PUBLIC SUPPORT FOR INNOVATION AND BUSINESS OUTCOMES

An Econometric Analysis

Research Paper for Innovate UK and National Measurement System

December 2019

Contents

Acknowledgements 2
Executive summary 3
1 Background to the Study 9
2 Study Approach and Data Sources 16
3 Understanding the Treatment 27
4 Impacts on economic activity 46
5 Understanding Productivity Impacts 74
6 Conclusions 96
References 98
Annexes 103

Acknowledgements

This work contains statistical data from ONS which is Crown Copyright. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

The study was conducted by Belmana Ltd, with the University of Middlesex. Researchers involved in the study were Prabhat Vaze, Sarah Aggarwal, Carolin Ioramashvili, Andrea Jansson, Ralf Martin and James Derbyshire.

The authors acknowledge the helpful comments and suggestions throughout the study from the project manager at the National Physical Laboratory, Mike King, and the steering board for the project: David Legg, Dan Hodges, Ian Alexander, Jose Arudo and Geeta Nathan (Innovate UK) and Benjamin Marriott (BEIS). The report also benefitted from comments made during a seminar at the Department for Business, Energy and Industrial Strategy and from two anonymous referees.

Executive summary

Innovate UK invested over £2.4b in more than 20,000 business innovation projects since 2004, with growing amounts of spend since 2009. The largest portion of the Innovate UK grant funded projects related to collaborative R&D. The National Physical Laboratory support, covering the period from 2002 to 2017, has over 59,000 instances of support completed by 2017. The interventions include free website downloads, contracted collaboration with NPL and paid-for research or measurement services. Further, the National Physical Laboratory is part of the National Measurement System (NMS) with partner bodies – the National Engineering Laboratory and the Laboratory of the Government Chemist – providing independent measurement services to innovative businesses.

This report presents results measuring the impact of the business support provided by Innovate UK and the NMS. It primarily uses data derived from the administrative systems in each organisation about the support provided by each body to businesses. This is then linked to firm-level data held in the ONS Secure Research Service, primarily the business data about employment, turnover and the results of surveys conducted by ONS as it compiles economic statistics on productivity, research and development, and innovation.

The research seeks to answer questions emerging from earlier work (BEIS, 2017), which focused on the employment growth seen in businesses supported by innovation funding from Innovate UK or the National Physical Laboratory. A key part of the new work is to explore what additional analysis can be undertaken by augmenting the ONS Business Structure Database used in BEIS (2017) with other public and survey data about businesses.

The approach

The research considers impact by comparing supported business performance with groups of similar unsupported businesses. Firm level characteristics are used to match the supported businesses to similar businesses within the ONS microdata, providing a counterfactual. The research has used additional data compared to BEIS (2017) to improve this matching, for example using indicators of innovation such as number of businesses holding patents.

Box A indicates the impact measures. The study has looked at employment, survival and turnover impacts of innovation support. The analysis has then looked at a wider set of impacts, including pay levels in the supported businesses and the quality of the patents.

BOX A: Measures of innovation support impact

Impact evaluation centres on measuring business performance, tracking both the changes seen in the supported businesses after an intervention and then comparing this to the performance of comparable businesses. The measures used in the study are described in this box.

The employment impacts are measured using the number of jobs in a business. The data sources provide annual snapshots of employment at a point in time and any annual change in this indicates the increase or decrease in the number of years of employment – or job years. If an increase of one job is seen in a business each year after support then, over three years, this equates to six job years, with the first year's additional job lasting three job years, the second two years and the new job in year three adding an additional single year. Such a measure is used in other studies (Grant Thornton et al 2013 looking at the London 2012 Olympics impact).

It has some advantages. It allows for the growth over time, as support causes a ramp up in activity. Also, there are useful economic measures consistent with the job year. BIS (2014) combined an annual estimate of GVA per job with job years estimates to value impacts. Recent Green Book guidance on measuring employment impacts has also highlighted how the additionality of employment is less likely in periods of full employment, and the job years measure is able to locate additional employment by when it occurs (HMT, 2019).

There is one key issue with the employment measure that this study had to tackle. Businesses benefiting from support could only be identified at the enterprise level, which could cover many different plants and establishments. Many of the UK's most innovative businesses are large, complex, multi-unit businesses. Linking any support to the whole enterprise will mean tracking employment across all establishments, which sometimes will mean the changes in employment seen will be imprecise. Adjustments are made to tackle this.

The ONS data from the Inter-Departmental Business Register (IDBR) tracks employment at points in time, but also the turnover over a year of a business. Such revenue data is deflated using two-digit SIC industry output deflators and provide an indicator of the growth in valued added. The real turnover growth is also divided by employment growth to provide a labour productivity metric. This is used in other studies (such as Goldman Sachs, 2019), because the measure tracks value added per employee. However, there are some known limitations for this proxy, such as missing improvements as less resources are used in producing the same level of sales.

A second proxy for productivity is the wages paid per week, as reported by businesses to the ONS in the Annual Survey of Hours and Earnings (ASHE). The pay bill takes a large share in business value added. An individual's wage is also a measure of the value of labour at the margin. Further, wages per week are well measured, especially as ASHE tracks the same jobholder over successive years so more robustly measuring any pay changes. The main issue with this measure – as it derives from a 1% sample of employees – is the coverage of employment, especially in SMEs.

The impact measures for innovation centre on the whether a business is R&D active or not, an indicator produced by ONS as it develops the sample frame each year for the Business Expenditure on R&D survey. ONS links together nine different sources to find businesses that are R&D active (Ker and Greenaway, 2012).

An impact measure to look at whether supported businesses invest in ideas that subsequently spread is derived by looking at whether businesses patent their innovations and how widely they are then cited. During patenting, inventors will cite the relevant previous patents that influence their intellectual property (thereby demonstrating an idea's novelty in departing from the cited patents). A simple measure is how often a patent is subsequently cited. The study also uses a citation index constructed for each patent. It looks at the influence of a patent by combining the citation index of the patents that then cite the patent being measured. This captures patents that may be modestly cited directly, but then have led to innovations that became highly influential. The impact measure then measures whether the supported business produce ground-breaking ideas.

Any difference in performance seen in the beneficiary businesses has, in some cases, then been compared to any change seen in comparable groups. Adjusting economic

impacts in this way permits an assessment of “deadweight”, in other words, what would have happened in the absence of the innovation funding. The analysis therefore allows an attribution of impacts to the support. In this study, the impact analysis can attribute to innovation support:

- Employment growth and real turnover growth, as the study has been able to track performance in these measures in supported businesses and then adjusted this for that seen in comparable unsupported businesses so that support effect can be isolated
- Survival impacts on employment recognising that support has protected jobs in firms receiving innovation funding, but also sometimes resulted in businesses surviving that otherwise ought to have closed.

However, this form of analysis has several challenges. A first is that, while a good picture of economic activity changes can be established, the extent to which productivity has changed due to support is hard to disentangle directly. A support measure with productivity impacts would need to estabish any changes in outputs per unit of input due to the intervention. Changes in the real turnover per employee could provide some evidence of this, but this will only proxy for productivity (i.e. value added per employee). For example, where innovation reduces the use of inputs (materials or energy used in production), the turnover based measure will miss this. There are also other measurement issues, such as whether innovation spillovers or improvements in the quality of outputs can be measured, which may be a key outcome of research and development.

The study has however then been able to widen the productivity proxies, analysing new measures quantifying using summary statistics about:

- The quality of the additional employment generated through support indicating this through wage premiums as people move between supported and supported businesses
- The quality of innovation outputs, especially patents, again looking at the average citation rank of patents from supported businesses compared to unsupported.

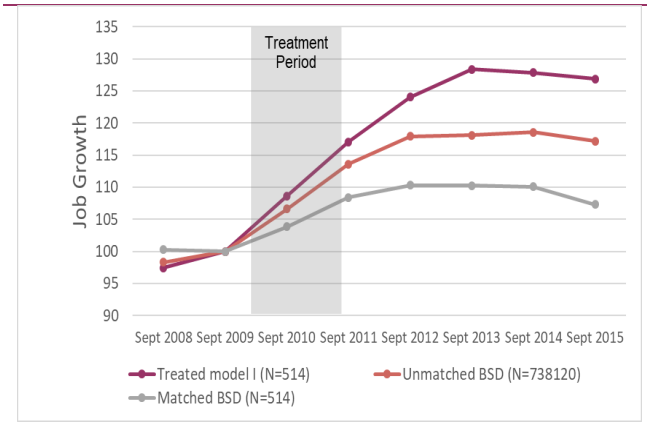
Finding a comparison group is difficult, as statistical matching needs data about unsupported businesses that are innovative and might be suitable for innovation funding. This study uses a multitude of methods and datasets to identify comparator businesses and then triangulates economic impacts to understand the contribution of innovation funding to the economy.

Study findings

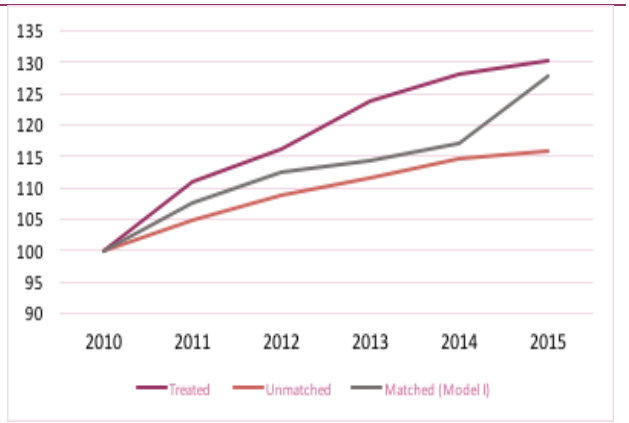
Overall, there is evidence of support increasing employment. Figure 1 shows how employment for one cohort of Innovate UK businesses show higher employment growth when compared to a comparison group; while the right panel shows a divergence in employment growth seen in NMS supported businesses above that seen in comparable

businesses. Other measures of economic activity, such as patenting and business survival, also increase in supported businesses.

Figure 1: Employment growth in supported businesses



Innovate UK treated in 2010



NMS – All regularly supported

The analysis undertaken builds on the BEIS (2017) work, differing in some ways to that analysis and then adding additional datasets to provide a wider range of analysis. The main departure from the earlier study is that impact effects have been assessed adjusting for businesses that receive support, by analysing a supported business from the first year of treatment. This allows the analysis to look at the level of additional economic activity, as well as the growth focus of the earlier work. To provide an analysis of the level of economic activity change due to support, another methodological departure is to include the business survival impacts of receiving support.

A general finding is that, while both employment and real turnover are growing, there is no evidence of significant impacts on the turnover per job. This is often seen as a proxy for productivity, but this measure has some known issues. For example, it misses productivity improvements though firms using fewer inputs while maintaining sales; also, sales growth may miss other changes due to innovation, such as increased quality of a product.

The research then has used new measures of the next stages in the productivity story. The analysis indicates there is a wage premium associated with jobs in supported businesses, an indication of high value jobs. Further, a new dataset has been used to analyse spillovers. Patents of supported businesses are more likely to be cited than other patents, and this is also the case using a new measure used for this study, which takes account of indirect citations.

Findings about Innovate UK support

Businesses supported by Innovate UK saw employment growth after support, creating around 170,000 job years, an increase of 7% in their economic activity, from 2009 to 2017. This is lower than the 11-14% estimated in BEIS (2017, p.49) because that growth estimate focuses on SMEs, excluding the large businesses and so reducing the denominator in growth estimates. This covers £1.2bn of funding made to businesses from 2010, so implying an expenditure of just over £7k per job year. The box on impact measures explains the job years concept. This job growth is broad based but driven by large businesses participating in collaborative R&D with grants of two years or more in length.

- Of the 170,000 years of employment generated, around half was also seen in a comparison group of businesses. Figure 1 indicates a single year but overall, 49-70% of the employment seen in the supported business was not observed in comparable businesses. This translates into about 101,000 of the job years seen in the Innovate UK supported businesses being additional.
- Investigating those receiving only one instance of support over the years, from 2009 to 2012, around 30,000 job years were created by around 2,000 businesses.
- For businesses receiving an R&D collaborative project, out of the 67k job years created in these treated businesses in the year 2010 to 2016 it is estimated that 44k were additional.
- Investigating the optimal dosage of support, all grant lengths have a positive impact with an optimal point being around two years in length.
- Support enhances businesses survival, with the comparable group of businesses seeing almost twice the number of business exits as the supported.
- Supported businesses are 2-6% more likely to patent post support than similar businesses.

Evidence of economic activity increase has then been complemented with analysis of indicators of productivity impacts:

- Real turnover in Innovate UK supported businesses has risen by 35-60% by 2015, with those treated first in 2011 achieving the highest growth. In the matched control groups, growth is 15-38%.
- As the productivity analysis uses a survey focusing on the largest businesses, the businesses supported by Innovate UK that have productivity data exclude the SMEs. Given this bias, the businesses supported by NMS and Innovate UK in this analysis are similar, in being large and with a high level of capital stock.
- Decomposition of labour productivity of Innovate UK supported manufacturers suggest – rather like the wider business population – there has been modest productivity growth, but there is evidence that for the supported businesses the growth in employment is correlated with productivity, i.e. labour resources have been reallocated towards more productive businesses.
- Innovate UK businesses register patents with higher average knowledge spillovers, particularly higher knowledge spillovers within Great Britain.

- Employees moving into Innovate UK supported businesses receive a £4 a week wage premium, but an even larger premium after leaving highlighting a gain in skills from working in Innovate UK supported businesses.

Findings about NMS support

Around 23,000 jobs years of employment have been created by 175 businesses who sought regular support from NMS from the period 2009-2017, an increase of 5.5% in economic activity, lower than the BEIS (2017) estimates of 10-15% again because of that study's focus on SMEs. Businesses supported infrequently by NMS showed a decrease in economic opportunity over the timeframe.

- Around 18,800 job years are additional as this is growth not seen in comparable businesses. Around 80% of job growth seen in NMS supported businesses is not seen in the comparator businesses.
- Survival rates for regularly supported are high with only 4% of businesses closing, compared to over 35% of the wider BSD, and 12% of a similar comparison group. After controlling for business size, the NMS supported businesses are more likely to survive.
- A second analysis has then been performed to understand the impact on job years as survival has been impacted by support. The additional impact is 4,404 job years.

Evidence of economic activity increase has then been complemented with analysis of indicators of productivity impacts:

- Real turnover growth in the businesses receiving NMS Support in 2012 is presented. There is a quick divergence in real turnover growth between the supported businesses and the matched and unmatched controls.
- Decomposition of labour productivity of NMS supported manufacturers suggest – rather like the wider business population – there has been modest productivity growth; the decomposition however indicates resources are shifting towards more productive businesses with employment tending to move to productive businesses.
- In general businesses supported by NMS generate patents with higher spillovers, with those spillovers increasing after 2001.
- Employees moving into NMS supported businesses receive a £50 a week wage premium, with the premium remaining significant after controlling for age and occupation of earners.

1 Background to the Study

This report presents results measuring the impact of the business support provided by Innovate UK and the National Measurement System. It primarily uses a dataset of the support provided by each body linked to firm-level data held in the ONS Secure Research Service.

The framework is grounded in approaches used to evaluate government policies (see HMT, 2011 and 2019; Innovate UK, 2018a). The next sections outline the original research questions for the study and then repeats some of the results from the earlier work analysing the two organisations' data.

The research seeks to answer questions emerging from earlier work, which focused on the employment growth seen in businesses supported by innovation funding from Innovate UK or the National Measurement System. A key part of the new work is to explore what additional analysis can be undertaken by augmenting the ONS Business Structure Database (BSD) used in BEIS (2017) with other public and survey data about businesses. The study has also benefitted from additional years of business data.

The research questions focus on issues that BEIS (2017) identified:

- **In identifying the counterfactual, can the matching to large firms be improved and robustness of the treatment/control comparison increased?** The framework centres on constructing more variables to understand selection, increasing the opportunities to identify a counterfactual. A key research question is what characteristics make uptake of support more likely and steps to address this included using past innovation funding support, past R&D activity and whether the business holds a patent in modelling support uptake. This helped to match the large, R&D intensive businesses more accurately.
- **Can the impact of business survival and differences observed between supported and comparable businesses be measured?** The framework integrates specific modelling to understand the survival of supported businesses in terms of a control. It then uses the differentials seen to adjust impacts to take account of differentials in survival rates in supported businesses over comparable businesses.
- **Can impact measures be improved?** The framework uses more of the productivity datasets in the ONS Secure Research Service. This introduces new variables on impacts, such as wage premiums and sales increase. With the ONS ARD dataset, the impacts can be widened to R&D, value added and investment, though sample sizes can be quite modest. Also, a new dataset about the spillovers from patents has been integrated into the analysis.

The rest of this chapter briefly describes the support data used for the study provided by Innovate UK and NMS. The chapter then looks at the logic models for the support

provided, with the models describing how the innovation support provided by the bodies translates into impacts.

UK Innovation Funding

Several key datasets are used in the econometric analysis: the ONS Business Structures Database (BSD), the Annual Respondents Database (ARD), the Business Enterprise Research and Development (BERD) survey, Intellectual Property Office lists of trademarks and patents. However, the core datasets have been about the beneficiaries of funding from Innovate UK and the National Measurement System.

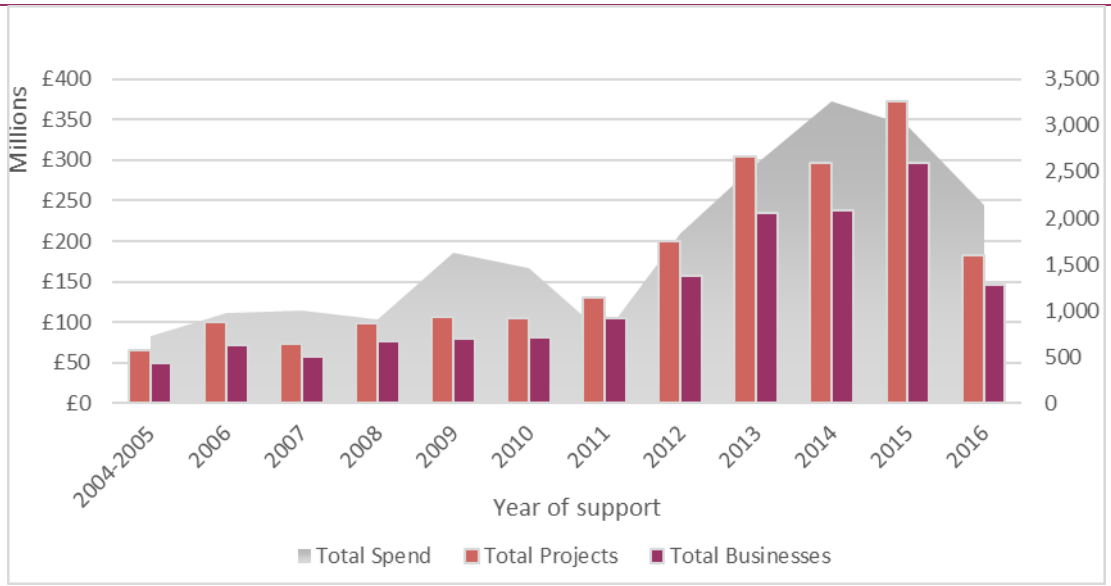
Innovate UK data

Innovate UK invested over £2.4b in more than 20,000 business innovation projects since 2004. A public dataset lists all the projects with a row for each participant, detailing the project, the amount, the project timings and from which Innovate UK competition the award was made (Innovate UK, 2018b). The data goes back to projects started in 2004, with the largest focus and amount of spend since 2009 onwards. Figure 1.1 shows the number of matched businesses per treatment year, their total grant spend in £m and the total number of projects each year for 2004-2016.

The figure focuses on the support provided by Innovate UK to businesses matched to ONS enterprise reference numbers within the Secure Research Service. Innovate UK supports R&D investment projects with grant funding and funds knowledge exchange interventions. The projects that are funded are conducted by a single entity or are collaborative; they can involve a mix of businesses and other bodies. This study focuses on business impacts, thus excludes the incidences of financial support where only non-business participants are involved (such as funding for Universities only). Excluded from this study are incidences of support where businesses are involved in Innovate UK knowledge networks. Details of the missing, unmatched spend are in the annex.

The year of support indicates the start date of the project. The spend is the actual amount Innovate UK spent, though the dataset provides the total amount allocated. While all grants require businesses to contribute to the total cost of the projects, this figure does not include the contribution, focusing only the Innovate UK funds spent.

Figure 1.1: Businesses, projects and total spend for Innovate UK 2004-2016



The figure exhibits two periods, with a notable change upwards after 2011. The rise in total support in the second half of the period reflects several drivers. A key one is that the completeness of the data increases after 2010, with the first half of the period including fewer of the government innovation programmes. The dataset was compiled relatively recently, and many older programmes' management information were not accessible to the analysts compiling the data.

The Innovate UK data lists the individual incidences of financial support. It includes a field containing the name of the support. In total, the dataset covers 23 products. The product descriptions, associated documentation and discussion with Innovate UK has been the basis for grouping the products, and hence condensing the incidences of support into three main types

The largest portion of the Innovate UK grant funded projects relate to collaborative R&D. This categorisation was relatively straightforward, with product descriptions and the Innovate UK data reflecting the collaborative nature of the funded projects. Innovate UK funding seeks to increase economic growth through business-led innovation, with collaborative projects stressing cross-organisation sharing of ideas, often between universities or other research bodies and businesses, and spillovers.

The second type of support covers the remaining products that provide significant grant funding. Here the funding often supports R&D that matures a product towards commercialisation, aligned to the readiness of a technology for market. There are Innovate UK products for different stages of innovation (e.g. "proof of concept"). This type of project has impacts directly on the beneficiaries.

In this second type of support is the innovation vouchers. A voucher is firm level support. However, they are relatively modest amounts of funding primarily for businesses to procure innovation advice. The impact analysis in this report has not included the vouchers, given the low size of the funding is unlikely to have material impacts. Where businesses access this support but then receive project funding subsequently, the business will enter into the impact analysis.

The third type of Innovate UK intervention – knowledge transfer – is usually not funding a research activity but provides the support systems needed to prepare businesses for innovation challenges. This type of support was relatively easy to differentiate because the product titles and description highlight the networking nature or skills development that is funded. This third type of support has not been included in the impact analysis in this report.

National Measurement System data

The National Measurement System dataset contains support data for firms from 2002 to 2017, with over 59,000 instances of support completed by 2017. The dataset includes three different possible interventions: free website downloads, official collaboration with NMS involving a written contract, and paid for contract research or measurement services.

The UK's national measurement infrastructure ensures a robust system of measurement and forms an essential component of being part of a global economy. At its core, it ensures that measurement in the UK is consistent with the global common system of measurement units, the International System of Units (SI). However, there are many areas of measurement such as in chemistry, biology and food science which are not currently directly linked to the SI system of units. Primary measurement underpinning these areas is established and disseminated through reference methods and/or materials. These areas are often where new measurement knowledge needs to be developed to tackle emerging needs.

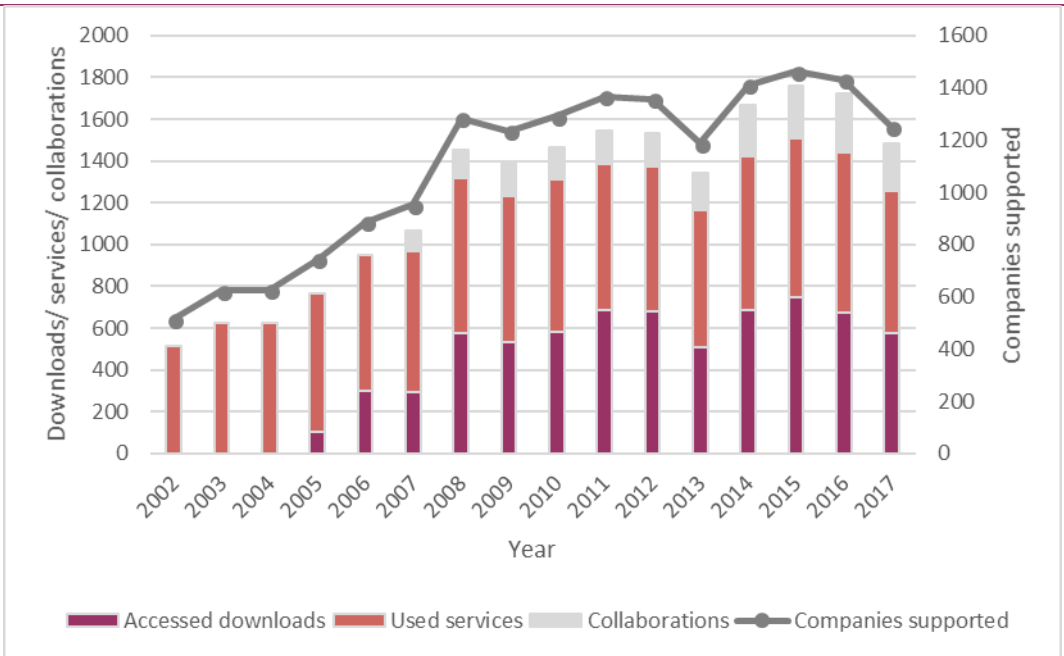
In most countries, there are designated institutes delivering specific measurement capabilities and are recognised internationally as the lead measurement organisation for a particular physical or other quantity. In the UK, the National Physical Laboratory (NMS) works in partnership with five designated institutes: LGC, formerly the Laboratory of the Government Chemist; NEL-TUV, formerly the National Engineering Laboratory; NGML (National Gear Metrology Laboratory); NMRO (National Measurement and Regulation Office); and NIBSC (National Institute for Biological Standards and Control).

NMS products and services for commercial partners and beneficiaries centre around measurement instrumentation. Standards for weights, measures, time, the quality of materials, etc have underpinned the functioning and regulation of markets and NMS is the national centre for this with an international standing. The Lab develops, maintains and allows use of technical infrastructure and this is a key service for businesses testing products. Businesses will pay for these measurement services, though the funding of the

NMS integrates a degree of subsidy/scale returns. Beyond this, the Lab provides training, enhancing skills of customers' workforce.

Where measurement challenges require it, NMS will collaborate with businesses to ensure their instruments, inspection systems and services are appropriate. Especially with innovative sectors and the businesses at the edge of technologies, collaborative research can provide NMS with the knowledge to make advances in their technology. Equally, the business partner can help to shape the measurement technique. Recent examples include work in geo positioning accuracy.

Figure 1.2: NMS businesses supported and types of support 2002-2017



Similar to Innovate UK vouchers, businesses that only accessed the free website downloads – data has been collected about this since 2007 – are not analysed within the report due to the lack of material impact expected. Instead, accessing downloads are used as a characteristic to identify similar businesses for comparison. Figure 1.2 shows the number of companies supported each year, and a breakdown by the three interventions. Unlike Innovate UK projects, the NMS interventions are generally completed within the year of the start date.

Logic of Innovation Support

The Innovate UK products once categorised can be analysed in terms of intervention logic. This is necessarily at a high level, synthesising across a diverse range of interventions and seeking to define generic models for each type.

Innovate UK and NMS Collaborative R&D

The first type of support funds collaborative research and development (CR&D), funding intended to help businesses, Higher Education Institutions (HEIs) and other bodies to explore jointly the technology issues and commercial feasibility of innovative ideas and develop prototypes. This type covers many Innovate UK products: Collaborative R&D, Legacy RDA Collaborative R&D, CR&D Bilateral, Feasibility Studies, BIS-Funded Programmes (removing centres), Fast Track, Large Projects and EU-Funded projects. It represents about 22,446 projects, worth £2.9bn. NMS undertakes a number of research studies also, often collaborating with businesses and research bodies about measurement issues and instrumentation related research.

Wider benefits into the innovation ecosystem is more likely in these projects because of the collaboration. It is expected that this will result in higher spillover effects than in support without a collaborative focus. Also, collaborations with Universities or the wider stakeholders make spillovers more likely. Indicators also look at the transmission mechanisms available. These are linked to networks of researchers.

Innovate UK Firm-Level R&D

This type of support focuses on individual businesses, where R&D enables innovation, but usually with a focus on developing products for the market. The rationale for such investments is usually that R&D tends to be underprovided by the private sector due to positive externalities that are not captured by market participants.

For Innovate UK, this type covers many products: GRD Proof of Concept/ Development of Prototype/ Proof of Market; Study; Legacy RDA Grant for R&D; Launchpad, and potentially SBRI. This type represents about 6,910 projects, worth £374m when including Vouchers. Vouchers represents 3,516 projects worth £12m and are not included in the impact analysis.

Projects are awarded grant support to fund research and innovation activities at different stages of the R&D process. This ranges from experimental studies at early stages in the development through to market testing and development of prototypes. This is expected to impact the time, after support, when outcomes and impacts will be visible. Where support is for projects nearer to market, business performance impacts should occur sooner after support than where R&D is at an earlier stage.

Categorising National Measurement System support

The National Measurement System provided a dataset of beneficiaries from 2002 to 2017, split into three categories of support: web downloads, invoices for services, and formal collaborations (above). Outside of this split, years of support and organisation type were provided. NMS works within a network of measurement service providers and further client level data was made available about the users of the Laboratory of the Government Chemist and the National Engineering Laboratory.

The Innovate UK logic models have been customised to the NMS beneficiaries. But there are some key differences in the types of support given to NMS businesses. The supported businesses are broken into several categories, primarily by intervention type: Collaboration, paid services and free online downloads. The businesses are divided into these categories based on the NMS intervention data with many firms overlapping across different categories. The basis for the division is the expectation that the three types of interventions have different timeframes and levels of outcomes.

Businesses that participate in multiple types of intervention are placed in the group of the more intensive support with the order above representative of most to least intense, somewhat representing dosage. Businesses in collaborations with NMS are put in a category regardless if they use the paid services and free downloads as well. This is because the collaborations are a significant intervention than the other two. As well, those using paid services and downloads are placed in the paid services group because of a similar logic.

This study

The two innovation support bodies – Innovate UK and NMS – have a long history of supporting businesses as they undertake R&D. Looking at how the support impacts businesses, in terms of the logic of the interventions, highlights some further considerations for this analysis. Firstly, there is a significant level of funding being provided – especially by Innovate UK – directly to businesses. The amounts suggest that impacts should be measurable and that the direct impacts on businesses should provide a starting point for the overall impact analysis. As R&D impacts generally occur with lags, as it takes time to turn ideas into commercially successful products and services, this impact analysis can capitalise on the long historical dataset.

The intervention logic then also highlights wider impacts. Innovation funding anticipates benefits beyond the directly supported business, encouraging collaboration with others and spillovers beyond the supported businesses.

This chapter also limits the scope for the study somewhat. Firstly, the focus is support for businesses and this removes a significant range of intervention, such as support to university research. Secondly, some particular products – such as innovation vouchers and knowledge networks – are not considered as support of sufficient size to expect impacts that would be measurable. Equally, for the NMS, the clients that download NMS materials may see impacts, but these are likely to be modest. However, in such cases, the businesses benefitting may be useful for the purpose of comparing the supported businesses.

2 Study Approach and Data Sources

This chapter describes the approach to analysing the impact of Innovate UK and National Measurement System support, primarily comparing the performance of supported businesses with that of comparable businesses identified through statistical matching.

The approach taken in this study is to compare treated businesses before and after support with a sample of non-treated businesses that are as comparable as possible to those that were supported. The approach identifies comparable non-beneficiaries using statistical matching. The increase in performance seen in the beneficiary businesses can then be compared to any change seen in the comparable group. Adjusting economic impacts in this way permits an assessment of "deadweight" or, in other words, what would have happened in the absence of the innovation funding.

A constraint on statistical matching is that it can only be undertaken on measured or observable characteristics. The most commonly used observable characteristics are the size, industry and location of businesses, variables that are a good start for the matching. The study uses the Business Structure Database (BSD), an annual snapshot of all UK businesses based primarily on tax registers that provides these variables for all employers and economically significant non-employers in the UK. The BSD is then enhanced by linking it to other firm-level data, such as the registers for intellectual property and ONS surveys.

Once a control group is identified, the BSD also allows tracking of business performance over time. BSD is a panel, providing turnover and employment each year and showing whether a business remains active. It is thus possible to analyse the changes seen in both the supported businesses and the comparable businesses identified through matching and estimate differences in growth seen. This approach is called difference-in-differences and has the advantage – if matching is successful – of understanding changes controlling for selection issues associated with support such as unobservable characteristics. For survival, the study focuses on comparing the rates of firm closure between the supported firms and the control.

Approaches to finding a Counterfactual

There is a large literature evaluating the effects of public innovation subsidies on a firm's R&D spending and performance. Cerulli (2010) provides a review covering key studies. One of the challenges with identifying the causal effect of the public subsidy is selection

bias, that any changes seen in the supported firms reflects inherent, unobserved productivity advantages rather than the effects of the public support.

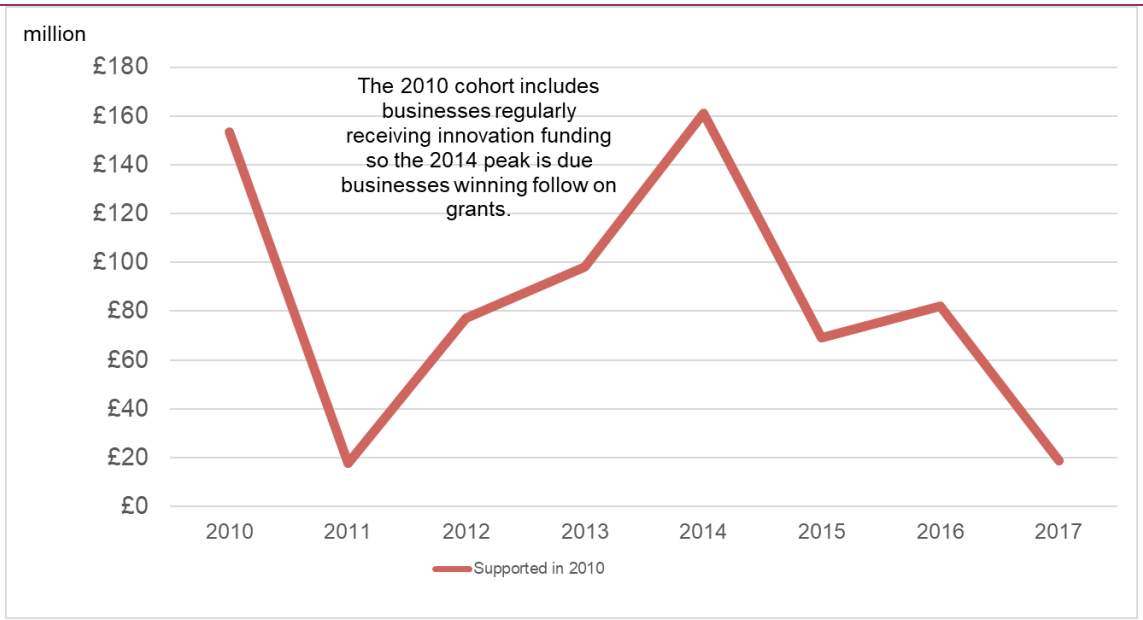
Modelling selection: Past studies

Cerulli (2010) reviews modelling strategies assessing a key challenge: private R&D expenditure as determined simultaneously either as part of a firm's investment planning or because the government 'picks the winner'. Cerulli offers approaches to building a selection model that accounts for the bias that results from this simultaneity, looking at possible empirical strategies.

The more recent literature with a focus on the firm level as surveyed by Zúniga-Vicente et al. (2014) takes the selection problem into account. Prominent methods that have been used to account for sample selection are matching methods (e.g. Czarnitzki and Fier, 2002; Almus and Czarnitzki, 2003), instrumental variables methods (e.g., Wallsten, 2000), and selection models (e.g., Busom, 2000; Hussinger, 2008). Most of the studies that account for selection issues report a positive effect of the subsidy on the subsidised firms' R&D expenses as well.

A key strand has been the characteristics used to understand selection. Firm level variables considered important are firm size, age, industry dummies, prior R&D expenditure, patent stock and foreign ownership (Czarnitzki and Licht, 2006; Hud and Hussinger, 2014; and Vanino et al., 2019). In UK work, Vanino et al. (2019) evaluating the effect of R&D grants by the UK research councils used knowledge intensive manufacturing and services sectors as well as regional variables related to regional productivity and R&D agglomeration areas. BEIS (2017) uses a firm's inclusion in the ONS Business Expenditure on R&D Survey sample as an R&D indicator.

Figure 2.1: Patterns of funding for those receiving Innovate support in 2010



Another issue often raised in identifying the causal effect of public R&D subsidies is the effect of repeat subsidies. With both NMS and Innovate UK, there are recurrences of recipient companies within the database. For NMS, the repeat usage is around 80% within three years. Within the Innovate UK database, 40% of firms receive multiple grants. Figure 2.1 tracks Innovate UK support to all businesses securing funding in 2010 in the subsequent years. This covers all businesses, including the large, R&D intensive businesses which lead numerous R&D projects. It indicates that continued funding is secured, perhaps necessitated by the nature of large innovation projects requiring successive years of funding.

Vanino et al. (2019) handles this issue by looking at the effect of a grant after the first instance of support, as well as looking at the effect of different amount grants, broken into four groups. Cerulli and Poti (2014) uses a dosage model, an approach common in epidemiology, to measure different levels of support. Both models use a cross-section of data and create a dosage of support based on a single year. Whereas many empirical studies see support as binary, with businesses either receiving the support or not, the dosage model compares impacts on similar treated businesses to others treated but with a different amount.

Dosage modelling allows for a continuous treatment variable so that different levels of support can be considered; it enables estimation across the entire treatment-outcome function, distinguishing from modelling that looks at an average level of treatment. Most importantly, it allows an analysis of the optimal dosage.

Looking at the scale of support requires an accurate assessment of the level of support. For example, it is important to be careful to ensure any comparison group is unaffected by some past support. BEIS (2017) create a matching pool removing those who receive a support measure at any time during the treatment and includes past support prior to 2008 as a control variable.

Selection modelling in this study

At the core of this study are lists of projects supported by Innovate UK and NMS. The innovation funding data provides identifiable businesses that benefit from different types of support. Most forms of support involve some (competitive) application process, with the funding body selecting the projects, or businesses self-selecting into support. Statistical methods to identify a counterfactual then seek to model the selection process to find businesses that statistically appear as if they are like the selected businesses except in not receiving the support.

For the selection modelling, propensity score matching (PSM) has been employed. PSM models selection by estimating a score for the likelihood of a business receiving support, controlling for observed firm characteristics such as industry, age and size. As the number of variables available in administrative datasets about businesses is relatively modest, key to the modelling has been to augment the business data with variables related to

innovation, such as whether a business conducts R&D or holds a patent. This strategy proves successful, increasing the robustness in the scoring. The matching then finds unsupported businesses with similar scores. A good PSM means that there is confidence that comparable businesses are identified, and robustness tests are used both to quantify how good the match is and to estimate how sensitive results are to different modelling.

This study explores whether selection into different types of support can be modelled jointly, or whether they should be analysed separately. As strategies they will differ according to the unsupported businesses that are used for matching.

- Matching to the full business register held by ONS, the Business Structure Database (BSD, described later). This has the advantage that the match pool is very large, making it likely that similar businesses can be found. However, the characteristics available are relatively few and unobservable characteristics, such as ambition that influence selection and growth outcomes could bias the results.
- Matching to businesses on similar innovation paths: The beneficiary data is complete, with many businesses' pathways through successive support discernible in the data. Businesses that have a similar history of innovation support but not securing the support being analysed provide a potential match pool and the use of past support in selection modelling controls for this.

Table 2.1 describes the models uses in the analysis. Each uses propensity score modelling (PSM) and an annex describes some of the assumptions for PSM to be suitable (such as common pre-support trends). There are a number of models employed and the three below are broadly reflective of the choices made in modelling. As noted in Angrist and Pischke (2013), PSM seeks to control for variables that affect the probability of treatment and, as this assignment is through the policy process and so of intrinsic interest, this modelling helps to understand decision-making. However, the choice of variables for the selection modelling then becomes important.

Broadly, the modelling has selected three themes to characterise assignment:

- Characteristics, especially size, age, location and industry: For each, continuous variables were used sometimes but also categories derived;
- Past innovation activity: Recent investments in R&D and holding intellectual property indicates a track record in delivering innovation, a feature of selection processes where some capability to manage projects is often tested. Also, this indicates a business as having ideas;
- Past support: As the dataset is comprehensive in terms of past support, this was used to model selection, allowing a score to take account of the pathway on innovation funding that is being taken.

Within the themes, selecting variables then involved exploring the characteristics of the supported businesses and how they differed from the wider business population. This particularly helped in constructing variables, such as for industry where categorising a

business as knowledge intensive, using Eurostat definitions of the sectors to include, correlated with support.

