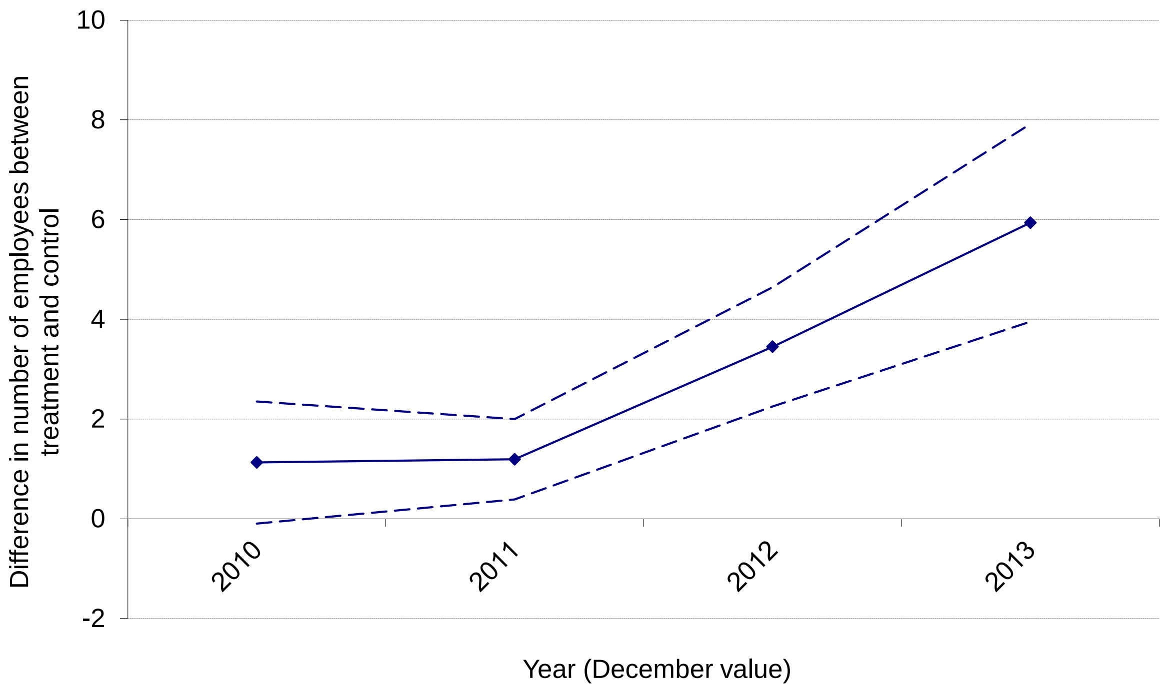

Figure 13 - Difference in employees between treatment and control groups. Dashed lines show upper and lower bounds of the 95% range

There is a DiD of around four employees between the treatment and control groups by December 2013 (over 10% of the mean number of employees). MAS advice can often involve altering recruitment strategies, so we might expect to see a difference in this variable if the policy is working effectively.

Although analysis of the growth in the number of employees is of interest, it is beyond the scope of this analysis and is not covered beyond the results presented in this section.

3.2 Experiment A - Distributional analysis

This section investigates whether there are differences in GVA DiD for subgroups of the treatment group in Experiment A. We also include the average grant funding amount for each subgroup to check whether differences can be attributed to differences in the scale of intervention. In cases where subgroups are big enough to provide statistically significant results we can use the average grant to assess the relative impacts for the different subgroups.

In most cases, subgroups are too small for the GVA DiD results to be statistically significant.

Impact by size

Tables 20 and 21 show the GVA DiD for the treatment group in Experiment A, split by number of employees and turnover.

Table 20 – Results by number of employees

Employees	Mean GVA DiD (£‘000s)	No. of obs	Average grant (£)
1. 0-4	27	190	2,540
2. 5-9	71	133	2,387
3. 10-24	180	251	2,637
4. 25-249	149	311	2,970
5. 250+	12959	2	300

Table 21 – Results by turnover

Turnover (£’000s)	Mean GVA DiD (£‘000s)	No. of obs	Average grant (£)
1. 0-249	26	152	2,455
2. 250-999	55	236	2,417
3. 1,000-2,499	135	201	2,920
4. 2,500-24,999	343	283	2,966
5. 25,000+	-607	15	1,549

There is no clear difference in average grant for the different subgroups, so we assume that the impact is proportional to the GVA DiD. Neglecting the outliers, there does appear to be a relationship between business size and GVA DiD.

Impact by region

Table 22 shows the GVA DiD for the treated group in Experiment A, split by region. As discussed in section 1, the number of L4 projects is fairly evenly spread across the regions.

Table 22 – Results by region

Region	Mean GVA DiD (£'000s)	No. of obs	Average grant (£)
East Midlands	443	86	2,796
East of England	111	75	3,179
London	355	66	2,676
North East	268	37	2,489
North West	235	75	2,774
South East	149	103	2,240
South West	-273	148	2,499
West Midlands	195	167	2,656
Yorkshire & Humber	225	118	2,860

Again, there is no clear difference in average grant for the different subgroups, so we assume that the impact is proportional to the GVA DiD. There is some variation in GVA DiD, with the South West doing relatively badly and the East Midlands doing well. However, we should be very cautious in drawing firm conclusions from this as the sample sizes are all too small for the results to be statistically significant.

Impact by L4 project type

Table 23 shows the GVA DiD for the treated group in Experiment A, split by L4 project type. As discussed in section 1, the distribution is highly skewed towards Strategy, Operational Improvement and Innovation projects.

Table 23 – Results by L4 project type

L4 project type	Mean GVA DiD (£'000s)	No. of projects	Average grant (£)
Operational Improvement	182	687	2,817
Innovation	77	223	3,404
ERDF	57	18	6,125
Strategy	149	398	3,032
Supply Chain	166	60	3,849
Expert Innovation	-200	13	3,375
Expert Finance	-170	4	2,917
Six Sigma	5,942	1	0

Again, there is no clear difference in average grant for the different subgroups, so we assume that the impact is proportional to the GVA DiD. There is some variation in GVA DiD but, again, we should be cautious in drawing firm conclusions when the sample sizes are all too small for the results to be statistically significant.

The results for Operational Improvement and Strategy projects are relatively significant. However, both of their GVA DiDs are similar to the £150,000 average GVA DiD and their mean grants are close to the £2,600 group mean, so we cannot draw any firm conclusions about whether either of these project types has an impact that is different from the average.

Impact by size of grant and number of repeat interventions

Table 24 shows the GVA DiD for the treated group in Experiment A, split into grant funding quintiles. Again, we should be cautious in drawing firm conclusions when the sample sizes are all too small for the results to be statistically significant. There does appear to be a relationship between amount of funding and GVA DiD, with those businesses allocated very small amounts of funding having a lower GVA on average over the period.

Table 24 – Results by total L4 grant funding amount

Grant funding range	Mean GVA DiD (£'000s)	No of obs
£225 - £1,400	-202	173
£1,400 - £1,947	154	178
£1,947 - £2,798	240	174
£2,798 - £3,900	67	172
£3,900 - £14,400	271	178

Table 25 shows the GVA DiD for the treated group in Experiment A, split by the number of repeat projects. As mentioned in section 1, the majority of businesses only undergo one or two projects. The subgroups of businesses receiving one or two interventions are large enough for the results to be statistically significant.

Table 25 – Results by number of repeat interventions

No of projects	Mean GVA impact (£'000s)	No of obs	Average grant (£)
1	54	514	1,713
2	208	245	3,306
3	159	90	5,290
4	173	22	7,685
5	35	2	9,759
6	-878	3	10,250

It is an interesting result that those businesses receiving two interventions end up with around four times the GVA DiD of businesses that only receive a single intervention, but the average grant funding is only double. This suggests that the impact for businesses developing a longer running relationship with MAS could be double that for businesses that only carry out a single project.

However, it is important to recall that we expect advisor selection-bias to be higher for MAS clients who receive repeat grants, because the requirements to demonstrate a high return-on-investment from earlier MAS funding. For this reason we are cautious about stating with certainty that multiple interventions offer greater value-for-money than one-off interventions, despite the results suggesting that this is the case.

Impact of L5 support

As described in section 1, MAS brokers additional Government support for some businesses through its L5 support service. As these businesses are likely to be receiving additional Government support, we might expect that they experience higher growth than those that don't receive L5 support. Table 26 shows the GVA DiD for the treated group in Experiment A, according to whether they receive this support or not.

Table 26 – Results by participation in L5 support

L5 support	Mean GVA impact (£'000s)	No of companies	Average Grant (£)
yes	123	408	2,597
no	177	470	2,767

There does not appear to be any clear difference between those that receive L5 support and those that don't. This may be because businesses not accessing L5 support are acquiring Government support independently. More qualitative evidence around the nature of L5 support could help to understand why that might be the case.

3.3 Experiment B – L2 vs. No MAS

As outlined in section 2.8, the aim of Experiment B is to find the GVA DiD between December 2013 and December 2011 between businesses that receive a MAS L2 review, but no L4 grant, and matched businesses who have never contacted MAS.

Matching

As the results of the matching analysis is broadly similar to Experiment A, we do not cover the full detail here but have included the covariate distribution charts in Annex 3. The KS tests for initial turnover, initial employees and birthdate give combined P values of 99.4%, 99.1% and 100.0% respectively.

GVA analysis

Figure 14 shows time series of the mean GVA for the treatment and control groups between December 2010 and December 2013.

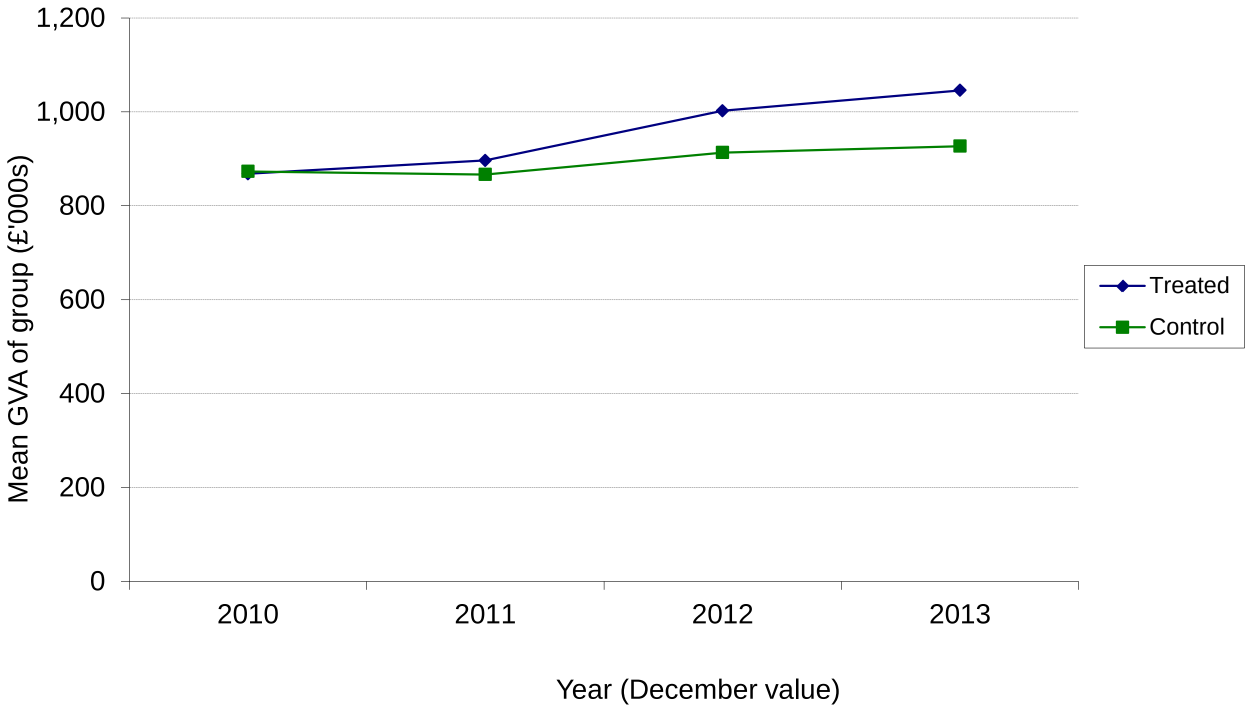

Figure 14 - GVA time series for treatment and control groups

The two groups appear to be closely matched up to 2012, at which point the mean GVA values diverge, with a DiD between 2013 and 2011 of around £90,000. Again, we do not ascribe the difference to the MAS intervention as we need to take account of selection biases, covered in section 4.

As with Experiment A, companies in the control group were more likely to go out of business. Nearly 11% of the original control group had gone out of business whereas only 3% of the treatment group had gone out of business over the same period. This result is discussed in more detail in the interpretation section.