Table 2.1: Selection Models used in Study

	Variable descriptions	Matching pool
Used in all models	Live local units, Patent owner, Trademark owner, 3-year prior average R&D spend, High knowledge intensive manufacturing and services, Foreign ownership dummy, Scale-up dummy, government region dummies, age category dummies	Annual BSD for year before support, with treatment variable 1 for treated (by different treatments), 0 for untreated. Missing values in some models where businesses excluded from counterfactual (e.g. subsequently treated) or treated (e.g. modelling single treated only)
Model I	Employment categories dummies, turnover categories dummies, previous Innovate UK grant/NMS service	BSD sample, keep all prior treated Innovate UK grant /NMS service recipients
Model II	Employment categories, turnover categories, t-1 log employment, t-1 log turnover, log turnover difference two years prior, log employment difference two years prior	BSD sample, removing all Innovate UK grant/NMS service recipients
Model III	t-1 log employment, t-1 log turnover	BSD sample, removing all Innovate UK grant/NMS service recipients

For small businesses, the BSD provides a large population from which to select a counterfactual. PSM works well where the researcher believes there are plenty of unsupported businesses that are comparable to the treated. However, if the whole BSD – with augmentation – has limited variables to model, a second strategy defines a sub-sample of the supported business that are close comparators and then use PSM to select from the sub-sample. So, for example, a common pool is if there are rejected applicants. For this study, as it relied on the published data, compiling the data about rejected applicants was not possible. Many schemes in the support databases are relatively old without applications data.

BEIS (2017) matched to a restricted pool of the businesses, by removing the businesses that had received any support restricting comparison groups to those untouched by innovation funding. However, there are analytical approaches that highlight using any evidence that a business is suitable for funding. This study's Model I therefore augments the base matching populations with businesses that have received some support in the past but not the current year of treatment, then matches on the fact of past support.

This tackles a bias when current support is provided to businesses that were past beneficiaries, in that the impacts of the past support may still be affecting performance. By introducing past support into the matching, this aspect can be balanced in the counterfactual. An alternative selection modelling approach – limiting the modelling to businesses that only receive one incidence of support – can complement this.

Outcome variables are detailed below. To some extent, once a good control group has been identified, impact can be measured with relative confidence. For several outcome variables, there will be a time-series of pre- and post-support performance. The control group is used to evaluate the additional impacts over-and-above deadweight on these variables. The evaluation frameworks recommend using difference-in-differences (DiD) to isolate these impacts. DiD is a version of a fixed effects model, where a comparison is made in the changes observed over time in the participating businesses and the control.

The results from a difference-in-difference analysis can indicate whether the growth performance of the supported business differs from that of the counterfactual, but then understanding how any difference reflects a different level of economic activity requires some further estimation. Box B in chapter 3 discusses some of the steps taken to develop impact estimates in level terms.

A key step has been to focus on looking at the economic activity of a trimmed sample of the supported businesses, as using all the businesses would overstate the level of economic activity touched by innovation support. This is primarily because both Innovate UK and NMS beneficiary lists, while identifying the supported business, can only be linked to data about the whole enterprise. In many large, multi-plant businesses that operate across sectors, this would then mean the support will be linked to too large a portion of the supported business.

Impact Measures

Measuring employment, sales and survival

BSD is a longitudinally linked set of annual snapshots of the Inter-Departmental Business Register, IDBR, from 1997 onwards. IDBR is the live register used by ONS as a sampling frame for business surveys. It includes all businesses that are registered with HMRC for VAT and PAYE and incorporates registration at Companies House or who otherwise come to ONS's notice. Employment data is believed to be accurate across all businesses – even if not surveyed – as data is updated from HMRC PAYE records. There are some known reporting issues because the focus of the ONS business register is securing the best estimate of the level of employment and turnover, while impact analysis focuses on growth in this.

The information maintained on the IDBR include employment, turnover, industry classification (SIC code), year of birth and location. Impact analysis requires the compilation of data into a panel and the data needed for a panel has been deposited by ONS in its Secure Research Service (SRS) as the Business Structure Database (BSD). Key impact variables compiled in SRS include business demography, turnover, and employment. Using several years of BSD this provides a time-series of firm level data, measuring key variables at enterprise unit level, the largest entity level for a businesses at UK level often aggregating across many plants and establishments for the largest businesses.

Employment may be analysed at the lower levels of reporting unit and local unit. The local unit is the ONS lowest level unit (for example a shop, workshop, factory, warehouse, office, mine or depot) situated in a geographically identified place. The reporting unit usually represents a large establishment or the smallest combination of local units that is an organisational unit producing goods or services, which benefits from a certain degree of autonomy in decision-making, especially for the allocation of its current resources.

One of the other measures held within the BSD is turnover, taken from VAT returns in most cases. VAT turnover differs from accounting measures, such as that presented in annual accounts, but has the advantage of catching new companies as they register for VAT. For smaller firms, the VAT turnover estimate is used to estimate initial employment, so some care has to be taken looking at the labour productivity measures for the new, small businesses. Whereas employment measures are available at local unit level, turnover is available at the enterprise level.

Understanding firm survival

For many incidences of support, the BSD holds many years of data after treatment allowing a more sophisticated modelling of any exit of businesses. It is a key area of investigation given the relationship between firm survival and the security of jobs, skills building up and the internalising of innovation. However, in terms of economic growth it is unclear whether a positive impact on firm survival is beneficial. Supporting plants that otherwise would have closed may hinder ‘creative destruction’ that creates growth in the economy by re-allocating productive inputs from low- to high-productivity firms. On the other hand, if the inputs become redundant and are not reallocated, then there is a positive contribution to growth in the economy.

The economic theory concerning the impact of government subsidies on grants is relatively clear. A subsidy increases discounted expected profits so that plants that would otherwise close choose to remain in operation (Moffat, 2013). However, as argued by Moffat, government grants are often conditional on meeting an activity target. Thus, a subsidised investment may necessitate a simultaneous increase in the use of other factor inputs leaving the discounted expected profits unchanged.

Survival analysis corresponds to a set of statistical approaches used to investigate the time it takes for an event of interest to occur. For the survival model, propensity score matching is combined with the Cox Proportional Hazards model. This is a similar approach to previous studies covering regional grant assistance (Girma et al., 2005; Moffat, 2013). The idea is that in creating the matched sample, the “treated” and “controls” have broadly similar values of important covariates.

Firm survival is then modelled. The key metric of interest is the hazard ratio, defined as the instantaneous risk that a firm will survive in the following period given it as survived up to just before that period. This is visualised in the Kapler-Meier plots that, on the horizontal axis, represents time, and on the vertical axis shows the probability of surviving, i.e. the

proportion of businesses surviving, for both the supported businesses and the matched control. The plots provide a measure of how likely a supported business is to survive after support, but with a benchmark from a group of businesses assessed as comparable. Hazard ratios are estimated for the two groups – the supported businesses and comparable non-recipients. Two hazard ratios can be compared and, if equal, there is no difference in the hazard rate between the two groups, while a hazard ratio of less than one implies that the "treated" firms have a lower hazard relative to the "control" firms.

Hazard modelling yields survival curves presented in the findings later. Using these, the absolute reduction in the probability of firm closure within a timeframe of interest after receiving support can be estimated and the differential rate of exit between the matched groups of businesses multiplied by the economic activity of the survivors is the economic impact.

Dosage modelling

In public policies to support business R&D companies are awarded a different amount of support and sometimes, as well as interest in the binary fact of receiving support or not, a question is whether impact varies with the level of R&D support exposure (or "dose"). A further dimension is where a business receives more than one public financial support. Dose-response functions are used in clinical settings to check the effects on patient outcomes to different levels of treatment.

The dose-response is articulated using a curve that maps the level of support to the impact it has. Code has been developed seeking to estimate the relationship so that the balance across intervals for the dose, with this study using Stata code gpscore2 (Guardabasico and Ventura, 2014). Where treatment is continuous, it allows an investigation of the form of the function of average treatment effects over all possible values of the treatment levels and to study whether effects change when the level of treatment changes. Different econometric approaches (Marino et al., 2011) on public programme evaluation using a dose-response frame has found a convergent result of decreasing relation between the level of subsidy and the outcome (capital or R&D capital). Cerulli and Poti (2014) also highlight there is a notable substitution between private and public funds occurring for high intensity of public subsidy.

Innovation measures: Intellectual property and spillovers from knowledge

Patents and trade marks confer intellectual property rights. Economic theory gives insights into why intellectual property exists and how it stimulates innovation. Patent rights act as instruments to incentivise innovation and promote the diffusion of new knowledge. Hall et al (2013) find a correlation between patents and innovation using the ONS microdata linked to the patent register but make important caveats. They note that the links are specific to some sectors and that the innovativeness observed is not a precursor to employment growth. Recent reviews have (Hall and Harhoff, 2012) argued that innovation knowledge is costly to create, lengthy to develop and difficult to imitate. Therefore, market-

based mechanisms, relying on standard competitive forces, should be sufficient for inventors to recover their initial investments, with no need for special legal rights (Munari, 2017).

BIS (2014) highlights how innovation results in market spillovers as businesses commercialise the products of research and development. There has been an effort in innovation policy to move from so-called closed innovation systems, where benefits are concentrated in businesses that do research, to more open systems. In open innovation systems, spillovers are a key outcome. BIS (2014) highlights that where a technology has wide applications, or is embedded in a relatively open architecture, this increases the likelihood of spillovers.

Wider spillover effects will be contingent on the project focus. If an R&D project is focused on a specific sector, the impacts may extend into the sector, because the R&D occurs in an open innovation system. A second set of spillovers may be into supply chains or into the local economy as R&D is shared into the wider economy. Also, collaborations with Universities or the wider stakeholders make spillovers more likely. Indicators also look at the transmission mechanisms available.

Spillovers are linked to networks of researchers. Citations are often used as a sign of knowledge spillovers and, for this research, a new dataset has been used described in Dechezleprêtre, Martin and Mohnen (2014, DMM hereafter). The data considers the citations for each patent, but analyses this in a novel way.

Two measures are used to provide a picture of the intensity of knowledge spillovers – citations per patent and a PatentRank index. There have been several studies focusing on the citations per patent (Jaffe et al., 1993; Jaffe and Trajtenberg, 1999; Caballero and Jaffe, 1993; Jaffe and Trajtenberg, 1996). In the UK data, citations per patent range from 0 to 1,470 with a mean of 4.9. These measures do make corrections for some of the problems of citations, such as removing subsequent self-citations by the inventor. A concern with citation counts is that it may miss indirect citations. Some patents can be very influential because the patent that cites them are then highly cited. Ground-breaking patents are sometimes modestly cited due to the small size of an industry at the time of creation, whilst subsequent patents are much-cited (Maslov and Redner 2009).

DMM (2014) account for the network of innovation using a ranking index, termed the PatentRank, calculated through a Google PageRank algorithm. These algorithms are most commonly known for determining the relevance or importance of webpages modelling the behaviour of a web-surfer, who surfs the web by randomly clicking links on the visited pages but periodically gets bored and jumps to a random page altogether. The modelling needs to avoid sink webpages/patents, ones that are never cited, as these would lead to endless loops. However, for a patent, it provides a weighted sum of the PatentRanks of all patents citing the patent, where the weights depend on the number of citations made by these citing patents.

This index has the advantage of capturing ground-breaking patents that may be more modestly cited but then subsequently amass significant numbers of indirect citations signifying its high contribution to innovation. DMM also attempts to value innovation spillovers in monetary terms.

Impact measures for large businesses

For larger businesses, the coverage of official surveys can be quite good. This means that beyond the employment and turnover collected about businesses in administrative systems, other performance characteristics are collected. In this study, the ONS Annual Business Survey, which is compiled each year into a panel called the Annual Respondents Database (ARD) and the ONS Business Expenditure on Research and Development (BERD) has been used. The data are derived from random, stratified surveys. Both have low sampling ratios for SMEs but provide considerable evidence about the largest businesses.

The BERD data are available since the mid-1990s. For BERD, the sample is drawn from a continually updated register of known R&D performers in England, Scotland and Wales. In addition, businesses in Northern Ireland are surveyed by the Department of Finance (DoF) and their estimates added to those collected to form UK totals. As part of the 2015 survey, approximately 5,400 (that is, 4,000 Great Britain and 1,400 Northern Ireland) questionnaires were sent to businesses known to perform R&D.

The ARD starts in 1970, with the most recent year of data, the 2016 ABS, being the largest business survey conducted by ONS in terms of the combined number of respondents and variables it covers (62,000 Great Britain questionnaires, with around 600 different questions asked). It is the main resource for understanding the detailed structure and performance of businesses across the UK. It is necessary to match the BERD to the ARD, as the ARD contains the value added, capital and employment data necessary for analysing productivity. Matching the two data sets is based on a ‘reporting unit’ identifier.

The research also uses the FAME database, derived from firms’ financial information obtained from company annual reports. The financial data include the usual items, such as sales, profits, wages and salaries, assets, standard industrial classification (SIC), and R&D expenditure. The accounts also include balance sheet details, such as the net capital stock of the businesses.

All financial data are deflated with using published price and deflator series. Some quality checks are required, such as the presence of high growth rates in the variables, due to firms merging or de-merging, or major adjustments in accounting procedures.

Wages in supported businesses

The Annual Survey of Hours and Earnings (ASHE) lends itself to conduct research on wages of supported businesses. ASHE is a dataset on individuals, randomly chosen by their National Insurance (NI) numbers.

Surveys are completed by their employers from payroll data, so quality and response rates are high. Included are all employees registered for PAYE taxes whose NI numbers end in two specific digits. Therefore, it is a panel and individuals can be tracked even if they change employers. Variables collected include weekly earnings and hours worked in a reference period, occupation (4-digit SOC code), managerial and supervisory duties, place of work, gender, and age. The survey with business identifies is available for 2004-2016.

The measure of earnings used is gross weekly earnings deflated by the consumer price index. As ASHE is a 1% sample of all employees, many employees of supported business are included, however usually just with a few employees for each business. While the wages available for an individual business may not be very representative of that businesses, the random nature of the survey means that the average should be representative of the average in each of those businesses. In other words, while only observing the CEO in one business, and the graduate trainee in another, averaging across those businesses approximates the true average wage of each of those businesses.

Concluding comments

The approaches outlined in this chapter are diverse. They focus on an analytical stage to compile a complex dataset of business characteristics and performance linked to the project beneficiary lists. There are then three analytical stages, covered in the next three chapters.

A first focuses on modelling the selection into innovation support, seeking to tease out what makes the businesses that are supported different from the wider population and then identifying a comparison group.

The second stage then analyses the supported and comparison businesses. Given the nature of the support, a first analysis focuses on what increased economic activity is seen in the supported businesses. This is in terms of increased employment or the avoidance of business closure. It also covers R&D activities. The analysis also considers how activity varies by the level of support. Firm growth, survival and decisions to research could be affected indirectly from the support, as unsupported but similar businesses may be adversely affected by their competitors' being supported. A caveat to this analysis is that it does not reflect such spillovers from the intervention.

The final stage of the analysis considers the productivity impacts of the support. Productivity establishes whether any increased economic activity involves better use of resources and is a difficult area to determine. Approaches can be direct, looking at output per unit of input but this requires a complete picture of business resource use. Productivity proxies, such as the earnings of employees and patent citations, are used in this study.

3 Understanding the Treatment

Innovate UK and the National Measurement System intervene with businesses through different types of support. Over time, the range of measures provided have changed. This chapter categorises the support that is provided by each body. It describes the characteristics of those receiving the different types of support. The chapter then gives results on modelling the characteristics of the businesses that receive the different types of support, a key stage in identifying the counterfactual.

As indicated in chapter 1, the databases of support from the Innovate UK and NMS record over 70,000 incidences of support. There is a huge range of support, in terms of types of support, scale of projects, nature of collaboration etc. Further, while the data covers a long period, the support products change, reflecting the changes in the scope of funding and the priorities of policy. There is a need to process the data about different support so that it is as comparable over time as possible.

This approach departs from the previous BEIS (2017) method in some dimensions. Firstly, the analysis seeks to maintain the sample of supported businesses. The earlier study focused in on certain types of supported businesses and then exhaustively analysed these. Here, the approach uses multiple approaches to customise the modelling for the mix of businesses that receive innovation funding.

Key steps have been taken to integrate into analysis the pathways of support which businesses have taken. BEIS (2017) carefully sought to compare the supported businesses with ones that had not been supported by the two bodies; also, the modelling is undertaken for each year. This has some advantages, especially in allowing the control group to be guaranteed to be unsupported. While the approach taken here also models each year separately, the past history of support is integrated into models similar to BEIS (2017).

However, the methods used are similar, with this chapter's focus of understanding selection into support using statistical matching. The next section motivates how the modelling is customised for different support types, with the rest of the chapter presenting results about the modelling.

Categorising Innovate UK Support

The business data used in this study requires that the projects are compiled in terms of supported businesses. In categorising the support, it is important to be able to define the categories to link to the business data. The modelling has necessarily had to simplify across different types of support ("products" are the different types of Innovate UK support; NMS has a diverse range of innovation support measures), as these change over

time reflecting changes in policy priorities and each product might be shaped by those priorities. Further, as analysis is at a firm-level, with many businesses having multiple support, the ability to discern the impact of different products is less.

A first step is to aggregate the amounts reported for a business within each year, addressing the issue that a single business can receive more than one grant within any given year. This is often the case for larger businesses. The Innovate UK products are then split into groups with variables defining the groups also created:

- Three groups based on the length of the grant-funded projects. This uses the categories in Innovate UK's management system but also creates variables about the length of support;
- Two groups based on whether the business received a single award in the dataset or multiple awards. Also, variables are created about the past support received, especially tracking if a business has been a past beneficiary in a year when a new project is initiated.

These definitions have some advantages. Firstly, they are relatively stable, allowing matching pools to be defined consistently over time. While there is rich evidence in the database about the competition that each award was won through and a description of the project, these are inconsistent over time, and sometimes difficult to map out even for the same year. Secondly, they appear to be economically meaningful, with the first being a proxy for the level of support and the second proxying for the inclination of a business to seek innovation support.

Support by Length of Grant

Innovate UK management systems categorise the portfolio of grants into three groups – under 12 months, 12-24 months and more than 24 months. Table 3.1 shows that the average value of Innovate UK support correlates with the three categories for the length of the grant. In awards made in 2012, with the average larger but the median amount close to the 12-24 grant size, the table highlights that there are very large grants in the 24 months or more category.

Table 3.1: Innovate UK participants and spend by length categories, 2004-17

	All Support	Length of projects		
		<12 months	12-24 months	>24 months
Ave Value of Support (£)	£133,308	£50,653	£153,931	£299,689
Number of projects	32,790	11,994	8,238	12,558
Live local units	12	5	8	20
Patent owner	22%	17%	28%	25%
Trademark owner	8%	4%	7%	13%
Knowledge Intensive Services Sector	27%	29%	33%	27%
Knowledge Intensive Manufacturing Sector	6%	5%	7%	8%
Foreign owner	13%	7%	10%	20%
Scaleup business	4%	4%	5%	4%
Employment (log)	3.01	2.3	2.9	3.69
Real turnover (log)	9.22	8.25	9.03	10.14
Number of businesses	1,138	552	263	378

Table 3.1 also analyses the Innovate UK support data linked to the Business Structure Database (BSD). This allows an analysis of the size, sector and recent growth of the businesses that secure the support by the support categories.

Not only does grant length correlate to grant spend, but patterns emerge about the types of businesses supported for the different length categories. Businesses receiving grants lasting less than 12 months are smaller, and less R&D active than those receiving longer grants. The size of business increases with the length of grant, both in live local units, and employment and turnover.

Table 3.2 below disaggregates this categorisation by the Innovate UK product type. It highlights how the product types, as can be expected, are relatively flexible about the duration of support that they provide, with all products providing support of a range of values. Only GRD Proof of Market and Vouchers have 99% of their projects within a single size category.

Table 3.2: Innovate UK projects by type and length of projects, 2004-17

	Under 12 Months		12-24 Months		24 Months +	
Innovate UK product types	Grants since 2004	Mean amount per grant	Grants since 2004	Mean amount per grant	Grants since 2004	Mean amount per grant
BIS-Funded Programmes	77	£542,336	222	£409,840	1,022	£1,028,981
Collaborative R&D	1,316	£119,309	4,304	£180,223	9,451	£228,303
EU-Funded	124	£34,854	178	£217,429	436	£194,252
Fast Track	218	£38,747	59	£32,974	3	£47,217
Feasibility Studies	4,426	£46,272	1,400	£58,299	157	£136,394
GRD Development of Prototype	127	£134,837	565	£191,017	74	£183,288
GRD Proof of Concept	343	£81,314	477	£88,456	10	£73,907
GRD Proof of Market	637	£22,552	5	£19,219		
Knowledge Transfer Partnership	419	£68,422	840	£84,244	1,093	£104,796
Large Project					200	£848,135
Launchpad	53	£61,768	60	£74,696	5	£71,984
Legacy RDA Collaborative R&D			4	£376,955	29	£1,997,639
Legacy RDA Grant for R&D	11	£79,037	60	£111,079	30	£223,736
Small Business Research Initiative	629	£118,474	52	£792,987	44	£1,754,765
Study	103	£66,902	9	£474,399	4	£1,812,524
Vouchers	3,511	£5,010	3	£5,000		
Total for all grants	11,994	£50,653	8,238	£153,931	12,558	£299,689

Single versus multiple grant recipients

Around 60% of Innovate UK beneficiaries receive only one grant. The other 40% receive several grants over the years. An important policy issue is whether providing multiple grants to the same business over the years is more effective than a single grant.

Looking at the amounts in 2017, Table 3.3 shows that the average grant for those receiving multiple grants is almost four times as large as the amount going to businesses receiving single grants. This is skewed by some large businesses receiving a multitude of grants within one year.

Table 3.3: Businesses receiving single/multiple Innovate UK projects, 2004-17

	Any grant	Single Grant	Many Grants
Average amount per project	£149,201	£52,560	£197,112
Median amount per project	£52,328	£35,960	£76,252
Live local units	12	4	16
Patent owner	22%	11%	28%
Trademark owner	8%	3%	11%
Knowledge Intensive Services Sector	27%	23%	30%
Knowledge Intensive Manufacturing Sector	6%	4%	8%
Foreign owner	13%	9%	16%
Scaleup business	4%	4%	5%
Employment (log)	3.01	2.32	3.40
Real turnover (log)	9.22	8.46	9.66
Number of businesses	1138	417	721

Comparing one-off and multiple grant recipients, the table also shows that those receiving multiple grants are generally larger and patent more.

Cohorts of Innovate UK support

Analysing Innovate UK support for individual years needs to take account of the way that businesses receiving multiple grants would then be included for every year they receive an Innovate UK grant. There is a risk of businesses appearing several times, and any effect from prior Innovate UK grants being included in years that they receive additional grants. Table 3.4 highlights the approach taken to establish impacts so that annual snapshots can be linked together without this double counting.

The table categorises businesses in terms of the year of first support, starting in 2010, and then looks at the support provided to these businesses in that first year of support and subsequent years. So, the businesses winning a first project in 2011 received £53m in that year. However, the receipts continue in the following years reflecting both businesses receiving payments for a project that extends over several years and winning further projects.

Table 3.4: Innovate UK funding following first projects, 2010-13

Business support	Project funding (£m)							
	Total	2010	2011	2012	2013	2014	2015	2016
First Support before 2011	677	154	18	77	98	161	69	82
First Support in 2011	180	-	53	24	32	24	22	16
First Support in 2012	182	-	-	74	26	30	28	12
First Support in 2013	183	-	-	-	101	25	39	11

A supported business enters the analysis (i.e. is treated) in the first year of the first project before 2013. The 2010 cohort, the first row in Table 3.4, is a little different in catching all the businesses supported in and before 2010. In total, the table contains almost all the £1.2b of Innovate UK funding made available to businesses receiving first support up to 2013.

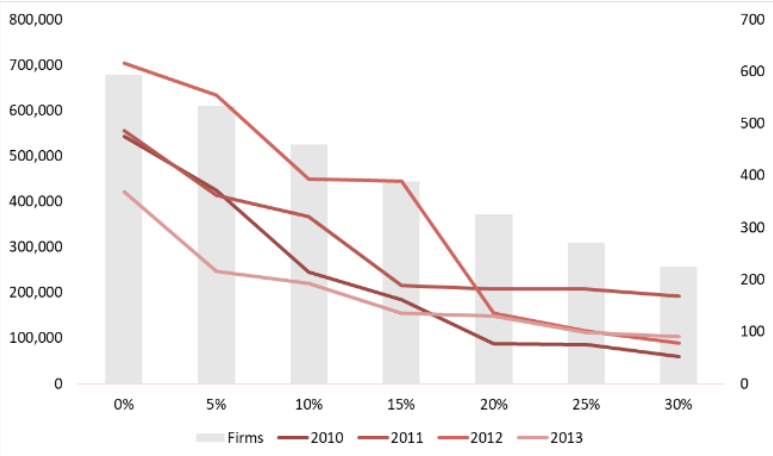
BOX B: Challenge of analysing the level of impact

The businesses supported by Innovate UK and NMS are diverse and the focus of the approach here – statistical matching and difference-in-difference – is to seek to estimate growth differences between supported businesses and the matched control. However, translating this in levels, i.e. the number of jobs or economic activity, then needs further steps.

After matching, the size of businesses is likely to differ between the supported and control. In the annex, this is explored for employment where – after matching – the supported counterfactual businesses on average employ about a third more than the businesses supported by Innovate UK in one year (2013). Sometimes, this can vary in the opposite direction, but the effect is that the change in the level of employment in the supported business will be a third greater in 2013, purely because of the starting position differing. This makes employment levels quite volatile.

A second area is a bias, due to the imprecision with which support provided to businesses is defined for large businesses. Many businesses that receive innovation support are very large, operating across multiple sites. The exact part of these firms that have been supported is not recorded in the beneficiary lists provided and yet is likely that the supported R&D will be specific to a part of business and not the entire UK enterprise.

UK economic statistics are based on sources that can identify individual plants: surveys are usually conducted by ONS at a lower level than the entire enterprise, with large plants being enumerated as separate reporting units. However, for large companies with multiple reporting units, the beneficiary data is linked at the larger enterprise level, considerably raising the level of employment (and economic activity more generally) that falls within the scope for measures for the supported firms.



The annex explains the approach taken to remove these two biases. For the first, levels of economic activity have been tracked only in the supported businesses, with the DiD estimates used to adjust this for the performance (and not the levels) of the counterfactual (similar to Goldman Sachs, 2019).

For the second bias, the supported businesses have been trimmed, excluding businesses that were supported but were either extremely unlikely or likely to have received support, as modelled by the propensity score. Large, multi-plant businesses are being treated almost every year at enterprise level as one of the reporting unit receives support, manifesting itself in these businesses having a high propensity score. The figure indicates how this affects the coverage of employment, each line indicating employment falling as trimming increases for four Innovate UK treatment years, with the x-axis indicating the trimming rate. The trimming rate was selected at a point where the employment stabilises, 25%.

The table also indicates the level of funding in each cohort. The cohort funded prior to 2011 for the first time is the largest by expenditure. Aggregating all businesses that received some support before 2011 means this cohort of businesses dominates the project database. This is because the large, complex, innovative UK businesses are highly likely to have successfully applied for innovation support early. Later cohorts are then businesses that are new to Innovate UK funding, usually SMEs.

However, first support in the next three years maintains a high amount of project funding. The businesses first supported in 2011 received £53m in 2011. As projects extend over several years, businesses subsequently receive £16m-£32m annually. Choosing 2011 is sensible, because – as seen in Figure 1.1 – this is prior to the ramp up in Innovate UK funding.

The focus of the analysis has been these four cohorts. This excludes the businesses that received support after 2013. This cut-off was chosen primarily to allow for enough post-support years of data to track impact.

Categorising National Measurement System support

The National Measurement System provided a dataset of beneficiaries from 2002 to 2017, split into three categories of support: web downloads, invoices for services, and formal collaborations. Outside of this split, years of support and organisation type were provided. The supported businesses are broken into intervention type: collaboration, paid services and free online downloads. The support products are divided into two categories placing collaborations and paid services into one category and downloads are separated. The former is considered to be treated while the latter – as downloads are very low intensity support – are used in the selection model as an identifier of businesses with an interest in NMS services.

With the large amount of businesses receiving more than one incidence of support, this study attempts to deal with this issue of repeat recipients in two ways. To control for the effect of previous treatment, one of the models (Model I) adds the prior treated into the matching and controls for prior treatment. The selection modelling shows that prior treatment is a very important factor for receiving future treatment. A second way of addressing this persistence challenge used in the Innovate UK analysis has been to analyse the beneficiaries in a year and focus in on those businesses that receive support for the first time in that year.

However, the NMS dataset is smaller in terms of the businesses supported. There are about 500 businesses and about half of these appear across many of the years of the data. Analysing the data by year of support leads to quite modest numbers of businesses. Further, there is a strong persistence in the relationship between NMS and companies that uses its services so that as much as half of the beneficiaries in any year had some contact with the NMS in the previous year. This means that the approach used in the Innovate UK cannot be replicated exactly.

The approach on defining a treatment variable is to consider support over the period 2009-17. Treatment for the selection modelling reflects whether the businesses interact regularly over the period or sometimes over the period. Regular treatment is defined as there being an incidence of NMS support for the businesses in over 85% of the years that the business is in the ONS data (so a business that begins in the period will only be tested for the years after its entry). The threshold served to split the treated businesses into two equally sized groups. The businesses that are treated sometimes then have more than one incidence of support but are not regular users.

Table 3.5 indicates some summary statistics for a single year, to tease out the difference between regularly supported businesses and the intervention type.

Table 3.5: Businesses supported by NMS by frequency and type, 2013

Variable	All Treated	Regularly treated	Downloads
Live local units	13.1	13.9	12.9
Patent owner	20%	17%	22%
Trademark owner	16%	21%	15%
Average R&D spend	£12,160	£26,190	£2,741
Knowledge Intensive Services Sector	11%	7%	12%
Knowledge Intensive Manufacturing Sector	21%	27%	21%
Foreign owner	28%	35%	29%
Employment (geomean)	73.7	106.6	78.6
Real turnover	£77.1m	£131.7m	£53.6m
Number of businesses	595	262	602

The businesses that are regularly supported by NMS are the largest and more innovation active. They are also more likely to be in knowledge intensive manufacturing. This is indicated in the middle column, comparing with the businesses that receive any grant type support and the businesses that make use of the download products.

Selection Modelling for Innovation Support

Innovation support is given each year to thousands of businesses, either by Innovate UK financing research projects or the measurement services provided by NMS. A first step in identifying what impact the support has is to consider the drivers for selection into support. Selection modelling seeks to understand the factors that correlate with receiving support, focusing on characteristics that are observable and measurable, but then uses statistical methods to understand whether the modelling captures all the factors as organisations are selected for innovation funding support.

Identifying Selection Variables

It is important for the selection model to represent variables that are economically significant in the selection process for receiving an Innovate UK grant or using an NMS

service. It is important to understand what sorts of characteristics would help a business to win an Innovate UK grant or use an NMS service. A few key selection characteristics for both services are:

- Prior innovation activity, represented in prior patents, trade marks and R&D spend
- Being within high knowledge intensive sectors, both services and manufacturing
- Prior growth in employment and turnover, proxied by being a scale-up business
- Business ability to absorb and manage a new project represented through live local units, age, turnover and employment categories or logs, and ultimate foreign ownership

Since the Innovate UK grant process is competitive, and Innovate UK chooses businesses based on their application strength, this makes it important to look at the ability to win grants. For Innovate UK, selection modelling incorporates the ability to write/win a grant, represented by being a prior recipient of Innovate UK grants. NMS has different characteristics that are relevant to using its services, and a few that are used include: prior usage of NMS services; downloading content from the NMS website; and a UKAS certification in two specific areas.

These characteristics create the propensity score that is used to match businesses who have received Innovate UK grants or NMS services to similar businesses who have not. Estimation of the propensity score is through a probit model, used to model binary outcome variables. In the probit model, the inverse standard normal distribution of the probability is modelled as a linear combination of the predictors. Matching uses nearest neighbour.

Table 3.6 and 3.7 below highlight the strength of key variables within some of the selection models tested in the study. Table 3.6A focuses on the selection model for Innovate UK in 2012. Generally, size (as measured by employment), the knowledge intensity of the business' industry and innovation activity are strong predictors for Innovate UK support. The model's negative relationship to real turnover needs more exploration, but the overall model has several other variables that correlate with size.

Model I differs from Model II and III in including into the selection model variables about past support and includes in the matching population businesses that have some history of Innovate UK support. This departs from earlier studies, such as BEIS (2017), because it was recognised that inclusion of past support would bias impact measurement. However, Table 3.7A highlights that exclusion of the variable may lessen the power of the selection modelling, with some prior support being a strong predictor of future support.

Table 3.6A: Probit models of main selection variables for Innovate UK, 2012

	Model I		Model II		Model III	
	beta	z-stat	beta	z-stat	beta	z-stat
Live local units	0**	2.25	-0.001	-1.38	0.001***	4.41
Patent owner	0.52***	13.13	0.709***	20.82	0.727***	21.65
Trademark owner	0.064	0.97	0.124**	2.17	0.238***	4.50
Average R&D spend (log)	0.077***	9.72	0.152***	22.11	0.162***	24.88
Knowledge Intensive Services Sector	0.463***	14.64	0.509***	17.30	0.533***	18.78
Knowledge Intensive Manufacturing Sector	0.477***	6.94	0.472***	7.59	0.451***	7.45
Foreign owner	0.176***	3.70	0.15***	3.56	0.221***	5.70
Scaleup business	0.005	0.09	-0.047	-0.91	0.061	1.21
Employment (log)			0.269***	8.84	0.099***	9.26
Real turnover (log)			-0.037***	-2.87	-0.045***	-6.50
Employment categories	Significant		Significant		Significant	
Turnover categories						
Prior Innovate UK grant	2.329***	52.82				
Two year turnover difference (log)			0.019	0.70		
Two year employment difference (log)			0.021	1.65		
Count of Businesses	481,725		417,520		478,336	
Psuedo R2	0.34		0.18		0.17	

* show the significance level; *** significant at the 1% level; ** at 5%; * at 1%. Standard errors are robust.

For NMS in Table 3.6B, generally size (as measured by employment), the knowledge intensity of the business' industry and innovation activity are strong predictors for regular support. Model I again differs from Models II in including into the selection model variables about past NMS support and includes in the matching population businesses that have some history of support.

Table 3.6B: Probit models of main selection variables for NMS, 2012

	Model I		Model II	
	beta	z-stat	beta	z-stat
Live local units	0.00*	0.08	0.00	0.28
Patent owner	-0.34***	0.00	-0.08	0.40
Trademark owner	-0.08	0.50	0.03	0.79
Average R&D spend (log)	0.00	0.88	0.05***	0.00
Knowledge Intensive Services Sector	0.03	0.86	0.06	0.60
KI Manufacturing Sector	0.26**	0.01	0.85***	0.00
Foreign owner	-0.17*	0.08	0.15*	0.06
Scaleup business	-0.36	0.31	-0.27	0.23
Employment (log)	0.24***	0.00	0.13**	0.02
Real turnover (log)	-0.03	0.51	0.03	0.38
Employment categories	Significant		Significant	
Turnover categories				
Age categories	Insignificant		Insignificant	
Prior NMS support	2.87***	0.00		
In BERD sample year before	0.03	0.82	0.11	0.28
NMS Download year before	0.57***	0.00	1.67***	0.00
UKAS Accreditation	0.83***	0.00	1.78***	0.00
Prior IUK support	0.12	0.26	0.60***	0.00
Constant	-5.84***	0.00	-4.83***	0.00
Count of Businesses	453,176		451,604	
Psuedo R2	0.70		0.54	

* show the significance level; *** significant at the 1% level; ** at 5%; * at 1%. Standard errors are robust.

The variables that provide a proxy for a business' need to use NMS services are strong in the selection modelling. The industry category – knowledge intensive manufacturing and services – is a very strong predictor of treatment. Most material is the past use of NMS services or using UKAS accredited bodies' services. This latter variable was derived by linking a list of businesses that have been clients of the measurement services of a group of providers aligned strongly in terms of the services provided to NMS.

Table 3.7A: Model I selection models for Innovate UK, 2008-16

	Year of Selection into Support							
	2009	2010	2011	2012	2013	2014	2015	2016
Observations	620,406	736,818	576,055	561,313	567,175	575,696	584,469	608,978
Adj R-Squared	0.10	0.12	0.13	0.10	0.13	0.12	0.12	0.09
Treated	195	184	236	302	327	281	384	254
Patent year before	0.28***	0.25***	0.23***	0.33***	0.48***	0.39***	0.54***	0.35***
Trade mark year before	0.03	0.15	0.22*	0.25***	-0.05	0.09	-0.17	0.03
High Knowledge Intensive	0.24***	0.43***	0.63***	0.44***	0.36***	0.53***	0.44***	0.42***
R&D in 3 years before	0.02	0.02	0.02**	0.02	0.02	0.02*	0.02	0.02
High KI Manufacturing	0.21	0.21***	0.21***	0.21***	0.21***	0.21***	0.21***	0.21***
Foreign owned	0.24**	0.24***	0.24	0.24***	0.24***	0.24	0.24*	0.24***
Scale-up in period before	0.08	0.08	0.08*	0.08	0.08	0.08	0.08	0.08
Employment year before	0.03*	0.03	0.03	0.03	0.03	0.03	0.03	0.03
Real turnover year before	0.03	0.03	0.03*	0.03	0.03	0.03	0.03	0.03*
Emp growth period before	0.14*	0.14	0.14	0.14	0.14*	0.14	0.14	0.14
TO growth year before	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
BERD sample year before	0.39***	0.39***	0.39***	0.39***	0.39***	0.39***	0.39***	0.39***
InnUK project period before	1.45***	1.45***	1.45***	1.45***	1.45***	1.45***	1.45***	1.45***
Constant	-4.18***	-4.18***	-4.18***	-4.18***	-4.18***	-4.18***	-4.18***	-4.18***

Innovation activity variables are strong indicators within both models for NMS, particularly patent ownership, but the co-efficient is negative. In later results, the balance achieved indicates that the selection model matches to the businesses with relatively high patenting activity, but that the fact of patenting at the margin after other variables have been incorporated is negative in Model I.

Tables 3.7A and 3.7B consider models by year of support. In each, the pool of unsupported businesses will exclude those that are treated for the first time after the modelled year. The tables indicate that selection models are stable over time, with almost identical parameter estimates for successive cohorts.

Table 3.7B: Model I selection models for NMS regular treated

	Year of Selection into Support				
	2010	2011	2012	2013	2014
Observations	433729	438346	453176	463748	484074
Adj R-Squared	0.71	0.71	0.70	0.70	0.63
Treated	250	245	256	250	244
Patent year before	-0.36***	-0.39***	-0.34***	-0.36***	-0.18*
Trade mark year before	-0.05	-0.03	-0.08	-0.04	-0.08
High Knowledge Intensive	0.00	-0.01	0.00	0.00	-0.01
R&D in three years before	0.02	-0.03	0.03	-0.03	0.02
High KI Manufacturing	0.23**	0.37***	0.26**	0.24**	0.21**
Foreign owned	-0.05	-0.16	-0.17*	-0.07	0.24
Scale-up in period before	0.06	0.24	-0.36	-0.16	0.08
Employment year before	0.19***	0.14*	0.24***	0.25***	0.03**
Real turnover year before	0.00	0.01	-0.03	-0.03	0.03
Emp growth period before	0.19**	0.10	0.12	-0.22*	0.14
Turnover growth year before	0.01	-0.01	-0.02	0.15***	0.00
NMS support before	2.76***	2.88***	2.87***	2.74***	0.39***
In BERD sample year before	0.02	0.05	0.03	0.10	1.45
NMS downloads	0.70***	0.66***	0.57***	0.64***	-4.18***
UKAS Accreditation	1.01***	0.86***	0.83***	0.92***	0.00***
Innovate UK support before	0.15	0.09	0.12	-0.05	0.00
Constant	-4.42***	-4.87***	-5.84***	-8.60	0.00***

Regressions also involve regional dummies and dummies for employment and turnover categories; * significance at 10%; ** significance at 5%; *** significance at 1%

Further models have been explored. Models I and II are the models used in the majority of the Innovate UK results and these have been completed for all years and different treatment types. Model III is run as a robustness test in year 2012, removing some subsets of larger businesses to ensure results are not biased to outliers. For NMS models I and II are used.

Robustness of selection models

Robustness tests of PSM include considering whether – after matching – the supported businesses and the matched counterfactuals are statistically similar. This is tested using balance tests after matching. A key characteristic that matching needs to produce is that the counterfactual group – prior to support – were on a similar trend to the treated in the key economic impact measures. If this is not the case, there is a concern that some unobserved characteristics remain, and these have put the supported businesses on a different growth trajectory prior to support. In this report, figures showing growth will usually include a year before the support to indicate whether this is a problem.

Visual checks can identify whether the matching has covered the range of propensity scores. Further, estimates of how robust results are to unobservable bias (Rosenbaum bounds tests) can check the sensitivity of impacts to alternative assumptions about the matching.

Balance after matching

The selection variables are used to build a group of businesses that are similar to the supported businesses but have been unsupported by Innovate UK and NMS using one-to-one matching with replacement. Balance tests look at whether, having been matched, the counterfactual is similar to the supported businesses. This tests for the so-called conditional independence assumption of propensity score matching (see annex D).

Below in the Tables 3.8 and 3.9 is the balance of the variables before and after the matching. For brevity, Tables 3.8 and 3.9 focus only on a small number of models, that look at being supported in 2012. The tables show that the matching process effectively creates an unsupported group that is similar to the group of supported businesses.

Before matching, the BSD sample of businesses had very different characteristics than the supported businesses. For example, only 5% of the BSD sample was within the knowledge intensive services industry, whilst 27% of the supported businesses were within this industry group. The three stars indicates that the difference between the supported businesses and the BSD sample before matching is significant at the 1% level.

Looking at the after matching group across all models, there is no significant difference between the treated and untreated group. While only the table for 2012 is presented, this remains consistent across the years 2009-2014 for Innovate UK.

Table 3.8: Balance tests for Models I and II, Innovate UK, 2012

	Model I			Model II		
	Treated	Untreated, after matching	Untreated, before matching	Treated	Untreated, after matching	Untreated, before matching
Live local units	11.93	15.51	1.54***	3.18	2.87	1.3***
Patent owner	0.23	0.25	0.02***	0.23	0.24	0.02***
Trademark owner	0.08	0.10	0.02***	0.07	0.08	0.02***
Average R&D spend (log)	1.79	1.71	0.11***	1.73	1.61	0.1***
Knowledge Intensive Services Sector	0.27	0.29	0.05***	0.27	0.28	0.05***
Knowledge Intensive Manufacturing Sector	0.07	0.06	0.01***	0.06	0.08	0.01***
Foreign owner	0.15	0.17	0.04***	0.14	0.14	0.04***
Scaleup business	0.05	0.04	0.05	0.05	0.04	0.05***
Employment (log)				2.88	2.77	2.35***
Real turnover (log)				9.12	8.95	8.88***
Prior Innovate UK grant	0.40	0.39	0***			
Two-year turnover difference (log)				0.10	0.08	0.09
Two-year employment difference (log)				0.08	0.09	0.09
Count of businesses	1,185	1,185	417,639	1,145	1,145	404,125

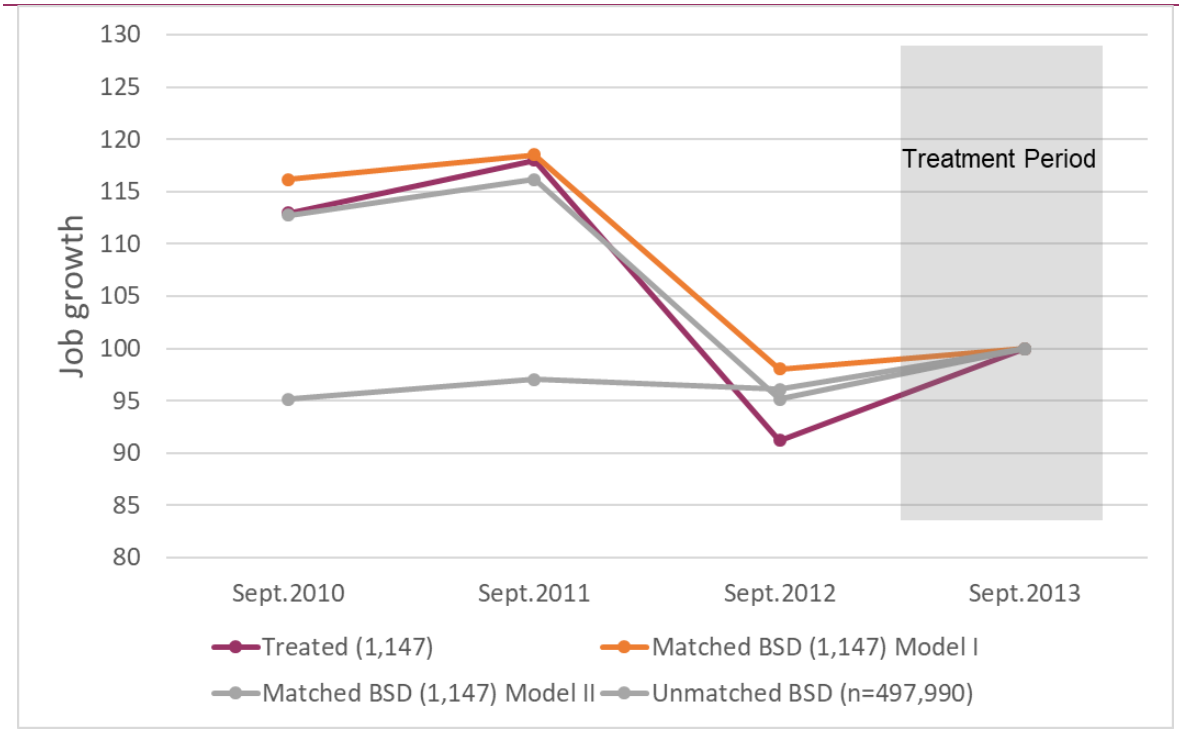
* show the significance level; *** significant at the 1% level; ** at 5%; * at 1%. Standard errors are robust.

While not included in the table, government regions, employment, turnover and age categories were all tested for balance. Out of four models, and 82 tests, on only 6 occasions did a significant difference at the 5% level remain after the matching.

An important test for matching is looking into pre-treatment trends in employment and turnover. If businesses were growing at very different rates before treatment, this means the comparison group is not similar to the treatment group. Pre-treatment trends are tested checking if growth in employment before support differs in the supported businesses from the matched group.

Figure 3.1 indicates the balance in pre-treatment employment growth for the businesses supported through Innovate UK in 2013. The lines plot similar changes in employment for the supported businesses and the matched businesses. Further, the unmatched BSD is quite flat in employment growth suggesting that matching has successfully been able to balance on pre-support growth rates. In the later figures, several years of pre-support are not provided, but one year of growth is usually included (to make graphs easier to read). However, this analysis indicates a general finding that most modelling matched businesses on statistically similar growth paths before support.

Figure 3.1: Pre-treatment employment trends



NMS recipients have a different makeup than Innovate UK beneficiaries, they are larger and more R&D active. However, here matching has also created balance between the treated and untreated. When prior supported businesses are included, the matching leads to a slightly larger, more R&D active group, although not statistically different.

Table 3.9: Balance test for Models I and II, NMS, 2012

	Model II			Model I		
	Treated	Untreated, after matching	Untreated, before matching	Treated	Untreated, after matching	Untreated, before matching
Live local units	13.40	13.97	1.71***	13.39	15.15	1.75***
Patent owner	0.23	0.28	0.02***	0.23	0.27	0.02***
Trademark owner	0.16	0.16	0.02***	0.16	0.16	0.02***
Average R&D spend (log)	2.85	2.66	0.11***	2.82	2.95	0.12***
Knowledge Intensive Services Sector	0.12	0.13	0.05***	0.12	0.13	0.05***
Knowledge Intensive Manufacturing Sector	0.21	0.19	0.01***	0.20	0.22	0.01***
Foreign owner	0.29	0.31	0.04***	0.29	0.31	0.04***
Scaleup business	0.03	0.03	0.06***	0.03	0.03	0.06
Employment (log)	4.43	4.32	2.47***	4.42	4.26	2.55***
Real turnover (log)	11.39	11.23	9.02***			
Prior Innovate UK grant				0.87	0.87	0***
Two-year employment difference (log)	0.12	0.10	0.16			
Employment categories		Included			Included	
Turnover categories		Included			Included	
Age categories		Included			Included	
Region categories		Included			Included	
Count of businesses	559	559	349,831	557	557	351,049

* show the significance level; *** significant at the 1% level; ** at 5%; * at 1%. Standard errors are robust

Further robustness tests

Robustness tests of PSM include visual checks to identify whether the matching has covered the range of propensity scores, estimate how robust results are to unobservable bias (Rosenbaum bounds tests), and look at the balance of characteristics in the businesses and the proposed control group.

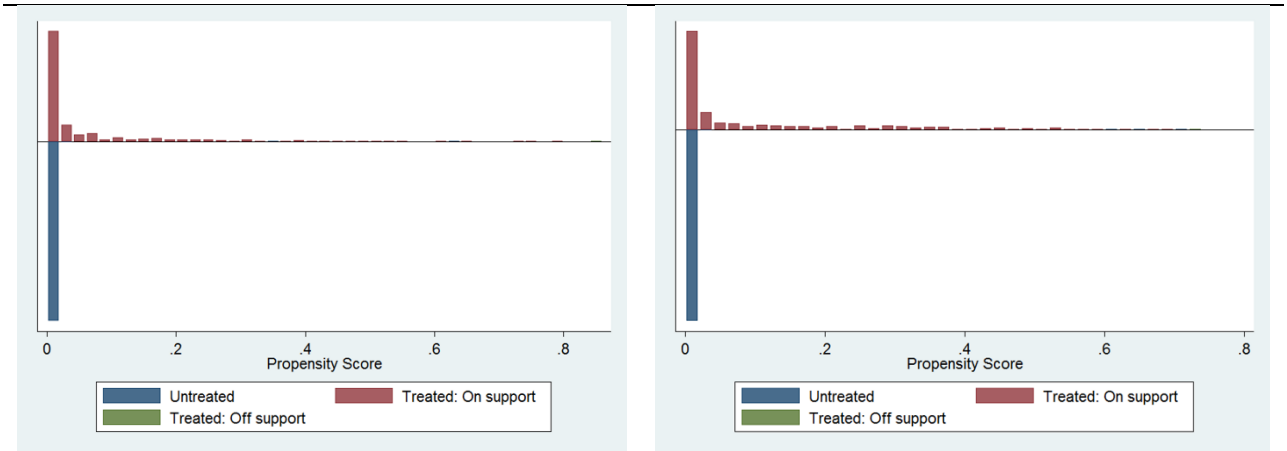
Figure 3.2 plots out the propensity scores for one of the NMS models and one of the Innovate UK models. This tests for the common support assumption of propensity score matching, that a comparable business can be found.

The matching to the BSD involves some difficulty establishing common support across all the supported. To some extent this can be anticipated, as the way that both Innovate UK and NMS operate means there is complete coverage for the largest and most R&D intensive businesses. A noticeable feature in testing for common support is that it becomes less of a problem when the matching is undertaken on subsets of the businesses that remove these large businesses. For example, focusing on businesses that receive their first ever support has fewer difficulties on this robustness issue.

Where selection modelling finds businesses that cannot be matched to a comparable unsupported business, the propensity score matching then proceeds without that business. The sample is effectively trimmed, focusing on businesses where a match was possible. Angrist and Pischke (2008) guide the evaluator to use trimming more purposively, advocating using the propensity score to focus in on businesses near the

boundary score, where businesses had the scores for either just being treated or not treated having modelled selection. The next chapter explores this aspect more, as the propensity scores are picking up sets of businesses that are far away from this boundary and their inclusion in analysis biases results on impact.

Figure 3.2: Propensity score plots for NMS and Innovate UK selection models



NMS 2009 Innovate UK 2010

Less than five businesses, at the right hand extreme, were demed off support, with the vast majority of businesses satisfactorily matched.

In Table 3.10 and in annexed Table C3, the Rosenbaum tests are presented. After a selection modelling has been estimated and the counterfactual determined, the bounds tests consider the level of weakening of the matching quality associated with a specified impact measure becoming insignificant.

A parameter gamma measures the degree of departure from random assignment of treatment. Two subjects with the same observed characteristics may differ in the odds of receiving the treatment by at most a factor of gamma. In a randomized experiment, randomization of the treatment ensures that gamma is one. The test then seeks to model an observational study recognising that without randomisation an unobserved characteristic may determine entry into the support. If gamma is two and two subjects are identical on matched covariates, then one might be twice as likely as the other to receive the treatment because they differ in terms of an unobserved characteristic (Rosenbaum 2005).

The tables indicate what happens to the R&D expenditure change following NMS support (analysed in the next chapter). When gamma is one, this impact measure is generally significant, as the supported businesses' growth is significantly higher than that seen in the counterfactual.

The table indicates this property is maintained for quite high levels of gamma. The point estimates column starts with the estimated impact with gamma 1. Then, as gamma

increases, the bounds around this value rise. However, they do not incorporate zero, i.e. an insignificant treatment effect, until quite high gammas. It indicates that the unobserved characteristics would have to affect the propensity to be treated by 20%, (gamma is 1.2) for the significance of the results to be lost.

Table 3.10: Rosenbaum tests for sensitivity of NMS impact on R&D spending

2009					2010				
	Significance of selection effect		Point estimate			Significance of selection effect		Point estimate	
Gamma	Positive	Negative	Upper	Lower	Gamma	Positive	Negative	Upper	Lower
1	0.00	0.00	35.95	35.95	1	0.00	0.00	35.26	35.26
1.1	0.00	0.00	25.08	51.00	1.1	0.00	0.00	19.40	45.81
1.2	0.00	0.00	18.00	67.92	1.2	0.00	0.00	10.68	65.57
1.3	0.00	0.00	12.00	85.56	1.3	0.00	0.00	1.86	80.26
1.4	0.00	0.00	5.80	106.29	1.4	0.00	0.00	0.00	98.01
1.5	0.01	0.00	0.00	124.96	1.5	0.01	0.00	0.00	124.96
2011					2012				
1.1	0.00	0.00	13.26	13.26	1.1	0.04	0.04	48.18	48.18
1.2	0.00	0.00	5.15	27.61	1.2	0.14	0.01	15.04	106.50
1.3	0.00	0.00	0.00	42.22	1.3	0.30	0.00	0.00	154.46
1.4	0.01	0.00	0.00	66.50	1.4	0.50	0.00	0.00	199.61
1.5	0.05	0.00	0.00	81.51	1.5	0.69	0.00	0.00	247.76
2013									
1	0.01	0.01	65.51	65.51					
1.1	0.05	0.00	31.50	95.93					
1.2	0.13	0.00	11.37	154.53					
1.3	0.27	0.00	0.00	206.50					
1.4	0.44	0.00	0.00	285.01					
1.5	0.61	0.00	0.00	398.49					

4 Impacts on economic activity

This chapter estimates the impact on economic activity in the businesses following innovation funding. Economic activity is primarily the employment change seen in the supported businesses. Supported businesses are growing significantly faster in terms of employment than the matched control group. Supported businesses also have higher R&D investment and patent more than the matched control group post support.

The last chapter indicated how comparison groups can be developed for the businesses supported by Innovate UK and NMS. Using difference-in-differences and multivariate estimation techniques, the additional impacts on economic activity can be estimated comparing supported businesses against the counterfactual groups of matched businesses. This chapter presents the results, focusing on employment as a measure of economic activity and then turning to patenting and R&D investments of businesses.

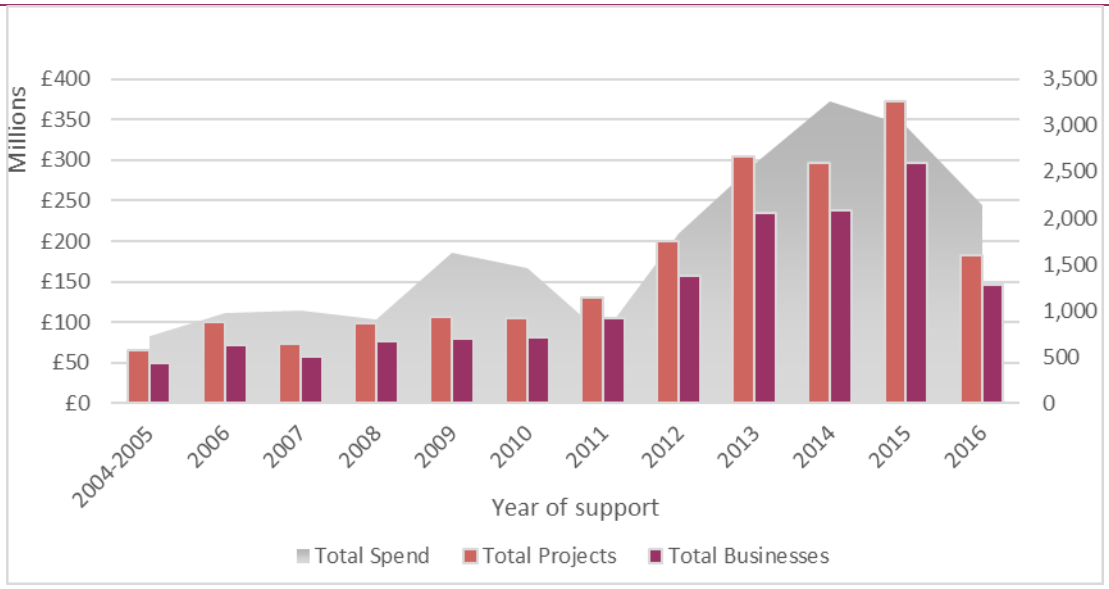
Employment Impacts of Innovation Support

Following the matching on treatment, the ONS BSD allows supported businesses, unsupported businesses and matched unsupported businesses to be tracked in terms of their employment and turnover. The BSD also documents when a firm closes. A first analysis focuses on the supported businesses, looking at the growth of jobs after support. This then forms the basis for estimating what employment growth would have occurred anyway to establish additional economic activity.

Measuring gross economic activity for Innovate UK

The businesses supported by Innovate UK can be tracked after they have been supported in the BSD. The BSD will indicate if their employment and turnover has grown and by how much; it will indicate if a business closes. For the analysis of post-support impact, a first decision is to decide over which period to observe supported businesses.

Figure 4.1: Businesses, projects and total spend for Innovate UK 2004-2016



Note: Annual spend on left-handside axis; Total projects each year higher than businesses supported as some businesses secured more than one incidence of support

Figure 4.1 (identical to chapter 1) was used to make this selection, with the focus for impact analysis being the businesses supported between 2009-13 with these then being tracked until 2016. The figure indicates this is a period of ramp up in the provision of innovation support. As innovation support is often provided to businesses that have received support in the past, this ramp up – as it occurs – involves both businesses new to Innovate UK support and some that would have received support in the past.

Gross economic activity is measured in terms of the employment in the supported businesses. Businesses that receive public support will see an increase in economic activity after this, usually taking on workers but also investing in R&D and registering intellectual property. A focus on employment to understand the scale of business activity increase is justified on two grounds. Firstly, employment is well-measured. The ONS data is largely based on payroll data and businesses report their workforce numbers through the tax system quickly and accurately. Secondly, especially for

simpler businesses with a small number of plants and a focused research project, it can be expected that a successful application to Innovate UK will translate into jobs.

There are two corrections made, however, to ensure that the gross jobs estimate does not double count businesses and the large complex businesses do not bias the measures. In the previous chapter, the selection modelling focused on an annual analysis looking at supported businesses in each year. However, where a business is supported in more than one year, in the gross measures, its activities are tracked from the first time the firm receives support. Secondly, to remove the extreme businesses in terms of size primarily, the propensity score was used to focus only on businesses with a propensity score close to the cut-off for selection. Supported businesses with a high or low score were removed from the gross jobs calculation, focusing on the half of the businesses in the 25th to 75th percentiles of the propensity score distribution.

Table 4.1 below presents the gross job growth for the businesses treated in four years. The year 2010 was the start of the analysis and all businesses supported could be included. This means that the 262 businesses in the sample includes ones supported in 2010 that had been supported in 2004-9. However, in the following years, only businesses that had not received support in the previous years were added into the employment totals.

Table 4.1: Jobs in supported business trimmed using propensity scores

Year of Support	Annual Employment estimates							Businesses supported
	Sep-09	Sep-10	Sep-11	Sep-12	Sep-13	Sep-14	Sep-15	
2010	87,170	88,713	86,618	91,199	96,140	99,563	101,199	262
2011	72,049	76,043	75,117	80,252	83,368	86,225	87,824	240
2012	222,499	212,011	211,245	220,989	224,216	233,466	234,262	288
2013	184,621	189,126	187,407	194,873	197,484	208,533	209,521	243

Figure 4.1 highlights that there is an increase in the number of businesses supported in 2011, 2012 and 2013. This growth means that the exclusion of all businesses that had previously received support is not overly restrictive. Some of the businesses entering into this pool are quite large, especially in 2011 and 2012.

By 2013, the businesses that have been supported and are in the trimmed sample, have 587,000 employees. While all are not growing in employment, overall their employment growth is positive for all four cohorts. Table 4.2. then tracks employment growth in businesses after support. A year after the support, businesses that were beneficiaries in 2012 had 9,744 more jobs, rising from 211,245 jobs to 220,989 (in Table 4.1).

Table 4.2: Job increases in supported businesses trimmed using propensity scores

Year of Support	Gross additional job years created						
	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Total
2010	1543	-2095	4581	4941	3423	1636	40,412
2011	-926	5135	3116	2857	1599	0	32,572
2012	9744	3227	9250	796	0	0	67,954
2013	2611	11049	988	0	0	0	30,921
Total							171,859

These represent the gross change in economic activity seen in the businesses and Table 4.2 provides an estimate of the years of employment that this growth represents. Around 170,000 years of employment had been created in the supported businesses in the years after support. The employment years in the supported businesses is around 2.5 million – estimated by adding the jobs in each year after support in Table 4.1 (e.g. employment from September 2012 to September 2015 is summed to estimate total years for businesses supported in 2012). So, supported businesses have seen an increase of economic activity of 7%. Many of the businesses are quite large, with the total employment in the sample being around 2% of UK employment.

Gross impacts by type of support

Tables 4.1 and 4.2 indicate changes in economic activity using the preferred measures, totals that trim the supported businesses removing the largest businesses. The gross employment impacts of the support provided by Innovate UK can be analysed by the types of support. Table 4.3 calculates the job years created after collaborative R&D support, which is a subset of the businesses that are seen in the previous analysis.

Table 4.3: Gross job years in businesses engaging in Collaborative R&D, 2009-12

Year of support	Gross job years created						
	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Total
2009 (N=461)	1,913	-2,598	5,683	6,129	4,246	2,029	17,402
2010 (N=398)	-876	4,857	2,947	2,702	1,513	-	11,143
2011 (N=398)	10,116	3,351	9,603	827	-	-	23,897
2012 (N=393)	2,601	11,003	984	-	-	-	14,588
Total							67,030

The table focuses on the years of employment created, calculating the change in employment from year of support in the treated businesses one year after, two years after and three years after. This shows a ramp up in employment, but also highlights

that there are some very large businesses that receive grants. This table is before the removal of 40 very large firms decreasing the job growth by 33,000 job years.

Table 4.4: Gross job years created by Innovate UK grant duration in 2012

Duration of support	Year of support	Year 1	Year 2	Year 3	Total
All treated	2012 (N=1,186)	22,980	44,961	68,357	136,298
12 months or less	2012 (N=430)	2,915	3,381	5,399	11,695
12-24 months	2012 (N=304)	3,520	1,401	25,433	30,354
24 months or more	2012 (N=450)	19,738	42,407	39,350	101,495

This is confirmed by Table 4.4. The largest amount in gross jobs is coming from businesses receiving grants over 24 months in length. As shown in the earlier summary statistics, those receiving longer grants are also larger businesses, which often see larger gross increases in jobs than smaller businesses.

Economic activity and NMS Support

Selecting the period to track economic activity for NMS beneficiaries follows a similar approach to that used for Innovate UK. Again, ideally a period can be chosen to identify supported businesses and appropriate comparison groups that is timed so that there are sufficient post-support years of data to estimate impacts. However, a problem is that there is a high level of recurrence in businesses that receive NMS support, with many businesses using NMS service in all years making defining an intervention point impossible.

Figure 4.2: Businesses, projects and total spend for NMS 2002-2017

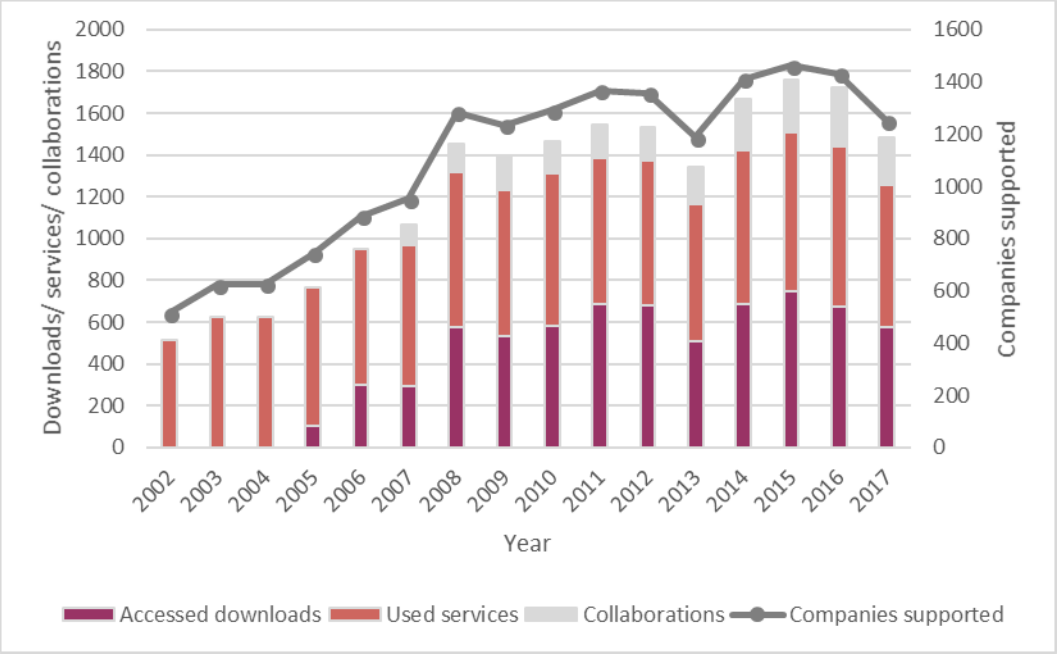


Figure 4.5 indicates how the number of businesses receiving NMS “used services” has been quite stable over the period 2002-17. There has been a ramp up in downloads of NMS material as NMS online capabilities were introduced in the mid-2000s. Collaborative research between NMS and businesses started in 2007. By the period analysed in this chapter, after 2008, however, all three forms of services had levelled in terms of businesses supported.

As innovation support is often provided to businesses that have received support in the past, this stable period reflects relatively few new businesses entering into NMS support. Further, there is an overlap in the businesses between those supported by Innovate UK.

As with Innovate UK, gross economic activity is measured in terms of the employment in the supported businesses. Whereas with Innovate UK, significant support should translate into jobs, this assumption is a little less appropriate for NMS’s intervention logic than is the case for Innovate UK. NMS support is indirect, such as the subsidy related to its measurement services, rather than direct giving of grant for R&D.

There are three corrections made for NMS. Two are similar to the approach for Innovate UK: to ensure that the gross jobs estimate does not double count businesses, where a business is supported in more than one year, in the gross measures, its activities are tracked only once. The second correction is the removal of extreme businesses, in the propensity to receive support. As in Innovate UK, supported businesses with a high or low propensity score were removed from the gross jobs calculation, focusing on the half of the businesses in the 25th to 75th percentiles of the propensity score distribution.

The third correction is to remove Innovate UK supported businesses, focusing on those that are included in the Innovate UK analysis for this study. There are about 140 businesses that NMS have supported but also have received Innovate UK support by winning funding for a research project. Businesses receiving only small innovation grants such as Innovation Vouchers and then supported by NMS are retained for the NMS analysis. The overlap between Innovate UK and NMS businesses – unsurprisingly – includes many large businesses. These businesses have multiple innovation streams and are so quite likely to secure Innovate UK funding for some of their research, while also have research that uses NMS measurement services.

Table 4.6 below presents the gross job growth for the businesses treated regularly with NMS support after 2009. If all 358 businesses are included, then their combined employment is almost half a million, peaking at the end of the period at 476 thousand jobs.

Table 4.6: Jobs in all supported businesses and trimmed using propensity scores

Regularly Supported by NMS	2009	2010	2011	2012	2013	2014	2015	Change in jobs	Obs
All Regularly NMS Supported	430,166	446,285	449,330	441,885	440,512	446,780	476,173	46,007	358
NMS Supported, Trimmed	126,279	128,829	132,549	131,692	131,933	132,916	131,117	4,838	217
All Regularly NMS Supported & Unsupported by IUK	85,641	85,448	86,245	91,861	92,778	96,335	94,285	8,644	257
NMS supported, trimmed and Unsupported by IUK	58,160	57,622	59,554	61,693	62,684	66,019	64,961	6,801	175
Sometimes Supported by NMS	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**Change in jobs**	**Obs**
All Supported	896,844	898,920	886,251	879,627	880,815	879,343	875,927	-20,917	1616
NMS Supported, Trimmed	352,453	333,129	347,767	349,338	344,185	342,676	329,539	-22,914	874
All Supported & Unsupported by IUK	322,273	319,786	318,523	313,672	316,930	319,363	314,645	-7,628	1269
NMS supported, trimmed and Unsupported by IUK	145,565	142,241	150,503	149,886	143,720	143,559	139,623	-5,942	735

The table then focuses on subsets of these businesses. There are 217 businesses which have propensity scores between the 25th and 75th percentiles. These are the businesses that are not very likely to receive support or highly likely, trimming to focus in on typical NMS supported businesses. While the business count falls by about half, the employment covered falls much more. This is because this segmentation removes large businesses that are highly likely to routinely use NMS support.

The preferred measure of gross economic activity excludes businesses that have received Innovate UK grants and do not have extreme propensity scores. This is the "NMS support, trimmed and Unsupported by Innovate UK" row. Here 175 businesses' employment is tracked. They grow so that by 2017 they have 6,801 more jobs than they had in 2009.

The lower panel of the table focuses on the 1,616 businesses that receive NMS support only sometimes. This accumulates about 250 businesses each year in that the supported businesses will receive NMS services only intermittently in this period. Again, the table proceeds to focus on businesses that are likely to sometimes receive NMS support and those that have not received Innovate UK funding for projects.

An observation for this set of businesses is that there is evidence of jobs being reduced in this set of businesses and – in terms of employment – economic activity

is decreasing. This is consistent with later analysis where the sometimes supported appear to have low additional activity suggesting their behaviour is similar to comparable businesses.

Table 4.7: Additional jobs in NMS supported businesses

Regularly Supported	2009	2010	2011	2012	2013	2014	2015	Job Years
Gross employment for NMS supported trimmed	58,160	57,622	59,554	61,693	62,684	66,019	64,961	430,693
Job years added each year		-538	1,932	2,139	991	3,335	-1,058	23,573
Additional job years		-429	1,542	1,707	791	2,661	-844	18,809
Sometimes Supported	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**Job years**
Gross employment for NMS supported trimmed	145,565	142,241	150,503	149,886	143,720	143,559	139,623	1,015,097
Job years added each year		2,076	-12669	-6624	1188	-1472	-3416	-80181

Table 4.6 presents the gross change in economic activity seen in the businesses and Table 4.7 provides an estimate of the years of employment that this growth represents. Around 23,000 years of employment had been created in the supported businesses in the years after support. The employment years in the supported businesses is around 0.43 million – estimated by adding the jobs in each year after 2009 in Table 4.6.

This conversion to job years is a means to gauge overall activity over the period and focuses on the NMS supported business, trimmed using propensity scores and not in receipt of Innovate UK funding. The second row then considers the growth in job years over the period. So, supported businesses saw a fall in jobs of 538 in 2010 and this is then multiplied by six years to understand the effect of this fall during the period. This fall is quickly revered in 2011 as 1,932 jobs were added in these last five years. Over the period, 23,573 job years have been added by the supported businesses in 2009-15, an increase of economic activity of 5.5%.

Additionality of impacts of Innovate UK funding

To understand how much of the gross job increases is attributed to the Innovate UK grants (and later NMS services), the growth in employment seen in the supported businesses is compared to that observed in the counterfactual. Rather than undertaking this calculation using numbers of jobs, the growth rate is estimated for the two sets of businesses and the proportion of growth that is seen in the supported businesses but not seen in the counterfactual is calculated (an additionality ratio). A similar approach is used in Goldman Sachs (2018).

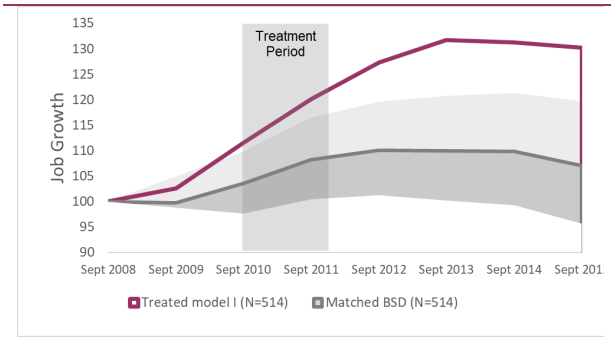
Figure 4.3 plots the growth trajectories for each of the four cohorts using model I to match a control group. To focus on growth in employment, the levels are indexed to

100 at two years before support, so that the pre-treatment growth can be inspected also. The figure indicates the period in which businesses would have initiated their Innovate UK project as a box. Also, the plots include confidence levels, as a fan around the control group.

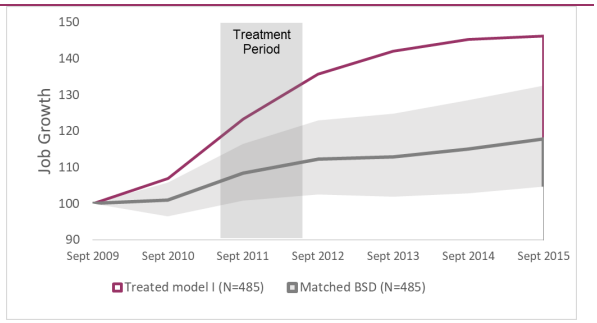
After support, beneficiary businesses see an increase in their economic activity as measured by employment. The pickup in growth is marked and – by the end of the period – the supported businesses have had several years of growth greater than comparison businesses. However, for the two early cohorts with more years of post-treatment data, there is some flattening out of growth in later years.

In terms of assessing what might have happened without support, firstly, matching makes the pre-treatment growths balance in one cohort (2013) but tends to increase the pre-support growth difference in other cohorts. However, the plots indicate pre-support group is not significantly different between supported and the matched control except in one year: tests indicate one of the differences in pre-support employment growth rates (2010) is significant. As noted above, 2010 is a difficult year to model, as estimation does not focus on the businesses that are new to Innovate UK grants, as is the case in the 2011, 2012 and 2013 cohorts.

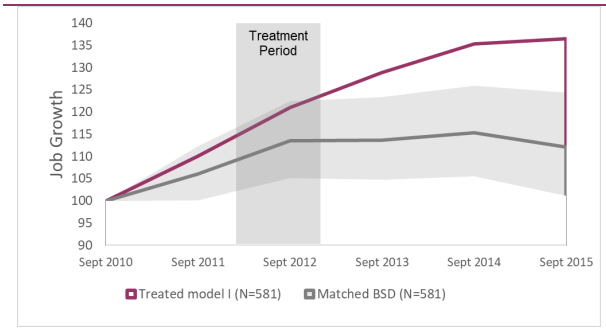
Figure 4.3: Employment growth in Innovate UK supported businesses



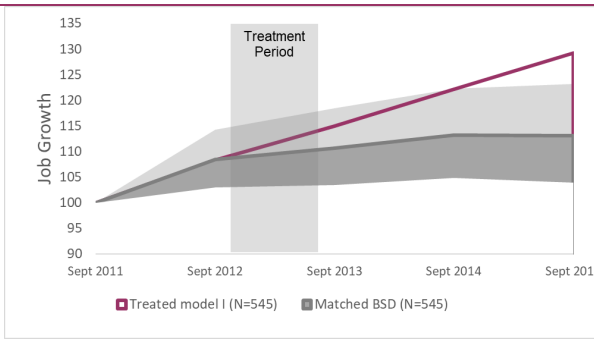
Treated in 2010



Treated in 2011



Treated in 2012



Treated in 2013

Note: Employment indexed at 100 two periods before support so a year of pre-support employment growth is presented; the grey areas indicated growth that does not differ from the matched control at 5%.

After support, the employment growth observed in the businesses that receive Innovate UK grants is higher than those in the matched BSD control groups. The four cohorts of supported businesses have higher employment growth than the counterfactuals. For those treated in 2010, by 2015, the employment growth has been 23% higher; for the 2011 cohort, the difference is 30%; for 2012 the difference is 24% and for the last cohort, the difference is 16%.

Attributing Employment Growth to Innovate UK grants

Attributing some of the gross jobs increases seen in the supported businesses to the support can be done in a variety of ways. The difference-in-difference of employment is a measure of the impacts per business and – if multiplied by the number of supported businesses – provides a measure. However, this measure can be quite sensitive to any differences in the sizes of the supported and matched businesses in levels after matching. Box B in chapter 3 explores this.

An alternative way is to calculate the share of gross jobs growth that is attributable to support using the difference in difference estimates on employment change. Table 4.7 applies these shares to the gross job years estimated in Table 4.2. To arrive at the proportion of gross jobs that are net additional, the gross job years created are

multiplied by an additionality ratio. This additionality ratio is estimated using the growth difference between supported and matched unsupported businesses. The additionality ratio varies across cohorts, ranging from 49% to 72%. Overall, almost half of the gross job years growth seen in the supported businesses is not observed in the comparison group. This means that around 101,000 job years could be additional economic activity.

Table 4.7: Additional job years in Innovate UK supported businesses, 2009-12

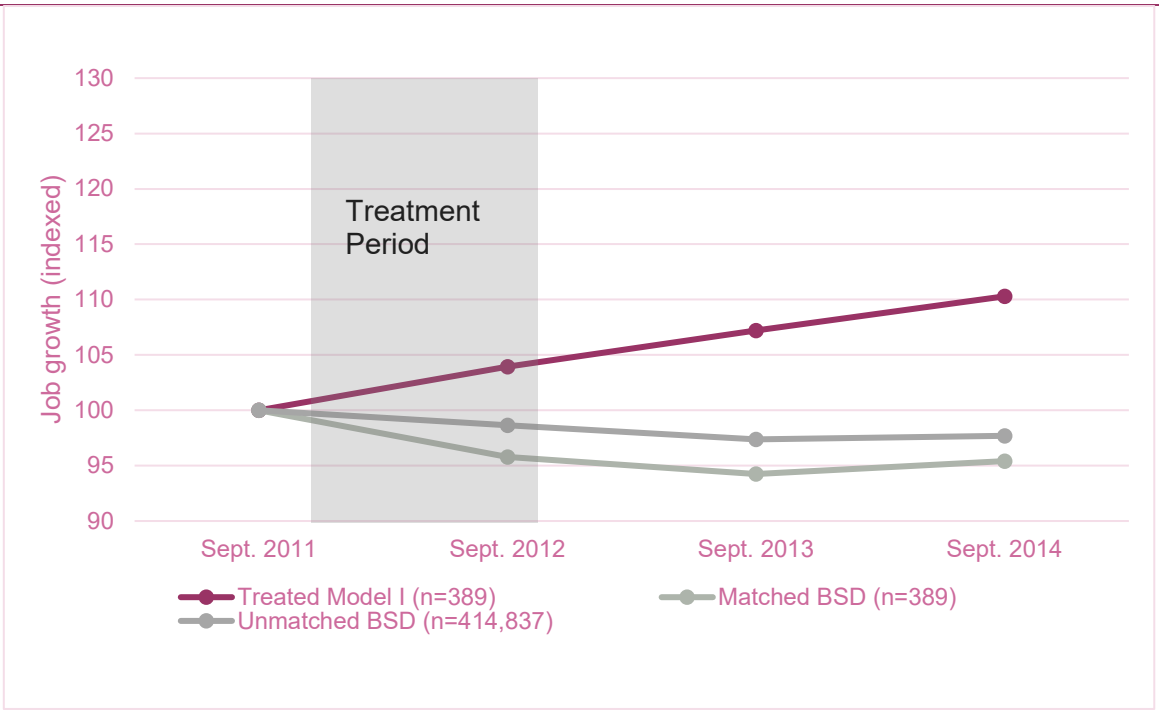
Year of support	Businesses Supported	Job Years	Additionality Ratio	Net Additional Job Years
2010	262	40,412	71.8%	29,002
2011	240	32,572	52.1%	16,986
2012	288	67,954	59.5%	40,446
2013	243	30,921	48.6%	15,036
Total		171,859		101,470

Single Grant Recipients

The analysis above has not sought to control for the Innovate UK grants received by businesses after first support, by focusing on the year of first treatment and accepting that businesses will subsequently receive funding, either in the second year of a project or as new projects are won. All effects from future grants are included in the outcomes presented above.