Figure 15 shows how the difference evolves over time (dashed lines show the upper and lower bound of the 95% range of the distribution).

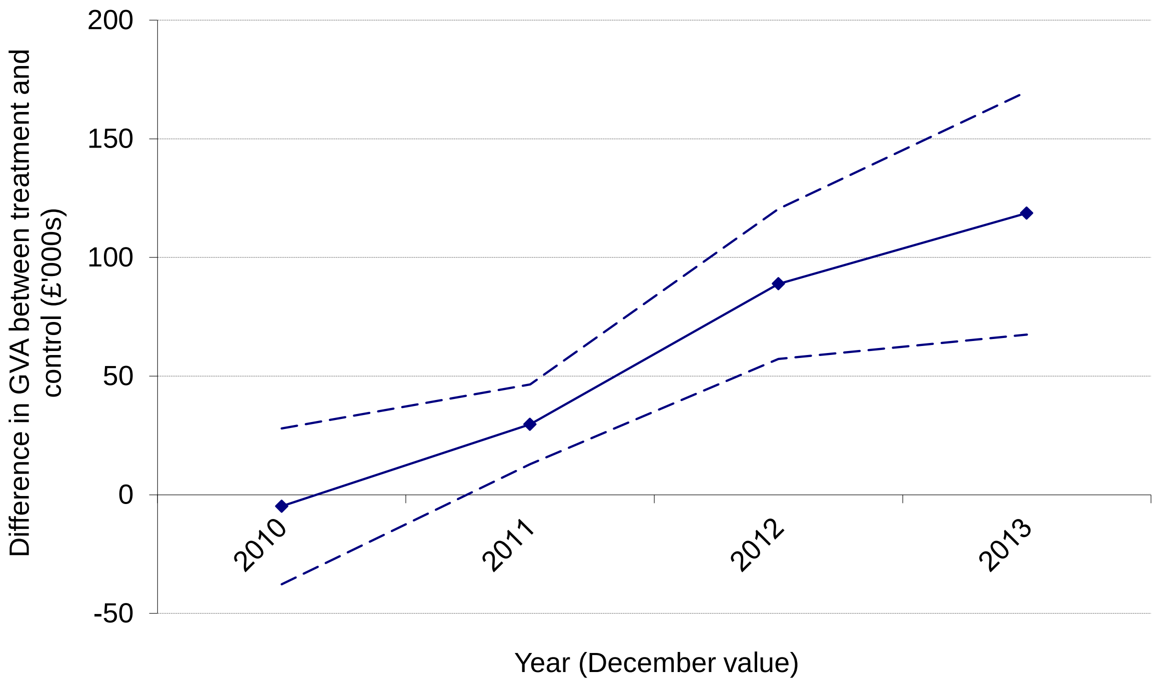

Figure 15 - Difference in GVA between treatment and control groups. Dashed lines show upper and lower bounds of the 95% range.

Table 27 shows the results of a paired t-test on the difference in GVA between December 2011 and December 2013 for treatment and control groups (the difference-in-difference test).

Table 27 – Results of paired T-test for Experiment B. Number of observations = 2922

(£ '000s)	**Mean GVA DiD**	**St. Err.**	**St. Dev.**	**95% range lower bound**	**95% range upper bound**
Treatment GVA difference	149	18	996	113	185
Control GVA difference	60	20	1099	20	100
Difference-in-difference	89	25	1360	40	138

Note that the standard deviation is very large due to a small number of very large outliers. This is confirmed by looking at the much more modest 95% range.

Table 28 shows the statistical significance of the T test and demonstrates than there is a significant difference at a 1% level.

Table 28 – Statistical significance in paired T-test for Experiment B

Statistical significance of paired T-test	
Pr(T < t)	1.00
Pr(|T| > |t|)	0.00
Pr(T > t)	0.00

Again, we have also run a regression analysis using the Experiment B populations, with GVA growth as the dependent variable and the matching covariates as dependent variables. Table 29 shows the main results of this regression analysis.

Table 29 – Results of regression analysis for Experiment B

(£ '000s)	Coef.	Std Err	T value	P>T	95% range lower bound	95% range upper bound
GVA difference-in-difference	98	25	3.92	0.00	49	147

Using this method the GVA DiD is estimated to be £100,000, which is broadly similar to the £90,000 estimated using the matching approach.

Employees analysis

Figures 16 and 17 show the difference in the number of employees between the treatment and control groups over the period December 2010 to December 2013.

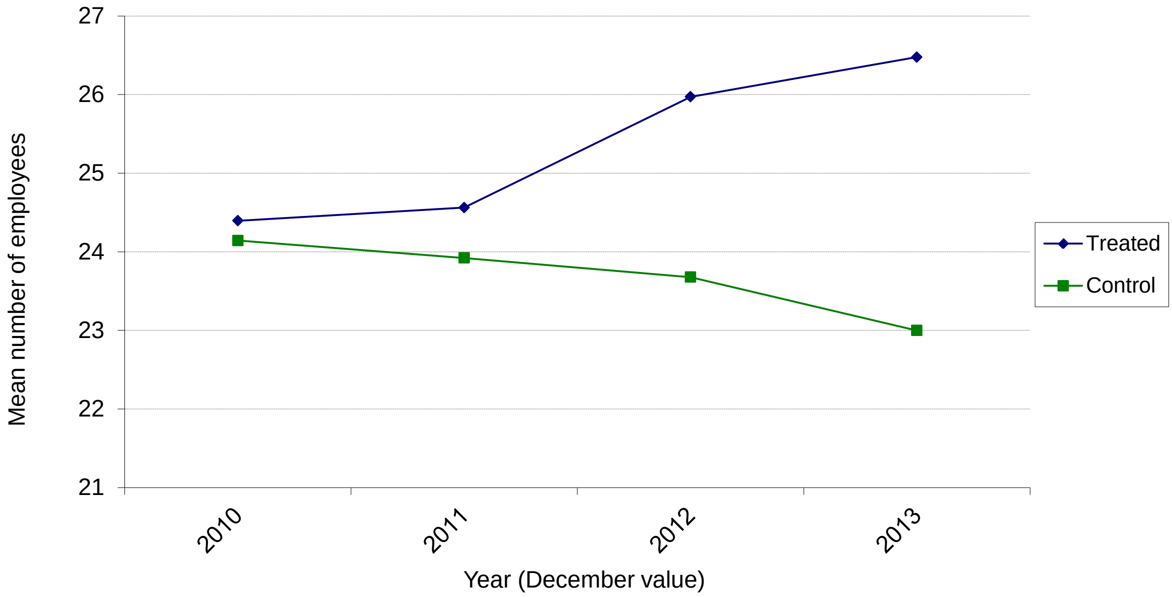

Figure 16 - Employee time series for treatment and control groups

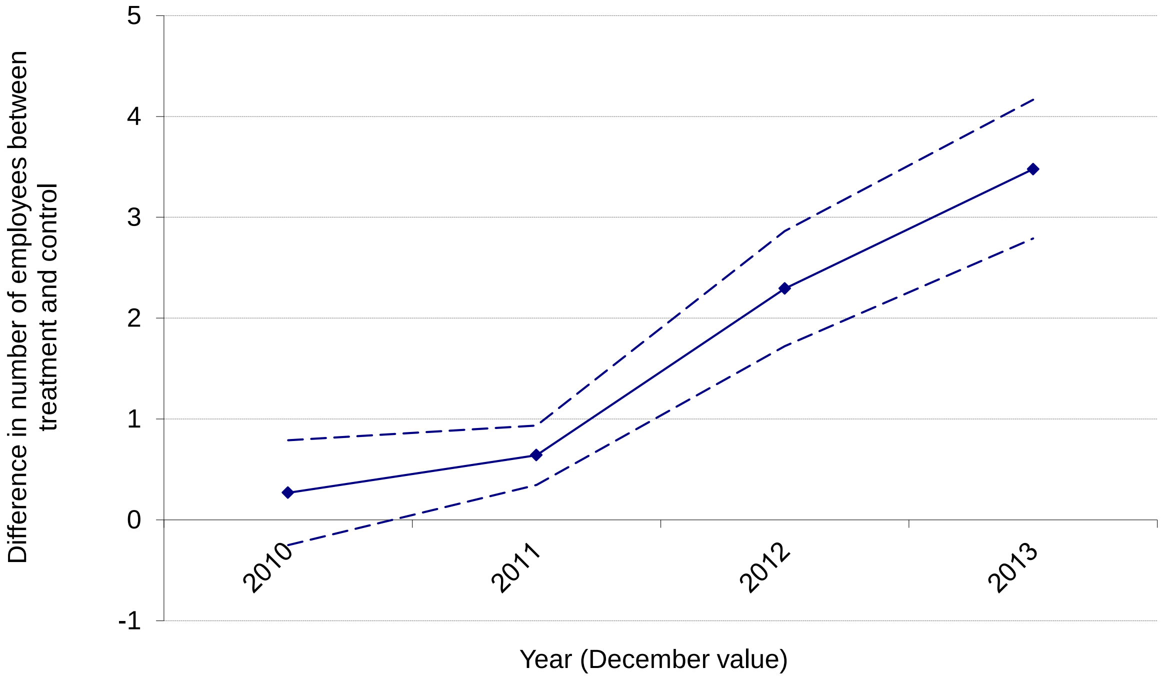

Figure 17 - Difference in employees between treatment and control groups. Dashed lines show upper and lower bounds of the 95% range

As with Experiment A, there is a DiD of around three employees between the treatment and control groups by December 2013 (again, over 10% of the mean number of employees). This is discussed in more detail in the section 4.

3.4 Experiment C – L4 vs L2

As outlined in section 2.8, the aim of Experiment C is to find the GVA DiD between December 2013 and December 2011 between businesses that receive a MAS L4 grant funded project and matched businesses who receive a MAS L2 review, but no L4 grant.

Matching

As the results of the matching analysis are broadly similar to the other experiments, we do not cover the full detail here but have included the covariate distribution charts in Annex 3. The KS tests for initial turnover, initial employees and birthdate give combined P values of 79.0%, 99.9% and 100.0% respectively.

GVA analysis

Figure 18 shows time series of the mean GVA for the treatment and control groups between December 2010 and December 2013.

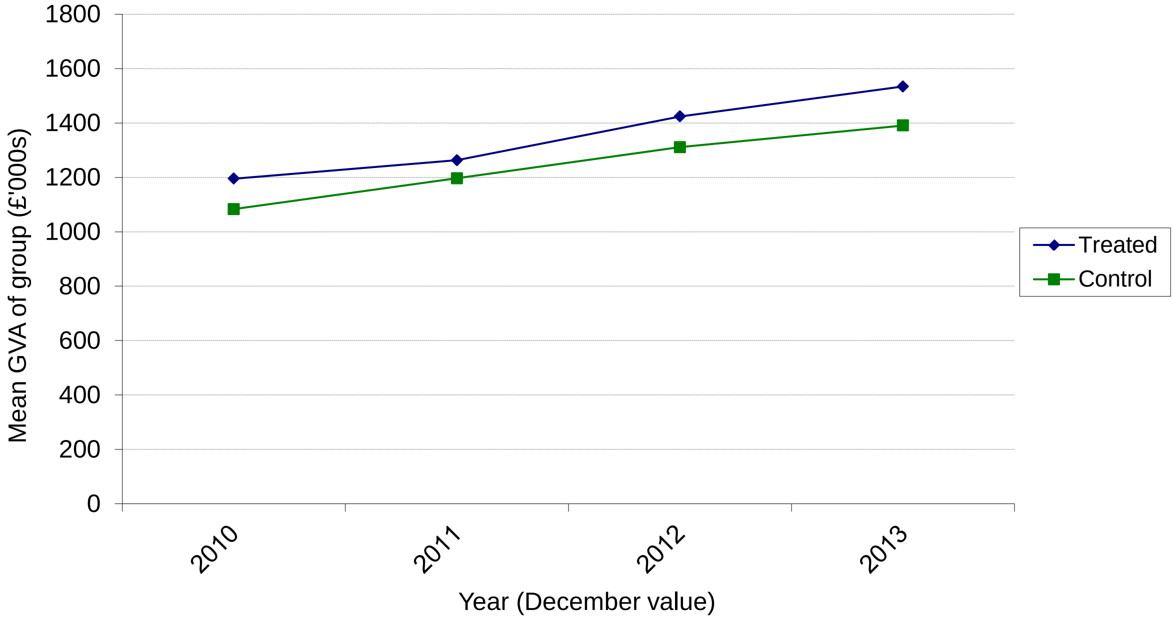

Figure 18 - GVA time series for treatment and control groups.