To test whether this assumption is valid, results are also shown for those businesses who receive only one grant from Innovate UK across all years. Since there is no risk of double counting or compounded treatment effects these results can be presented over multiple years. Similar to Figure 4.3 initially, businesses receiving one Innovate UK grant have positive employment growth after receipt of the grant, while those similar businesses who never receive an Innovate UK grant have negative growth.

Figure 4.4: Businesses who receive an Innovate UK grant in 2012 only



Next, gross job years and additional net jobs are presented for those who have received a single grant. While there was a gross job decrease across the whole of the BSD, the businesses treated by Innovate UK had less of a gross job decrease than similar treated firms.

Table 4.8: Gross and net additional job years created by businesses supported by Innovate UK grants, 2009-2012, no prior treatment

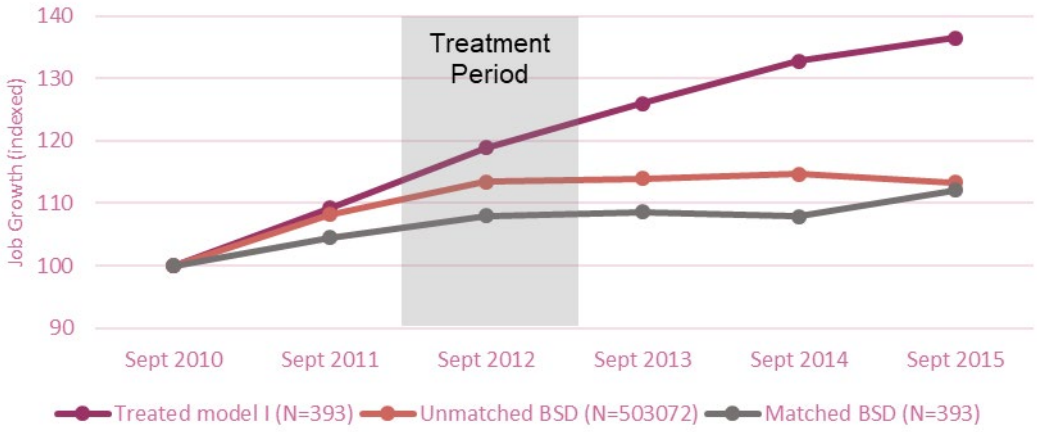
Year of support	Gross additional job years created						
	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Total
2009 (N=514)	1,543	-2,095	4,581	4,941	3,423	1,636	14,029
2010 (N=485)	-926	5,135	3,116	2,857	1,599	-	11,781
2011 (N=581)	9,744	3,227	9,250	796	-	-	23,017
2012 (N=545)	2,611	11,049	988	-	-	-	14,648
Total							63,475

Year of support		Net additional job years created						
	Add. Ratio	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Total
2009 (N=514)	72%	1,107	-1,503	3,288	3,546	2,457	1,174	10,068
2010 (N=485)	23%	-211	1,169	709	650	364	-	2,681
2011 (N=581)	44%	4,312	1,428	4,094	352	-	-	10,186
2012 (N=545)	52%	1,356	5,739	513	-	-	-	7,608
Total								30,543

Further analysis which was done focused on support identified as Collaborative R&D. This shows the same trend, with treated businesses growing faster than comparable businesses in the BSD. Table 4.8 shows that this group of Innovate UK

beneficiaries have seen a gross job creation of between 12k and 23k depending on the year of treatment, and a high level of additionality across the cohorts. Out of the 67k job years created in these treated businesses in the year 2010 to 2016 it is estimated that 44k were additional.

Figure 4.5: Businesses Engaging in Collaborative R&D 2012 only

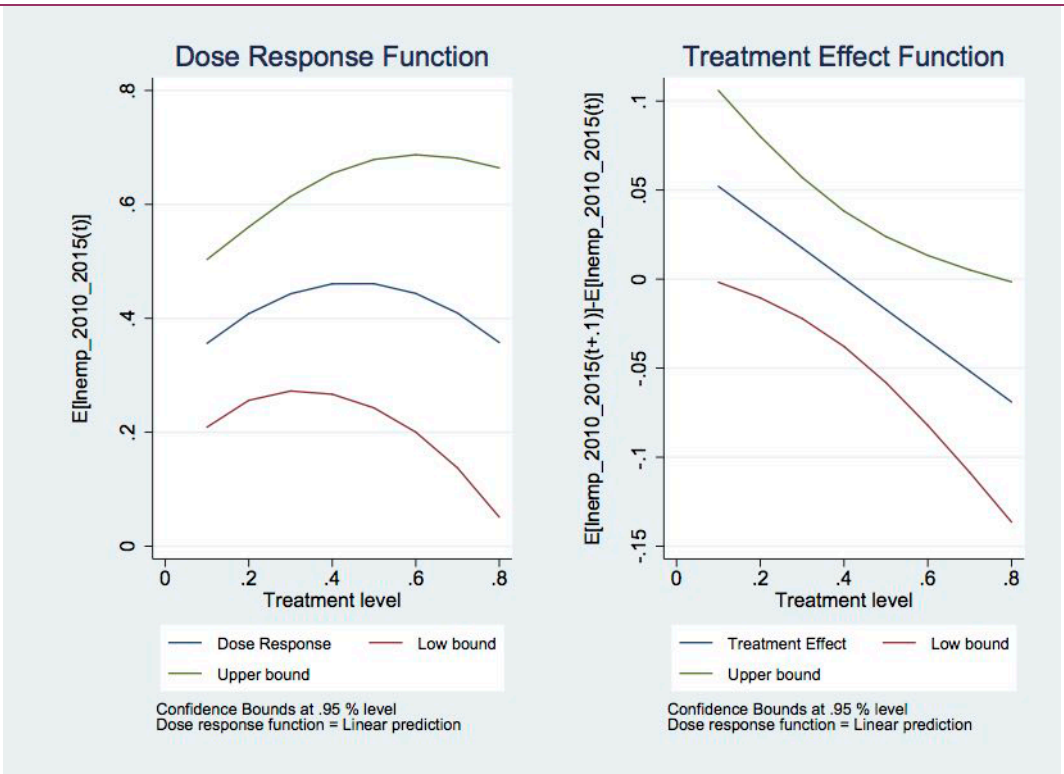


Dosage modelling

Innovate UK projects vary by grant size and the length of the project. This means that – at a business level – impacts are likely to vary both because of the fact of receiving support and the scale of the projects that are funded. Dosage modelling can be used to reveal the causal relationship of interest: a dose-response function is estimated.

The dose here is the length of projects, normalised so that the shortest project takes the value zero and the lengthiest project has a unit dose. Alternative dosages, such as project expenditure were considered. However, the measure is affected by firm size in a complicated manner. Larger firms tend to receive larger projects, but the correlation is not linear with size or easy to model. Analysis of the distribution of the project length suggested outliers was not a problem with the maximum lengths of around four years as expected and distributional peaks at 6 months, 12 months and whole years.

Figure 4.6: Treatment Impact by Dose Size for Innovate UK, 2011 and 2012



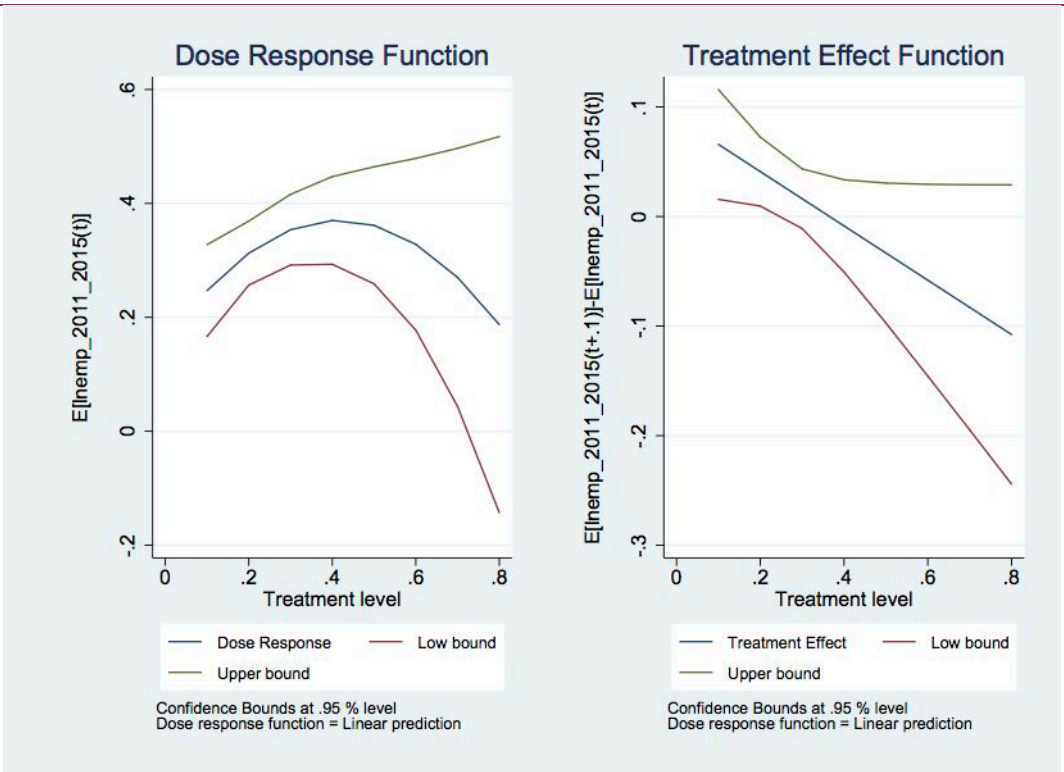


Figure 4.6 indicates the dose response function and treatment effect function for 2011 (top) and 2012 (bottom) for Innovate UK projects. To estimate the functions, the supported businesses were the focus and the matching was on similar variables

to the selection modelling. However, here the econometric modelling is focused on understanding what determines project length. The estimated function maps the length of supported projects to the economic activity change observed, measured by employment change over 2010-15 for the 2011 cohort and 2011-15 for the 2012 cohort. The central line in each figure is the function with upper and lower bounds estimated also.

Firstly, there is a positive impact on employment change for all doses. It is higher in 2011 than in 2012. Also, both cohorts of support indicate possible optimum lengths of project, with dose response peaking at about two years project length. The derivatives of the dose-response function, which indicate the increase in the growth rate in percentage points resulting from a marginal increase in the dose, is presented in the right panels. The optimal point is where the x-axis is crossed.

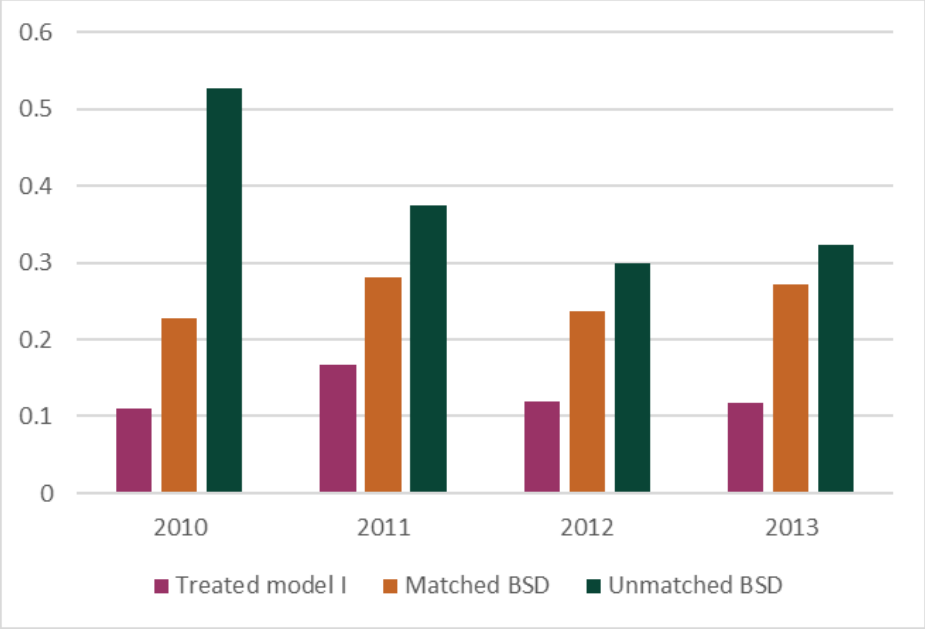
Various other dosage proxies – such as the amount of support – was found to be very difficult to model satisfactorily. Dosage modelling, by focusing on the level of support, can then use measures that are correlated with the size of business, which in turn is linked to any treatment effect. To some extent, the use of the duration of the support is less affected by this. However, as the appendix indicates, a size effect on this dosage measure remains.

Firm survival impacts

Support provided to businesses affects their chance of survival. All but 10% of businesses that received support in 2010 survived until 2017, while in the wider business population, average business survival is much less common, with just over half of businesses disappearing from the Business Structure Database in the eight years. This differential survival rate could result in additional economic activity, occurring due to the support resulting in fewer businesses closing than other wider would be the case.

Figure 4.7 indicates the closure rates for the matched comparison businesses using Model I, which are higher than for the supported businesses. It provides results for four cohorts of supported businesses. The survival rates of the BSD businesses – as expected – exhibits a rise as the graph moves to more recent years, reflecting that a share of businesses is likely to close each year. The dataset analysed here is a panel, with each cohort analysed without new businesses replenishing the sample. It is therefore fairly predictable that each year will see a large number of the businesses in the BSD exit. Survival rates also vary for different cohorts, but this is likely to reflect the variation in the businesses that are supported and then matched to the supported.

Figure 4.7: Businesses Closure Rates by Year of Support



The variation in the survival rates of the samples of businesses that are supported each year is likely to be due to the characteristics of the supported businesses differing from the wider BSD. So, as the businesses that Innovate UK support are larger or older than the wider BSD, their survival is likely to be higher than the wider BSD and these determine survival rates.

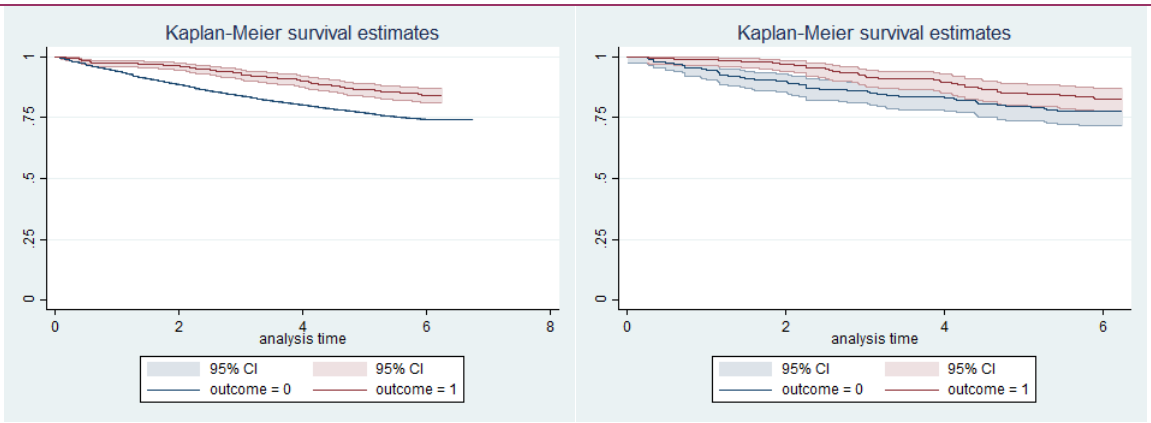
To control for this, the Box-Cox approach was used to model survival rates for the supported businesses, the wider BSD and the matched businesses using Model I. At its centre is a Cox proportional hazard model to assess whether support has an impact on business survival while controlling for other characteristics which affect survival on the two different matched samples. The hazard function is the probability of closure in a period having survived until that period, driven by some baseline hazard adjusted by characteristics such as age, size and industry (Box and Cox, 1964).

Figure 4.8 shows the time to closure for supported and unsupported businesses. These are Kaplan-Meier survival curves[1] firstly comparing supported businesses with all businesses (on the left panel) and the treated with the matched sample (on the right panel). The matched sample was created by the propensity score matching using the selection model earlier. The estimated survival curves are always higher for the businesses that received support which shows that successful applicants always have a higher probability of surviving than matched. Therefore, we can be somewhat confident that the support does have an effect on survival.

[1] The Kaplan-Meier estimate of the survivor function is given by (Kaplan and Meier, 1958): $\hat{S}(t) = \prod_{j|t_j \leq t}(\frac{n_j - d_j}{n_j})$,where nj is the number of businesses not closed at time tj and dj is the number of businesses closed at time t_j.

On the left panel, before matching, the survival rates for the unsupported businesses is the lower line, significantly lower than the survival rates for the supported businesses, with this difference being highly significant.

Figure 4.8: Businesses Survival Functions with and without matching, 2011



On the right-hand side is the survival rate estimated for the supported businesses and the matched counterfactual, using Model I. The matched control group is less likely to survive at each year following support. However, there is more overlap suggesting that matching has balanced on characteristics that determine survival. Some imbalance remains however.

A second analysis has then been performed to understand the impact on job years as survival has been impacted by support. This can go in two directions. On the positive side is the higher chance of a business surviving, resulting in job years increase. This is the case in 2010 (1,462 job years), 2012 (2,542 job years) and 2013 (2,085 job years). However, if supported businesses are larger than the matched control group, then these job years may be lessened where businesses that closed despite receiving support are large. In 2011, this offsets the survival premium and job years are reduced by 2,100.

Patent growth

To assess impacts beyond employment and turnover the analysis also estimated the impacts on patenting of support. The results showed a larger growth of patents in business beneficiaries post receipt of Innovate UK grant as compared to similar businesses that had not received the support.

Looking at patenting behaviour in the years 2009 to 2014 (Table 4.9), 8 to 15% of businesses supported by Innovate UK applied for a patent afterwards up to four years later. Only 5-10% of businesses similar to those supported by Innovate UK applied for a patent in the same timeframe. This shows that Innovate UK grant beneficiaries apply for patents at a higher rate (5-10%) than other similar businesses.

Table 4.9: Percentage of businesses applying for a patent four years later

	Innovate UK businesses	Similar businesses	Difference	T-stat	Count
2009	12.5%	4.9%	7.6%***	4.83	633
2010	14.3%	5.8%	8.5%***	5.08	638
2011	14.0%	4.2%	9.9%***	7.00	812
2012	15.2%	5.8%	9.4%***	7.57	1193
2013	11.2%	4.8%	6.4%***	6.98	1740
2014	8.5%	3.5%	5.0%***	6.32	1771

*** significant at the 1% level; ** at 5%; * at 1%. Standard errors are robust

Table 4.10 looks at this difference across several different models, and provides a weighted average. All differences are significant at the 5% level except the three highlighted in grey. The models discussed earlier highlight that the results are robust across different takes of the data. Additionally, they highlight that the lower bound of the additional patent registration is 2-5% over the years. Model III offers a lower bound of the patent effect by allowing prior treated Innovate UK firms into the matching pool.

Table 4.10: Additional patenting of Innovate UK businesses by model

	2009-2014	2009	2010	2011	2012	2013	2014
Model I	7.3%**	7.6%**	8.5%**	9.9%**	9.4%**	6.4%**	5.0%**
Model II	7.2%**	7.7%**	9.7%**	7.8%**	9.8%**	6.9%**	4.5%**
Model III	3.4%**	4.8%	3.0%	2.6%	5.2%**	3.7%**	1.9%**
Model IV					9.2%**		
Model V					9.4%**		

*** significant at the 1% level; ** at 5%; * at 1%. Standard errors are robust

These results are skewed because Innovate UK businesses receiving more than one grant are picked up multiple times, once for each grant they receive. For example, if a business receives a grant in 2009 and then again in 2011, and applies for one patent in 2012, its effect will be counted in 2009 and 2011. This effect is dealt with in two ways.

Firstly, Model I controls for prior treatment. This means that for all those businesses who have received prior treatment, there will be a similar number of firms who received prior treatment in the matching pool. The earlier effect should be washed out by this variable.

Secondly, the analysis split the treatment group into businesses that only receive a single grant over the years, and those that receive multiple. Although it is difficult to define a treatment effect for the group with multiple treatments the results for single

treatments offer a clear result for the segment of businesses with a single Innovate UK grant over the years.

Table 4.11 shows that between 2-7% businesses with a single Innovate UK grant apply for patents four years after receiving Innovate UK grants. Only 1-3% of similar businesses apply for patents in a similar time period. This shows an effect of 2-6% depending on the year of grant receipt.

Table 4.11: Percentage of businesses applying for a patent four years later, 2009-2011

	Innovate UK businesses	Similar businesses	Difference	T-stat	Count
2009	4.4%	1.5%	2.9%*	1.76	204
2010	5.7%	3.1%	2.6%	1.24	193
2011	3.0%	1.9%	1.1%	0.84	266
2012	6.8%	1.3%	5.6%***	4.01	395
2013	4.3%	2.2%	2.1%**	2.23	725
2014	2.3%	0.6%	1.7%***	2.71	701

Single grant recipients only; *** significant at the 1% level; ** at 5%; * at 1%. Standard errors are robust

Additionality of impacts of NMS support

This section replicates the Innovate UK analysis for NMS, comparing the growth in employment seen in the supported businesses to that observed in the counterfactual.

Attributing Employment Growth to NMS Support

To understand how much of the gross job increases are attributed to the NMS services, the growth in employment seen in the supported businesses is compared to that observed in the counterfactual. Rather than undertaking this calculation using numbers of jobs, the growth rate is estimated for the two sets of businesses and the proportion of growth that is seen in the supported businesses but not seen in the counterfactual is calculated (an additionality ratio).

Table 4.12: Estimates of additionality in jobs growth in supported businesses

Model to estimate DiD	Additionality ratio	Diff-in-Diff	Growth in Treated	Growth in Control	Obs
Regular support from NMS					
Model I	0.80	0.18***	0.23	0.05	696
Multivariate on regular treat and matched control	n/a	0.21***	n/a	n/a	696
Multivariate on regular treat	n/a	0.24***	n/a	n/a	306,993
Multivariate on trimmed using PS	n/a	0.20***	n/a	n/a	1,656
Multivariate trimmed sample with InnUK	n/a	0.23***	n/a	n/a	306,993
Multivariate with InnUK	n/a	0.21***	n/a	n/a	1,656
Sometimes supported from NMS					
Model I	-	-0.03	0.05	0.09	2,990

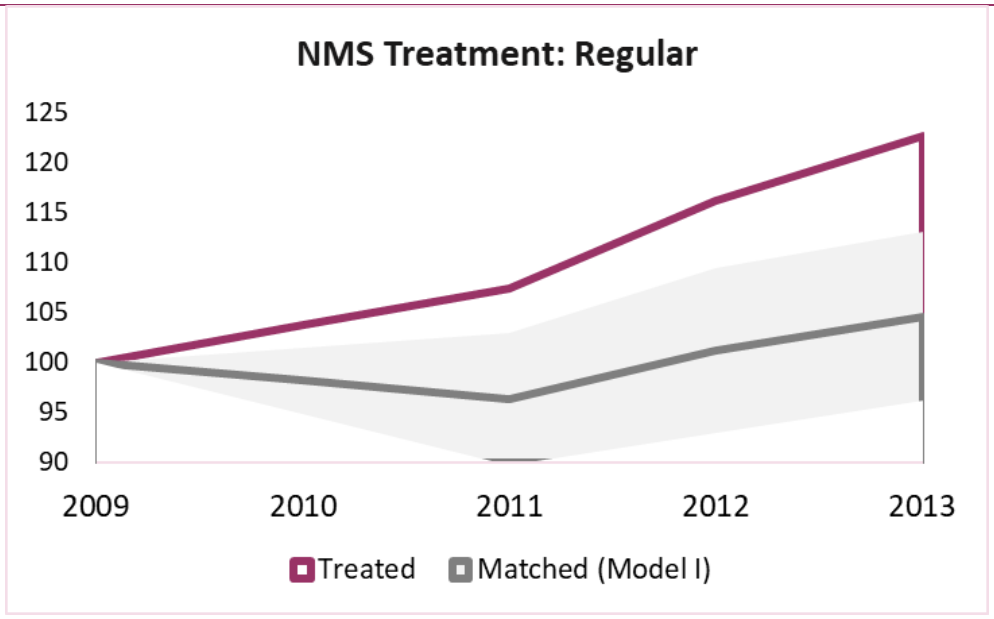
*** significant at the 1% level; ** at 5%; * at 1%. Standard errors are robust; Multivariate analysis only presented to compare alternative estimates of the DiD term, not used to calculate additionality.

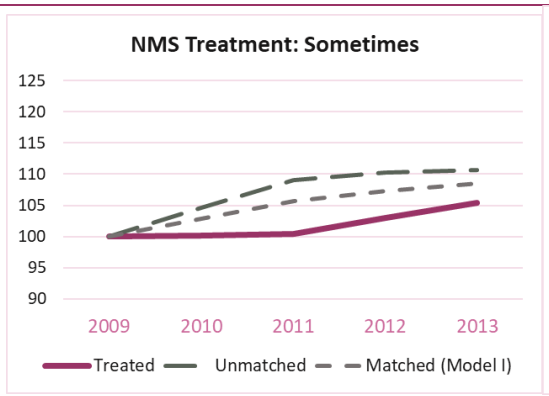
Table 4.12 indicates different estimates for the regularly supported businesses and the key column is the estimated difference-in-difference. This indicates that the growth seen in the supported exceeds that in the unsupported by 18-24% for the three years at the start of the 2010-17 period. The first estimate uses the 348 treated, which grow at 23% and compares with 348 unsupported businesses, which grow at 5%, implying that 80% of the growth seen in the supported businesses is not observed in the comparison groups. This ratio is then used to estimate that around 18,809 job years are additional for the regularly supported businesses in Table 4.5.

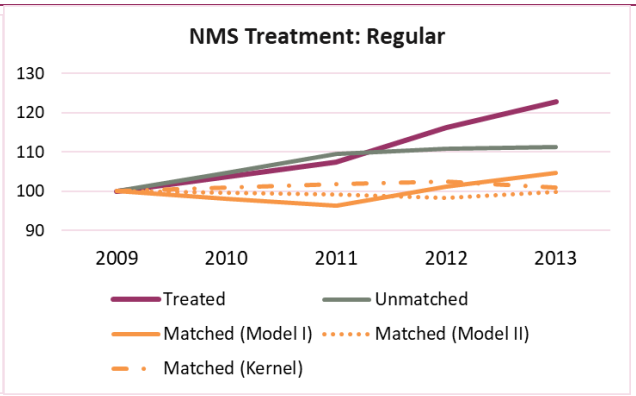
Figure 4.9 plots the growth trajectories for those regularly receiving NMS support in the top panel illustrating the confidence level and indicating a significant divergence from the matched control group. The lower panel repeats the analysis for the regularly treated, but also considers those only sometimes in receipt. The lower panel presents growth in the unmatched BSD, Model I and Model II to match a control group. It also presents a line for the control group using a kernel estimation technique. To focus on growth in employment, the levels are indexed to 100 at 2009.

Regular NMS beneficiary businesses see a 22% increase in their economic activity as measured by employment. In terms of assessing what might have happened without support, firstly, matching tends to increase the growth difference.

Figure 4.9: Employment growth in NMS supported businesses







The results for the sometimes treated indicates there is no evidence of additional jobs. The bottom left panel in figure 4.2 highlights how employment growth in these businesses is insignificantly different from control groups of businesses. Attributing some of the gross jobs increases seen in the supported businesses to the support can be done in a variety of ways, but it relies on there being evidence of a significant difference-in-difference term.

Table 4.12 also indicates some of the sensitivity tests applied to the difference-in-difference terms estimated for the sometimes treated. In particular, the growth seen in the supported is regressed on variables used in the selection model and a dummy taking one for those businesses regularly treated. There is a consistency in the different multivariate models that have been used.

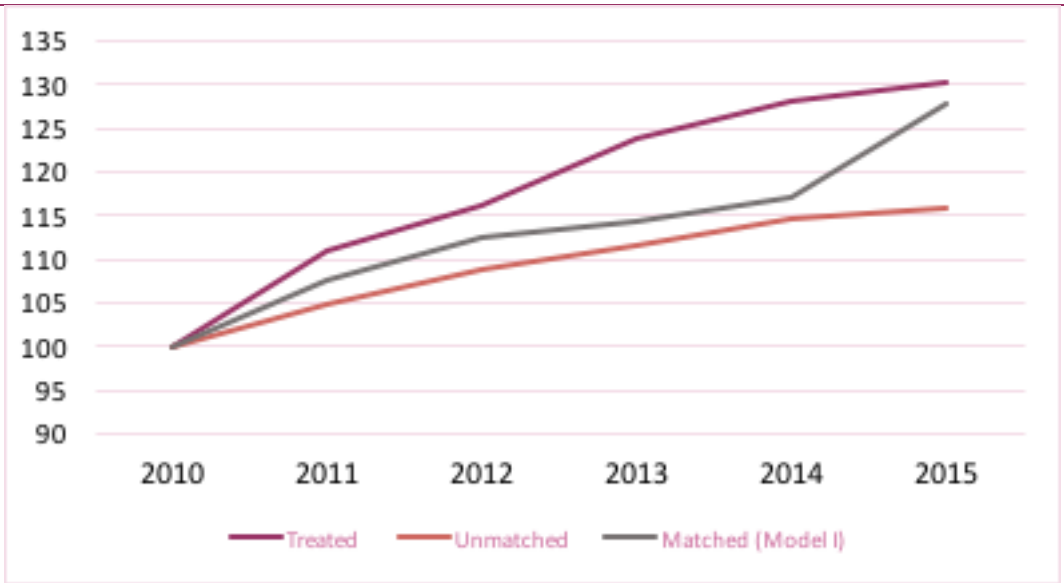
The first model undertakes this analysis focusing on the supported businesses and the matched control, so the sample is 696 businesses. This tests that the difference-

in-difference estimate is not changed should there be any unbalance after matching. So, by this regression incorporating the selection modelling variables (size, age, industry and past support and innovation activity), it can check that the estimate does not correlate with any selection criteria that have not been adequately modelled.

The propensity score is used in the third multivariate model. Here only the businesses with the 25% and 75% percentiles of the propensity score are included, covering both some of the supported and some of the unsupported businesses. The final regressions add to the overall multivariate regressions variables about Innovate UK support, testing whether the employment growth correlated with the three years of support after the start of the 2010-17 period (i.e. 2010-13).

Figure 4.10 plots out the employment growth that the significant difference-in-difference estimates reflect. While the difference-in-differences focus on three years of growth only, it indicates these differentials persist for much of the period.

Figure 4.10: Employment growth for NMS regularly supported vs comparators

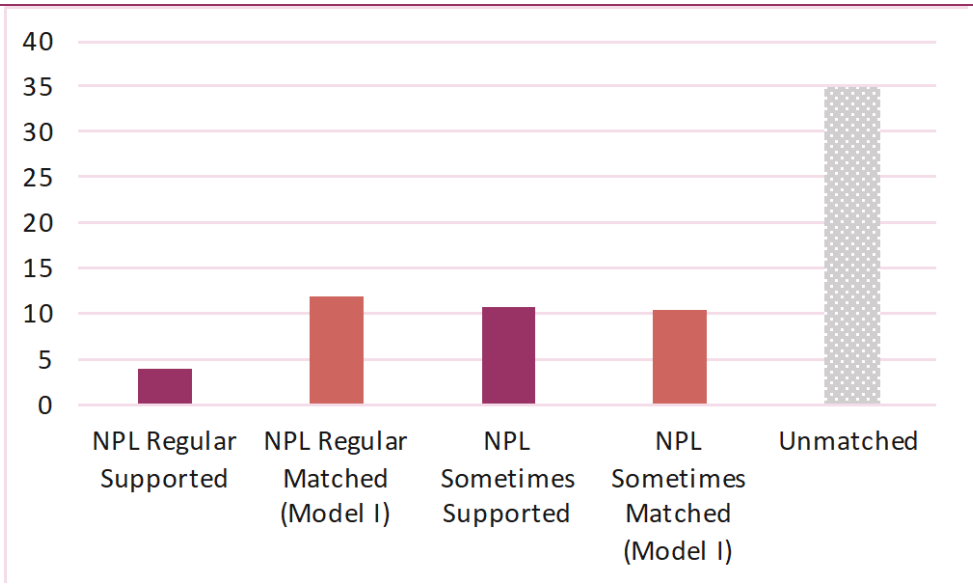


Firm survival impacts

Support provided to businesses affects their chance of survival. Of the businesses that regularly receive NMS support from 2010, all but 4% of businesses survived until 2017, while in the wider business population, average business survival is much less common, with just over a third of businesses disappearing from the Business Structure Database in the seven years. This rate of closure is higher for the businesses that sometimes use NMS services, where 10% close. This differential survival rate could result in additional economic activity, occurring due to the support resulting in fewer businesses closing than otherwise would be the case.

Figure 4.11 indicates the closure rates for the matched comparison businesses using Model I. It is noticeable that, while regularly supported businesses have a higher survival rate than the businesses matched to them, this is not the case for those sometimes receiving Innovate UK support. For these businesses, there is less of a survival effect with the matched and sometimes supported having almost identical closure rates, though each is much lower than the wider BSD.

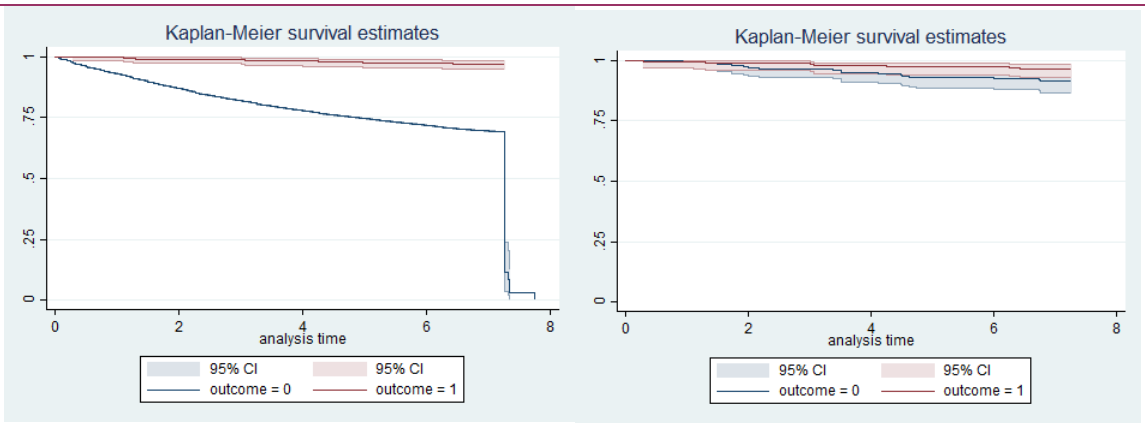
Figure 4.11: Businesses Closure Rates by 7 years



The variation in the survival rates of the samples of businesses that are supported each year is likely to be due to the characteristics of the supported businesses differing from the wider BSD. So, as the businesses that NMS support are larger or older than the wider BSD, their survival is likely to be higher than the wider BSD and these determine survival rates.

To control for this, the Box-Cox approach is used to model survival rates for the regularly supported businesses, the wider BSD and the matched businesses using model 1. This is plotted in Figure 4.12. On the left panel, before matching, the survival rates for the unsupported businesses is the lower line, significantly lower than the survival rates for the supported businesses, with this difference being highly significant.

Figure 4.12: Businesses Survival Functions for regularly supported



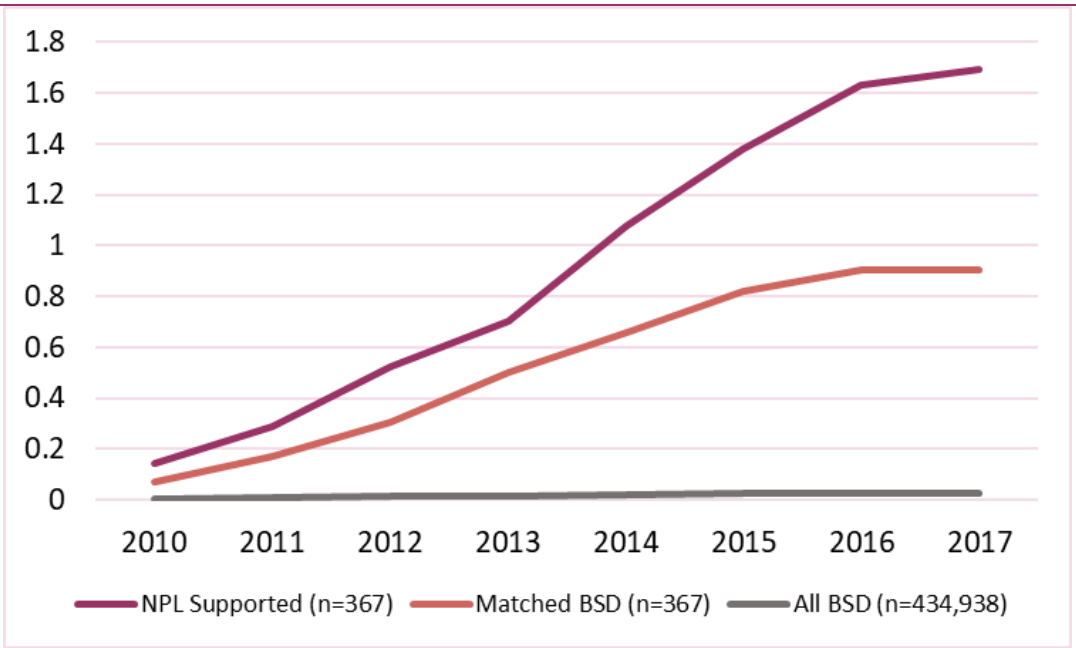
On the right-hand side is the survival rate estimated for the supported businesses and the matched counterfactual, using Model I. The matched control group is less likely to survive at each year following support. However, there is more overlap suggesting that matching has balanced with characteristics that determine survival. Some imbalance remains, however.

A second analysis has then been performed to understand the impact on job years as survival has been impacted by support. This can go in two directions. On the positive side is the higher chance of a business surviving, resulting in an increase in job years. There are 5,122 job years estimated due to this. However, support has resulted in 718 job years from businesses that would have closed without support. The net total impact is 4,404 job years.

Patent impacts

The businesses supported by the NMS are generally research active and so the registration of intellectual property is quite common. The number of patents registered for all businesses was linked to the BSD and then various measures were used to look at the supported businesses in comparison with other businesses. Figure 4.14 plots the average increase in patents per business from 2009 base. It indicates this for the supported businesses and the matched control by model I.

Figure 4.14: Change in patents for NMS supported and comparison groups



During the period after 2009, a gap emerges between supported businesses and the comparator groups in the number of new patents. The widening increases in the second half of the period. By 2017, the supported businesses have registered nearly two new patents on average, while the unsupported comparable businesses have registered about half this. The wider business population registers very few patents.

The modelling for the NMS was somewhat different for Innovate UK, in not looking at impact for support for different cohorts of businesses, rather looking at businesses regularly supported in the period. This has meant the measures used for Innovate UK (patents four years after support) are not possible for the NMS businesses. However, this measure does show that the quantum of patents in these businesses are different to the comparison groups, though because the variation in patenting is high, the differences are not statistically significant at standard confidence levels.

Concluding comments

Around 170,000 years of employment had been created in the Innovate UK supported businesses in the years after support, 2010-13. Supported businesses have seen an increase of economic activity of 7%. Many of the businesses are quite large, with the total employment in the sample being around 2% of UK employment.

The share of gross jobs growth that is attributable to support is estimated using the growth difference between supported and similar unsupported businesses. This means in different cohorts 22% to 72% of gross jobs might be additional and, overall, almost half of the gross job years growth seen in the supported businesses is not observed in the comparison group. This means that around 101,000 job years could be additional economic activity.

These estimates are then tested using alternative categories of support, such as focusing on businesses that are only supported once by Innovate UK and considering the subset of businesses that are supported by collaborative R&D projects. Consistent estimates are derived.