The two groups are not that well matched in 2011 and the mean GVA gap increases by the end of 2013, with a DiD between 2013 and 2011 of around £80,000. Again, we should not ascribe the whole difference to the MAS intervention as we need to take account of selection biases, covered in section 4.

In this experiment there is no clear difference in the percentages of treated and untreated businesses that go out of business over the period.

Figure 19 shows how the difference evolves over time (dashed lines show the upper and lower bound of the 95% range of the distribution).

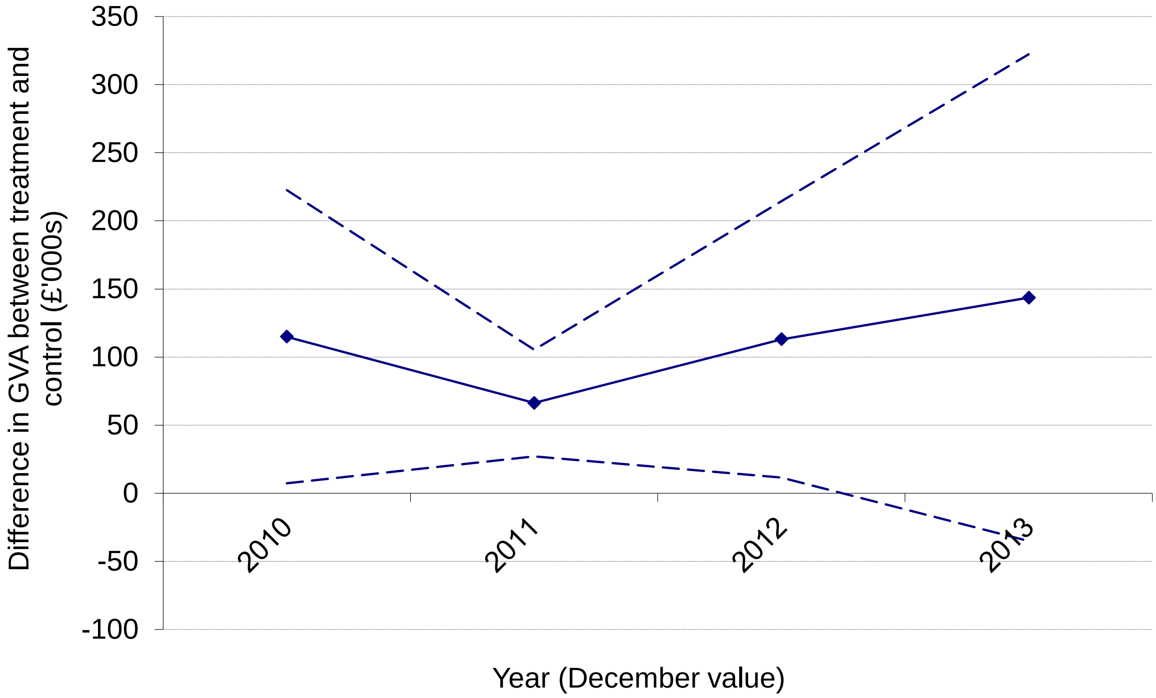

Figure 19 Difference in GVA between treatment and control groups. Dashed lines show upper and lower bounds of the 95% range

Table 30 shows the results of the“average treatment effect in the treated” (ATT) test on the difference in GVA between December 2011 and December 2013 for treatment and control groups (the difference-in-difference test).

Table 30 – Results of average treatment effect analysis for Experiment C. No. of observations = 597

(£ '000s)	Mean (ATE)	St Err	z	P>|z|	95% range lower bound	95% range upper bound
GVA difference-in-difference	79	96	0.82	0.41	-110	268

Again, we have carried out a regression analysis using the Experiment C populations, with GVA growth as the dependent variable and the matching covariates as dependent variables. Table 31 shows the main results of this regression analysis.

Table 31 – Results of regression analysis for Experiment C

(£ '000s)	Coef.	Std Err	T value	P>T	95% range lower bound	95% range upper bound
GVA difference-in-difference	75	53	1.40	0.16	-30	179

Using this method the GVA DiD is estimated to be £75,000, which is broadly similar to the £80,000 estimated using the matching approach.

Employees analysis

Figures 20 and 21 show the difference in the number of employees between the treatment and control groups over the period December 2010 to December 2013.

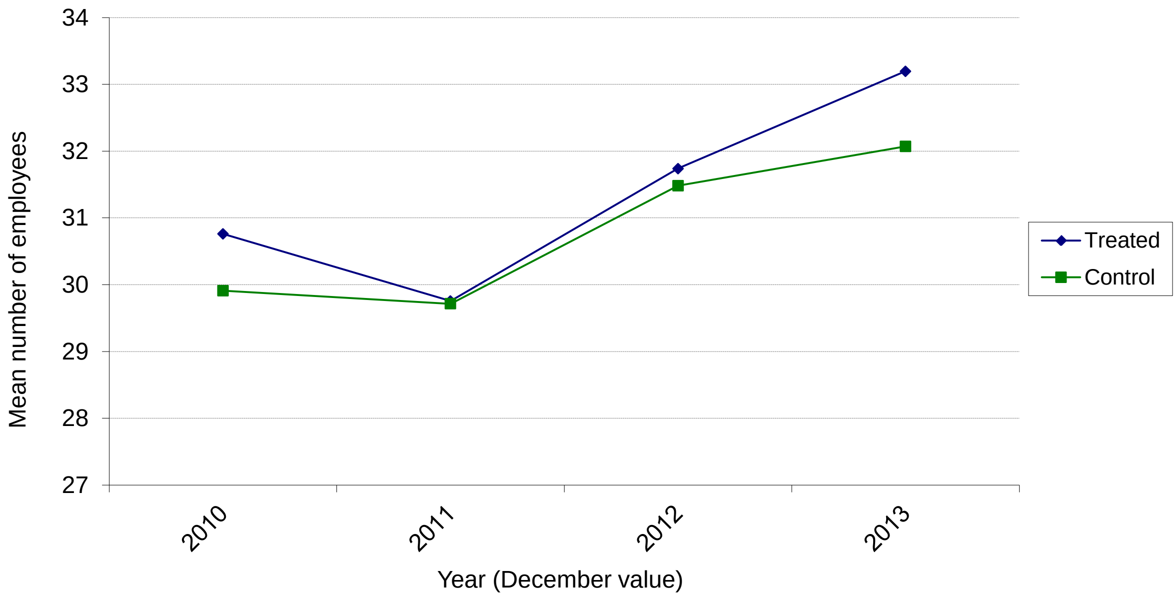

Figure 20 - Employee time series for treatment and control groups.

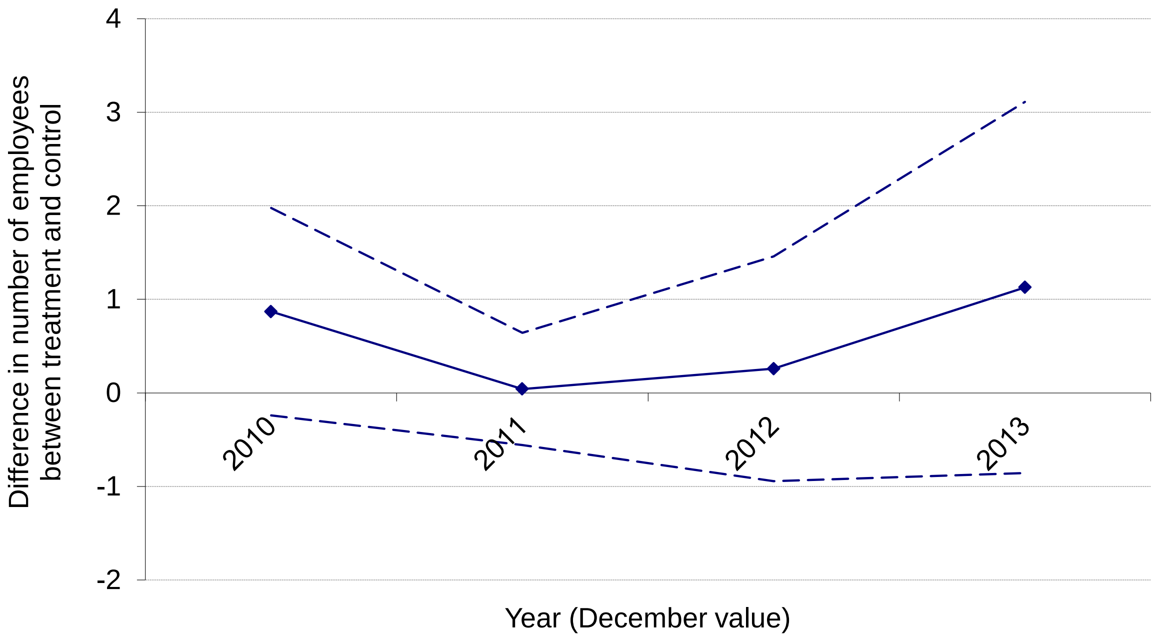

Figure 21 Difference in employees between treatment and control groups. Dashed lines show upper and lower bounds of the 95% range

Here, there is a DiD of around one employee between the treatment and control groups by December 2013 but the result is not statistically significant.

3.5 Experiment D – L4 vs later L4

As outlined in section 2.8, the aim of Experiment B is to find the GVA DiD between the December 2013 and December 2011 between businesses that receive a MAS L4 grant funded project in 2012 and matched businesses who receive a MAS L4 grant funded project in 2013.

Matching

As the results of the matching analysis are broadly similar to the other experiments, we do not cover the full detail here but have included the covariate distribution charts in Annex 3. The KS tests for initial turnover, initial employees and birthdate give combined P values of 79.2%, 94.1% and 97.9% respectively.

GVA analysis

Figure 22 shows time series of the mean GVA for the treatment and control groups between December 2010 and December 2013.

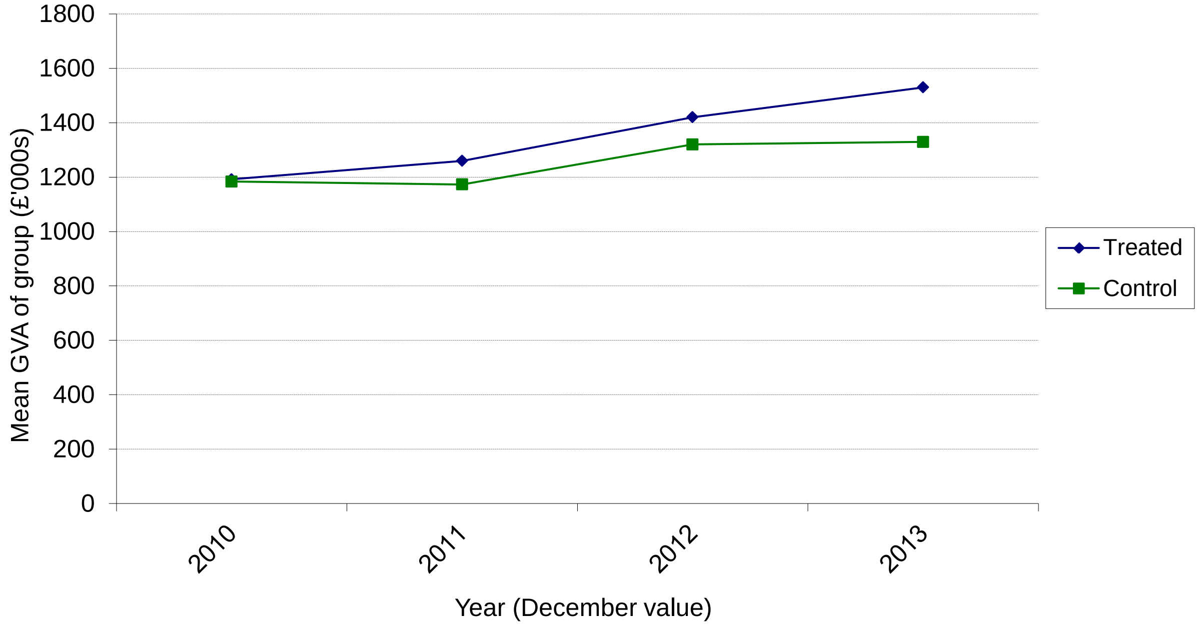

Figure 22 - GVA time series for treatment and control groups

The two groups appear to be closely matched up to 2012, at which point the mean GVA values diverge, with a DiD between 2013 and 2011 of around £125,000. Again, we should not ascribe the whole difference to the MAS intervention as we need to take account of various biases, covered in section 4.

In this experiment there is no difference in the percentages of treated and control businesses that go out of business over the period. In fact, control businesses are forced by our treatment criteria to to still be operating in 2013 in order to carry out a L4 project in that year.

Figure 23 shows how the difference evolves over time (dashed lines show the upper and lower bound of the 95% range of the distribution).