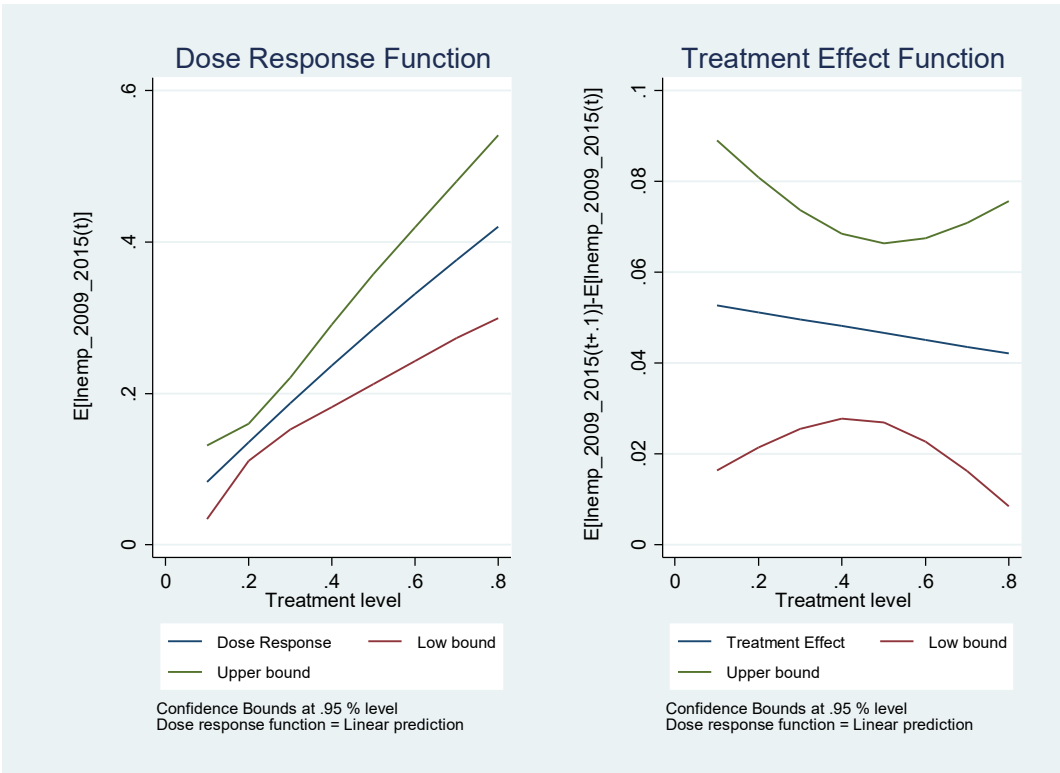
Analysis has looked at how impacts vary by the length of the supported projects. This uses dose modelling. Firstly, there is a positive impact on employment change for all doses. It is higher in 2011 than in 2012. Also, both cohorts of support indicate possible optimum lengths of project, with dose response peaking at about two years project length.

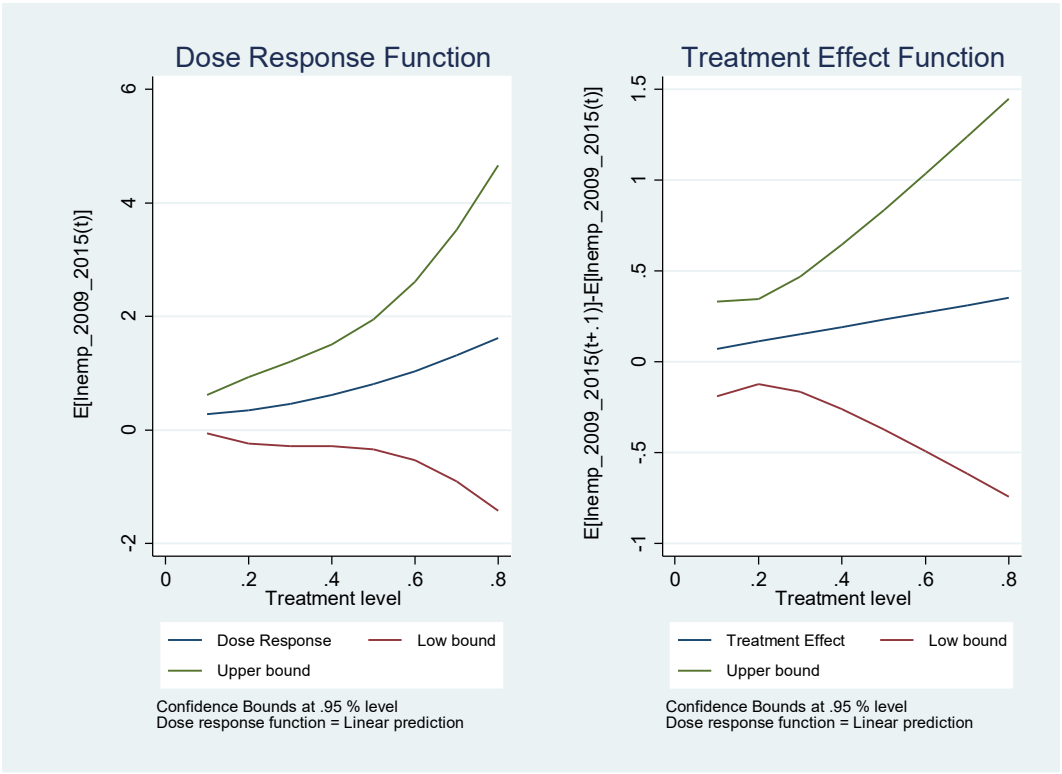
NMS dosage modelling

Figure 4.13 considers the dose response for NMS. Here, the dose is different to that used in Innovate UK looking at the number of times that services were purchased from the NMS and the amount spent. The employment impact is again modelled, and the modelling focuses on the businesses that received some support exploring how impact varies at different levels of support.

The top panel considers the number of times services are purchased during the period 2008-16. The average is 2.5, with many businesses having only a single incident of using measurement services. The results in dose response functions and treatment effects functions indicate that impact on employment does not diminish with increased treatment levels, with dose response increasing across the span of levels. The marginal effect graph in the top right panel is positive across all doses, with this being significantly different from zero across the range also.

Figure 4.13: Treatment Impact by Dose Size for NMS, 2008-16





The lower panel presents evidence where the value of the services purchased is modelled. Here the modelling is not statistically significant with the confidence interval always including zero impact, consistent with other analysis of the services type of treatment. The impact of the NMS value of purchased services on employment impacts proves more difficult to model than the number of incidences. This is consistent with other results where the value of support is used as the dose. The statistical difficulty is that the value measures are often a poor proxy for the size of the support given, primarily because the measure correlates strongly with the size of the business. Experimenting with adjusting for the size of the business did not mitigate this, the number of incidences of support being a preferred measure that proves to be less affected by the size of the business.

5 Understanding Productivity Impacts

Business support that fosters innovation has direct effects on business performance, such as on output and productivity. A measure of productivity may be the real sales per employee as BSD tracks turnover. Alternatively, wages of employees, as these are also closely linked to productivity, may indicate higher productivity levels in businesses supported in their innovation. This chapter explores the turnover effect of support and the wage rates in businesses supported by Innovate UK and NMS.

Businesses might train staff, hire additional staff to carry out the R&D or invest in capital that helps staff become more productive. Upskilling the workforce or giving them better tools to work with would all be expected to raise productivity and pay. While wages are not perfectly correlated with productivity, there is a strong correlation between the two (Hellerstein et al., 1999).

Analysing impacts beyond increases in economic activity has become a focus for evaluations, with appraisal highlighting the importance of productivity impacts. HMT Green Book (HMT, 2019) notes the importance of looking at evidence about productivity using earnings as a proxy. There is a recognition that only about half of the productivity effect is passed on to workers in the form of higher earnings (Dearden et al., 2005). Firms pass productivity shocks through to employees if employees can use the threat of outside job options to force a renegotiation of their wage (Postel-Vinay and Turon, 2010).

For innovation policy, productivity impacts may accrue beyond supported businesses, through spillovers. Public funding for R&D is often designed so that knowledge can be shared, and the benefits of new technologies widened (BIS, 2014) and this analysis has benefited from new data allowing spillovers from patenting to be quantified. The following gathers evidence for any of these effects for businesses supported by Innovate UK and NMS.

Turnover Impacts of Innovation Support

Businesses receiving Innovate UK support have generally seen higher real turnover growth rates than those observed in the wider BSD and matched control groups. This trend is true for all but one year of support. In the 2010 cohort, the real turnover growth of the supported businesses is similar to the matched comparison group.

Figure 5.1 illustrates the growth trajectories of businesses receiving support in 2012, but not before, compared to the matched and unmatched BSD using the preferred selection model. This is representative of the growth trends and differentials seen across all of the treatment cohorts 2009-2013, and the individual graphs are annexed in full (Annex D). In

supported businesses, real turnover has risen by 35-60% by 2015, with those treated first in 2011 achieving the highest growth. In the matched control groups, growth is 15-38%.

The four cohorts of supported businesses have higher employment growth than the counterfactuals. For those treated in 2010, by 2015, the employment growth has been 23% higher; for the 2011 cohort, the difference is 30%; for 2012 the difference is 24% and for the last cohort, the difference is 16%. However, for real turnover, the rates seen in the supported businesses do not exceed that seen in the comparison groups by the same amounts as in employment.

Overall, therefore, while the employment growth and the real turnover growth is higher in the supported than the counterfactual, the growth in real turnover per employee is generally lower in the supported businesses than the control.

Figure 5.1: Turnover growth in Innovate UK supported businesses 2012

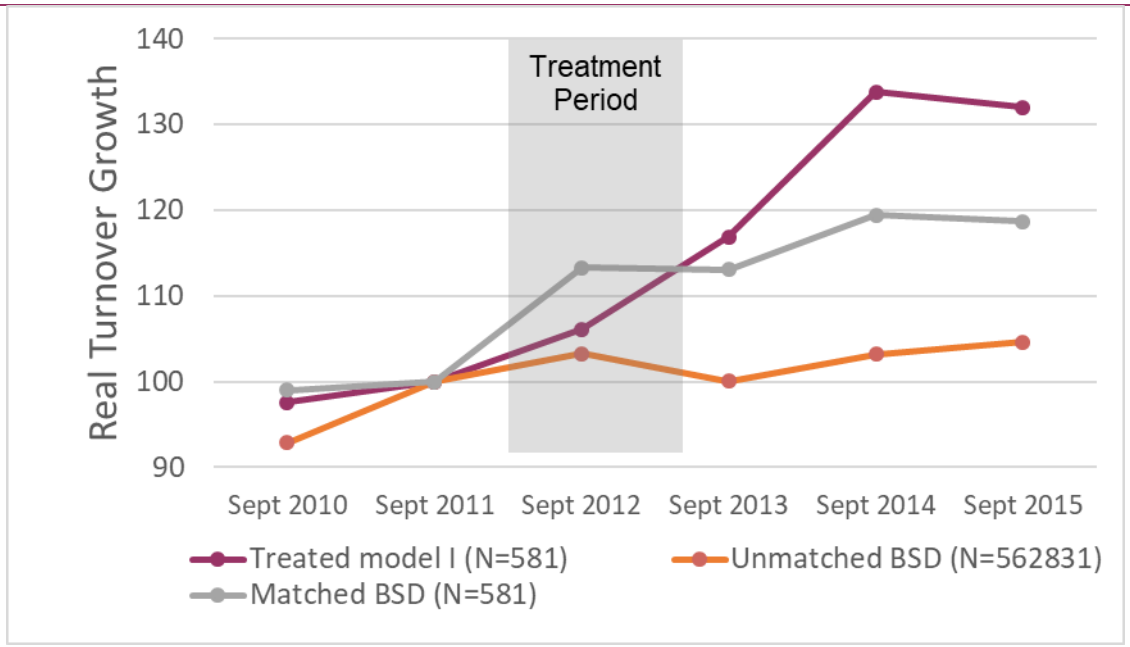


Table 5.1 presents these overall growth differentials, derived from the level of employment and real turnover seen in each cohort by 2015. It indicates that the real turnover per employee growth differential between treated and the counterfactual is negative for the 2010, 2011 and 2012 cohorts being as much as 21 points in the first of these. Only in 2013 do the supported businesses have real turnover per employee growth that is higher than that observed in the matched control group.

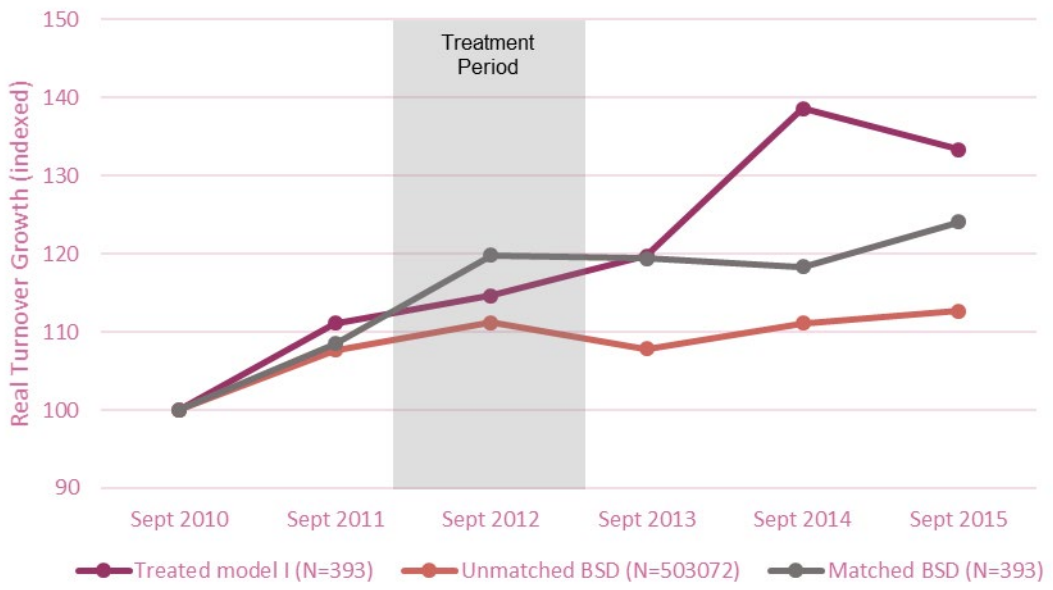
Table 5.1: Employment and real turnover growth differentials for Innovate UK beneficiaries

Supported in		2010	2011	2012	2013
Growth Period		2009-15	2010-15	2011-15	2012-15
Treated	Employment in 2015	131.2	136.5	136.5	129.1
Treated	Turnover in 2015	138.0	158.8	135.3	138.9
Matched BSD	Employment in 2015	109.8	115.0	112.1	113.1
Matched BSD	Turnover in 2015	138.0	135.0	119.8	115.3
Difference	Employment in 2015	21.4	30.4	24.3	16.0
Difference	Turnover in 2015	00	23.7	15.4	23.6

Figure 5.1 also indicates a second feature of trends seen in real turnover growth – a slight delay between treatment period and observed turnover change. This may reflect the lags between the timing of the R&D activities, and this translates into increased sales. There will be time taken to finalise the research and to commercialise innovation outputs.

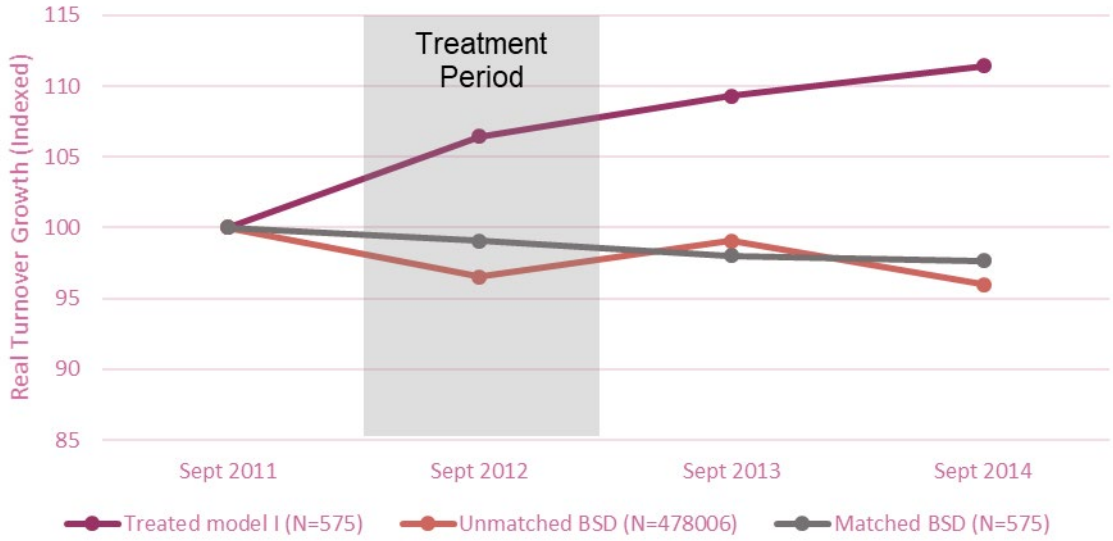
Both these trends can be seen for the subgroup of treated businesses engaging in Collaborative R&D, although the difference starts to close in the Financial Year 2014/15. Figure 5.2 shows the real turnover growth in the subset of Innovate UK businesses supported in 2012 through a collaborative R&D project. These projects tend to be lengthier and so the timing of the pick-up in turnover growth may reflect the time taken for the project to be completed.

Figure 5.2: Turnover growth in businesses engaging in Collaborative R&D 2012



Real turnover growth in the businesses receiving NMS Support in 2012 is presented in figure 5.3. There is a much quicker divergence in real turnover growth between the supported businesses and the matched and unmatched controls, which both experiencing negative growth rates.

Figure 5.3: Turnover growth in businesses collaborating with or using an NMS service, 2012



Gross value-added impacts

The previous results have focused on modelling using the ONS Business Structure Database. This section looks at impacts beyond turnover and employment through data from respondents to the ONS survey the Annual Business Survey, ABS.

ONS Annual Respondents Database

The ABS is a large-scale, official survey conducted each year. ONS then has compiled successive years of the survey and linked each by the business into the Annual Respondents Database (ARDx). The focus here is the period 2009-16, a period after the recession and in which the survey instrument was relatively stable. Prior to 2009, the ABS was known as the Annual Business Inquiry with some changes made during the transition to the new survey.

Sample sizes inevitably become smaller as there will be an incomplete overlap between being a respondent of a sample survey and being a business benefiting from innovation funding. Between 2008-16, almost 100,000 business reporting units responded to the ABS at least once. In individual years, there are about 25,000 respondents of which about 12,500 would have responded in the year before.

However, those respondents that are beneficiaries provide a great deal of detail about business performance. The ARDx contains data about sales, employment, purchases, remuneration to employees and capital investment. These responses also lead to an estimate of gross value added, preferred as a measure of output because it is less effected by variation in the procurement businesses make as they produce output.

For this study, the ARDx has also been combined with FAME balance sheet data about the plant and vehicles owned by UK businesses. The annex describes the approach taken to use this data, but essentially the initial balance sheet metrics of businesses as reported in their accounts is then combined with the ARDx investment data to provide capital stock measures. Several imputations are necessary to provide an annual series, but this allows some of the growth in value added to be attributed to firm-level investments in plant and vehicles.

Table 5.2: NMS and Innovate UK beneficiaries in the ARD, 2008-16

Measure		All	Beneficiaries	
			Innovate UK	NMS
Real Gross Value Added	£'000 mean	22,387	177,142	173,120
Real Turnover	£'000 mean	88,777	546,033	577,222
Employment		138	1,097	949
Real Purchases	£'000 mean	65,358	369,079	397,690
Capital stock plant	£'000 mean	4,423	60,269	82,991
Capital stock vehicles	£'000 mean	638	5,148	6,911
Industry shares				
Manufacturing		14%	35%	54%
High-Tech		3%	8%	10%
Low Paid		25%	13%	6%
Number of Reporting Units		28,016	1,362	1,204

Note: Real measures in 2010 prices; averages for reporting units so number of units will be higher than enterprise numbers for beneficiaries

Table 5.2 highlights that the gross value added amongst the reporting units of the enterprises that were ever supported by Innovate UK or NMS is much higher than that of all businesses in the ARD. They have a higher turnover and purchases in levels terms. However, in relation to employment, these output and input measures are not all that different, suggesting that the per employment values are similar. So, reporting units in the supported enterprises are larger, more likely to be in high technology or manufacturing industries and less likely to be in a low paid sector.

Using Productivity Decomposition to Explore Reallocation

In this section, productivity decomposition methods are applied to the ARD panel. The focus is the manufacturing businesses.

Tables 5.3 and 5.4 indicate decomposition results for Innovate UK and NMS respectively. Two productivity decompositions, Griliches and Regev (1995, hereafter GR) and Foster, Haltiwanger and Kirzan (2008, hereafter FHK) are used. Each decomposition seeks to disaggregate the total growth in productivity into the components attributable to different sources of productivity growth.

Both decompositions focus on three sources. The first is the productivity effects within a firm. These are internal improvements that raise the productivity of the businesses. For example, access to innovation funding will allow a business to invest in productivity

improving investments. The second part of the decomposition, made up of two terms in the FHK decomposition and one for GR, is often considered as “external” drivers and is the reallocation of resources to different businesses as a result of productivity differences. In Disney et al. (2003) and FHK, these are considered to be restructuring impacts. In presenting results, the between term is calculated by adding together the two measures of the FHK decomposition, while GR only provides a single measure of between effects.

The entry and exit terms are the final component. As noted in FHK, the entry/exit measures are generally quite sensitive to the economic cycle. Because the panel constructed for this study includes the effects as the UK emerged from recession. [NOTE: We need to redo this for a more stable period}

Tables 5.3 and 5.4 indicate the results of decomposing labour productivity, measured by real value added per employee with value added double deflated. The first results are for the manufacturing sector but identifying beneficiaries of Innovate UK; the second looks at NMS supported businesses. The period is 2009-16, focusing on the supported businesses and the rest of the manufacturing reporting units. The tables then unpack the productivity performance. In the top of each table is the decomposition of productivity indices where firm level productivity is weighted using employment with the lower panel weighting by value added.

The decompositions “add up”, so that the different disaggregations both total to the overall productivity growth when weighted by employment and value added (which will differ). The table does not have any evaluation analysis: it is merely comparing what is happening in the businesses that are the target for the support measure with the productivity performance of wider economy.

The productivity impacts of supported businesses form a significant portion of overall manufacturing productivity as the largest businesses are all in the supported category. By employment, NMS supported businesses are a quarter of employment and Innovate UK about a third. The value-added shares are a little higher.

In the period, there was a general modest fall in labour productivity: over eight years the fall was about 3%. The decompositions then consider whether – at a business level – the allocation of resources was towards more productive uses.

Some general trends can be observed. The first column – within – considers the productivity change in the businesses that continued throughout the period, about 40% of businesses. For Innovate UK, within productivity has grown in the period for the supported businesses whereas it has declined in the wider ARD businesses. Table 5.4 indicates that the NMS supported businesses are also outpacing the wider ARD businesses, but this is positive only for the indices where value added is used to weight the firm level productivity.

Table 5.3: Manufacturing Labour Productivity Decomposition for Innovate UK

Design	GR/ FHK	Within	B/w	Entry	Exit	Total
Employment weighted, 2009-16						
Unsupported	GR	-1.48%	0.12%	-2.32%	1.11%	-2.56%
IUK Supported	GR	-0.13%	0.30%	0.02%	-0.80%	-0.61%
Total		**-1.61%**	**0.42%**	**-2.30%**	**0.32%**	**-3.17%**
Unsupported	FHK	-0.80%	-0.69%	-2.92%	1.95%	-2.46%
IUK Supported	FHK	0.19%	-0.10%	-0.16%	-0.64%	-0.71%
Total		**-0.61%**	**-0.79%**	**-3.08%**	**1.31%**	**-3.17%**
Value added weighted, 2009-16						
Unsupported	GR	-1.20%	-0.28%	-1.02%	0.30%	-2.20%
IUK Supported	GR	0.13%	0.13%	-0.62%	-0.39%	-0.75%
Total		**-1.06%**	**-0.15%**	**-1.64%**	**-0.10%**	**-2.95%**
Unsupported	FHK	-0.96%	-0.59%	-1.50%	0.91%	-2.13%
IUK Supported	FHK	0.32%	-0.11%	-0.80%	-0.22%	-0.82%
Total		**-0.64%**	**-0.70%**	**-2.30%**	**0.69%**	**-2.95%**

The other columns focus on reallocation of resources. The between term answers the question about whether value added, or employment change, is correlated with productivity, so that firms growing in productivity tend to attract employment and value added. Generally, for the GR decomposition, this correlation occurs for the supported businesses while for the FHK decomposition – while the correlation is negative – the impacts are more modest than is the case for the wider ARD.

The two indices differ somewhat in the timing of when reallocation is measured. The GR index looks at this at the mid-point in the decomposition period by averaging share weights at the start and end of the period; FHK use share weights in the base year to then compare what happens at the end of the period. Given the length of time involved in the decomposition and the first year being soon after the recession, the differences in the decompositions is unsurprising. However, both indicate that broadly resources are from the middle of the period, moving in a manner that correlates with productivity and that this is towards supported businesses.

The entry-exit columns are most affected by the recession period. Much of the negative growth in productivity is attributable to the level of productivity of entrants being below the economy average. The period was one of considerable entry and so employment and value added that was associated with the entrants tended to depress productivity growth because the entrants were at the lower end of the labour productivity level.

For the supported businesses this effect is far less pronounced. Entrants that received support tend not to have productivity levels below the average. Earlier, it was noted that the supported businesses tended not to exit. The decomposition suggests that where exit

occurred, this was depressing productivity, with the exits being businesses with average or higher productivity.

Table 5.4: Manufacturing Labour Productivity Decomposition for NMS Support

Design	GR/ FHK	Within	B/w	Entry	Exit	Total
Employment weighted, 2009-16						
Unsupported	GR	-1.31%	0.07%	-2.38%	1.19%	-2.42%
NMS Supported	GR	-0.30%	0.35%	0.08%	-0.87%	-0.75%
Total		**-1.61%**	**0.42%**	**-2.30%**	**0.32%**	**-3.17%**
Unsupported	FHK	-0.80%	-0.57%	-2.97%	2.00%	-2.34%
NMS Supported	FHK	0.19%	-0.22%	-0.11%	-0.69%	-0.83%
Total		**-0.61%**	**-0.79%**	**-3.08%**	**1.31%**	**-3.17%**
Value added weighted, 2009-16						
Unsupported	GR	-0.96%	-0.23%	-1.16%	0.42%	-1.92%
NMS Supported	GR	-0.11%	0.08%	-0.48%	-0.52%	-1.02%
Total		**-1.06%**	**-0.15%**	**-1.64%**	**-0.10%**	**-2.95%**
Unsupported	FHK	-0.77%	-0.50%	-1.62%	1.01%	-1.88%
NMS Supported	FHK	0.12%	-0.20%	-0.68%	-0.31%	-1.07%
Total		**-0.64%**	**-0.70%**	**-2.30%**	**0.69%**	**-2.95%**

Both tables have used weights, so that they analyse the growth in productivity for the whole of manufacturing grossing up the individual firms in terms of their sample weights. While the comparison between supported and unsupported businesses is not within an evaluation framework (such as treated vs control), they do account for the performance of manufacturing in this period. One caveat to this analysis is the very large churn in businesses during the period.

Analysis of patent spillovers

Around 30,000 businesses in the UK have produced 260,000 registered patents from 1990 to 2014. Approximately 2,625 Innovate UK supported businesses produced 81,000 of these patents since 1990, with 17,800 patented at least one year after a business received its first instance of Innovate UK support.

Table 5.5 shows the split of patents registered by a UK business from 1990 to 2014. Industry divides patents with over 89,000 patents coming from the manufacturing industry, the knowledge intensive industries account for approximately 45,000 patents since 1990, with 34% of these coming from the highly knowledge intensive manufacturing industry.

Table 5.5: Patents registered by UK businesses

	All UK businesses		Innovate UK supported	
	Patents	Businesses	Patents	Businesses
All Industries	260,000	30,099	81,485	2,625
Highly Knowledge Intensive Industry				
Manufacturing	15,978	876	7,775	247
Services	30,744	4,558	12,547	624
Manufacturing Industry Not Highly KI	74,737	6,171	29,669	881
Other industries	144,022	18,494	36,043	873

While patents alone can be a sign of a business' innovation, they can also help identify knowledge spillovers by exploring whether patents are cited by other patents. Citations are used within every patent application to define the pre-existing knowledge upon which the invention was built. Patent citation can be direct (a patent is cited by another) or indirect (a patent is cited in patent that is then cited). Measures that incorporate this latter form of citation can be a more accurate measure and the patent ranking approach in Dechezleprêtre, Martin and Mohnen (2014, DMM) is used here.

Patent ranking is a proxy for the spillovers of an innovation and the ranking can also distinguish between citations from within and outside the UK. Citations from inside the UK would locate knowledge spillovers in the UK, potentially of greater policy importance, as it would lead to national economic benefits.

The DMM global patent index ranges from 0 to 49.99, with a mean of 0.19, and 48% of patents having a score of 0; i.e. these are innovations that are never cited. The Great Britain Patent index is smaller[2] and ranges from 0 to 5.90 with a mean of 0.02 and 59% of patents with a score of 0; i.e. those are innovations that are never cited (indirectly or directly) by an innovator in the UK.

An additional measure of value is used – number of countries in which the same invention was registered. This is referred to as the 'patent family' – or the family of patents registered across the globe for a single patent. The patent family ranges from 1 to 239 and indicates the number of countries in which a patent is registered, with an average of 3.5. This dataset focuses on patents from UK registered businesses alone.

Spillovers from patents in Innovate UK supported businesses

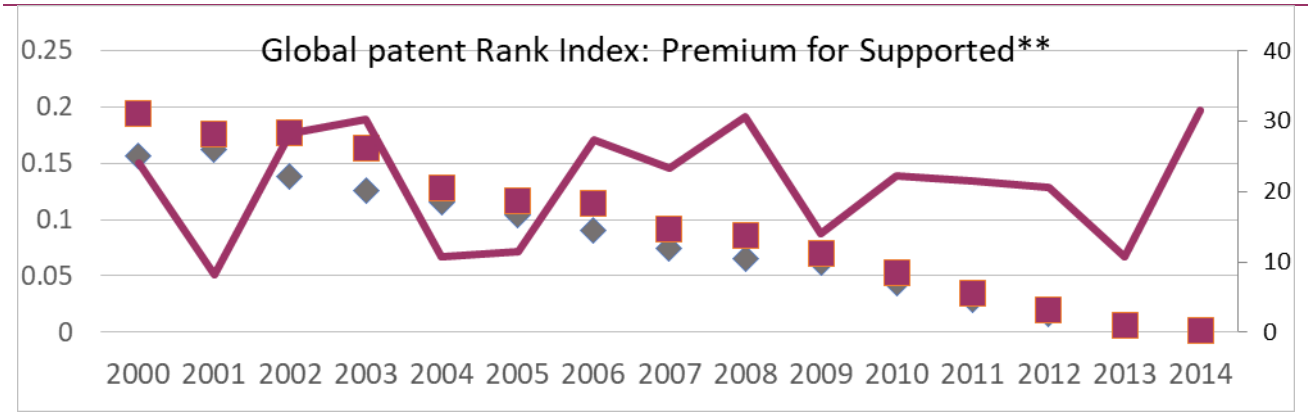
Innovate UK businesses register patents with higher than average knowledge spillovers. Looking at figure 5.4 below, the purple points show that using all three measures of knowledge spillovers, the supported businesses have patents with higher averages. This results to a premium in each measure, the percentage by which spillovers of the supported

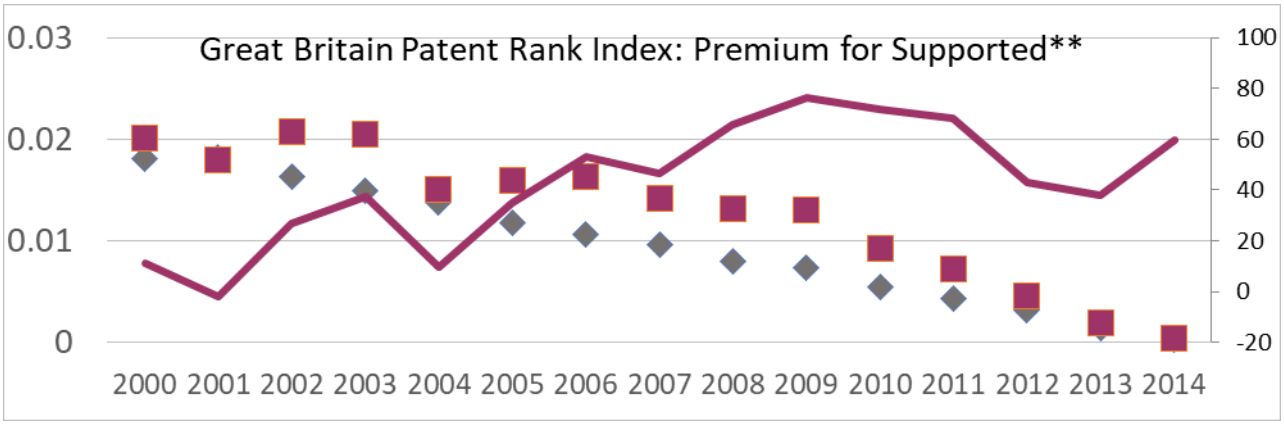
[2] This is by definition as it assigns a value of zero to any direct or indirect spillover that is to the benefit of an innovator outside of the UK.

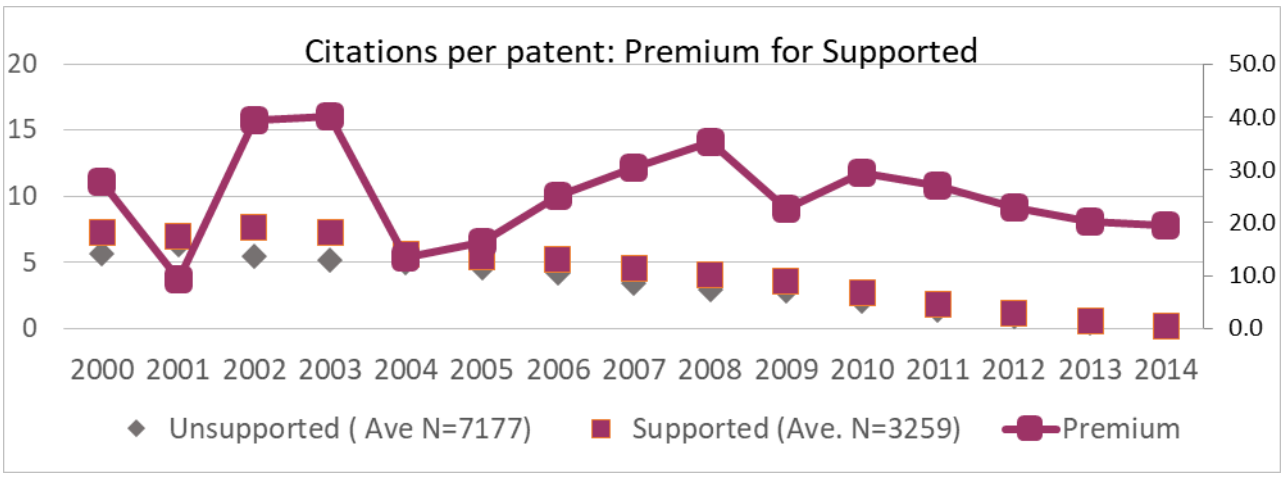
businesses' patents exceeds the unsupported, the purple line. The premium value uses the right-hand axis, a percentage premium measure, which – as the patent rank tends to decrease in later years – remains more stable as the ratio of the citation index for unsupported and supported is estimated. Whilst premiums are generally positive for businesses receiving any grants from Innovate UK (started in 2004), there is evidence that after the launch of Innovate UK, the businesses who received support registered patents with higher GB knowledge spillovers. This is seen by the increase in the average Great Britain Patent Rank in 2006 to 2009.

The differences between the treated and untreated are significant, with t-statistics above the threshold for 10% significance in all years for the global index rank except 2013. For the Great Britain Patent Rank, the difference between the supported and unsupported was insignificant before 2002, and then remains a significant difference beyond this. Citations similar to the Global Patent Rank are significantly different from 1990 onwards.

Figure 5.4: Patent citation index for Innovate UK supported businesses







The graph shows the mean Patent Rank scores of patents registered in the year by businesses supported by Innovate UK at some point between 2004-2017 and the mean score of patents registered by unsupported businesses.

Table 5.6 below splits the time periods for registered patents into pre-Innovate UK (1990-2006) and post Innovate UK (2007-2014), showing that after Innovate UK support, businesses are producing patents with higher Great Britain knowledge spillovers than before. The table shows that pre-Innovate UK there was a small difference in knowledge spillovers for patents from businesses who are unsupported, and those who will be supported in the future. After Innovate UK is launched, supported businesses started

producing patents with higher knowledge spillovers within Great Britain. All differences, whilst small, are significantly different as can be seen by the t-statistics. Of the 26,003 patents registered from 2007 to 2014 by businesses who receive support at some point by Innovate UK, 17,800 are after the receipt of support.

Table 5.6: Patent rank before/after launch of Innovate UK

Time period	Global Patent Rank	GB Patent Rank	Family size per patent	Citations per patent	Patent Count
2007-2014					
Unsupported	0.038	0.005	3.107	1.839	53,218
Supported	0.045	0.008	3.777	2.287	26,003
Difference	0.007***	0.003***	0.670***	0.448	
T-statistic (dif)	8.94	16.84	25.19	12.24	
1990-2006					
Unsupported	0.235	0.025	3.375	5.753	126,229
Supported (in the future)	0.281	0.027	4.151	7.098	55,482
Difference	0.046***	0.001***	0.776	1.345	
T-statistic (dif)	11.65	3.54	32.93	15.45	

This analysis presents overarching patents – rather than patents per business, but the results are similar for patents per business. Additionally, splitting the analysis by industry leads to a similar trend as above within the knowledge intensive and manufacturing industries.

Out of the 81,000 patents produced by Innovate UK businesses, 69,000 of these patents were produced by businesses who had a collaborative R&D project with Innovate UK and in fact of the 17,800 patents produced at least one year after an Innovate UK project, 17,500 are from Collaborative R&D projects.

Spillovers from patents of NMS supported businesses

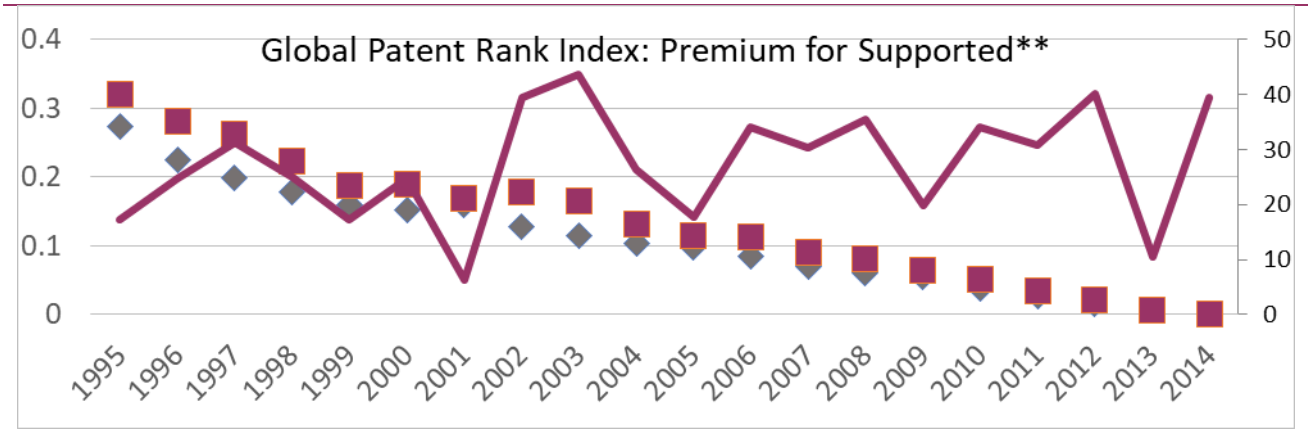
Patents from NMS supported businesses are highly concentrated in the manufacturing sector, with the majority (29,625) within the highly knowledge intensive manufacturing industry. Table 5.7 highlights that only 8,200 patents are from businesses outside the highly knowledge intensive and manufacturing industries.

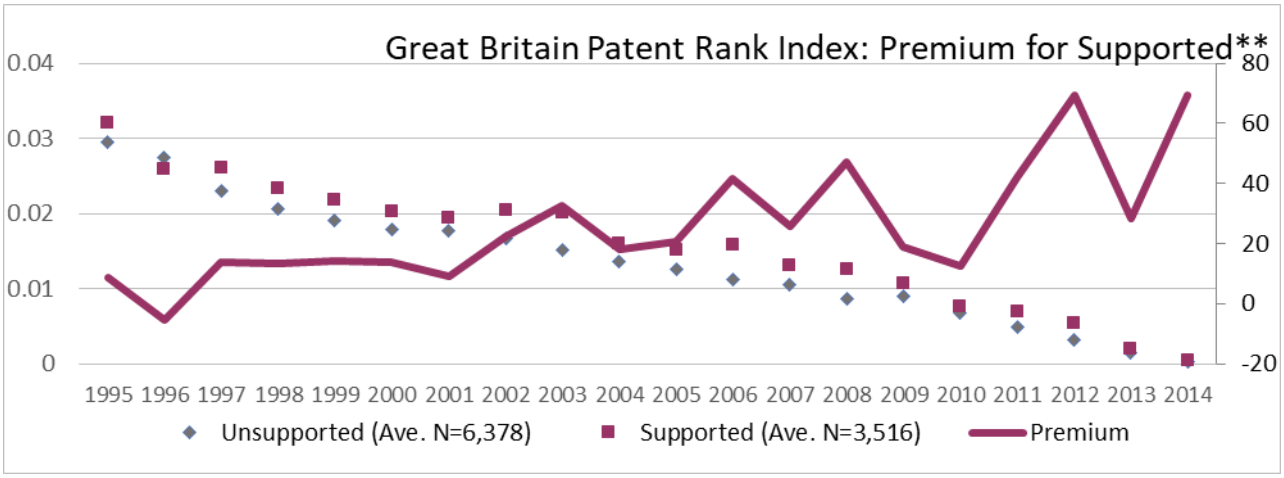
Table 5.7 Overall Industry split of patents from NMS supported businesses

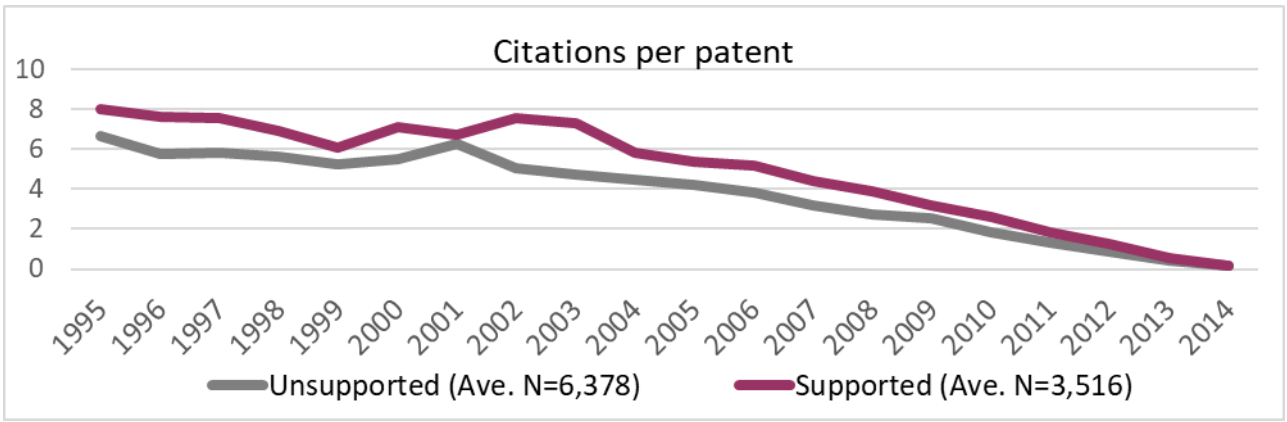
		Patents
All Industries		87,999
High Tech - Knowledge Intensive Industry		
	Manufacturing	29,625
	Services	21,625
Manufacturing Industry		
	Not highly KI	28,463
Other industries		8,286

The time trend within the NMS data strongly highlights the change around 2001 when NMS started supporting businesses. In general businesses supported by NMS generate patents with higher spillovers, with those spillovers increasing after 2001.

Figure 5.5: Patent citation index for NMS supported businesses



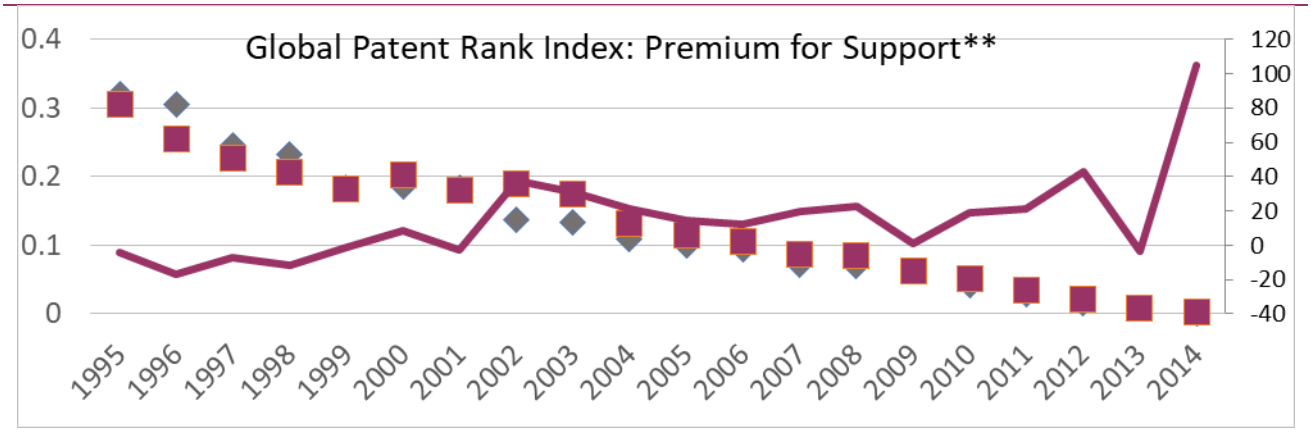


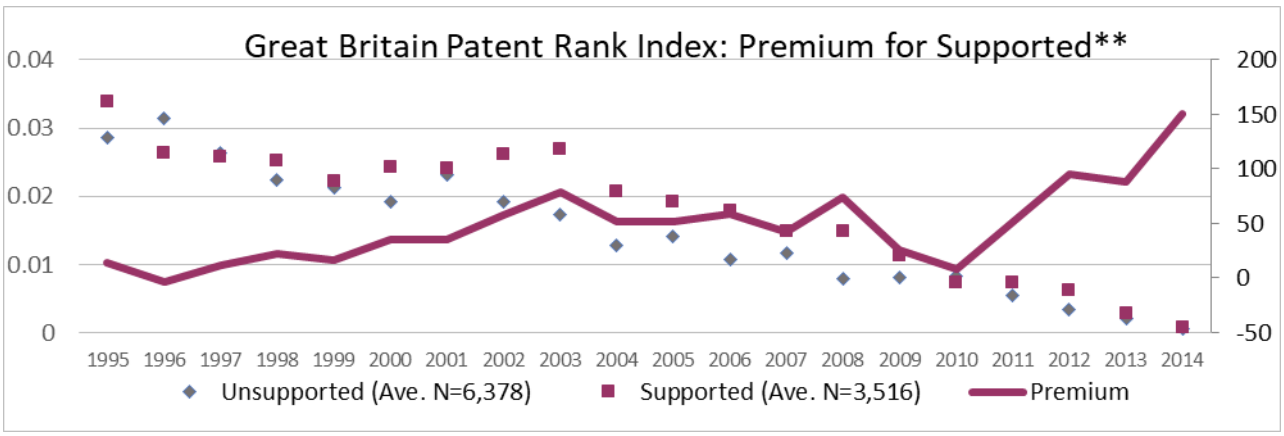


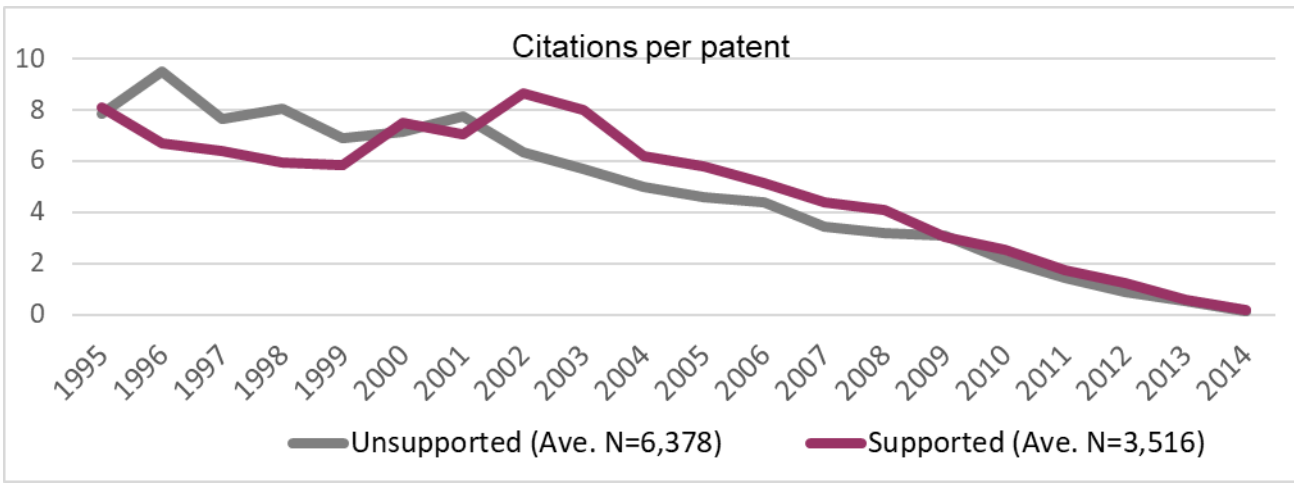
The graph shows the mean Patent Rank Index scores of patents registered in the year by businesses supported by Innovate UK at some point between 2004-2017 and the mean score of patents registered by unsupported businesses.

Since NMS is very industry-specific, the following figures breaks patents down for businesses within the highly knowledge intensive manufacturing industry. The change post 2001 appears more pronounced for this industry, perhaps highlighting the effect that NMS support had within the industry as a whole in producing more knowledge spillovers.

Figure 5.6: Patent citation index for NMS supported highly KI manufacturing businesses







The graph shows the mean Patent Rank scores of patents registered in the year by businesses supported by Innovate UK at some point between 2004-2017 and the mean score of patents registered by unsupported businesses.

Impact of support on earnings

The previous analysis showed that Innovate UK and NMS support has a positive impact on employment at supported businesses. Microdata on individual employees from the Annual Survey of Hours and Earnings (ASHE) allows analysis of the earnings effects of the support. ASHE is a 1% sample that is drawn based on the last two digits of an employees' National Insurance number, drawn from HMRC's PAYE system. Surveys are then sent to the employers of the selected employees who complete details about the wages, including overtime, bonuses etc, and the employee's hours and occupation. Earnings statistics are high quality as sample sizes are high and have a good coverage of employee jobs in Great Britain. Due to the panel nature of the data, it is also possible to track people over time, even as they change jobs or have a spell out of work.

There are several avenues in which to analyse the data. They all rely on comparing wages before and after support, and against the control group defined using PSM.

Firstly, average wages per firm can be computed each year. Even if individual workers join and leave a company, average wages can be calculated in a pseudo-panel, and changes in workforce characteristics can be accounted for. Wage growth in supported firms can then be compared against wage growth in matched unsupported firms. However, changes in the workforce may be an outcome of support itself. For example, a firm may hire more scientists or engineers to perform the R&D they have been awarded funding for. In that instance, the causal effect of the support may be less clear: does the support have an independent effect on productivity, or does it only come about because of the new hires? Without support, those workers may have never been hired in the first place, so this would just be another channel through which support may have an effect.

Another possibility is to compare only wages of those who stay with one firm during the whole period, before and after support. This makes it possible to measure the effect of innovation support on the productivity of individual workers. Worker fixed effects can be used to control for individual characteristics that are stable over time, and comparison against the matched control group controls for general wage growth, e.g. due to experience and seniority.

The next section focuses on individual-level effects from switchers, following D'Costa and Overman (2014). A treatment effect can be identified from job switchers. Those are people joining the treated firm around the time of the treatment. Wage growth of switchers to treated firms, higher than that of switchers to non-treated firms, would indicate that treatment has a positive effect on earnings, even controlling for the fact that some workers may have been specifically hired in response to the support. D'Costa and Overman do also use the second approach, but in the present analysis this was not undertaken.

Worker characteristics in the supported businesses

Using ASHE allows taking a closer look at the workforce characteristics of businesses supported by Innovate UK and NMS. Table 5.8 shows that employees at Innovate UK supported businesses earn slightly more than other employees on average, but they also work slightly longer hours. They are slightly younger than the average employee and predominantly male. Although they are younger, they have more work experience in general and in their current job and spent less time out of employment.

Surprisingly, there are fewer employees in high- and medium-skill occupations and more in low-skilled occupations in Innovate UK businesses than in ASHE. This might be due to a quirk in the UK labour market, in that an increasing proportion of low-skilled workers are self-employed or work small numbers of hours, so that they are not included on ASHE, for example self-employed contractors in cleaning or construction, or gig-economy workers, such as Uber drivers. In innovative businesses, even employees in elementary occupations may be vital to the business, so these jobs are kept in-house rather than outsourced to an agency.

Table 5.8: Summary statistics of employees of Innovate UK supported businesses

	Innovate UK supported	Voucher only	ASHE
Real gross weekly pay (logs)	6.0	6.1	6.0
Basic weekly hours	32.8	35.7	32.6
Total weekly hours	34.5	36.9	33.6
Age	39.9	40.3	40.8
Female	39%	29%	53%
Full-time	74%	84%	72%
Total years of work experience	12.1	10.9	11.2
Years out of employment	7.4	9.0	8.3
Length of current employment spell	6.9	4.6	5.4
More than one job	1%	2%	1%
High skilled (SOC 1-3)	28%	43%	36%
Medium skill (SOC 4-8)	48%	37%	44%
Low skill (SOC 9)	15%	13%	12%
Number of observations	315,000	1,156	1,662,787

The contrast between employees of NMS supported businesses and the general ASHE population is larger. Average earnings are higher than among the Innovate UK supported, and hours are longer. The average age is higher, and two-thirds of employees are male. Employees are also more skilled with 42% in high-skilled occupations, compared to 34% of employees on ASHE in general.

Table 5.9: Summary statistics of employees of NMS supported businesses

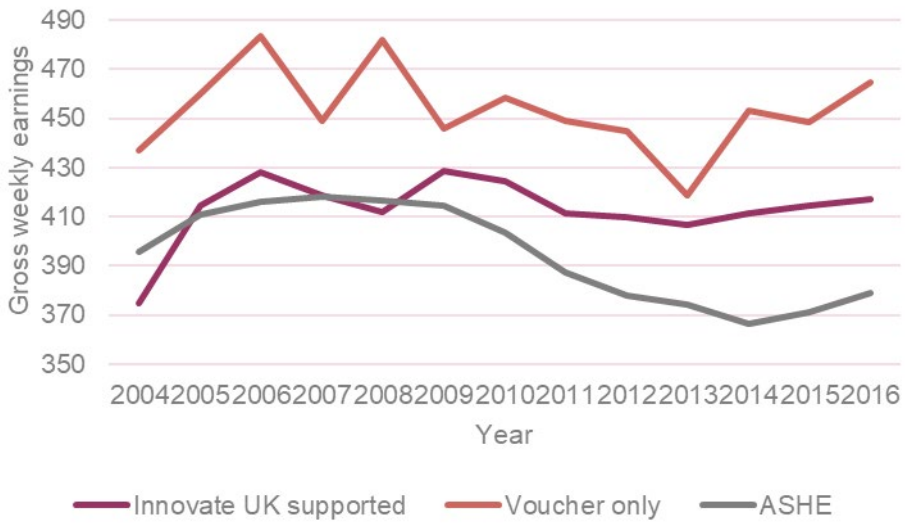
	NMS supported	Downloads only	ASHE
Real gross weekly pay (logs)	6.3	6.2	5.9
Basic weekly hours	35.0	34.2	31.7
Total weekly hours	36.6	35.7	32.8
Age	41.2	39.8	40.5
Female	33%	35%	53%
Full-time	86%	83%	69%
Total years of work experience	13.1	12.5	11.0
Years out of employment	7.8	7.1	8.2
Length of current employment spell	7.1	7.0	5.5
More than one job	0.6%	0.5%	1.4%
High skilled (SOC 1-3)	42.1%	34.3%	34.1%
Medium skill (SOC 4-8)	43.3%	48.4%	44.4%
Low skill (SOC 9)	6.1%	9.2%	13.4%
Number of observations (all years)	123,814	63,944	1,839,612

Earnings premium

To get a first impression of wages at businesses supported by Innovate UK and NMS, Figures 5.7 and 5.8 plot average earnings at businesses that were supported at some point in the period. Figure 5.7 shows that businesses supported by Innovate UK paid on average higher wages than other UK businesses. This is true in particular since the start of the financial crisis around 2008, where the average UK employee took a wage cut, but wages at Innovate UK supported businesses remained steady.

However, this cannot be read as an impact of Innovate UK support, as businesses would be included in this before receiving support. Rather, it shows that supported businesses are operate in activities with higher value-added and productivity, and a more skilled workforce than the average UK business. The figure also shows earnings by employees at businesses that only received a voucher. The earnings are much higher than those in businesses receiving other support as well as other employees in ASHE.

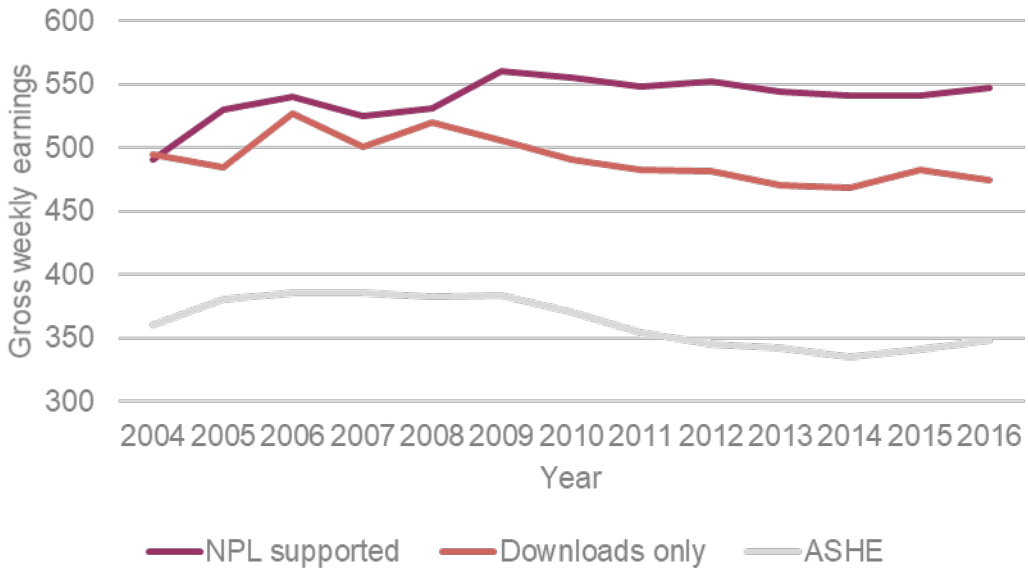
Figure 5.7: Average earnings in Innovate UK supported businesses



Note: Averages were calculated in logs and then exponentiated to reduce the impact of exceptionally high earnings.

This is more evident when looking at businesses supported by the NMS. The gap between average weekly wages paid by NMS supported businesses and the rest of the population on ASHE is almost £200. Again, this is not surprising given the highly educated, older, and predominately male workforce at NMS supported businesses.

Figure 5.8: Average earnings in NMS supported businesses



Note: Averages were calculated in logs and then exponentiated to reduce the impact of exceptionally high earnings.

The panel structure of the ASHE data can be used to better understand the effect of support on wages. The next figures look at the wage effect of job switching to or from a supported business. This is a popular technique when analysing ASHE, as it allows to control for individual characteristics, that are otherwise difficult to observe in data, for

example a degree in physics or an entrepreneurial ability (D'Costa and Overman, 2014; Gibbons et al., 2014).

Figure 5.9 compares earnings of job switchers to and from Innovate UK to earnings of employees moving between other businesses and those who do not switch jobs. Switchers to and from Innovate UK supported businesses are considered if the switch occurred in the year of first support by Innovate UK or any year thereafter. It also considers people missing for up to one year for the survey after a job move, as this can be related to the timing of the survey. Individuals missing for more than one consecutive year are not considered. The purple bars in Figure 5.9 refer to average earnings of individuals before a job switch, and the orange bars to earnings after the job switch, each deflated so that analysis is in real terms. For comparison, average earnings differences for individuals not switching jobs are also provided.

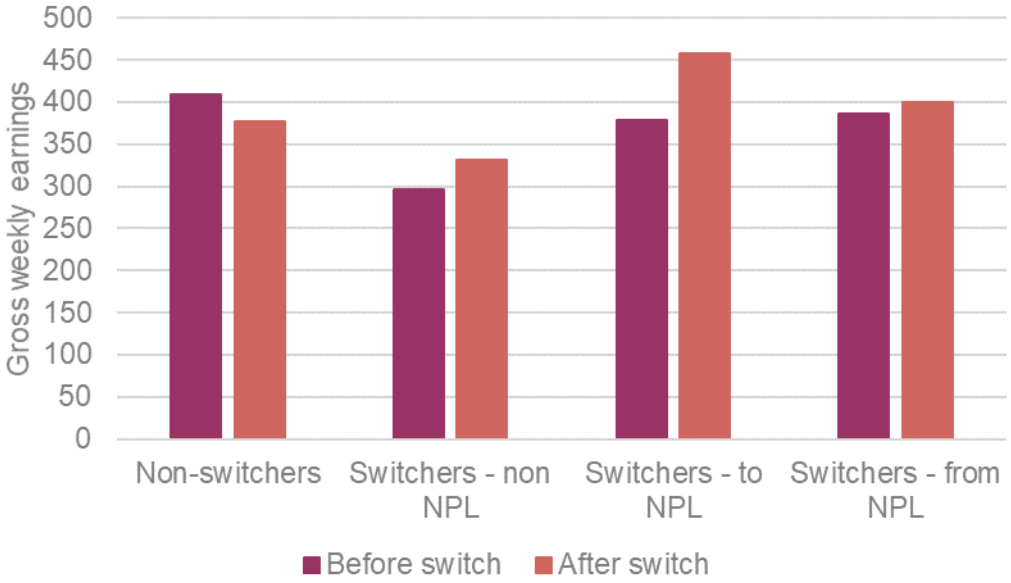
Figure 5.9 shows that job switchers generally earn less than those who stay put. This a usual result from tracking individuals over time in ASHE. However, changing jobs is related to an increase in earnings, which is larger for employees both joining and leaving businesses supported by Innovate UK. The significance of these differences is tested later, particularly to incorporate the possibility that the businesses are inherently different. However, Figure 5.9 highlights that switching into Innovate UK supported businesses provide similar wage effects to switching between non-supported businesses. The final set of individuals – switchers out of supported businesses – provides evidence that spells in these businesses produces a wage premium. Again, any causality is not determined as there will be selection issues in any analysis (e.g. supported businesses may be small, using non-wage remuneration, that then translates in wages in the business the employee moves too).

Figure 5.9: Earnings of Innovate UK job switchers



The effect is stronger for individuals switching to a business supported by the NMS. Employees gain more than £50 in weekly wages after switching to an NMS supported business. However, the wage gains are much smaller when leaving one of those businesses.

Figure 5.10: Earnings of NMS job switchers



Multivariate analysis

It is quite likely that employees at businesses supported by Innovate UK or NMS are quite different from the general population. Multivariate regression analysis can be used to control for the effect of observable characteristics, such as age and occupation on earnings. Including a variable for the effect of Innovate UK or NMS support in these regressions can identify the effect of the support.

Table 5.10 reports the impacts measured from different regression specifications and for different treatments. The table only includes the treatment coefficient, full results can be found in the annex E. There is one category of support for NMS, and multiple different ones for Innovate UK. The first two specifications use Ordinary Least Squares (OLS) estimation. In the first of these two specifications, an individual's skills are accounted for by two high-skilled and medium-skilled dummies (with low-skilled as the baseline), derived from occupation codes. In the second specification, these are replaced by more detailed O*NET variables[3].These are only available from 2010 onwards, so that a shorter time series is available for estimations using the O*NET data. The second set of regressions uses fixed effects (FE) estimation. This exploits the panel structure of earnings, in that it

[3] O*NET is a coding scheme that rates occupations on various physical and cognitive skills as well as the use and importance of different knowledge domains. There are 70 variables, measuring both the level and importance of skills and knowledge. These were reduced into 26 variables using factor analysis. It is based on US survey data, but can be mapped to UK Standard Occupation Classification (SOC) codes.

includes a fixed effect for each individual, and so controlling for unobservable personal characteristics.

Table 5.10 shows treatment effects and their significance. In the first two regressions, almost all the coefficients are significant. However, some of the Innovate UK treatments are estimated to have a small but negative effect. The effects are so small that they are zero for multiple digits behind the comma, and only have statistical significance due to the large sample size. As the dependent variable is logged earnings, the coefficients can be interpreted as percentage changes. The estimates show that support by NMS increase wages by 5 to 7 per cent and Innovate UK support increases wages by around 5%.

The second set of regressions, including individual fixed effects, shows slightly different results. The coefficients are generally smaller and less significant. This suggests that some positive selection is going on: individuals with higher unobserved ability are more likely to work at Innovate UK or NMS supported businesses. This finding is expected, as innovative businesses and their employees are likely to be different from other businesses, such as ambition and entrepreneurialism.

Table 5.10: Treatment effects from multivariate regression

	Baseline	O*NET	Baseline FE	O*NET FE
NMS support	5%***	7%***	3%***	4%***
Innovate UK				
Any grant started in that year	1%***	2%***	0%	0%**
Lead of a collaborative R&D project	6%***	6%***	1%***	1%***
Collaborator of a R&D project	1%***	3%***	1%***	1%***
Only one project started in this year	0%***	0%***	0%**	0%***
Multiple projects started in this year	4%***	5%***	1%***	1%***
Logic Model 1: Collaborative R&D	2%***	3%***	1%***	1%***
Logic Model 2: Single Firm grant	0%***	0%***	0%	0%***
Grant less than 10K	0%***	0%***	0%**	1%**
Grant between 10k and 200k	2%***	1%	0%*	0%
Grant over 200K	8%***	7%***	1%***	0%
12 months or shorter project	4%***	3%***	0%	0%
12-36 months project	0%***	0%***	0%	0%
36 months project	1%***	3%***	0%	1%
Logic Model 3: KTP	4%***	4%***	1%**	1%**

Note: Dependent variable: Earnings one year after treatment. ***Coefficients statistical significant at 1% level, ** 5% level, and 10%.

Unlike the PSM results presented before, no explicit steps have been taken to ensure the causality of these results. Earnings at supported businesses might have grown faster also in the absence of support. Unlike results at the businesses level, there is less of a concern that this is driven by unobserved business characteristics, such as managerial ability or ambition, as individual employees are unlikely to have influence on applications to Innovate UK or participation in NMS support, particularly in larger organisations. While it is

possible that innovating businesses attract more productive employees, this should not affect the regression results including fixed effects, as unobservable individual characteristics are controlled for.

The different results by type of support broadly align with expectations. Significant and high support effects are seen where the business is a lead in collaborative R&D or the business is tagged as having multiple projects. The effect increases with project size also, with grants over 200k have employees with a 1% premium on wages taking account of individual specific effects.

Concluding comments

This chapter has sought to understand the impacts of innovation support on productivity. Looking at real turnover growth, the businesses that are supported have generally performed better than comparison groups. In supported businesses, real turnover has risen by 35-60% by 2015, with those treated first in 2011 achieving the highest growth. In the matched control groups, growth is 15-38%.

However, the employment growth seen in the supported business is high in comparison to the control group. The real turnover per employee growth differential between treated and the counterfactual is negative for the 2010, 2011 and 2012 cohorts being as much as 21 points in the first of these. Only in 2013 do the supported businesses have real turnover per employee growth that is higher than that observed in the matched control group.

Innovate UK businesses register patents with higher average knowledge spillovers. For all three measures of knowledge spillovers, the supported businesses have patents with higher averages. Whilst this was the case before businesses received any grants from Innovate UK (started in 2004), there is evidence that after the launch of Innovate UK, the businesses who received support registered patents with higher GB knowledge spillovers. In general businesses supported by NMS generate patents with higher spillovers, with those spillovers increasing after 2001.

Changing jobs is related to an increase in earnings, which is larger for employees both joining and leaving businesses supported by Innovate UK. This suggests that individuals both gain a one-off premium from working to those businesses, as well as gaining skills and experience that help them in their next job.

The effect is even stronger for individuals switching to a business supported by the NMS. Employees gain more than £50 in weekly wages after switching to an NMS supported business. However, the wage gains are much smaller when leaving one of those businesses, suggesting that the learning effects are smaller, or knowledge and skills are less transferable to other businesses.

6 Conclusions

The two innovation support bodies – Innovate UK and National Measurement – have a long history of supporting businesses as they undertake R&D. Looking at how the support impacts businesses, in terms of the logic of the interventions, highlights some further considerations for this analysis. Firstly, there is a significant level of funding being provided – especially by Innovate UK – directly to businesses. The amounts suggest that impacts should be measurable and that the direct impacts on businesses should provide a starting point for the overall impact analysis. As R&D impacts generally occur with lags, as it takes time to turn ideas into commercially successful products and services, this impact analysis can capitalise on the long historical dataset.

The intervention logic also highlights wider impacts. Innovation funding anticipates benefits beyond the directly supported business, encouraging collaboration and spillovers beyond the supported businesses.

The approaches outlined in this report are diverse. They focus on an analytical stage to compile a complex dataset of business characteristics and performance linked to the beneficiary lists of the two organisations. There are three analytical stages:

- A first focuses on modelling the selection into innovation support, seeking to tease out what makes the businesses that are supported different from the wider population and then identifying a comparison group.
- The second stage then analyses the supported and comparison businesses. Given the nature of the support, a first analysis focuses on what increased economic activity is seen in the supported businesses. This is in terms of increased employment or the avoidance of business closure. It also covers R&D activities. The analysis also considers how activity varies by the level of support.
- The final stage of the analysis considers the productivity impacts of the support. Productivity establishes whether any increased economic activity involves better use of resources and is a difficult area to determine. Approaches can be direct, looking at output per unit of input but this requires a complete picture of business resource use. Productivity proxies, such as the earnings of employees and patent citations, are used in this study.

This study has used a variety of methods, which individually provide some of the picture of the impacts of innovation support. Individual approaches have strengths and weaknesses, but the research has provided a chance to put together a complex evidence base and look across this. This firstly tackles the difficulty of identifying comparable businesses that did not receive innovation support. Using the comparison group, an increase in economic activity due to the support has been identified. This has been tested in terms of timing,

scale and avoiding the double counting that can be problematic given multiple incidences of support for the same businesses.

The study has then explored whether there is evidence that the increase in economic activity has been productive. Using the wages earnt by the employees of supported businesses and establishing a wage premium is indicative of high-quality jobs in the supported businesses. New analyses have included looking at spillovers and this indicates that the patents of the supported businesses are of high quality, being more likely to be cited.

References

Almus, M. and D. Czarnitzki (2003) “The Effects of Public R&D Subsidies on Firm’s Innovation Activities” Journal of Business and Economic Statistics, 21(2).

Angrist, Joshua and Jorn-Steffen Pischke (2009) Mostly Harmless Econometrics: An Empiricist’s Companion.

BEIS (2017). The impact of public support for innovation on firm outcomes. BEIS Research Paper Number 3.

Bia, M. and Mattei, A. (2008). A Stata package for the estimation of the dose–response function through adjustment for the generalized propensity score. Stata Journal 8: 354–373.

BIS (2014) ‘An Economic Analysis of Spillovers from Programmes of Technological Innovation Support’. Report to BIS by ICF GHK, March.

BIS (2016) “Exploring the feasibility of a productivity based approach for evaluating business support interventions” BIS Research paper 295.

Box, G. E. P., and D. R. Cox. 1964. An analysis of transformations. Journal of the Royal Statistical Society, Series B 26: 211–252

Busom, I. (2000) “An Empirical Evaluation of the Effects of R&D Subsidies” Economics of Innovation and New Technology, 9(2).

Caballero, R. and Jaffe, A. (1993). How high are the giants’ shoulders: An empirical assessment of knowledge spillovers and creative destruction in a model of economic growth. In NBER Macroeconomics Annual 1993, Volume 8, pages 15–86. MIT press.

Caliendo, M. and S. Kopeinig (2005) “Some Practical Guidance for the Implementation of Propensity Score Matching” IZA Discussion Paper 1588, May.

Cerulli, G., 2010. Modelling and measuring the effect of public subsidies on business R&D: A critical review of the econometric literature. Economic Record, 86(274), pp.421-449.

Cerulli, Giovanni and Bianca Potì (2014). Public support intensity and company R&D performance: Evidence from a dose-response model. 54th Congress of the European Regional Science Association: "Regional development and globalisation: Best practices", 26-29 August 2014, St. Petersburg, Russia, European Regional Science Association (ERSA), Louvain-la-Neuve.

Czarnitzki, Dirk and Andreas Fier (2002) “Do Innovation Subsidies Crowd Out Private Investment? Evidence from the German Service Sector”, ZEW Discussion Papers, No 02-04

Czarnitzki, D. and Licht, G. (2006). Additionality of public R&D grants in a transition economy: the case of Eastern Germany. Economics of Transition, 14(1), pp.101-131.

Czarnitzki, Dirk, Bernd Ebersberger and Andreas Fier (2007) “The relationship between R&D collaboration, subsidies and R&D performance: Empirical evidence from Finland and Germany”, Journal of Applied Econometrics, December.

D’Costa, S. and H.G Overman. (2014). The urban wage growth premium: Sorting or learning? Regional Science and Urban Economics, 48, 168-179.

Dearden, L., H. Reed and John van Reenen (2005). The impact of training on productivity and wages: Evidence from British panel data. CEP Discussion Paper no. 674.

Dechezleprêtre, A., Martin, R. and Mohnen, M. (2014). Knowledge spillovers from clean and dirty technologies: A patent citation analysis. Centre for Economic Performance Discussion Paper, No 1300, September.

DiPrete, T.A., and M. Gangl (2004). “Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments.” Sociological Methodology, vol. 34, issue 1 pp. 217-310.

Disney, R., J. Haskel and Y. Heden (2003) "Restructuring and Productivity Growth in UK Manufacturing" The Economic Journal, 113(489): 666-694.

Du J. and Bonner K. (2016), Decomposing UK aggregate labour productivity and growth: 1998-2013 using the ONS business structure database data, ERC Research Paper No. 48.

Essama-Nssah, B. (2006). Propensity score matching and policy impact analysis: A demonstration in Eviews. World Bank Policy Research Working Paper, no. 3877.

Foster, L., J. Haltiwanger and C.J. Krizan (2001) "Aggregate Productivity Growth: Lessons from Microeconomic Evidence". In Hulten, C., Dean, E.R., and Harper, M.J. (eds) *New Developments in Productivity Analysis*, University of Chicago Press.

Fryges, Helmut, Alexander Vogel and Joachim Wagner (2015) “The impact of R&D activities on exports of German business services enterprises: First evidence from a continuous treatment approach”, The World Economy.

Gibbons, S., H.G. Overman and P. Pelkonen (2014). Area disparities in Britain: Understanding the contribution of people vs. place through variance decomposition. Oxford Bulletin of Economics and Statistics, 76, 5, 745-763.

Girma, Sourafel and Holger Gorg (2005) “Foreign Direct Investment, Spillover and Absorptive Capacity: Evidence from Quantile Regressions”. Kiel Institute for World Economics Working Paper, June.

Goldman Sachs (2018) “10,000 Small Businesses: Building Small Business Britatin”, https://www.goldmansachs.com/citizenship/10000-small-businesses/UK/news-and-events/building-business-britain-f/report.pdf

Grant Thornton, Ecorys, Loughborough University, Oxford Economics, Future Inclusion (2013) “Report 5: Post-Games Evaluation Meta-Evaluation of the Impacts and Legacy of the London 2012 Olympic Games and Paralympic Games”, Report for Department for Culture, Media & Sport Prepared.

Griliches, Z. and H. Revev (1995). "Firm Productivity in Israeli Industry: 1979-1988". Journal of Econometrics, 65(1): 203.

Guardabasico, B. and M Ventura (2014). “Estimating the dose–response function through a generalized linear model approach,” The Stata Journal

Hall, B. and D. Harhoff (2012). Resent Research on the Economics of Patents. NBER Working Paper Series 17773: 1-53.

Hall, B.H., C. Helmers, G. von Graevenitz and C. Rosazza-Bondibene, (2013). A study of patent thickets. Research commissioned by the Intellectual Property Office.

Hellerstein, J.K., D. Neukmark, and K.R. Troske, (1999). Wages, productivity, and worker characteristics: Evidence from plant-level production functions and wage equations. Journal of Labour Economics, 17, 3, 409-446.

HMT (2011). The Magenta Book: Guidance for Evaluation. London, UK: HM Treasury.

HMT (2019). The Green Book: Appraisal and Evaluation in Central Government. London, UK: HM Treasury.

Hud, Martin and Katrin Hussinger (2014). "The impact of R&D subsidies during the crisis," ZEW Discussion Papers 14-024, ZEW - Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research.

Hussinger, K. (2008). R&D and Subsidies at the Firm Level: An Application of Parametric and Semi-Parametric Two-Step Selection Models, Journal of Applied Econometrics 23, 729-747.

Innovate UK (2018a). “Evaluation Framework: How to assess impact on business and the economy”, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/681741/17.3253_Innovate_UK_Evaluation_Framework_RatherNiceDesign_V2_FINAL_WEB.pdf

Innovate UK (2018b). Innovate UK Funded projects between 2004 and 1 February 2018. (https://www.gov.uk/government/publications/innovate-uk-funded-projects).

Intellectual Property Office (2017). Open Data: Domestic UK Applications. (https://www.gov.uk/government/publications/ipo-trade-mark-data-release).

Kaplan, E. L. and P. Meier. (1958). "Nonparametric estimation from incomplete observations". J. Amer. Statist. Assoc. 53 (282): 457–481.

Lechner, M. (2001). Identification and estimation of causal effects of multiple treatments under the conditional independence assumption. In Econometric Evaluation of

Labour Market Policies, ed. Lechner, M. and F. Pfeiffer, pp. 43-58. Heidelberg: Physica-Verlag.

Imbens, G. W. (2000) The role of the propensity score in estimating dose-response functions. Biometrika, 87, 706–710.

Imbens, G. W. (2004) Nonparametric estimation of average treatment effects under exogeneity: a review. Rev.Econ. Statist., 86, 4–29.

Jaffe, A. B., M. Trajtenberg, and R. Henderson (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. Quarterly Journal of Economics, 108(3):577–598.

Jaffe, A. B. and M. Trajtenberg (1996). Flows of knowledge from universities and federal labs: Modeling the flow of patent citations over time and across institutional and geographic boundary. NBER Working Papers 5712, National Bureau of Economic Research.

Jaffe, A. B. and M. Trajtenberg (1999). International knowledge flows: Evidence from patent citations. Economics of Innovation and New Technology, 8(1-2):105–136.

Ker, Daniel and Matthew Greenaway (2012) "Coverage of the Business Enterprise Research and Development Survey", Online note. http://www.ons.gov.uk/ons/guide-method/method-quality/specific/business-and-energy/research---development-surveys/information-note--coverage-of-the-business-enterprise-r-d-survey.pdf

Lechner, M. (2001) Identification and estimation of causal effects of multiple treatments under the conditional independence assumption. In Econometric Evaluation of Labour Market Policies (eds M. Lechner and F. Pfeiffer). Heidelberg: Physica.

Marino, M., P. Parrotta and D. Sala (2011). New Perspectives on the Evaluation of Public R&D Funding, Working paper 11-02, Department of Economics, Aarhus School of Business, Aarhus University.

Maslov, S., and S. Redner (2008). "Promise and Pitfalls of Extending Google's PageRank Algorithm to Citation Networks." Journal of Neuroscience 28.44 11103–11105.

Moffat, J. (2013). Regional selective assistance in Scotland: Does it make a difference to plant survival? Regional Studies, no. 49, vol. 4, pp. 568-581.

Morgan, S., & Winship, C. (2014). Counterfactuals and Causal Inference: Methods and Principles for Social Research (Analytical Methods for Social Research). Cambridge: Cambridge University Press.