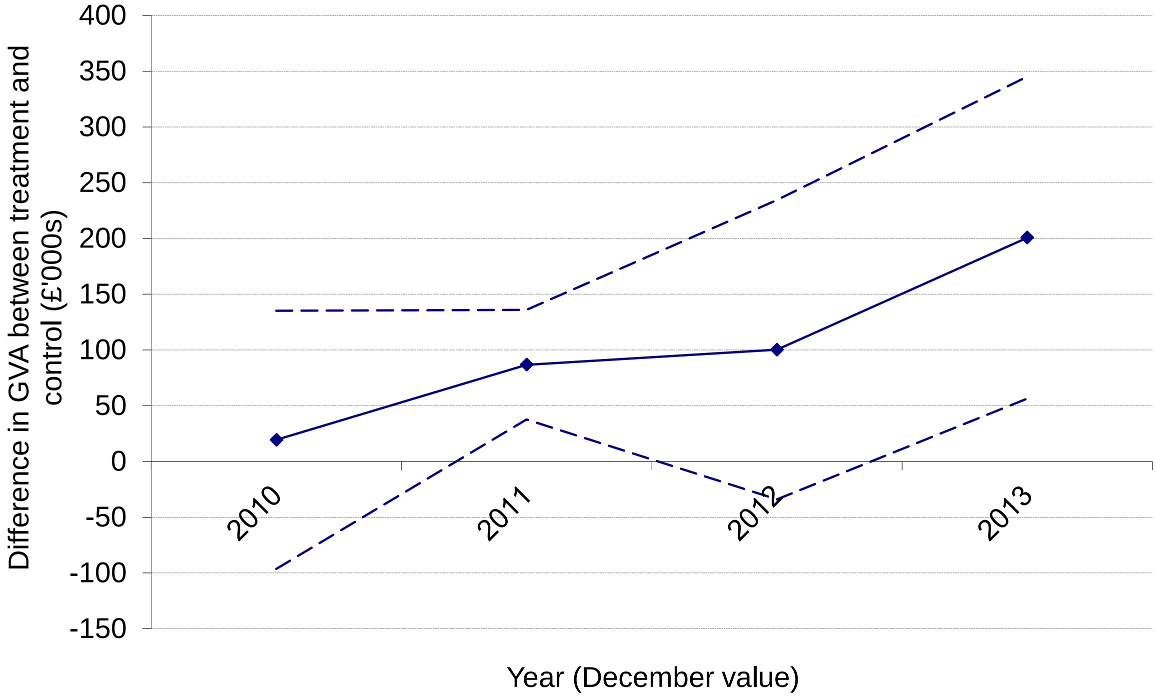

Figure 23 Difference in GVA between treatment and control groups. Dashed lines show upper and lower bounds of the 95% range.

Table 32 shows the results of the "average treatment effect in the treated" (ATT) test on the difference in GVA between December 2011 and December 2013 for treatment and control groups (the difference-in-difference test).

Table 32 – Results of average treatment effect analysis for Experiment D. No. of observations = 597

(£ '000s)	Mean (ATT)	St Err	z	P>|z|	95% range lower bound	95% range upper bound
GVA difference-in-difference	125	77	1.62	0.10	-26	275

Again, we have carried a regression analysis using the Experiment D populations, with GVA growth as the dependent variable and the matching covariates as dependent variables. Table 33 shows the main results of this regression analysis.

Table 33 – Results of regression analysis for Experiment D

(£ ‘000s)	Coef.	Std Err	T value	P>T	95% range lower bound	95% range upper bound
GVA difference-in-difference	92	54	1.72	0.09	-13	198

Using this method the GVA DiD is estimated to be £90,000, which is broadly similar to the £125,000 estimated using the matching approach.

Employees analysis

Figures 24 and 25 show the difference in the number of employees between the treatment and control groups over the period December 2010 to December 2013.

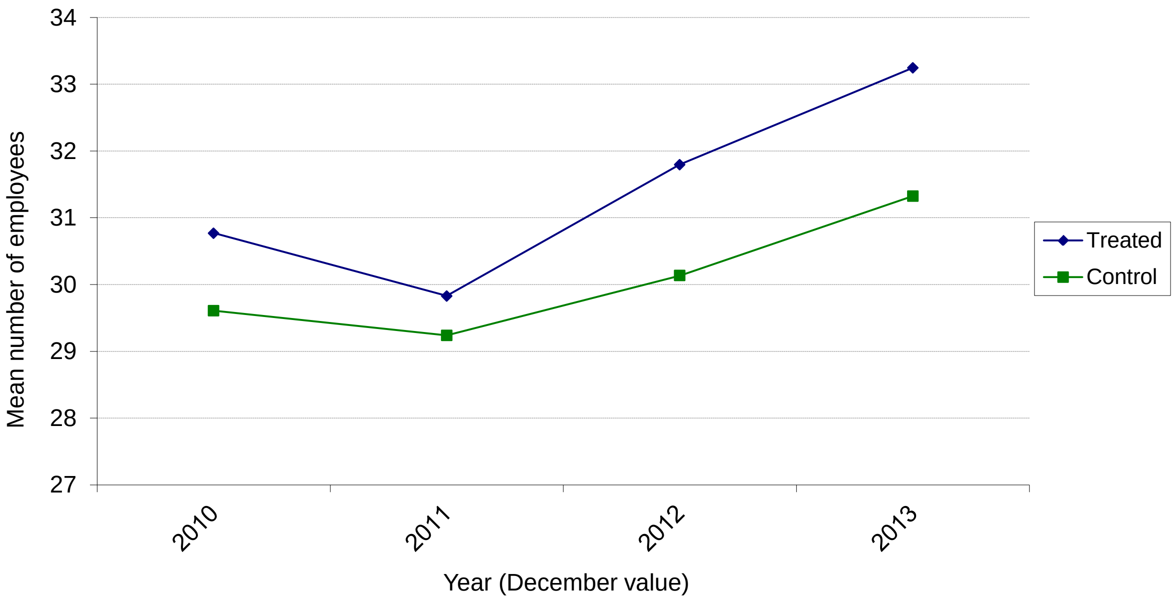

Figure 24 - Employee time series for treatment and control groups

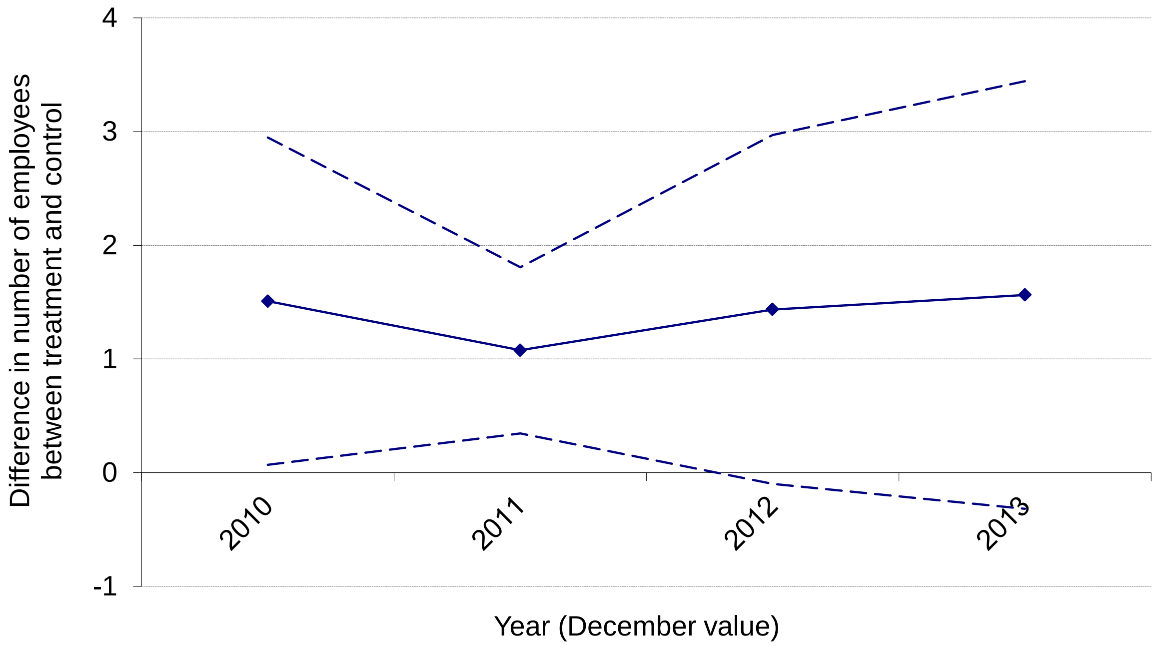

Figure 25 Difference in employees between treatment and control groups. Dashed lines show upper and lower bounds of the 95% range.

Here, there is a DiD of around one employee between the treatment and control groups by December 2013 but the result is not statistically significant.

3.6 Summary of results

Table 34 summarises the final GVA DiD results across the four experiments.

Table 34 – Summary of GVA DiD results and variance for Experiments A – D

(£ '000s)	Mean GVA DiD	95% range lower bound	95% range upper bound	Number of observations
Experiment A (L4 vs No MAS)	149	5	294	886
Experiment B (L2 vs No MAS)	89	40	138	2922
Experiment C (L4 vs L2)	79	-110	268	597
Experiment D (L4 vs later L4)	125	-26	275	597

In the next section we interpret these results and discuss the limitations of the methodology and associated uncertainty.

4. Interpretation

4.1 Selection bias

Empirical evidence

As discussed in section 1 and section 2, businesses receiving MAS grant funding are subject to both self-selection and advisor-selection bias. Section 2.8 outlined how the four experiments should allow us to isolate the effect of self-selection bias and potentially the effect of advisor-selection bias. The presence of these biases could overstate the potential impact of a scheme if not taken care of.

Experiment A should show the actual impact of MAS grant funding that does not correct for bias. The average GVA DiD compared to the matched control group, who have never contacted MAS, is £150,000. This difference is statistically significant to a 5% level.

Experiment B should enable us to isolate the effect of self-selection bias, assuming that the L2 review intervention has a minimal impact on GVA growth. The average GVA DiD compared to the matched control group, who have never contacted MAS, is £90,000. This difference is statistically significant to a 1% level. We assume that this additional £90,000 growth is due to unobserved differences in characteristics which may also explain self-selection. We interpret these unobserved differences as a high level of business “proactivity”, which could include the enthusiasm and the skills required to spend time researching and applying for Government funding. Here, we assume that these unobserved characteristics could generate the £90,000 difference in growth between the two groups of businesses. In doing so, we acknowledge that these assumptions are highly uncertain and based on logical argument rather than hard evidence.

One observable characteristic that is worth of mentioning in relation to self-selection bias is the likelihood of going out of business. Sections 3.1 and 3.3 showed that MAS clients are generally less likely to go out of business, regardless of whether they receive grant funding or not. This suggests that the qualities associated with self-selection bias (more proactivity and demonstrating more growth potential) are linked to a business’s ability to stay afloat.

Experiment C should demonstrate the sum impact of the advisor selection and the actual impact of MAS grant funding. The average GVA DiD compared to the matched control group, who have never contacted MAS, is £80,000. This difference is **not** statistically significant to a 5% level. However, the results can still be used to carry out a useful sense check. If we assume that Experiment B covers the self-selection bias and Experiment C covers both the advisor selection and the actual impact of MAS grant funding, then summing their impacts together should give them same result as Experiment A. As the sum of £90,000 and £80,000 is relatively close to £150,000, this provides some reassurance that our assumptions on self-selection bias are valid.

In Experiment D we attempt to isolate the impact of MAS grant funding by comparing groups of businesses that undertake L4 projects in different years. Although the experiment is powerful in theory, its practical implementation suffers from two main problems. The first is small sample size and a high number of duplicates in the matched groups, which leads to a high variance and a result that is not statistically significant. The second is the potential timing bias mentioned in section 2.8. Businesses may approach MAS for support at a particular time in their development as an organisation. MAS support could coincide with a change in leadership or strategy are with a package of additional investment intended to stimulate growth. The observed difference is £125,000, roughly the same size as in Experiment A, which suggests that this timing issue could influence both the self-selection and advisor-selection bias. This suggests that the business "proactivity" that determines self-selection and the business qualities that cause MAS advisors to award L4 funding may both be highly time-dependent. However, we cannot draw any strong conclusions on this point as the results are not statistically significant.

We conclude that, despite the empirical evidence provided by this analysis, we are unable to isolate the effect of advisor-selection bias, and have to rely on theoretical assumptions on deadweight to estimate the net impact of MAS grant funding.

Assumptions on advisor-selection bias

Unfortunately, the final stage of this analysis relies on assumptions that lack a hard evidence base. Due to this lack of evidence, we have used a wide range of values to estimate the scale of advisor-selection bias.

We have consulted the BIS central evaluation team and guidance used to evaluate the allocation of similar pots of money as part of the Regional Growth Fund policy. We have arrived at the assumption that 50-75% of the GVA growth, after controlling for self-selection bias, would have occurred anyway and is therefore attributed to advisor-selection bias. We believe that a high percentage is warranted in this case as the advisor-selection bias is increased by the fact that businesses must demonstrate an even higher potential for growth (through a high forecast of return-on-investment) to receive a second or third round of grant funding.

Implementing a Randomised Controlled Trial

Although this impact analysis lacks robust evidence to assess the scale of advisor-selection bias, this could be overcome in future through the implementation of a Randomised Controlled Trial (RCT). MAS clients undergoing a L2 review could be split into two groups, one in which businesses are allocated L4 grant funding as normal by MAS advisors and a second group where they are allocated MAS grant funding at random. The difference in performance between the two groups would constitute a reasonably good estimate of advisor selection bias.

However, in reality this would measure the combined impact of two separate effects. The first is the true advisor-selection bias, whereby advisors simply pick the best businesses that would have grown anyway. The second is the possibility that MAS advisors are particularly skilled at matching specific improvement projects to specific businesses, which the randomisation would not replicate. So, there is a risk that through this approach we would discount some of the value that MAS advisors add. Complementary qualitative evidence might help to assess whether this is really an issue.

MAS advisors were asked about the feasibility of implementing an RCT in the interviews described in section 1.6. Although the advisors recognised the potential benefits of this type of analysis, they expressed concerns that communicating the approach to clients could prove a potential obstacle.

4.2 Time dependence

In section 2.2 we estimated that the average MAS intervention considered in this analysis has around 15 months to give rise to an impact by December 2013, which is a relatively short period of time for the business to implement a project and experience the associated benefits. It is possible that the observed impact will persist in subsequent years and that the cumulative benefit to the economy will be higher than this analysis implies. There is significant uncertainty over whether this additional benefit will materialise and we recommend that this analysis be repeated in subsequent years to determine whether this is the case.

4.3 Economy wide effects

The results presented in section 3 cover the direct, average impact of MAS support on individual businesses. These results cannot simply be scaled up to estimate the overall impact of MAS on the economy because we have not taken account of the way MAS clients might interact with other UK businesses. A full policy evaluation should incorporate the estimation of the full range of indirect effects, typically referred to as “additionality”. The usual components of additionality are deadweight, displacement (and substitution), leakage and multipliers[22].

Deadweight can be considered at an individual business level, whereas the other three components involve interaction with other businesses and so require analysis of the wider economy. We have assumed that the majority of the deadweight is accounted for in our construction of the counterfactual and the matched comparison group of businesses, and the associate net GVA estimates, as well as our estimation of the selection biases in the difference-in-difference analysis presented above. However, we do not have sufficient evidence to account for the effects of the other three components on the impact of MAS. A brief summary of the potential impact these effects might have on MAS is given below.

- *Displacement* – GVA displacement effects will be important if growth in the treated businesses is associated with the business increasing its share of the UK market to the detriment of their UK competitors. The scale of the displacement effect is usually policy specific, so using the findings from the analysis of one policy to evaluate another is not advisable unless the two policies are very similar in nature. The 2010 DTZ evaluation of MAS found no clear evidence of displacement and assumed that it was negligible. In the interviews described in section 1.6, MAS advisors suggest that displacement effects are small because MAS clients typically make niche products, with very few UK competitors. However, it is difficult to take account of these views in the quantitative analysis as they are non-specific and are likely to be subject to bias and uncertainty. Measuring displacement effects directly is very difficult and typically involves implementing a survey methodology[23], which is beyond the scope of this analysis.

---

[22] BIS OCCASIONALPAPER NO.1, 2009, “Research to improve the assessment of additionality” - http://webarchive.nationalarchives.gov.uk/20090609003228/http://www.berr.gov.uk/files/file53196.pdf

[23] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/77592/Displacement_Final_Report.pdf

- *Leakage* – Leakage is the phenomenon whereby some of the benefits experienced by funded businesses are diverted outside the UK economy. This could be through importing new input materials or recruiting new employees from overseas. The 2010 DTZ evaluation of MAS found no clear evidence of displacement and assumed a GVA leakage of 10% but there was no clear evidence provided to justify the choice of this value.

- *Multiplier* – A multiplier is taken into account when considering spillover benefits to the supply chain and local economy of providing public funding to businesses. This can be particularly important when funds are being redistributed to stimulate regional economies outside London and the South East. The 2010 DTZ evaluation of MAS found no clear evidence of displacement and assumed a GVA multiplier of 1.4 but there was no clear evidence provided to justify the choice of this value.

4.4 Economic impact of MAS grant funding

Economic impact

In section 4.1 we arrived at a GVA DiD, after controlling for self-selection bias of £60,000 over the period of interest. We also decided on using the assumption that 50-75% of this difference was due to advisor-selection bias. This gives a final average GVA DiD benefit per business of **£15,000 - £30,000** over the treatment period.

Costs

The costs are split into two components:

- **MAS grant funding** – Here we are considering the impacts for the treatment group in Experiment A. They receive an average MAS L4 grant of £1700.

- **MAS administration costs** – The MAS budget was around £18.0m per year over this period, with around £4.5m of this being spent on L4 grants. We take the non-grant related budget between 2012 and 2013 and divide it by the number of businesses undertaking a L2 review over that period. This gives an average administration cost per business of £2,160

- Summing these costs together gives an average total cost per business of £3,860 over the treatment period.

Benefit-cost ratio

A common measure of the cost effectiveness of a given pot of public spending is the benefit-cost ratio (BCR). Although we have quantified the average GVA impact and public costs per MAS client, we would also need to quantify economy-wide effects such as displacement, leakage and multipliers to estimate a BCR. As outlined in section 4.3, we are not able to quantify the scale of these effects with the available data and therefore avoid estimating a BCR in this analysis. Estimating the scale of these effects for MAS funding would be an important area for future analysis.

4.6 Uncertainties and issues with the approach

The analysis presented in this report has provided some useful insights into the ways in which businesses interact with and benefit from MAS. However, there are a number of issues with this approach which could affect the robustness of the conclusions drawn from the analysis. Here we acknowledge these issues so that they might be dealt with in future.

Technique for estimating Gross-Value Added

One potential flaw in this approach is the technique used to estimate GVA from turnover. We assume that GVA is proportional to turnover within a given sector, which may hold on average, but is unlikely to hold for subgroups within that sector. It is likely that, within a sector, the relationship between GVA and turnover depends on certain characteristics of the business and the type of GVA growth taking place. Two particular issues are described below.

- Are businesses are growing or declining? As the businesses are growing faster than the background manufacturing sector, on average, it may be that using the average GVA to turnover relationship is not an appropriate approach.

- Are businesses using the MAS intervention to boost sales or cut costs? Boosting sales would boost turnover and GVA in a similar way. Cutting costs would boost GVA but the effect on turnover is unclear. If cost reductions give rise to a reduction in price that makes the business more competitive, this could lead to a sales boost that is big enough to offset the initial price reduction.

The conclusions of this analysis are weakened by these uncertainties. Future analysis of the scheme should investigate whether improved estimation methods or additional sense-checks are available. The GVA validation exercise that is currently being carried out, described in section 2.2, will help to test the extent to which this technique is flawed.

Unobserved differences

Quasi-experimental techniques are at their most effective when the only characteristic that distinguishes the control group from the treatment group is that they do not receive the treatment. By matching the control group to the treatment group we have controlled for observable differences between the two groups. We have attempted to control for some of the unobserved differences by estimating the additional GVA growth associated with self-selection and advisor-selection. However, there may be other unobserved differences that we have failed to account for, which may influence the final estimated GVA impact.

We have had to use a relatively small number of covariates as we are restricted by the data available in the IDBR. If more covariates became available in future, it would be useful to repeat this analysis including them in the matching as this might help to control for additional unobserved differences. One downside of including more covariates is that it would not be computationally feasible to use Nearest Neighbour Matching and we would have to resort to Coarsened Exact Matching or Propensity Score Matching, which would reduce the closeness of the matches and may lead to unbalanced matching.

Impact of other Government support

Although we have been able to identify a subset of businesses that access other forms of Government support through the MAS L5 service, we do not have detailed information on the nature or value of the additional support they receive. It may be the case that we are “double counting” benefit for businesses that are receiving Government grants through multiple schemes. Further work is required to link databases for these different support schemes to the IDBR so that we can assess the sum total of Government grants going to each business. This is an important area for future work as it has the potential to improve the evaluation of a range of small-business support policies.

5. Conclusions and Recommendations

Note: Any estimates of benefits are provisional – as noted in the recommendations a full impact assessment needs to be undertaken a few years after the scheme has ended, in order for the full impact of the benefits to materialise.

We have presented a methodology for estimating the average business-level impact of MAS grant fund over the period January 2012 – December 2013. The methodology uses data linking, matching methods and difference-in-difference analysis to identify a control group of businesses, that don't receive MAS support. We then use a series of different quasi-experiments to estimate the additional GVA growth associated with self-selection bias. Finally we estimate the scale of additional GVA growth associated with the selection bias introduced by MAS advisors when choosing businesses for grant funding. As there is no robust evidence available to assess advisor-selection bias we use a wide range of values. We avoid extrapolating these results up to estimate the aggregate economic benefit of MAS as we lack the data to estimate the different components of additionality.

We demonstrated that this methodology is an improvement on that used in previous evaluations of MAS as we have been able to identify a counterfactual and avoid using self-forecast or self-reported growth. However, we also described some significant uncertainties and methodological issues that weaken the conclusions of the analysis and help to identify the most important areas of future work.

The key results, outstanding policy questions and recommendations for future work are outlined below.

Empirical evidence of self-selection bias

One of the most important results in this analysis is that businesses that undertake a telephone review with a MAS advisor have an average GVA growth that is £90,000 higher than matched businesses that have never contacted MAS, over the treatment period. Although there may be some economic benefit to this telephone review, it is unlikely to be large, so we attribute this additional growth to unobserved differences between the treatment and control group. We propose that these differences are those characteristics associated with self-selection, such as "proactivity". It is also possible that that the inherent nature of the policy design could introduce some bias affecting the outcomes of our findings.[24] MAS advisors pick businesses that demonstrate high growth potential to receive grant funding.

[24] There is little empirical evidence to allow us to understand the scale of this so called 'advisor selection bias' although we could apply benchmarks from other schemes such as the Regional Growth Fund.

Estimate of the average business-level impact of MAS grant funding

In this analysis, we focused our attention on L4 grant funded projects, which are the most popular and most in-depth service offered by MAS. Experiment A demonstrated a significant difference in GVA growth of £150,000 between businesses that undertake L4 projects and matched businesses that have never contacted MAS. However, this did not take account of selection biases. In Experiment B we estimated that the additional GVA growth associated with self-selection is £90,000. We then estimated that 50-75% of the remaining growth was associated with advisor-selection bias (we used a wide range of values as we lack any robust evidence to support this assumption). This leads to a final, average net GVA benefit per business of **£15,000 - £30,000** over the treatment period. The average MAS funding per business over this period is £3,860, including an average grant of £1,700. We avoid explicitly calculating an overall benefit-cost-ratio for the scheme due to a lack of further evidence on other components of additionality.

Repeat grant funding - enhanced benefits or enhanced advisor-selection bias?

One of the stated aims of this analysis was to identify whether certain types of support provided a higher benefit than others. We investigated this in the distributional analysis presented in section 3.2 but the majority of results were not statistically significant due to small population sizes.

One result that was statistically significant was that those businesses carrying out two L4 projects had roughly double the benefit per pound of grant funding compared to businesses that only carried out a single L4 project. If this result is valid, it suggests that the cost effectiveness of MAS could be improved by focusing on repeat rather than one-off interventions. However, the qualitative evidence presented in section 1.6 suggests that advisor-selection bias is more significant for businesses that receive repeat grant funding, i.e. advisors allocate repeat funding to those businesses that have already proved themselves to be growing. This casts doubt on the apparent difference in cost-effectiveness and means that no robust policy recommendations can be provided. A Randomised Controlled Trial (RCT) may provide more robust evidence on the scale of advisor-selection bias.

Recommendations

The main aim of this analysis was to set out a methodological framework for evaluating MAS (and other small business support policies) rather than offer detailed policy recommendations.

This analysis has identified a series of areas of future work that could significantly improve the impact evaluation of small business support policies. The main recommendations for future impact analyses are summarised below.

- Validate or improve the methodology for estimating GVA using turnover data from the IDBR.