Munari, F. (2017) Review of Literature on the Use and Impact of IPRs at the Firm Level: Patents, Trademarks and Designs. Report to the UU Intellectual Property Office.

Peel, M.J., and G.H. Makepeace (2009poste). Propensity score matching in accounting research and Rosenbaum bounds analysis for confounding variables. Available at SSRN: http://ssrn.com/abstract=1485734.

Postel-Vinay, F., and H. Turon (2010). On-the-job search, productivity shocks, and the individual earnings process. International Economic Review, 51, 3, 599-629.

Ravallion, M. (2008). Evaluating anti-poverty programmes. Chapter in the Handbook for Development Economics. Amsterdam, NL: Elsevier.

Riegler, R. (2012). Fragmentation and integration: New evidence on the organisational structure of UK firms. PhD Dissertation, University of Nottingham.

Ritchie, F., D. Whittard, and C. Dawson. (2014). Understanding official data sources. Final Report for the Low Pay Commission.

Rosenbaum, P. and D. Rubin (1983). The central role of the propensity score in observational studies for causal effects. Biometrika, vol. 70, pp. 41–55.

Sarkisyan, A., B. Casu, A. Clare, and S. Thomas (2009). Securitization and bank performance. Mimeo.

Smith, J., and P. Todd (2005). Does matching overcome Lalonde's critique of non-experimental estimators? Journal of Econometrics, vol. 125, pp. 305-353.

Vanino, E., Roper, S. and Becker, B. (2019). "Knowledge to money: Assessing the business performance effects of publicly-funded R&D grants." Research Policy 48(7): 1714-1737.

Wallsten, S. (2000) "The Effects of Government-industry R&D programs on private R&D: The case of the Small Business Innovation Research Program", RAND Journal of Economics, 31(1):82-100

Zúniga-Vicente, J. C. Alonso-Borrego and F. Forcadell (2014) "Assessing the effect of public subsidies on firm R&D: A survey", Journal of Economic Surveys

Annexes

A: Business Structure Database

There are advantages and disadvantages to using the BSD as the data spine for this study. Ritchie et al. (2014) look at several available employment datasets. A main advantage of the BSD is its completeness. On employment, the IDBR provides a near-census as it incorporates all businesses operating a PAYE system.

The BSD therefore provides a time-series of employment data for employers. Alternative data sources, such as those based on company accounts, do not have employment for the smallest businesses so would not provide evidence on a significant portion of the Catapult beneficiaries.

As well as its comprehensiveness, BSD has been selected because of its potential for linking. The BSD identifies each business through a referencing system to which other surveys and administrative data can be linked.

A further advantage of the BSD is its timeliness as it is an extract from a live register that is routinely updated. The variables extracted from the IDBR into BSD include employment and this comes from a variety of sources. HMRC provide employment estimates for all businesses based on the number of individuals on payrolls as income tax is collected. Further, the annual Business Register and Employment Survey (BRES) collects data on employment and type of activity at the local unit level (individual plant). Some of the ONS monthly surveys that collect employment data are also used to supplement the IDBR for firms that do not appear in BRES. Finally, if there is no PAYE record or survey data, employment is imputed from VAT data, provided by HMRC, averaged over a year to the middle of each year.

There are some issues with using BSD. A first set focus on the quality of the data when used to track businesses. IDBR is primarily a register for surveying and so there is a premium attached to the quality of the data about a business in a particular year, rather than over time. So, to meet the needs of a survey sampling frame, the IDBR may make changes to support surveying, but inadvertently affecting the time-series. Recent ONS (Lai, 2019) analysis highlights that the replacement of the Annual Business Inquiry with the Annual Business Survey in 2009 may have affected the panel series. A complication was that the change occurred in the Great Recession, when many businesses closed, so that the changes observed were complicated by the mixing of survey and economic changes. This is on-going research as ONS invests in improving the BSD.

Outside of IDBR changes, there are some issues resulting from the use of administrative data in producing a business register. Table A1 indicates some of these. These fall under three broad headings – timeliness of the different data sources, period of observation and the different sources of local level data. Timeliness issues do not affect this study materially. For this study, with its focus on small businesses, the local unit issue is also of less significance because the research is not looking at regional or local impacts.

The period of observation issue has meant steps have been taken to ensure observations before and after support can be identified. The employment numbers in the 2017 BSD dataset, published in December 2017, will be taken from the IDBR in March 2017. The IDBR estimates are based on actual employment numbers at a variety of points in time, but the most common would be September 2016. This means that, where the Innovation Funding was provided post September 2016, it will not be possible to see employment or turnover impacts yet.

Table A1: Issues with using the Business Structure Database

	Timeliness of data	Period of observation	Different sources for local unit
Background	Register continually updated to be the most accurate record of employment at any point in time. However, this reflects a number of different influences and definitions.	BRES is carried out with reference to a date in September; ONS 'freezes' the IDBR in March.	BRES surveys 80,000 stratified businesses so that large companies (over 500 employees) are almost always likely to be included; for businesses that are not surveyed, local unit data is imputed using the returned information in BRES.
Issue	PAYE records are provided quarterly and reflect the total employment over the last four quarters, with unknown churn effects. BRES data comes from a survey and is assumed more reliable and so "held on to" despite more up-to-date PAYE data.	BSD 2014 includes BRES data from Sept 2013; Monthly survey data up to early 2014; PAYE data covering the year up to 2014. The employment data will have a different timing depending on where it comes from.	For the largest companies, enterprise and local unit information is from BRES and specific about local units, but problems in complex organisations; for smaller companies, employment comes from PAYE or imputed from VAT; in some cases (new companies, or those without a PAYE scheme) even enterprise data may be estimated.

There will be a pipeline effect for businesses receiving support at different times of the year. For example, those receiving support during the first half of financial year 2013/14 may have a measured employment change in BSD 2014 because they would have received support before September 2013. In the analysis, the year of the employment estimate is referred to rather than the BSD dataset from which the data is derived (i.e. employment estimate 2014 uses BSD 2015 data).

The turnover data largely comes from HMRC VAT feeds as well surveys and imputations from PAYE data. Though the source for the data differs, the lag on turnover is like that in employment, up to approximately one year with estimates of the average turnover in the year to September entering the BSD towards the end of the following year. Further, to disentangle the effect of the policy from inflationary pressures on businesses' turnover (Du

and Bonner, 2016), the turnover data was deflated using deflators based on two-digit standard industrial calculation 2007 official GDP estimates.

Annex B: Matching datasets

As a first step in building a complete dataset of business level outcomes and characteristics, businesses within all the public and support databases are linked to Companies House numbers (CRNs), which can then be linked to ONS enterprise reference numbers (entrefs) within the government Secure Research Service (SRS).

Matching to Companies House numbers

Some databases already contain CRNs, while others only provide company names. Some work was previously undertaken by Innovate UK to identify CRNs for supported businesses. However, not all businesses were linked, and data linking was used to supply CRNs for more businesses. This process comprised six broad steps. These were applied and improved iteratively. Special care has been taken to make sure those businesses receiving the most grants or registering the most patents and trademarks are matched correctly.

Identifying the Companies House number has several stages because – given the relatively long time-series to the data – the linking has to tackle a few challenges:

- The Innovate UK and NMS beneficiaries were linked to the register in around 2015, and the businesses supported some years before the year of linking may have closed and so left the register; a few older vintages of the Companies House register were therefore used to fill gaps
- For businesses that received support but registered after 2015, it was possible to link to CRNs using more recent registers.

The legal entity of a company may be different to a business unit that is surveyed by the ONS. While most companies are simple, single entities, larger businesses often consist of many, even hundreds of entities, organised in a group structure, that are treated as one or a small number of reporting units by the ONS for survey purposes. To increase the chances of matching to survey responses in the ONS Secure Research Service (SRS), supported businesses were also matched to their domestic ultimate owner (DUO) as provided on Bureau van Dijk's FAME database.

Matching rates to Companies House – Innovate UK

Innovate UK data contains a company identifier for the majority of organisations within the database – the company registration number in UK Companies House. As a project participant database, every participant is listed for each project. Additionally, an 'Enterprise size' variable identifies the organisations as companies of different size or non-companies (academic institutions, public bodies, etc).

Among the organisations without company registration numbers, 471 were identified by the 'Enterprise size' variable as companies. These were matched by name, and an additional 128 CRN numbers were added to the database.

The Knowledge Transfer Partnerships (KTP) list the university not the company as the organisation, and do not have the company registration number in the database. These names had to be extracted from the title of the project and matched by name and postcode to the Companies House register to find CRNs for the businesses partnering with universities in KTPs.

Missing values

There are 3,100 lines of projects in the Innovate UK database with no company number. Of these 2,199 are academic organisations according to Innovate UK's database descriptor 'Enterprise size'. The rest of the projects with no company number are split between 252 projects by public focused or non-UK organisations (Catapult, public service organisations, public service research establishments, and research and technology organisations, charities), 433 projects by private enterprises (large, medium, small, micro, sole trader, start-up) and 215 projects by 'unknown' enterprises.

To investigate the enterprise level data for econometric purposes, the academic, public focused or non-UK organisations were dropped. These organisations, even if they can be matched, have complex accounting structures and reporting requirements. This leaves missing identifiers for 433 projects by private enterprises and the 215 projects by 'unknown' enterprises. Of these 649 projects, word tagging identified an additional 45 'university' projects, 7 'Catapult' projects, 78 public service organisation projects and 3 public service research establishment projects, leaving 504 projects with possible enterprises without Companies House numbers, with 471 unique companies. Among the 471 unique unidentified companies, 439 companies amount to £8.65m of funding, with the other 32 companies reporting £0 spend to date. These 471 unidentified companies were input into FAME for matching, and an additional 128 CRN numbers were identified.

To understand why the Companies House number is missing, a random sample was tested. For a random sample of 20, Companies House numbers were found via web research and searching the Companies House register online for ten. Four of these companies had different names in the database (similar or changed) and two additional had different addresses. One appears to be incorporated after the grant, and another is an Isle of White company, with the final company being a university enterprise. Of those where research could not find a CRN, one was a confirmed sole trader, another a university research programme, two were outside the UK, and the others could not be found for unknown reasons.

Matching rates to Companies House numbers – Trademarks and Patents

Over 302,000 trademark and 30,000 patent owners were identified as companies. Following the iterations outlined below, about half of the corporate owners could be matched to company registration numbers.

Table B1: Match rates for trademarks and patents

Stage of matching	Trademarks	Patents
Total number of owners	302,554	30,324
Total number of TMs/ patents	903,418	121,899
Matched by name and postcode	80,061	-
Matched by name	140,334	16,683
OpenRefine matches	3,647	-
Ultimate owner matches	14	687
Clerical matches	26	92
Total matched TMs/ patents	461,972	86,408

Transfer to the SRS

The datasets matched to CRNs were transferred into the ONS SRS, where they were matched to entrefs. Not all businesses can be matched, as some have ceased trading since they received Innovate UK support, and for some the right legal unit surveyed by the ONS could not be identified. It was especially tricky to match the largest companies in the sample to the right entity, so several attempts were made and CRNs for several legal entities in the group were sent, until a match was found.

An initial sample of 11,903 organisations were sent into the ONS SRS from the Innovate UK database, of these 1,235 were not matched, with 70 identified as non-companies (academic or public institutions). The other 10,668 CRNs were matched to 10,347 entrefs. Trademarks and patents had a lower rate of entref matching.

Missing spend from Innovate UK

The 1,165 CRNs that did not match to any enterprise reference numbers in the BSD represent around £38m of Innovate UK grant spend. The table below outlines the enterprise classifications for the missing grant money from the 1,165 CRNs identified as companies. The majority (£27m of £38.3m) of companies not being matched are micro or small traders. They could be unmatched because they have not registered or are not operating with any PAYE employees.

Table A2: Spend not linked in the SRS to enterprise reference numbers

Enterprise Category	Amount of grant missing from analysis (£)
CIC	79,225
Charity	468,305
Large	3,197,408
Medium	1,542,260
Micro	17,200,000
Small	10,300,000
Sole Trader	23,868
Start-Up	449,944
Unknown	4,979,718
Total	38,300,000

Matching highlights mergers and acquisitions

Sometimes companies with separate CRNs in the database match to the same enterprise reference number within the ONS. For the Innovate UK data while 10,148 of the CRNs matched to single entrefs, 520 CRNs matched to 210 entrefs. This is likely due to mergers, acquisition or companies requesting to submit one return for multiple subsidiaries. These companies are tagged in the database, but all analysis is done by the entref, not the initial CRN within the database.

Some companies have changing entrefs over time, often due to mergers. This appears in the data, although it is limited. Since the only way to see this is with multiple years of matching to the CRN, there is no comparison in the wider BSD population. The earlier entref can be used to match to earlier years of surveys, increasing the overall matching rate to surveys.

Annex C: Econometric methods

The approach used in understanding the impact of Innovate UK and NMS support identifies statistically a counterfactual, a set of businesses that are similar to the supported businesses. The performance of the supported businesses after the intervention is compared to the counterfactual and any difference used to assess the economic impacts of the support. The annex focuses first on the impact measurement and then considers the statistical processes used to determine the comparison businesses. It highlights use of:

- An additionality ratio in estimating what impact is additional
- Trimming the supported businesses to focus on a sub-sample of the businesses
- Balance tests in establishing the quality of the matching
- Propensity score matching and testing the match robustness.

Impact estimation

When the supported businesses have been identified in the dataset, the attention turns to the estimation of impact and assessing this in comparison to a control group. The general approach is to look at the performance of the supported businesses from a period before support to some years afterwards and then to adjust this for what might have happened anyway. Most analysis occurs in terms of the changes seen in performance. Where growth is higher in the supported businesses than a comparable set of businesses, this is evidence of additional impact.

However, of interest is the level of economic activity to which any differential growth rate can be applied. Assessing economic activity has to be careful to avoid double counting. A first consideration is over multiple years the same businesses are supported more than once. In the analysis the overlap of business support over years is taken into account in two ways.

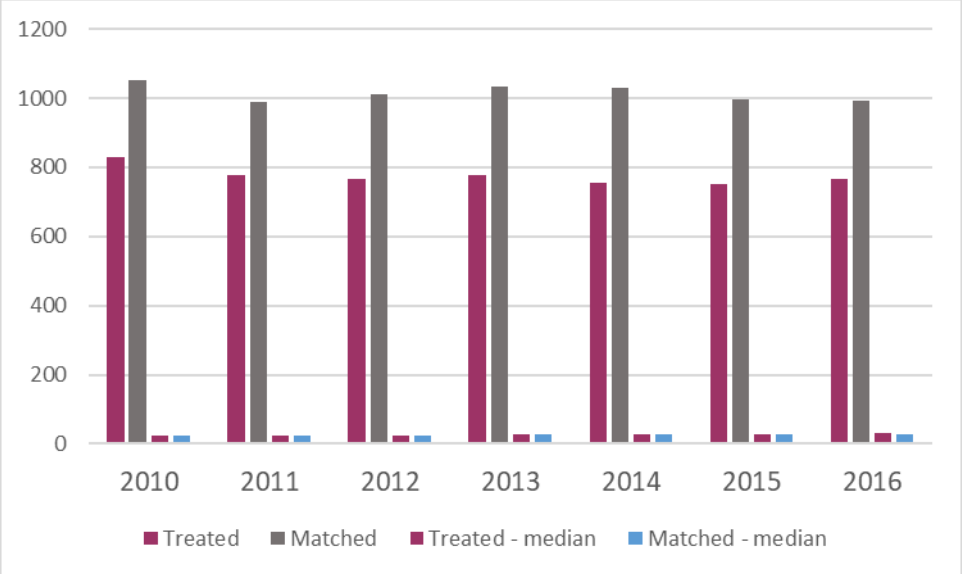
For NMS, the analysis has accepted that repeated support is the norm, a consequence of the operating models of businesses that use measurement services regularly. Here, the analysis has been undertaken where the treatment is in terms of whether over a period support is provided almost every year or not (as the regular treatment modelling); or whether it is used less frequently but at least once (as treated sometimes modelling). Then the analysis matches at the start of a period in 2009, allowing performance analysis over the next eight years.

For Innovate UK, analysis is undertaken each year 2010-13, with supported businesses being identified each year and then the treatment variable is one if first support occurs in that year. Businesses are then removed from analysis in later years avoiding double counting. A final adjustment is take account of the overlap in NMS and Innovate UK beneficiaries, only analysing these once.

A problem in the business data is that the size of businesses has a long tail of very large businesses. Figure C1 presents summary statistics about this, the simple mean of

employment for the treated and the matched comparison businesses for those supported for the first time by Innovate UK in 2013. It also presents the median.

Figure C1: Employment in business supported by Innovate UK, 2013



The figure highlights two problems. Firstly, where employment growth is applied to the jobs seen in the supported businesses, it will be affected by the treatment of the largest businesses. In 2013, there were about 700 businesses supported by Innovate UK for the first time, so the employment in the supported and counterfactual is 1.5m jobs. A problem is that – where R&D has been supported in a major company's research unit – the economic data held about its performance will cover multiple other units.

The skewed distribution in business size means there may be inaccuracies or sometimes biases to:

- Measures of firm level impact: as treatment is provided to businesses where the median employment is around 25 employees, the typical supported business is very different to the average employment of a supported business.
- Data about the incidences of support: the fact that an enterprise has been supported may miss that different entities within a large enterprise, operating quite different research portfolios, may be benefitting from different incidences of support that – in the data – appear to go to the same (enterprise level) beneficiary
- Assessing the level of economic impact: Figure C1 highlights the issue that after matching, employment change in the level of jobs in the supported businesses cannot by compared to the change seen in the counterfactual, as they start with very different average levels of employment, about a third higher in the 2013 cohort.

The third bullet motivates the use in this study of an additionality ratio, estimating the proportion of growth seen in supported businesses not seen in the counterfactual. This proportion is applied to the growth in jobs for the supported. This avoids the additional employment estimates being influenced by the (largely random after matching) differences in the size of businesses in the two groups.

The second bullet exposes the influence of size of a business on the propensity to seek support. There are likely to be two drivers for this, one productivity based but one more related to the diversity of activities in a large business. The former is that underpinned in the economics of innovation. Research activity will be related to the scale at which a business operates at, as scale increases the opportunities to enhance technology as businesses are more likely to be operating on the technology frontier. The selection model would then incorporate this driver to seek support.

However, the largest businesses are also likely to have multiple activities, with enterprise level size also reflecting this diversity. This is not necessarily productivity related and more that – in a large conglomerate – some part of the business is always likely to be innovating and so seeking support. Unless the data can be analysed at the unit level, the increase in the chance of seeking support purely because the organisation operates in many technologies makes analysis more difficult to justify in productivity terms.

There is evidence that this is a problem in this study, where whole enterprises have to be linked to supported data rather than the individual plants or units within large businesses. For example, treated businesses represent, in total, a very high proportion of UK employment, at over 5% of employment. This is probably more due to some large businesses having received support.

To tackle this coverage issue, the key step has been to trim the businesses used in assessing activity levels in the supported businesses (though not in calculated the portion that is additional). This trims out the largest businesses, which are likely to be driven to support more because the enterprise group is too large an organisational unit to model assignment of support for an activity in the business.

Propensity score matching

The following section will outline the formal PSM model. Analysis used the Stata routine *psmatch2*. Treatments have been categorised to the treatment variable $\mathrm{Treat_{it}}$ equal to 1 if an enterprise i receives support in period t. Let $\Delta \mathrm{y}^{1}_{\mathrm{i,t+1}}$ be the employment of enterprise i at time t+1 after receiving treatment in period t and let $\Delta \mathrm{y}^{0}_{\mathrm{i,t+1}}$ be the hypothetical performance of the same enterprise i at the same time t+1 had it not received treatment in period t. The employment effects of the treatment in enterprise i, called the average treatment effect on the treated, can then be expressed as follows:

$$\hat{\alpha} = E\left(\Delta y^{1}_{i,t+1} \middle| Treat_{it} = 1\right) - E\left(\Delta y^{0}_{i,t+1} \middle| Treat_{it} = 1\right) \qquad (1)$$

The second term in equation (2), - $\mathrm{E}(\Delta \mathrm{y}^{0}_{\mathrm{i,t+1}} | \mathrm{Treat_{it}} = 1)$, is the counterfactual mean or the hypothetical employment change in enterprises supported, had they not received the treatment. Since this term is unobservable, a proxy needs to be found for the counterfactual mean. In experimental studies, the selection problem is dealt with by random assignment of treatment, which ensures that every individual has ex ante the same chance of receiving treatment (Ravallion, 2003). Non-experimental studies try to replicate this by applying propensity score matching. The main idea behind this approach is to find a control group that is like the treatment group in all respects except the exposure

to the treatment (Ravallion, 2003). A group of enterprises must be found that is like those benefitting from the Catapult in all relevant characteristics, prior to the actual treatment. The estimation of the causal effect in this case becomes:

$$\hat{\alpha} = E\left(\Delta y^{1}_{i,t+1} \middle| Treat_{it} = 1, X_{i,t-1}\right) - E(\Delta y^{0}_{i,t+1} \middle| Treat_{it} = 0, X_{i,t-1} \quad (2)$$

where $\mathrm{E}\left(\Delta \mathrm{y}^{1}_{\mathrm{i,t+1}} \middle| \mathrm{Treat_{it}} = 1, \mathrm{X_{j,t-1}}\right)$ is the mean employment change at time t+1 of the enterprises being supported at time t; $\mathrm{E}(\Delta \mathrm{y}^{0}_{\mathrm{i,t+1}} \middle| \mathrm{Treat_{it}} = 0, \mathrm{X_{i,t-1}})$ is the mean employment change of the control group at time t+1; and $\mathrm{X_{i,t-1}}$ is a vector of observed conditioning covariates in the pre-treatment year t-1. By matching enterprises whose covariates are closely aligned in the pre-treatment year it is possible to derive the causal effect of support on employment or an outcome. A practical constraint arising from the application of such matching techniques is that exact matching across multiple covariate indices poses high demands to the data available. Rosenbaum and Rubin (1983) suggest that this problem of dimensionality can be significantly reduced by matching on a single index: the propensity score, or the probability of receiving treatment conditional on the relevant pre-treatment covariates. Using the propensity score, the equation for the average effect of support becomes:

$$\hat{\alpha} = E\left(\Delta y^{1}_{i,t+1} \middle| Treat_{it} = 1, p(X_{i,t-1})\right) - E(\Delta y^{0}_{i,t+1} \middle| Treat_{it} = 0, p(X_{i,t-1})) \quad (3)$$

where p is a propensity score conditional on $\mathrm{X_{j,t-1}}$. The average effect of support on employment is estimated as the difference between the mean employment change of treated enterprises and that of enterprises that had ex-ante similar likelihood of being supported but were not.

For consistent estimates of the employment effects, two key assumptions must hold: the conditional independence and the common support assumption. Conditional independence means that there are no unobservable differences between supported and non-supported enterprises after conditioning for $\mathrm{X_{i,t-1}}$, so that any systematic differences in outcomes can be attributed to the treatment (Imbens 2004, Smith and Todd 2005). The assumption can be stated formally as:

$$\left(\Delta y^{1}_{i,t+1}, \Delta y^{0}_{i,t+1}\right) \perp Treat_{i,t} \middle| X_{i,t} \quad (4)$$

where $\perp$ indicates orthogonality between two variables. This is a strong assumption, as there can still be differences after conditioning for the observable covariates available in the data. In some cases, this issue is addressed by using employment change as the outcome variable and looking at the differences in this outcome between treatment and control groups (Sarkisyan et al. 2009). This is known as difference-in-difference or double difference matching, where the first difference removes the unobserved heterogeneity and restores conditional independence and the second produces the impact estimates (Smith and Todd 2005, Essama-Nssah 2006). This method is used for impact estimation in this study.

The common support assumption requires an overlap in the distribution of covariates between the treated units and the control group members to make matching possible. This is stated formally as:

$$0 < \Pr\left(Treat_{it} = 1 | X_{i,t-1}\right) < 1 \quad (5)$$

If the two assumptions hold, the mean outcome of the non-treated enterprises acts as a counterfactual for the outcome trend beneficiaries would have shown in absence of the treatment.

Researchers have relied on propensity score matching, which uses a single variable (the propensity score) to undertake the matching. The propensity score is estimated by means of a probit model in which the dependent variable is a dummy equal to 1 in the year an enterprise receives treatment and 0 otherwise. A probit estimation of propensity to receive support is estimated, i.e.:

$$P(Cat_{it} = 1 | X_{i,t-1}, Z_{i,t-1}, Cat_{ti}) \quad (6)$$

where $\mathrm{Cat_{it}}$ is the dummy for a category of support; $\mathrm{X_{i,t-1}}$ includes all enterprise level variables that affect the probability of receiving treatment, $\mathrm{Z_{i,t-1}}$ controls for external factors that are likely to affect the probability of receiving treatment and $\mathrm{Cat_{ti}}$ are state dummies reflecting environmental aspects. One of the required conditions in the propensity score matching analysis is that the variables included in the model should not be affected by the treatment. To ensure this, the firm specific variables used in the model are lagged one year.

After having estimated the propensity scores for each enterprise in the Business Structures Database, first-time beneficiaries are matched with non-beneficiaries using nearest-neighbour (1:1 ratio) matching where the unit chosen from the pool of non-beneficiary businesses (i.e. an untreated enterprise j) as a match for beneficiaries (i.e. a treated enterprises i) is the one closest in terms of the propensity score. Given the size of the control group, the data is sorted randomly prior to the matching procedure to avoid systematic bias due to ties in the data. The matching procedure formally is:

$$\left|p_i - p_j\right| = \min_{k \in \{Treat=0\}} \{|p_i - p_k|\} \quad (7)$$

A common problem in PSM is the occurrence of bad matches (i.e. the nearest neighbour is not very near). Several tests, detailed below, have been employed to check the quality of matches. The procedure is run on each year of the sample, ensuring that new beneficiaries are matched to enterprises in the same year. The model is specified to restrict the matching pool to enterprises never receiving any support. Further, the common support assumption, discussed earlier, rules out the perfect predictability of support given the observed covariates. This ensures the existence of potential matches in the pool of non-treated enterprises from the BSD.

Testing the quality of the matching results

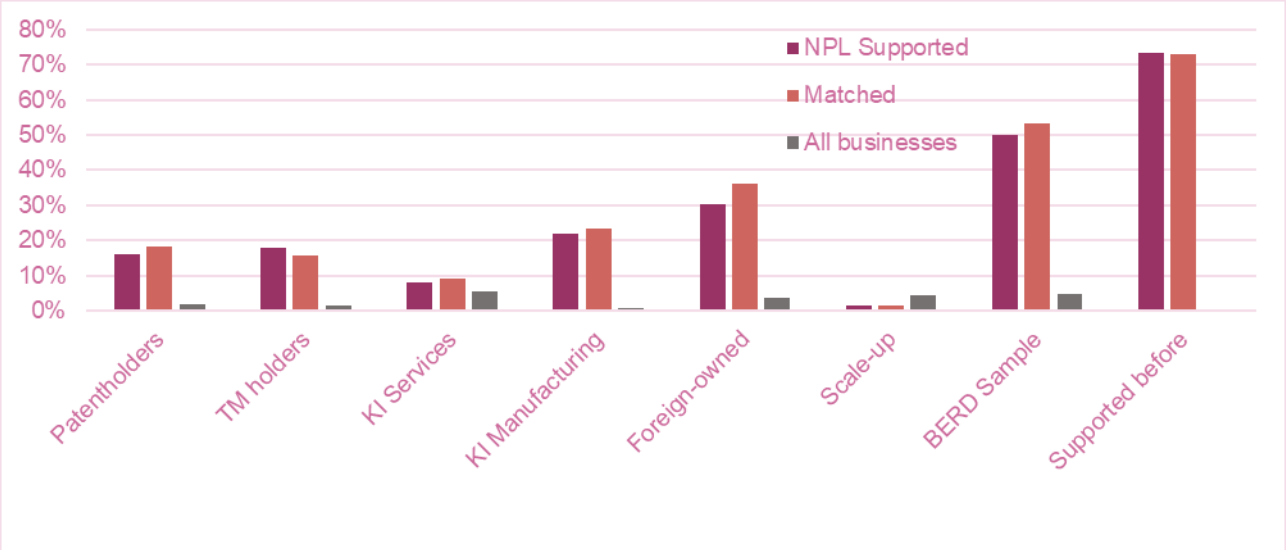
There are two sets of tests to indicate how well the businesses identified for the control group match the characteristics of the beneficiaries. A first set of tests – the balancing tests – look at whether the two groups are similar in terms of their characteristics. A second set of tests focuses on what may be missed in the selection modelling and looks at how sensitive the average treatment effect is to different assumptions what may be unobserved when matching to generate a control group.

Balancing tests

For all characteristics, the average values of the businesses in the control group can be compared to those of the treated. Analysis used Stata commands *pstest* and *kdensity*.

For characteristics that are continuous variables, such as age of businesses or the turnover, the distribution of the variable can also be compared, as well as the average. In the ideal case, the variance ratio of the treated and untreated groups would be one, implying that the variances are the same. An F-test can be used to assess whether the observed variance ratio differs enough from one to conclude that it the variances are significantly different from each other. Additionally, Kolmogorov-Smirnov (K-S) tests are used to test for equality of distributions.

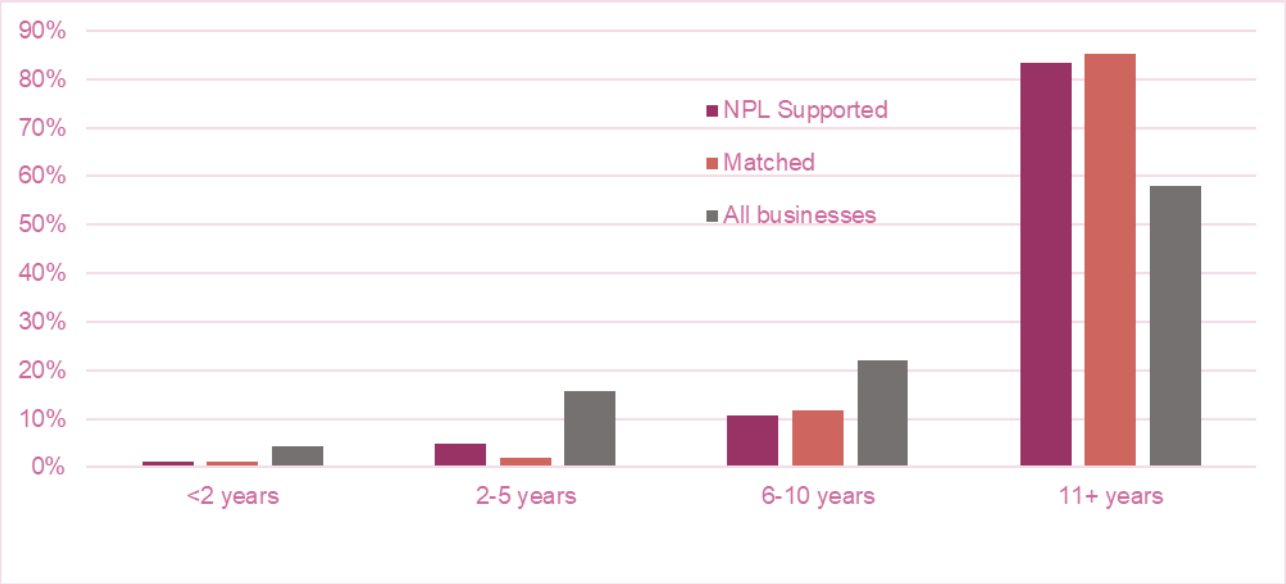
Figure C1: Characteristics of NMS-Supported businesses and Comparators



To test for the balancing achieved by PSM, the standard procedure is to look at the reduction in the difference of means among variables (either in absolute terms or as a standardized percentage) and at the variance ratio of treated over non-treated observations. To test for equality of means, a regression with a variable, for example turnover, is run on a dummy variable for treatment. If the dummy does not have a statistically significant effect, it can be assumed that the means are the same in the treated and the matched untreated group. Running this test, called PS-test before and after matching, it can be assessed what reduction in bias was brought about by PSM.

Table C1 looks at the characteristics of businesses regularly supported by NMS in the 2010-16 period, looking at whether they hold a patent or trademark, industry and ownership categories and whether the businesses were in the ONS Business Expenditure on R&D survey. The effect of PSM in achieving a balance is quite marked, as is the significant difference between NMS supported businesses and the wider business population. For example, almost 15% of NMS supported businesses hold a patent, while the percentage in the wider business population is about one. The PSM model means that the counterfactual businesses are slightly more likely to hold a patent. PSM also removes the imbalance seen in the age of the supported businesses in comparison to the wider business population.

Figure C2: Age of NMS-Supported businesses and Comparators



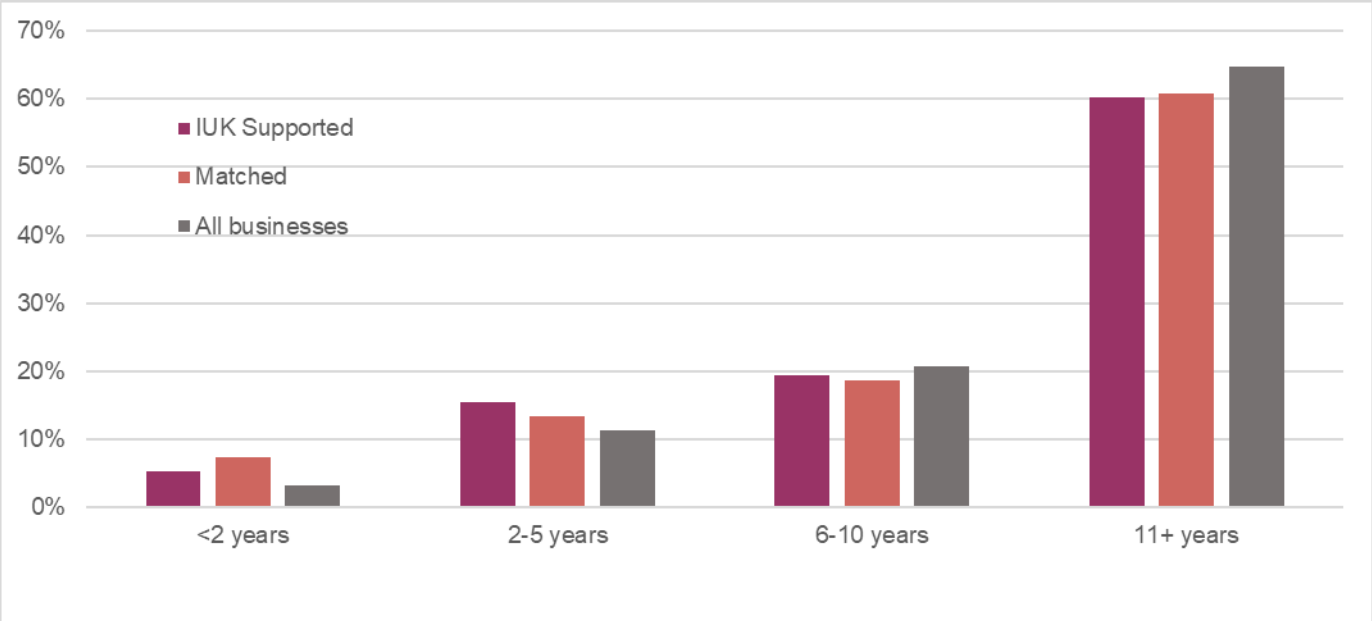
Figures C3 and C4 consider the balance results for Innovate UK in the 2013 modelling. It is noticeable that the Innovate UK businesses are similar to the NMS in terms of being more research active than the wider business population reflected in holding intellectual property, being in knowledge intensive industries and having been supported in the past. However, they are more likely to include businesses scaling up, having 20% annual sales growth in the three years before support and with 20 employees at least.

Figure C3: Characteristics of Innovate UK Supported businesses and Comparators, 2013



The results for age indicate that the PSM would have had to perform less balancing on age, as the businesses supported by Innovate UK for the first time in 2013 are similar in terms of age to the wider business population.

Figure C4: Age of Innovate UK Supported businesses and Comparators, 2013



Sensitivity tests

PSM tries to mimic a randomised control trial by matching treated businesses with those that are similar on all observable characteristics. In effect, the counterfactual contains businesses that have the same probability of being treated as those that received support.

However, the analysis may be biased if there are unobservable variables that drive selection into treatment as well as the outcome (employment growth, turnover growth etc). One possible variable is managerial capability and ambition. If more ambitious managers are more likely to seek support, the supported businesses would differ in an important respect from untreated businesses, and it may be the manager's ambition rather than the support that causes higher growth.

One approach to test for the potential impact of unobserved variables is the Rosenbaum-bounds method, using the Stata command *rbounds*. It assesses "how strongly an unmeasured confounding variable must affect selection into treatment in order to undermine the conclusions about causal effects from matching analysis" (DiPrete and Gangl, 2004). Different levels of hidden bias can be expressed in terms of the odds ratio, gamma Γ, of two matched observations being treated. If matching is unbiased, observations with the same observable characteristics have the same probability of being treated. When Γ=2, an unmeasured cofounding variable causes one observation to be twice as likely to be selected into treatment than the matched observation with the same observable characteristics (Peel and Makepeace, 2009).

The method then assumes that there is a known factor causing bias to the level of Γ, and that the treatment effect from this bias can be stripped out. Once this is done, whether the treatment effect remains significant is tested. In this fashion, starting with zero bias, the treatment effect can be computed and the effect of the assumption of ever larger bias then tested.

The Rosenbaum bounds estimation for different matching models used in this study. For different levels of Γ, it gives the upper- and lower-point estimates of the treatment effects, under the assumption of negative and positive selection bias, respectively. It also gives significance levels for these estimates under the null hypothesis that the true treatment effect is zero at a certain level of positive or negative bias. The upper- and lower-point estimates can be interpreted in terms of a – usually – increasing cone of possible values as Γ rises. Where the cone begins to include zero, then that is the level of bias in the matching that would render the treatment results insignificant.

The tables give the treatment effect: called Hodges-Lehman (HL) point estimate. Estimates should be the same at Γ=1, but differ slightly at zero bias from earlier difference-in-difference calculations because HL estimates the median, rather than the mean treatment effect. The point estimate of the treatment effect is based on the difference in the outcome variable, in this case log employment growth, for each matched pair. The differences between pairs, Δ_i are ranked, and for adjacent Δ_i, the average is computed. The median of these values gives the HL point estimate. P-values and confidence intervals around this estimate can be obtained using Wilcoxon sign ranking.

Alternative matching algorithms

There are different methods available to implement Propensity Score Matching. The most commonly used methods are nearest neighbour, radius and kernel matching. Nearest neighbour matching matches a treated observation to that untreated observation with the closest propensity score. Radius matching matches with all untreated observations whose propensity score is within a given radius of the propensity score of the treated observation. Finally, kernel matching defines the counterfactual as a weighted average of all other untreated observations, where the distance of their propensity scores weights observations. A related method, coarsened exact matching, does not estimate a propensity score but coarsens variables, usually by converting continuous variables into ordered categorical variables, and then matches to an observation that falls into the same bins for such variables. Because variables have been simplified, there is a finite number of bins – the permutations that the different coarsened variables take – and matching becomes possible.

The literature does not identify a preferred method, so that the researcher has to make a choice on the basis of appropriateness to the given data (Morgan and Winship, 2014). Nearest neighbour matching has been chosen for a number of reasons. First, this is the most straightforward method and requires no arbitrary choices (for example the size of a radius or the level of coarsening). Secondly, nearest neighbour performs well where the number of potential matches is high. This is the case with the data used here, with the Businesses Structures Database being a very large census of businesses. In contrast, kernel matching is recommended where there is a small sample from which control cases can be selected (Caliendo and Kopeinig, 2005).

A third aspect is the computational practicality. In nearest neighbour, the matching is performed on a single metric and searching for the closest score is then relatively straightforward. This is an important consideration when dealing with large datasets such as the BSD. Finally, and perhaps most importantly, nearest neighbour matching performed well in reducing bias between the supported businesses and those in the comparison group, which is the goal of PSM.

Robustness tests

Tests look at the balance of observable variables before and after propensity score matching. This shows weather the matching is working. If there was a statistically significant results before the matching, the psm should remove this and lead to a balanced treatment and control group. The balance across all variables is measured post matching.

The Rosenbaum test shows us how easy it is for the outcome results to change based on an additional unobservable variable used in the matching process. This allows us to see how robust the results are. The test is done for each outcome and model and leads to a few general conclusions. Model II using categorical employment and turnover variables is the most robust to unobservable variables. Also, the R&D outcomes are the least robust, and often lose significance with any adjustment. Employment and turnover outcomes are

more robust, but dividing the treatment into groups, decreases the sample sizes and makes them more fragile to additional variation.