- Repeat the quasi-experimental analysis annually to lengthen the time series and check for persistent impact.

- Investigate the implementation of an RCT for the allocation of L4 support following the L2 review to assess the scale of advisor-selection bias. This is currently being considered as part of the Business Growth Service evaluation framework, alongside implementing a standardised survey and regular data linking exercise between IDBR and business support.

Annex 1 Interviews with MAS advisors

The MAS programme delivery manual provides some usual context to understand the allocation of MAS support but there is a lack of detail in some areas. To gather additional evidence on this decision making process we conducted telephone interviews with a small number of (seven) MAS advisors out of a total population of 85 nationally. In order to obtain a representative range of views, the advisors interviewed had a range of experience and came from a mixture of regions.

Findings from the small number of consultations suggest that:

Allocating grants – The advisors gave a mixed response when asked how they choose which companies to allocate L4 grants to. Two of the advisors emphasised the importance of the business reaching the target ROI in deciding whether to allocate funding (ratios of 50:1 and 70:1 were mentioned). Three advisors suggested that the decision was based on a broader assessment of the business's potential to grow – two of these three stated that their intuitive "feeling" about the business was an important factor. The final two advisors suggested that businesses were often self-selecting as they have to be willing to provide sufficient funds to match the L4 grant funding, and many businesses are not willing to do this.

Deadweight – Here, we are using deadweight as shorthand for any growth that would have occurred in the absence of MAS support. In general, the advisors admit that there may be some deadweight but they think it is probably low, although they are not able to quantify this or provide any evidence. One advisor stated that MAS clients are typically looking around for opportunities, so may have been able to carry out a similar project without MAS support but it probably would have been less well targeted. Some of the MAS advisers also stated that they did not think deadweight was a key consideration and there was a role for MAS in terms of initiating projects where businesses would be slow to get started.

How businesses benefit – The advisors described a wide range of benefits that MAS offers to businesses. Recurring themes included;

- providing independent advice, without emotional attachment to the business;
- using expert knowledge to identify and prioritise business specific interventions;
- helping businesses to build and disseminate a coherent corporate strategy;
- allowing CEOs to discuss concerns with someone outside the business; and
- building a long-term relationship with business and adapting support over time.

The advisors also emphasised that they thought L2 reviews were beneficial for businesses, but that this was very difficult to quantify and disentangle from the impact of subsequent L4 support. Businesses often receive repeat L2 reviews and advisers thought this is beneficial in itself as a means of providing businesses with a progress check and continued guidance.

This evidence provides some of the detail missing from the MAS programme delivery manual and helps us to give some context to our econometric analysis. Although useful, it is worth noting that evidence from the interviews may well be subject to bias as MAS advisors may consciously or unconsciously emphasise the benefits of their advice and downplay any deadweight. It is difficult to translate this kind of evidence into an input that can be used in a quantitative analysis, but by understanding the interaction of businesses and MAS advisors at each stage of the process we are able to identify and describe a series of biases that are likely to influence our results. These biases are summarised in the next section.

Annex 2 – Detailed methodology

Experiment A – L4 vs No MAS

The aim of this experiment is to find the GVA DiD between the December 2013 and December 2011 between businesses who receive a MAS L4 grant funded project and matched businesses who have never contacted MAS. We expect the results to be influenced by self-selection bias, advisor-selection bias (selection for a L4 project) and the actual impact of the MAS support.

> **Treatment group** – Businesses receiving MAS L4 support between 01/01/12 and 30/06/12.
>
> **Control group** – Matched businesses who have never contacted MAS.

Experiment B – L2 vs No MAS

The aim of this experiment is to find the GVA DiD between the December 2013 and December 2011 between businesses who receive a MAS L2 review, but no L4 grant, and matched businesses who have never contacted MAS. Self-selection bias will have affect the results of this experiment but advisor-selection bias will not as the treatment group are not selected for a L4 project. As a L2 review usually constitutes a single telephone conversation to decide on L4 eligibility, we do not expect the impact of MAS support to be very high (although this assumption is somewhat uncertain). For this reason we propose to treat the entire GVA DiD in this experiment as self-selection bias, which allows us to isolate the effect of this bias and use the results to analyse the other experiments.

> **Treatment group** – Businesses receiving MAS L2 support between 01/01/12 and 30/06/12, but no L4 support in 2012 or 2013
>
> **Control group** – Matched businesses who have never contacted MAS

Experiment C – L4 vs L2

The aim of this experiment is to find the GVA DiD between the December 2013 and December 2011 between businesses who receive a MAS L4 grant funded project and matched businesses who receive a MAS L2 review, but no L4 grant. As both treatment and control groups are subject to self-selection bias, this bias should not influence the GVA DiD. We expect the results to be influenced by advisor-selection and the actual impact of the MAS support. By looking at the results of this experiment alongside those of Experiment A, we can carry out a check on the impact of self-selection bias estimated in Experiment B.

> **Treatment group** – Businesses with MAS L4 support between 01/01/12 and 31/12/12, but not afterwards
>
> **Control group** – Businesses with MAS L2 service between 01/01/13 and 31/12/14, but no L4 service in 2012 or 2013

Experiment D – L4 vs later L4

The aim of this experiment is to find the GVA DiD between the December 2013 and December 2011 between businesses who receive a MAS L4 grant funded project in 2012 and matched businesses who receive a MAS L4 grant funded project in 2013. This is a "phased" approach, where both treatment and control groups receive the same intervention but at different times. In theory this experiment is very powerful as it controls for both self-selection bias and advisor-selection bias, leaving only the actual impact of the MAS support. However, additional bias could be introduced due to the fact that businesses have chosen to seek MAS support at a particular time. For example, businesses may contact MAS following a change of leadership or direction of strategy or at a time when they are "ready to grow" – i.e. have other measures/investments in place to coincide with MAS funding. This type of timing bias means that we might expect to see higher growth for the earlier treatment group in this experiment.

> **Treatment group** – Businesses with MAS L4 support with MAS L4 service between 01/01/12 and 31/12/12 but not afterwards
>
> **Control group** – Businesses with MAS L4 support with MAS L4 service between 01/01/13 and 31/12/13 but not before

Methodology for Experiments A and B

1. Label treated and untreated businesses and filter groups of businesses not to be included in the comparison (i.e. those who have contacted MAS but not received L4 or L2 support in the relevant period).
2. Match exactly on sector (SIC 2007, 3 digit) by splitting the dataset into 1,000 sub-datasets.
3. For each sub-dataset find pairs of treated and untreated businesses, matched on initial turnover, initial employees and birthdate, using two steps:
    a. Coarsened-exact match to restrict matching to the region of common support (the overlap between treated and untreated businesses).
    b. Nearest-neighbour match within this region to find the closest untreated business to each treated business.
4. Save matching information into each sub-dataset and recombine all the sub-datasets.
5. Using a "Kolmogorov-Smirnov test", compare overall distributions for treated and untreated matched pairs on employees and birthdate to check matching quality.
6. Using a "Paired T-test", compare distributions for treated and untreated matched pairs on outcome difference variables to get final results.

Methodology for Experiment C and D

1. Label treated and untreated businesses and filter groups of businesses that are not to be included in the comparison (as described above).

2. Find an untreated business, matched on initial turnover, initial employees, sector (3 digit SIC code) and birthdate, for each treated businesses. For these experiments we just use nearest-neighbour matching to find the closest match.

3. Save matching information into each sub-dataset and recombine all the sub-datasets.

4. Using a "Kolmogorov-Smirnov test", compare overall distributions for treated and untreated matched pairs on employees and birthdate to check matching quality.

5. Using a the "average treatment effect" tools in the nnmatch routine[25] to compare distributions for treated and untreated matched pairs on outcome difference variables to get final results.

[25] A. Abadie, D. Drukker, J. Leber Herr, G. W. Imbens, 2004, "Implementing matching estimators for average treatment effects in Stata"

Annex 3 – Matching results

Experiment B

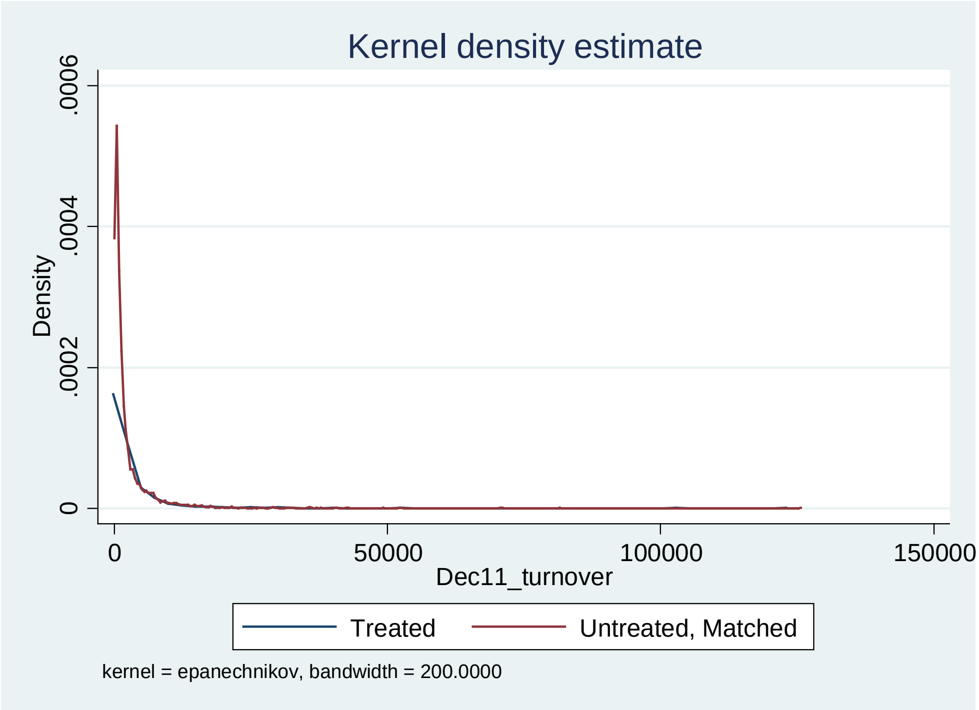

Figure B1 Turnover distributions for the treated and control groups, experiment B

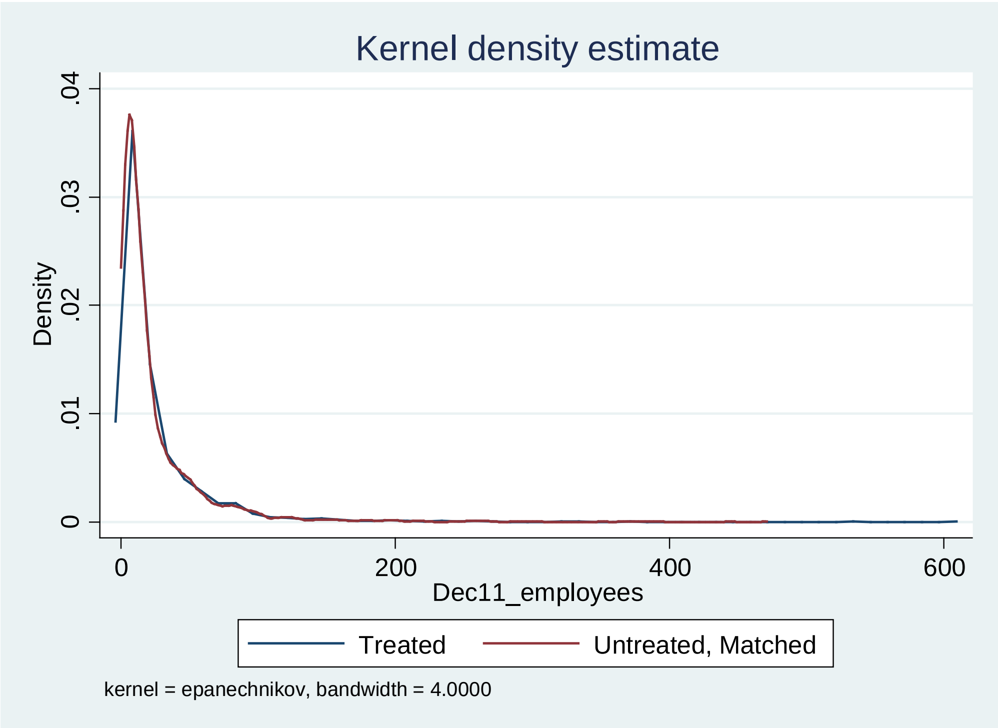

Figure B2 Employee distributions for the treated and control groups, experiment B

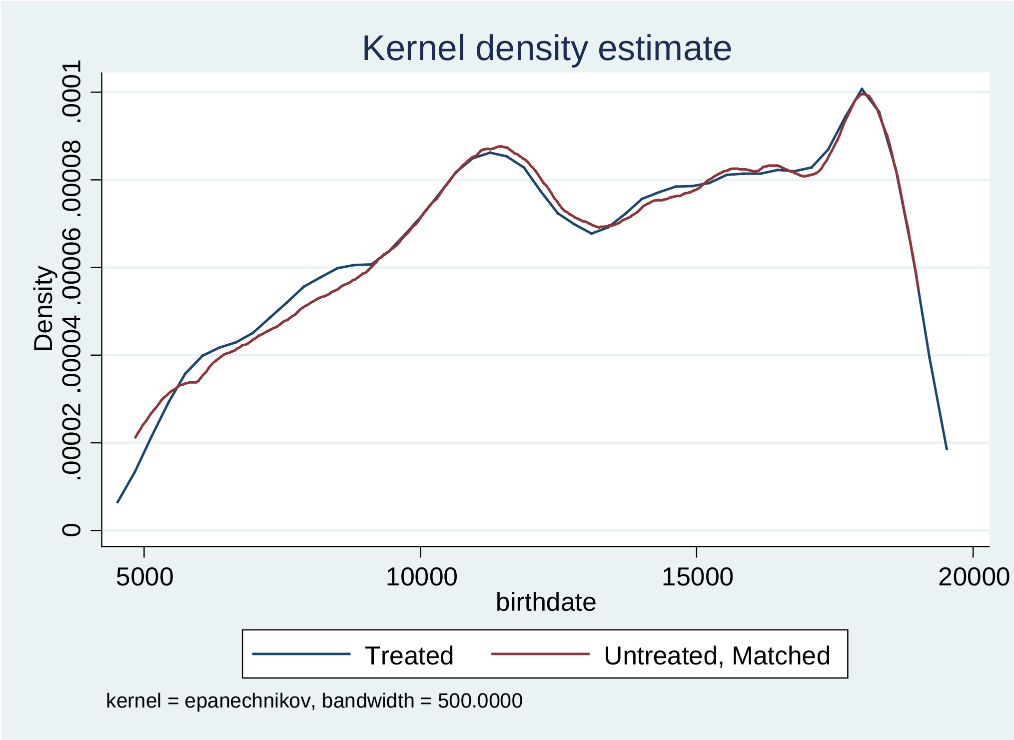

Figure B3 Birthdate distributions for the treated and control groups, experiment B

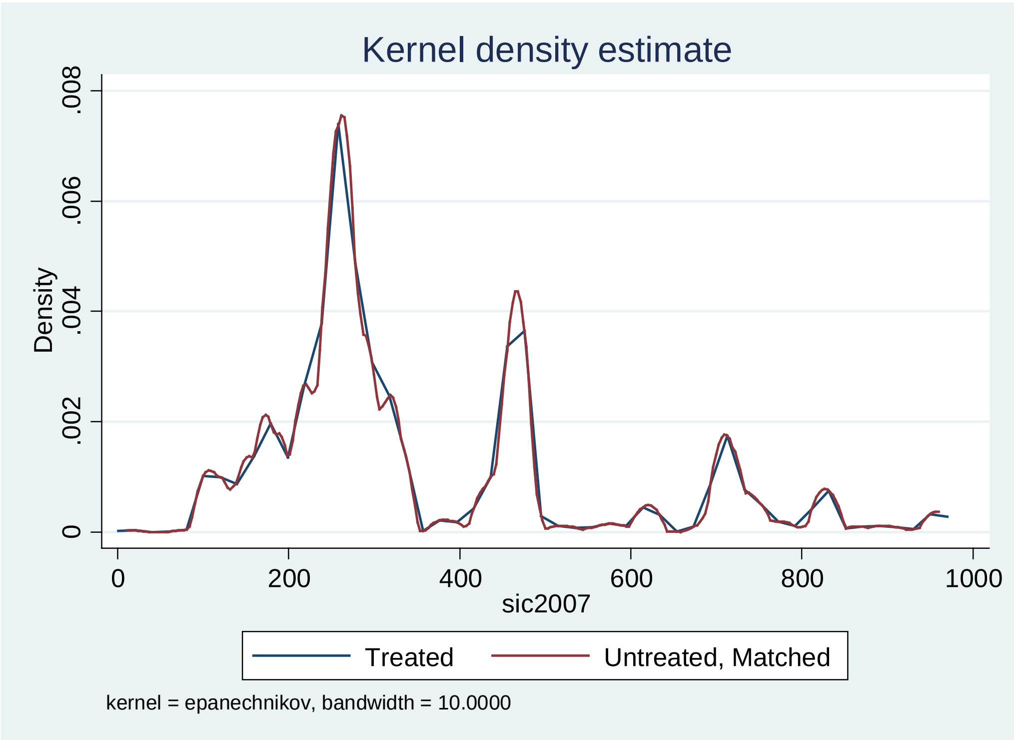

Figure B4 Industry sector distributions for the treated and control groups, experiment B

Experiment C

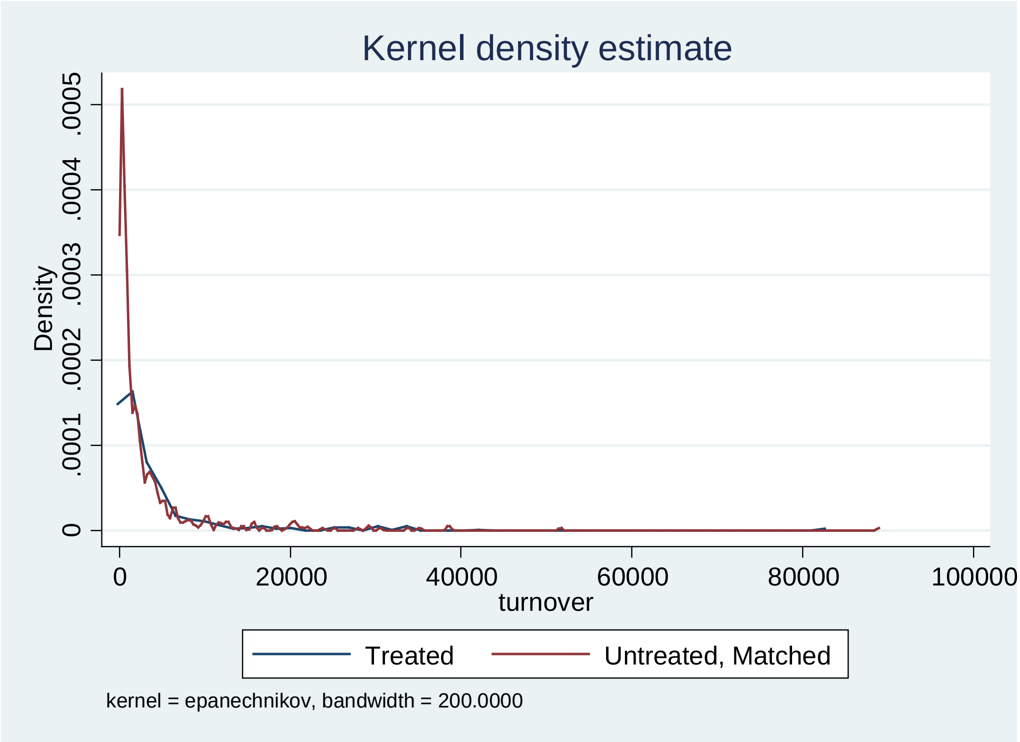

Figure C1 Turnover distributions for the treated and control groups, experiment C

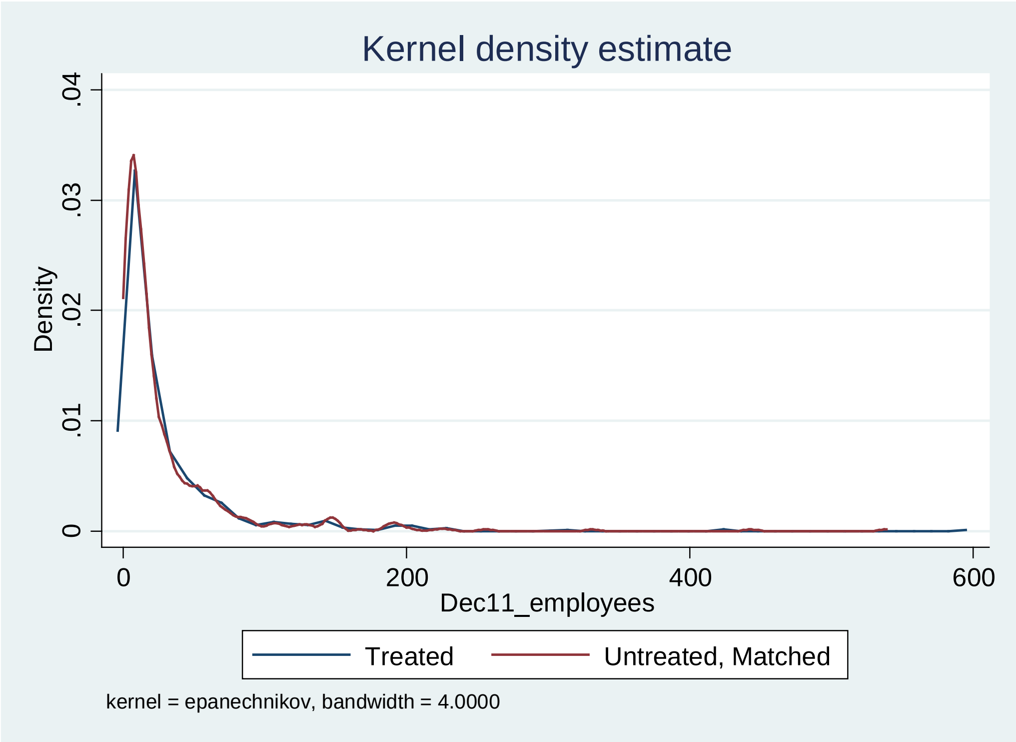

Figure C2 Employee distributions for the treated and control groups, experiment C

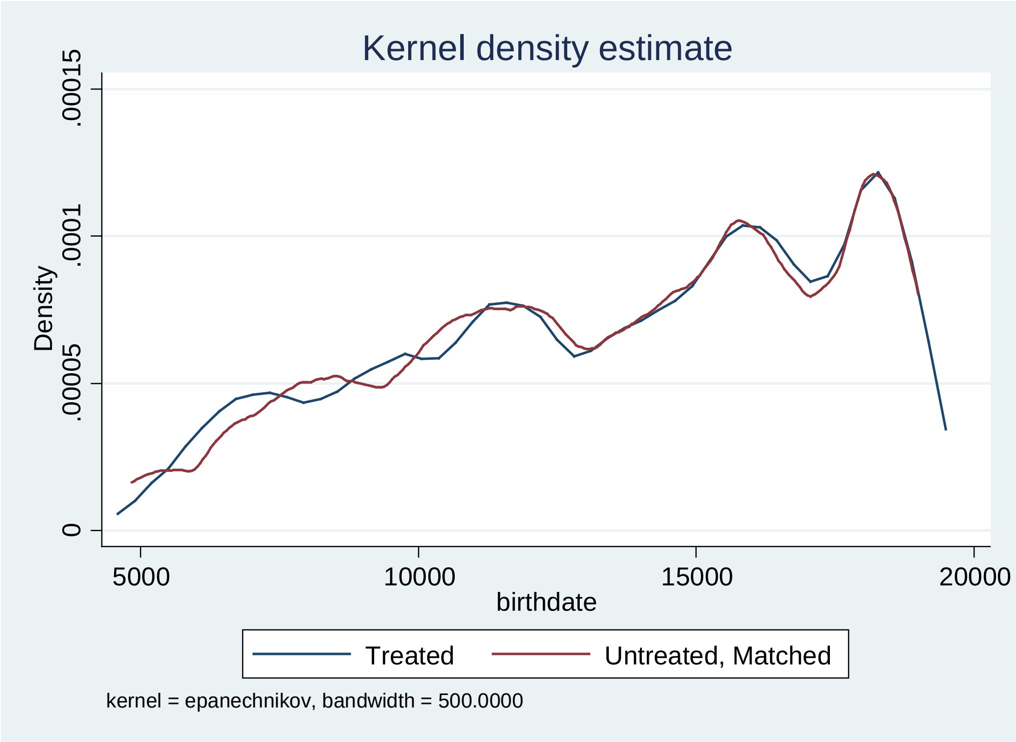

Figure C3 Birthdate distributions for the treated and control groups, experiment C

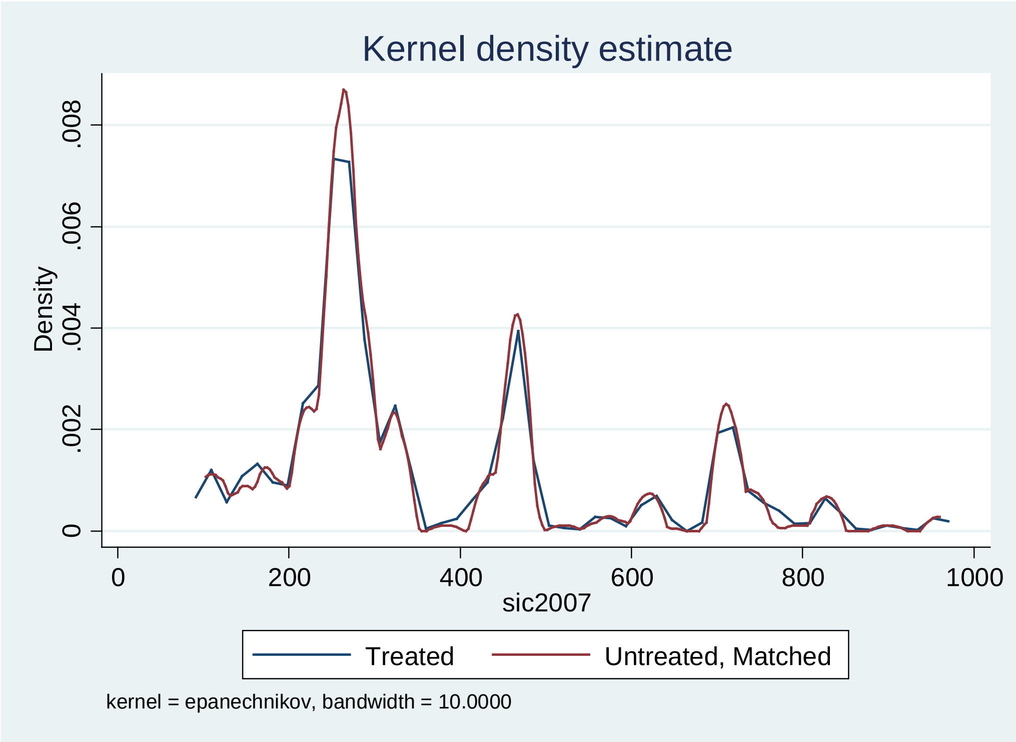

Figure C4 Industry sector distributions for the treated and control groups, experiment C

Experiment D

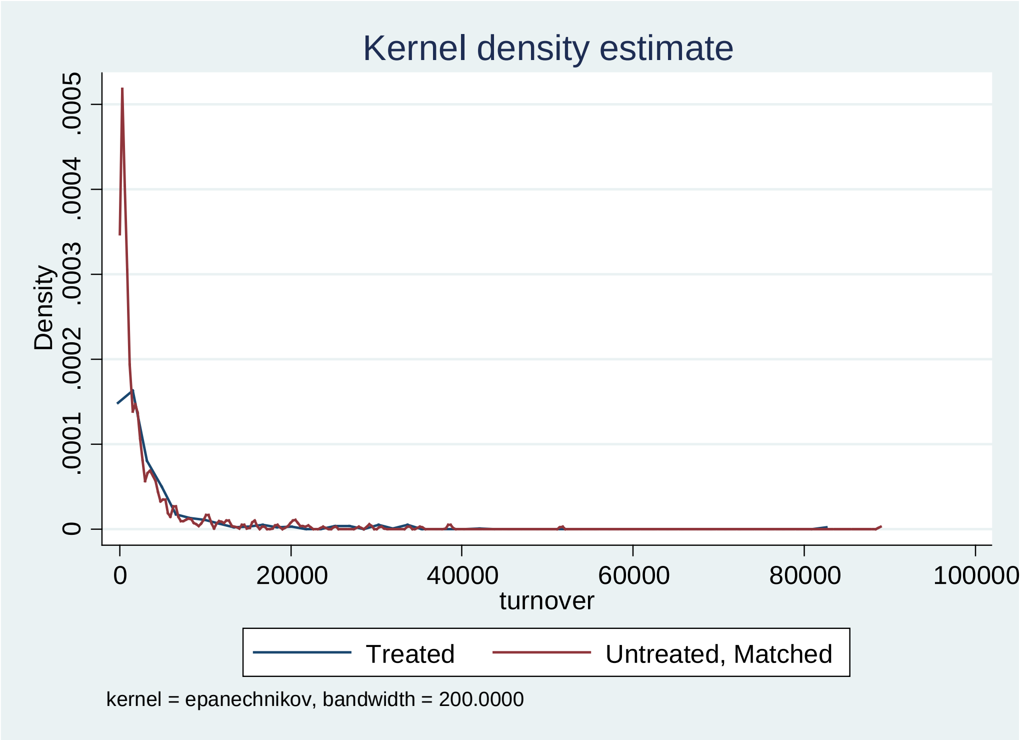

Figure D1 Turnover distributions for the treated and control groups, experiment D

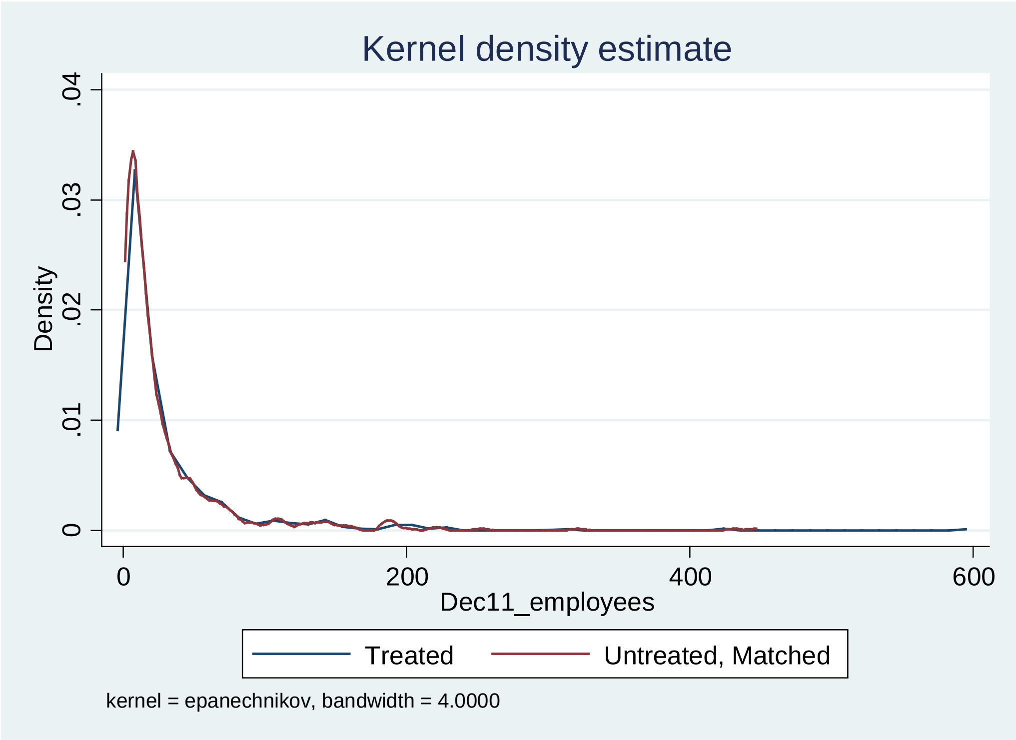

Figure D2 Employee distributions for the treated and control groups, experiment D

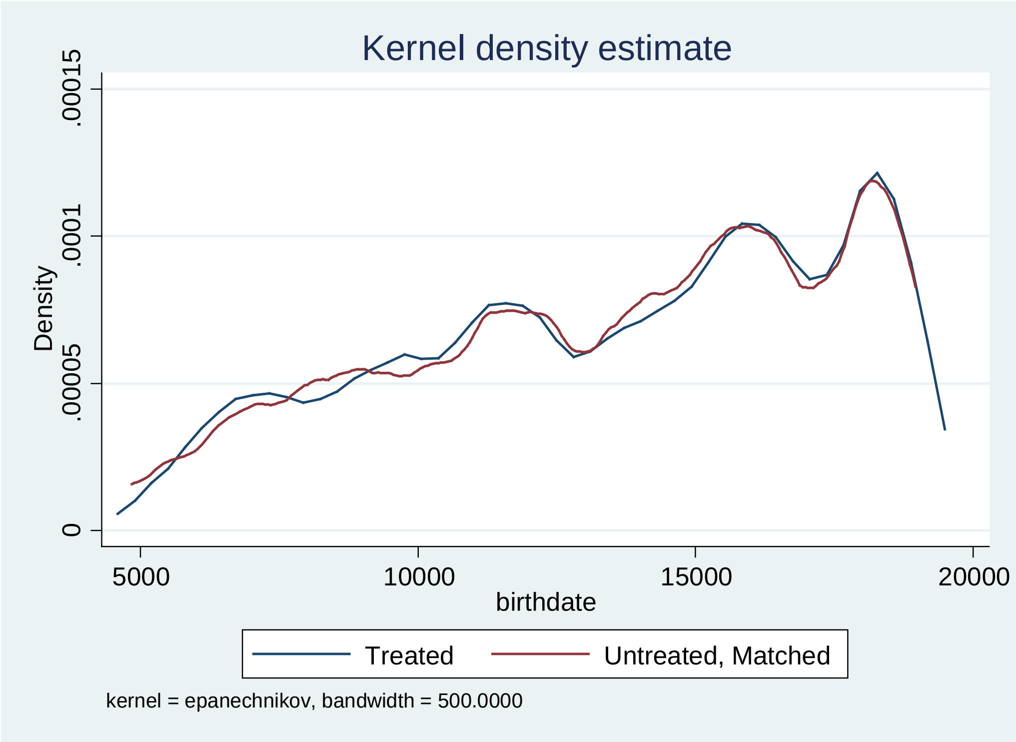

Figure D3 Birthdate distributions for the treated and control groups, experiment D

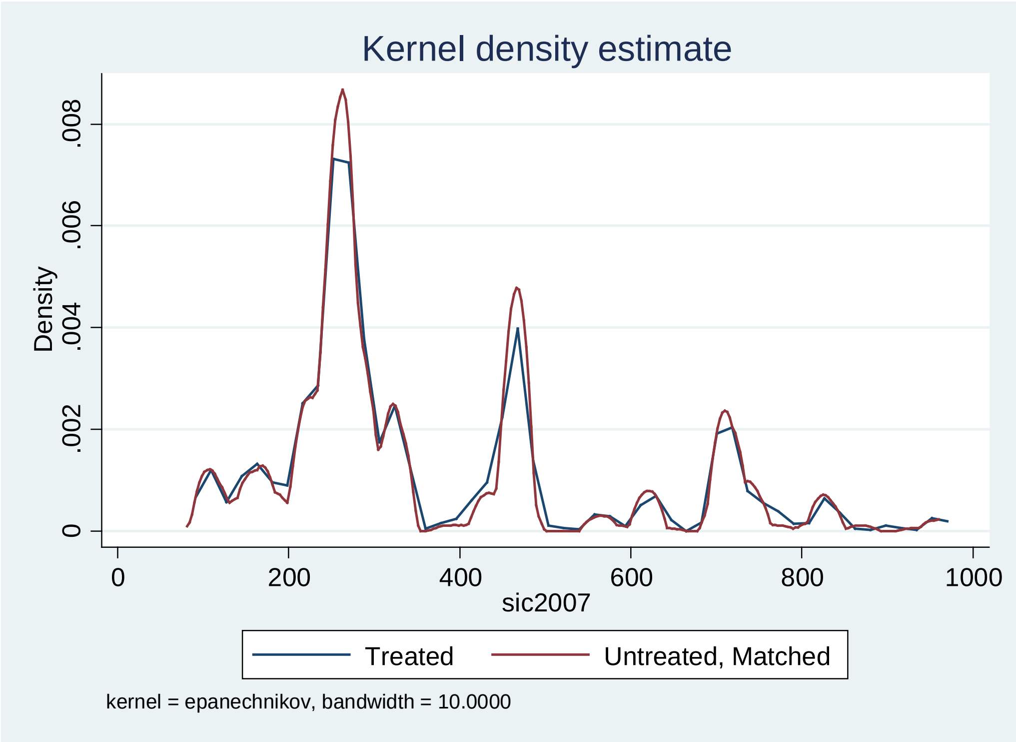

Figure D4 Industry sector distributions for the treated and control groups, experiment D

This publication available from www.gov.uk/bis

Contacts us if you have any enquiries about this publication, including requests for alternative formats, at:

Department for Business, Innovation and Skills
1 Victoria Street
London SW1H 0ET
Tel: 020 7215 5000

Email: enquiries@bis.gsi.gov.uk

BIS/16/7