The Ksmirnov test looks at the distribution of the matched control group and treatment groups along certain variables to see if the distribution is significantly different. For both employment and turnover, there seems to be an unequal distribution of the treated and control groups.

Removing outliers and trimming using the propensity score

Outliers in growth can affect the results by pulling the mean average up. To ensure that outliers are not pushing the results, log growth is used. Additionally, the specifications are run after removing the largest 1% of businesses, which allows us to remove any bias from the massive organisation in the outcomes.

In order to create matches by propensity score there are different methods for defining common support for the treated and control groups. One option is to keep all companies in the pool based on the highest and lowest propensity scores of common support. Another option is to look at quintiles of propensity support and remove any quintiles that do not have enough common support between matched and control companies. This option is known as a way of trimming the interior of the distribution if there is not enough shared distribution between the treated and control groups.

Testing discontinuity in selection models

Table C2 presents the estimates of a cross-section pseudo-Chow test to identify whether the explanatory variables in the selection modelling have different effects in three years for Innovate UK. The first-row tests whether a selection model that pools across all three years (2011, 2012 and 2013) would have the same co-efficients to one where the years are modelled separately. It indicates how the pooling is not appropriate, with different years having significantly different parameters in the modelling.

Table C2: Chow tests for Innovate UK selection models, 2011-13

	Test Statistic	
	Chi-stat	d.f.
Years tested		
All	107.59***	72
2012 vs 2013	41.33	36
2011 vs 2013	66.37***	36
2011 vs 2012	55.54**	36

* show the significance level; *** significant at the 1% level; ** at 5%; * at 1%. Standard errors are robust.

The other rows then consider the three years in pairs and this highlights that only for one year, 2011, is there a strong case for separately modelling the selection modelling. When

2012 and 2013 is pooled the test statistic indicates a hypothesis that the parameters are the same cannot be rejected.

This confirms other findings and the general approach taken to the modelling. The 2011 cohort is all the businesses that are supported in 2011, whereas the following years then only considers the businesses that have not been included in previous cohorts to avoid double counting the businesses from earlier modelling. So, in 2012 and 2013, these businesses are removed so that beneficiaries then only include ones receiving their first support after 2010 in 2012 and 2013.

The 2011 cohort then includes a set of businesses which tend to receive support every year, which can be characterised as the large, innovation focused businesses. As they receive support on many occasions, by 2012 and 2013, the newly supported businesses stabilise to a set of businesses that are new, smaller and seek support relatively infrequently. The Chow tests confirm this difference in selection modelling cohorts.

Dosage modelling

Fryges et al (2015) review approaches where there is a continuous treatment based on the generalised propensity score (GPS), testing for a causal effect of R & D activities. The GPS for continuous treatments is an extension of the propensity score methods described above for binary treatments (Rosenbaum and Rubin, 1983) and multi-valued treatments (Imbens, 2000; Lechner, 2001) which takes into account the selectivity in the grants awarding processes (Bia and Mattei, 2008; Guardabascio and Ventura, 2014).

The GPS method focuses on the Neyman-Rubin causality concept of potential outcomes. Whereas the binary treatment modelling focuses on the mean outcome of the group of treated entities compared to the situation where treated entities did not receive the treatment, the focus in a continuous treatment situation is the pairwise treatment effects along the full domain of potential treatment doses, the difference I outcome of the entities that received a particular dose of the treatment compared to a situation in which they received another dose.

The advantage of the GPS are:

- It allows for a continuous treatment variable so that different levels of support can be considered
- It enables estimation across the entire treatment-outcome function, distinguishing from modelling that looks at an average level of treatment
- Most importantly, it allows an analysis of the optimal dosage

This method estimates the dose-response function (DRF) through a generalized linear model which is more flexible in terms of underlying data structure. The estimation of the DRF thus follows from a two-step procedure in which we first estimate the parameters of the conditional distribution of the treatment given the covariates that explain the selection

into the treatment. In the second step, the conditional expectation for the outcome variable, given the conditional distribution of the treatment, is modelled.

Table C3: Maximum likelihood estimations of the GPS

Treatment variable	InnUK support, 2011	
	Coef.	Std. Error
Live local units	-0.001	0.00***
Patent year before	-0.059	0.10
Trade mark year before	0.078	0.13
R&D in three years before	0.000	0.20
High KI Services	-0.069	0.09
High KI Manufacturing	-0.122	0.15
Foreign owned	0.089	0.13
Scale-up in period before	0.189	0.24
Employment year before	0.087	0.05*
Real turnover year before	0.050	0.03
InnUK project period before	0.039	0.09
BERD sample year before	0.105	0.21
Constant	-1.618	0.20***
Observations	729	
Log pseudolikelihood	-325.244	

Table C4: Balancing tests for the GPS estimator

Variable	Treatment Interval 1 [.01,.05]		Treatment Interval 2 [.06,.19]		Treatment Interval 3 [.20,1]	
	mean diff	t-value	mean diff	t-value	mean diff	t-value
Live local units	5.170	0.733	8.418	1.002	-8.481	-1.190
Patent year before	-0.038	-1.127	0.026	0.688	0.001	0.033
Trademark year before	0.042	1.785	0.019	0.701	-0.025	-1.101
R&D in three years before	-0.002	-0.062	0.072	1.654	-0.047	-1.332
High KI Services	0.072	1.956	0.007	0.173	-0.047	-1.468
High KI Manufacturing	-0.020	-0.907	0.037	1.534	-0.016	-0.791
Foreign owned	0.031	1.156	0.032	1.076	-0.030	-1.199
Scale-up in period before	-0.028	-1.684	0.001	0.086	0.011	0.889
Employment year before	0.550	4.297	0.186	1.298	-0.333	-3.589
Real turnover year before	0.682	3.902	0.126	0.658	-0.334	-2.446
InnUK project period before	0.017	0.426	0.012	0.270	-0.028	-0.815
BERD sample year before	0.015	0.374	0.070	1.623	-0.052	-1.497

Tables C3 and C4 highlight some of the estimations underpinning the graphs in the report. They focus on the Innovate UK support in 2011 with the first table indicating the estimate of the scoring model. Then in table C4 the dosage is modelled in terms of three intervals. Like the propensity score matching methods, the balance of the treated against comparators is tested, but now rather than focusing on the treated, the focus is on selected intervals along the dosage levels.

Overall, the dosage modelling does balance on firm level characteristics (industry, innovation) but there remains a less convincing balance on the two size measures.

Table C1.1: Balance tables for Model III 2009-2012 before and after matching, Innovate UK

	2009			2010			2011			2012		
	Treated	Control		Treated	Control		Treated	Control		Treated	Control	
		Before	After		Before	After		Before	After		Before	After
Patent owner prior	0.19	0.01***	0.21	0.21	0.01***	0.24	0.21	0.01***	0.22	0.24	0.02***	0.27
Trade mark owner prior	0.14	0.01***	0.14	0.13	0.01***	0.15	0.10	0.01***	0.12	0.10	0.01***	0.12
R&D in 3 years before	1.97	0.06***	1.78	2.25	0.07***	2.08	1.95	0.08***	1.87	2.15	0.09***	2.03
High Knowledge Intensive	0.20	0.05***	0.21	0.26	0.05***	0.28	0.32	0.05***	0.34	0.27	0.05***	0.3
High KI Manufacturing	0.06	0.01***	0.05	0.07	0.01***	0.07	0.07	0.01***	0.07	0.08	0.01***	0.09
Foreign owned	0.21	0.03***	0.23	0.20	0.03***	0.21	0.13	0.03***	0.17**	0.17	0.04***	0.19
Employment year before	3.95	2.33***	3.8	3.70	2.29***	3.75	3.26	2.32***	3.48*	3.23	2.34***	3.25
Real turnover year before	10.57	8.85***	10.52	10.38	8.79***	10.5	9.55	8.86***	9.93**	9.60	8.87***	9.6
Employment growth 1 year prior	0.07	0.03***	0.09	0.11	0.07**	0.1	0.07	0.04	0.1	0.04	0.01*	0.02
Turnover growth 1 year prior	0.07	0.05	0.1	0.07	0.08	0.1	0.08	0.08	0.08	0.00	0	-0.01
Innovate UK project before	0.37	0***	0.37	0.43	0***	0.4	0.42	0***	0.38	0.45	0***	0.41*
Employment growth 1 year prior	0.07	0.03***	0.09	0.11	0.07**	0.1	0.07	0.04	0.1	0.04	0.01*	0.02
Employment growth 2 years prior	0.17	0.07***	0.15	0.16	0.13	0.17	0.16	0.12	0.19	0.08	0.06	0.05
Turnover growth 1 year prior	0.07	0.05	0.1	0.07	0.08	0.1	0.08	0.08	0.08	0.00	0	-0.01
Turnover growth 2 years prior	0.28	0.25	0.37	0.17	0.17	0.24	0.16	0.17	0.19	0.10	0.08	0.03
Productivity growth 1 year prior	- 0.02	0	-0.01	- 0.07	-0.02	-0.02	- 0.01	0.01	-0.05	- 0.05	-0.03	-0.03
Productivity growth 2 years prior	0.08	0.14	0.16	- 0.03	-0.01	0.02	- 0.02	0	-0.04	0.00	-0.01	-0.04
Observations		598			602			724			999	

Table C1.2: Balance tables for Model III 2013-2016 before and after matching, Innovate UK

	2013			2014			2015			2016		
	Treated	Control		Treated	Control		Treated	Control		Treated	Control	
		Before	After		Before	After		Before	After		Before	After
Patent owner prior	0.28	0.02***	0.29	0.27	0.02***	0.28	0.29	0.02***	0.28	0.29	0.02***	0.29
Trade mark owner prior	0.09	0.01***	0.08	0.10	0.01***	0.11	0.08	0.01***	0.08	0.09	0.01***	0.1
R&D in 3 years before	2.20	0.1***	2.1	2.33	0.1***	2.26	2.09	0.1***	1.93	2.54	0.11***	2.37
High Knowledge Intensive	0.27	0.05***	0.28	0.30	0.05***	0.32	0.27	0.06***	0.28	0.30	0.06***	0.31
High KI Manufacturing	0.07	0.01***	0.08	0.09	0***	0.11	0.07	0***	0.09	0.06	0***	0.06
Foreign owned	0.14	0.04***	0.14	0.14	0.04***	0.15	0.14	0.04***	0.18**	0.17	0.03***	0.2
Employment year before	3.11	2.4***	3	3.18	2.31***	3.09	3.21	2.35***	3.22	3.31	2.35***	3.27
Real turnover year before	9.50	8.89***	9.33	9.56	8.7***	9.43	9.53	8.75***	9.61	9.69	8.76***	9.66
Employment growth 1 year prior	0.10	0.06**	0.1	0.10	0.07***	0.1	0.08	0.08	0.08	0.10	0.05***	0.09
Turnover growth 1 year prior	0.11	0.04***	0.12	0.07	0.06	0.09	0.05	0.1**	0.06	0.10	0.1	0.14
Innovate UK project before	0.51	0.01***	0.5	0.58	0.01***	0.57	0.59	0.01***	0.59	0.65	0.01***	0.65
Employment growth 1 year prior	0.10	0.06**	0.1	0.10	0.07***	0.1	0.08	0.08	0.08	0.10	0.05***	0.09
Employment growth 2 years prior	0.14	0.08***	0.13	0.25	0.16***	0.17***	0.20	0.15***	0.14***	0.19	0.13***	0.16
Turnover growth 1 year prior	0.11	0.04***	0.12	0.07	0.06	0.09	0.05	0.1**	0.06	0.10	0.1	0.14
Turnover growth 2 years prior	0.18	0.05***	0.13	0.17	0.13	0.16	0.11	0.18**	0.13	0.20	0.22	0.23
Productivity growth 1 year prior	- 0.02	-0.05	-0.01	- 0.07	-0.03*	-0.05	- 0.06	0***	-0.04	- 0.03	0.03**	0.02
Productivity growth 2 years prior	- 0.01	-0.06*	-0.05	- 0.15	-0.07**	-0.05**	- 0.13	-0.01***	-0.03**	- 0.05	0.04***	0.02
Observations		1147			1140			1312			941	

Table C2.1: Selection Model I probits for all treated 2009-2016

	2009	2010	2011	2012	2013	2014	2015	2016
Observations	622,359	738,734	577,889	563,137	568,923	577,349	586,178	610,158
Adj R-Squared	0.21	0.24	0.22	0.23	0.24	0.25	0.23	0.25
Treated	598	602	724	999	1147	1140	1312	941
Patent owner prior	0.56***	0.58***	0.56***	0.63***	0.74***	0.68***	0.78***	0.74***
Trade mark owner prior	0.17***	0.22***	0.17***	0.17***	0.11*	0.18***	-0.03	0.07
High Knowledge Intensive	0.40***	0.50***	0.61***	0.50***	0.47***	0.58***	0.52***	0.51***
R&D in three years before	0.07***	0.07***	0.06***	0.06***	0.08***	0.07***	0.07***	0.09***
High KI Manufacturing	0.33***	0.49***	0.58***	0.55***	0.53***	0.71***	0.63***	0.40***
Foreign owned	0.17***	0.16***	0.02	0.15***	0.05	0.02	0.06	0.17***
Scale-up in period before	-0.05	0.11*	-0.01	0.14***	0.07	0.03	0.11**	0.09*
Employment year before	0.06***	0.03	0.11***	0.06***	0.03	0.08***	0.07***	0.05**
Real turnover year before	-0.02*	0.03**	-0.03***	-0.03***	-0.01	0.01	-0.01	-0.02**
Employment growth 1 year prior	0.06*	0.11***	0.03	0.02	0.08***	0.07**	-0.01	0.09***
Turnover growth 1 year prior	0.02	-0.04*	0.00	0.02	0.03*	0.01	-0.01	0.02
BERD sample period before	0.50***	0.58***	0.67***	0.71***	0.73***	0.68***	0.62***	0.53***
Innovate UK project before	-	-	-	-	-	-	-	-
Constant	-3.15***	-3.56***	-3.08***	-2.72***	-2.75***	-3.1***	-3.00***	-2.93***

Table C2.2: Selection Model II probits for all treated 2009-2016

	2009	2010	2011	2012	2013	2014	2015	2016
Observations	623,679	740,454	579,928	565,545	571,848	580,859	590,176	614,967
Adj R-Squared	0.29	0.33	0.31	0.32	0.35	0.38	0.37	0.38
Treated	598	602	724	999	1,147	1,140	1,312	941
Patent owner prior	0.41***	0.38***	0.35***	0.46***	0.54***	0.42***	0.52***	0.41***
Trade mark owner prior	0.06	0.17**	0.11	0.08	0.03	0.15**	-0.14**	0.04
High Knowledge Intensive	0.36***	0.46***	0.58***	0.43***	0.34***	0.45***	0.35***	0.34***
R&D in three years before	0.02	0.02	0.02	0.02**	0.03***	0.02***	0.02*	0.02***
High KI Manufacturing	0.21	0.21***	0.21***	0.21***	0.37***	0.21***	0.21***	0.21*
Foreign owned	0.24*	0.24***	0.24	0.24**	-0.01	0.24	0.24	0.24**
Scale-up in period before	0.08	0.08	0.08	0.08**	0.03	0.08	0.08*	0.08
Employment year before	0.03	0.03	0.03**	0.03	-0.02	0.03	0.03	0.03
Real turnover year before	0.03**	0.03	0.03***	0.03**	0.00	0.03	0.03	0.03
Employment growth 1 year prior	0.14	0.14**	0.14	0.14	0.07**	0.14**	0.14	0.14***
Turnover growth 1 year prior	0.00	0.00	0.00	0.00	0.03*	0.00	0.00	0.00
BERD sample period before	0.39***	0.39***	0.39***	0.39***	0.58***	0.39***	0.39***	0.39***
Innovate UK project before	1.45***	1.45***	1.45***	1.45***	1.58***	1.45***	1.45***	1.45***
Constant	-4.18***	-4.18***	-4.18***	-4.18***	-2.86***	-4.18***	-4.18***	-4.18***

Table C2.3: Selection Model II probits for single treated 2009-2016

	2009	2010	2011	2012	2013	2014	2015	2016
Observations	620,406	736,818	576,055	561,313	567,175	575,696	584,469	608,978
Adj R-Squared	0.10	0.12	0.13	0.10	0.13	0.12	0.12	0.09
Treated	195	184	236	302	322	281	384	254
Patent owner prior	0.28***	0.25***	0.23***	0.33***	0.48***	0.39***	0.54***	0.35***
Trade mark owner prior	0.03	0.15	0.22*	0.25***	-0.05	0.09	-0.17	0.03
High Knowledge Intensive	0.24***	0.43***	0.63***	0.44***	0.36***	0.53***	0.44***	0.42***
R&D in three years before	0.00	0.00	-0.03**	0.01	0.02	0.03*	0.02	0.01
High KI Manufacturing	0.19	0.58***	0.58***	0.53***	0.52***	0.74***	0.47***	0.36***
Foreign owned	0.17**	0.23***	0.02	0.26***	0.18***	0.01	0.11*	0.18***
Scale-up in period before	-0.17	0.10	-0.31*	0.05	-0.04	-0.13	0.04	-0.08
Employment year before	-0.06*	0.03	0.01	-0.03	-0.05	-0.02	-0.02	-0.02
Real turnover year before	-0.03	0.03	-0.03*	-0.02	-0.01	0.00	0.02	-0.03*
Employment growth period before	0.10*	0.05	-0.03	0.00	0.08*	0.02	0.03	0.05
TO growth year before	0.00	-0.01	0.01	0.03	0.03	-0.01	-0.01	0.03
BERD sample period before	0.54***	0.54***	0.71***	0.56***	0.65***	0.54***	0.45***	0.46***
Constant	-3.11***	-3.68***	-3.23***	-3.06***	-2.92***	-3.37***	-3.43***	-3.01***

Table C3 presents results from the Rosembaum bounds tests. PSM tries to mimic a randomized controlled trial. It asks: What would have happened to a particular business had it not received treatment? Therefore, it matches treated and untreated businesses that are similar on all observable characteristics, so that both have the same probability of being treated.

However, the analysis may be biased if there are unobservable variables that drive selection into treatment as well as the outcome, employment growth. One approach to test for the potential impact of unobserved variables is the Rosenbaum-bounds method. It assesses “how strongly an unmeasured confounding variable must affect selection into treatment in order to undermine the conclusions about causal effects from matching analysis” (DiPrete and Gangl, 2004).

Different levels of hidden bias can be expressed in terms of the odds ratio, gamma (Γ), of two matched observations being treated. If matching is unbiased, observations with the same observable characteristics have the same probability of being treated. When Γ=2, an unmeasured confounding variable causes one observation to be twice as likely to be selected into treatment as the matched observation with the same observable characteristics (Peel and Makepeace, 2009).

The method then does the following. It assumes that there is a known factor causing bias to the level of Γ, and that the treatment effect from this bias can be stripped out. Once this is done, it is tested whether the treatment effect remains significant. In this fashion, starting with zero bias, the treatment effect can be computed, and the assumption of ever larger bias tested.

The panels of Table C3 give the Rosenbaum bounds estimation for different matching models used in this study. For different levels of Γ, it gives the upper- and lower-point estimates of the treatment effects, under the assumption of negative and positive selection bias, respectively. It also gives significance levels for these estimates under the null hypothesis that the true treatment effect is zero at a certain level of positive or negative bias. The upper- and lower-point estimates can be interpreted in terms of a – usually – increasing cone of possible values as Γ rises. Where the cone begins to include zero, this is the level of bias where results are no longer robust.

The results for 2009 model I indicate the level of bias would have to quite high before the estimated firm level employment growth, used as the outcome variable, would be significantly different from that estimated were no bias assumed (0.10 when no bias is assumed, remaining 0.01 and 0.22 and positive as Γ increases). This suggests the 2009 results are robust, with Model II results being similar.

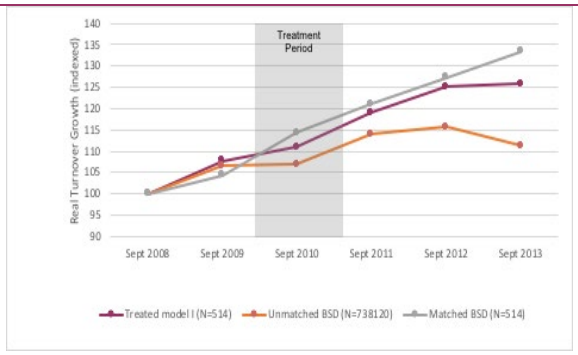
The results for other years are less clear cut. At high levels of bias, the PSM matches would give different results with regard to employment growth comparisons between the control and treated groups.

Table C3: Rosenbaum Bounds Test for Innovate UK Selection Models

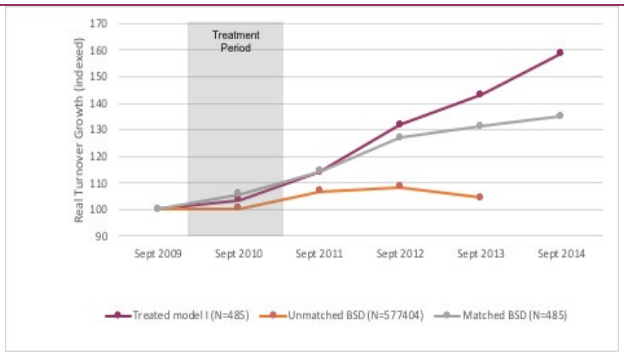
2009					**2010**				
Model I					**Model I**				
	Significance of selection effect		**Point estimate**			**Significance of selection effect**		**Point estimate**	
Gamma	Positive	Negative	Upper	Lower	Gamma	Positive	Negative	Upper	Lower
1	0.00	0.00	0.10	0.10	1	0.00	0.00	0.06	0.06
1.1	0.00	0.00	0.08	0.13	1.1	0.02	0.00	0.04	0.08
1.2	0.00	0.00	0.06	0.15	1.2	0.12	0.00	0.02	0.10
1.3	0.02	0.00	0.04	0.17	1.3	0.36	0.00	0.01	0.12
1.4	0.10	0.00	0.03	0.19	1.4	0.66	0.00	-0.01	0.13
1.5	0.28	0.00	0.01	0.20	1.5	0.87	0.00	-0.02	0.15
Model II					**Model II**				
1	0.00	0.00	0.12	0.12	1	0.00	0.00	0.09	0.09
1.1	0.00	0.00	0.10	0.14	1.1	0.00	0.00	0.06	0.11
1.2	0.00	0.00	0.08	0.16	1.2	0.01	0.00	0.04	0.13
1.3	0.00	0.00	0.06	0.18	1.3	0.08	0.00	0.03	0.15
1.4	0.02	0.00	0.04	0.20	1.4	0.25	0.00	0.01	0.16
1.5	0.10	0.00	0.03	0.22	1.5	0.52	0.00	0.00	0.18
2011					**2012**				
Model I					**Model I**				
1	0.00	0.00	0.08	0.08	1	0.00	0.00	0.05	0.05
1.1	0.00	0.00	0.06	0.10	1.1	0.04	0.00	0.03	0.07
1.2	0.01	0.00	0.04	0.12	1.2	0.30	0.00	0.01	0.09
1.3	0.08	0.00	0.02	0.14	1.3	0.71	0.00	-0.01	0.11
1.4	0.29	0.00	0.01	0.16	1.4	0.94	0.00	-0.02	0.13
1.5	0.59	0.00	0.00	0.17	1.5	0.99	0.00	-0.04	0.15
Model II					**Model II**				
1	0.00	0.00	0.06	0.06	1	0.00	0.00	0.09	0.09
1.1	0.00	0.00	0.05	0.09	1.1	0.00	0.00	0.06	0.11
1.2	0.05	0.00	0.03	0.10	1.2	0.00	0.00	0.04	0.13
1.3	0.24	0.00	0.01	0.12	1.3	0.05	0.00	0.03	0.15
1.4	0.56	0.00	0.00	0.14	1.4	0.25	0.00	0.01	0.17
1.5	0.82	0.00	-0.01	0.16	1.5	0.59	0.00	0.00	0.19
2013					**2014**				
Model I					**Model I**				
1.1	0.00	0.00	0.08	0.13	1.10	0.00	0.00	0.06	0.10
1.2	0.00	0.00	0.06	0.14	1.20	0.00	0.00	0.04	0.12
1.3	0.00	0.00	0.04	0.16	1.30	0.07	0.00	0.02	0.14
1.4	0.02	0.00	0.03	0.18	1.40	0.33	0.00	0.00	0.16
1.5	0.15	0.00	0.01	0.20	1.50	0.70	0.00	0.00	0.18
Model II					**Model II**				
1	0.00	0.00	0.08	0.08	1.00	0.00	0.00	0.06	0.06
1.1	0.00	0.00	0.06	0.11	1.10	0.00	0.00	0.04	0.08
1.2	0.00	0.00	0.04	0.13	1.20	0.03	0.00	0.02	0.10
1.3	0.03	0.00	0.03	0.15	1.30	0.21	0.00	0.00	0.12
1.4	0.22	0.00	0.01	0.17	1.40	0.59	0.00	0.00	0.14
1.5	0.59	0.00	0.00	0.19	1.50	0.89	0.00	-0.01	0.15

Annex D: Turnover Growth

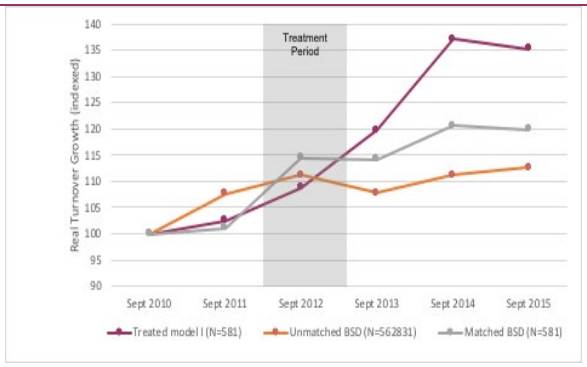
Figure D1: Real Turnover growth in Innovate UK supported businesses



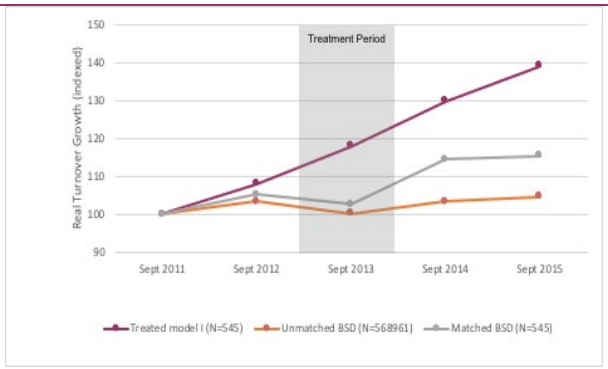
Treated in 2010



Treated in 2011



Treated in 2012



Treated in 2013

Annex E: Effect of business support on earnings

The Annual Survey of Hours and Earnings (ASHE) lends itself to conduct research on wages of supported businesses. ASHE is a dataset on individuals, randomly chosen by their National Insurance (NI) numbers. Surveys are completed by their employers from payroll data, so quality and response rates are high. Included are all employees registered for PAYE taxes whose NI numbers end in two specific digits. Therefore, it is a panel and individuals can be tracked even if they change employers. Variables collected include weekly earnings and hours worked in a reference period, occupation (4-digit SOC code), managerial and supervisory duties, place of work, gender, and age. The survey with business identifies is available for 2004-2016.

The measure of earnings used is gross weekly earnings deflated by the consumer price index. As ASHE is a 1% sample of all employees, there are many of the supported business included, however usually just with a few employees for each. While the individual wages available for an individual business may not be very representative of that business, the random nature of the survey means that averaging observed wages across businesses means this should be representative of the average in each of those businesses. In other words, the CEO in one business and the graduate trainee in another may be collected through ASHE and averaging across these businesses approximates the true average wage of each of those businesses.

To account for skewed wage distributions with outliers at the top, averages are calculated over weekly earnings in natural logarithms form, and these averages are then exponentiated again in the figures presented below, to help contextualising the numbers.

Tables E1 and E2 provide the sample sizes of individuals of working for businesses supported by Innovate UK and NMS respectively. The Innovate UK sample sizes include both employees at businesses that received only a voucher, as well as other support. In the remained of the analysis, voucher-only recipients will generally be excluded or reported separately. The NMS sample is further split into businesses that participated in collaborations or accessed paid services, and those that accessed downloads only. The latter group will be used as a comparison group, as it is likely that they perform activities similar to the other NMS supported businesses but did not receive any intensive support.

The tables give the average number of observations per business, the number of individual businesses, and the number of businesses with more than ten individual observations in each year. Businesses supported by Innovate UK are on average larger, so there are also more observations available on average. Both NMS and Innovate UK supported businesses are larger than other businesses represented on ASHE, where the average number of employees sampled is just over 2.

Table E1: Sample sizes for businesses supported by Innovate UK

	Observations per business		Number of businesses		Businesses with > 10 observations	
Year	IUK supported	ASHE	IUK supported	ASHE	IUK supported	ASHE
2004		2.0	<30	23675	<30	506
2005	42.0	2.9	121	45910	51	1518
2006	36.5	3.0	207	44653	80	1571
2007	20.3	2.8	146	39999	55	1342
2008	35.0	2.8	206	40138	60	1327
2009	45.7	3.0	232	46171	95	1585
2010	25.2	3.0	205	47771	66	1611
2011	40.2	3.0	195	50616	60	1689
2012	39.9	2.9	282	49164	85	1566
2013	26.0	2.9	334	49809	87	1610
2014	29.5	2.8	335	53085	83	1653
2015	28.8	2.6	375	55898	97	1598
2016	23.4	2.6	256	53911	65	1529

Table E2: Sample sizes for businesses supported by NMS

	Observations per business			Number of businesses			Businesses with > 10 observations		
Year	NMS supported	Downloads only	ASHE	NMS supported	Downloads only	ASHE	NMS supported	Downloads only	ASHE
2004	12.9		2.2	174		23839	40		539
2005	20.8		3.2	270	<30	46192	76	<30	1595
2006	22.8	12.2	3.2	213	59	44942	75	<30	1642
2007	16.8	12.2	3.1	206	45	40307	64	<30	1405
2008	13.7	16.9	3.0	237	86	40450	57	<30	1381
2009	21.3	24.6	3.2	239	99	46498	77	32	1661
2010	24.3	13.7	3.2	256	100	48121	78	<30	1686
2011	21.5	11.9	3.2	246	127	51011	74	<30	1769
2012	19.1	12.0	3.1	248	131	49570	71	<30	1654
2013	21.8	17.4	3.1	240	107	50240	74	<30	1703
2014	22.6	10.8	3.0	224	137	53556	56	<30	1737
2015	22.6	10.3	2.8	199	165	56386	57	<30	1679
2016	21.1	9.8	2.8	201	113	54451	53	<30	1607

ASHE also provides a breakdown of employment by actual place of work, rather than by the head office of the employer, which tends to lead to an over-representation of the South East. Table A.55 shows that employment of businesses supported by Innovate UK is fairly similar to the distribution of employment in general. There is a slightly lower representation of London, but slightly higher of the rest of the South East. Scotland and Wales are slightly underrepresented. However, it needs to be kept in mind that this still only represents the whole of the businesses supported by Innovate UK, not the specific branch or plant that

was allocated the support. However, the analysis gives a better picture of the true regional spread of the employment of those businesses.

Table E5: Regional employment share

Work region	Innovate UK supported	Voucher only	ASHE
North East	4%		4%
North West	12%	18%	11%
Yorkshire and the Humber	8%	14%	9%
East Midlands	8%		7%
West Midlands	9%	18%	9%
South West	9%		9%
East	10%	12%	9%
London	14%		15%
South East	14%	19%	13%
Wales	4%		5%
Scotland	8%	18%	10%
Total number of employees	315000	859	1662787

The picture is slightly different when looking at NMS supported businesses. Here, the South East is more dominant, and London more underrepresented. The dominance of the South East may also be related to the location of the NMS in the area, providing easy access to businesses. However, the Midlands are also slightly overrepresented, while there are less employees supported by NMS in Yorkshire and the Humber than in the general business population.

Table E6: Regional distribution of employees in NMS supported businesses

	NMS supported	Downloads only	ASHE
North East	5%	3%	4%
North West	12%	9%	11%
Yorkshire and the Humber	6%	10%	9%
East Midlands	10%	7%	7%
West Midlands	10%	7%	9%
South West	8%	7%	9%
East	9%	10%	9%
London	11%	19%	14%
South East	16%	16%	13%
Wales	5%	3%	5%
Scotland	8%	9%	9%
Total number of employees (all years)	123814	63944	1839612

Further results

The results above have the obvious shortcoming that they are comparing employees at Innovate UK and NMS supported businesses with a random sample of all employees. As seen in the summary statistics tables, employees at supported businesses are quite different from the general population. This section tries to mitigate this problem by controlling for observable and unobservable characteristics using regression analysis.

Table E7 shows regression results for NMS supported businesses. The dependent variable are log gross weekly earnings in constant GBP. The treatment variable, “NMS treated” is equal to 1 for all businesses that ever used NMS services or participated in a collaboration, from the year of their first collaboration. There is another NMS variable, “NMS downloads only”, that is equal to 1 for businesses that only accessed downloads from NMS. As this is a light-touch support measure, it will not be considered as a treatment here, but could give important information on industry affiliation with an interest in measurement services. Four different model specifications were estimated.

The first two specifications use Ordinary Least Squares (OLS) estimation. In the first of these two specifications, an individual’s skills are accounted for by two “high-skilled” and “medium-skilled” dummies (with low-skilled as the baseline), derived from occupation codes. In the second specification, these are replaced by more detailed O*NET variables. O*NET is a coding scheme that rates occupations on various physical and cognitive skills as well as the use and importance of different knowledge domains.

There are 70 variables, measuring both the level and importance of skills and knowledge. These were reduced into 26 variables using factor analysis. It is based on US survey data but can be mapped to UK Standard Occupation Classification (SOC) codes. However, the mapping is only available from 2010 onwards, so that a shorter time series is available for estimations using the O*NET data.

The second set of regressions uses fixed effects (FE) estimation. This exploits the panel structure of the wages, in that it includes a fixed effect for each individual. In these regressions, the effect of the NMS treatment is identified through changes, either a change from untreated to treated, or as an individual moves to a treated business. Again, there are two specifications, one with skill dummies and one with O*NET variables.

The results in Table E7 show a positive effect of the NMS treatment throughout. As the dependent variable is expressed in logarithms, these coefficients can be interpreted as percentage changes: The wage effect of NMS support is estimated to be between 3 and 7%.

Table E7: Multivariate estimation of the effect of NMS on wages

	OLS (1)	OLS (2)	FE (1)	FE (2)
NMS treated	0.05***	0.07***	0.03***	0.04***
NMS downloads only	0.03***	0.03***	0.01***	0.03***
Labour force experience	0.01***	0.01***	0.01***	0.00***
Age	0.05***	0.04***	0.07***	0.08***
Age sq.	-0.00***	-0.00***	-0.00***	-0.00***
Full time	0.80***	0.74***	0.62***	0.55***
High-skilled	0.62***		0.16***	
Medium-skilled	0.17***		0.06***	
Low pay	-0.14***	-0.11***	-0.08***	-0.08***
High tech	0.04***	0.01***	0.04***	0.00
Manufacturing	-0.00	0.01***	0.02***	0.04***
Scale up	0.02***	0.03***	0.01***	0.01***
Knowledge intensive	0.03***	0.06***	-0.00	0.01*
Constant	3.86***	4.30***	3.74***	3.62***
O*NET	No	Yes	No	Yes
Region effects	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes
R-sq.	0.67	0.70	0.40	0.37
adj. R-sq.	0.67	0.70	0.40	0.37
within R-sq.			0.40	0.37
overall R-sq.			0.59	0.60
between R-sq.			0.66	0.65
N	1754948	978771	1754948	978771

Employment and Turnover categories highly significant across models; these and some other variables suppressed for brevity

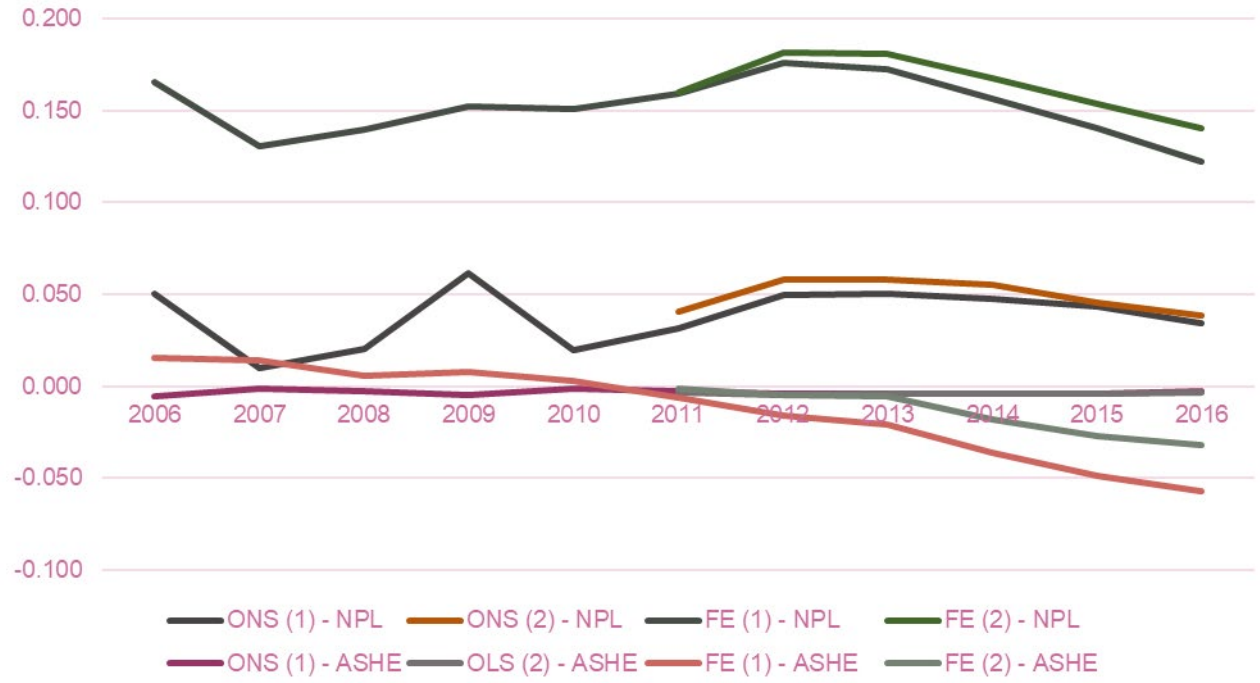
Note: Asterisks on coefficients indicate the statistical significance: * for p<.05, ** for p<.01, and *** for p<.001. Selected models also include O*NET variables, region and year effects, where individual coefficients are not displayed, but their presence is indicated at the bottom of the table.

Another way to analyse these results is to estimate the regressions without the treatment dummy and compare the regression residuals, i.e. the variation in wages not explained by the regression, across different groups. This is particularly useful for comparing the effect over time.

Figure E1 shows the regression residuals for employees in NMS supported and unsupported businesses between 2006 and 2016. For the regressions including O*NET data, the available timeseries is again shorter. The figure shows a clear earnings premium for employees working in NMS supported businesses. The fixed effects results are especially powerful. Controlling for observable factors such as occupation, and experience, as well as unobservable factors that are constant over time, employees at NMS supported businesses earn about 15% more than would be expected when comparing them to the wider employee population in the UK.

However, it has to be caveated that only individual fixed effects are included here. There may still be considerable selection bias relating to the business that chooses to work with NMS. On the business side, only employment, turnover, and some industry characteristics are included. Factors such as innovation and management motivation that might both influence using NMS services as well as employee remuneration are not included.

Figure E1: Regression residuals for NMS supported and unsupported businesses



Note: Model numbers corresponding to Table 7. The variable on the y-axis measures the divergence between actual and expected earnings based on the regression results.

Table E8 and Figure E2 repeat the same analysis for businesses supported by Innovate UK. As for the NMS data, four different model specifications are estimated. Also, the treatment variable is defined to be equal to 1 for businesses from their first treatment onwards. Therefore, the wage premiums can be read as pertaining to businesses that have ever been supported by Innovate UK. Depending on how long ago the support was received, one should be cautious in attributing the effect directly to the support. In the regression analysis in Table E8, the Innovate UK support is estimated to have a positive effect on earnings. While the effect is slightly smaller in magnitude than for NMS, it is much larger than in the simple analysis presented above. This is likely the case because Innovate UK supported businesses tend to be younger and smaller – both characteristics associated with lower wages.

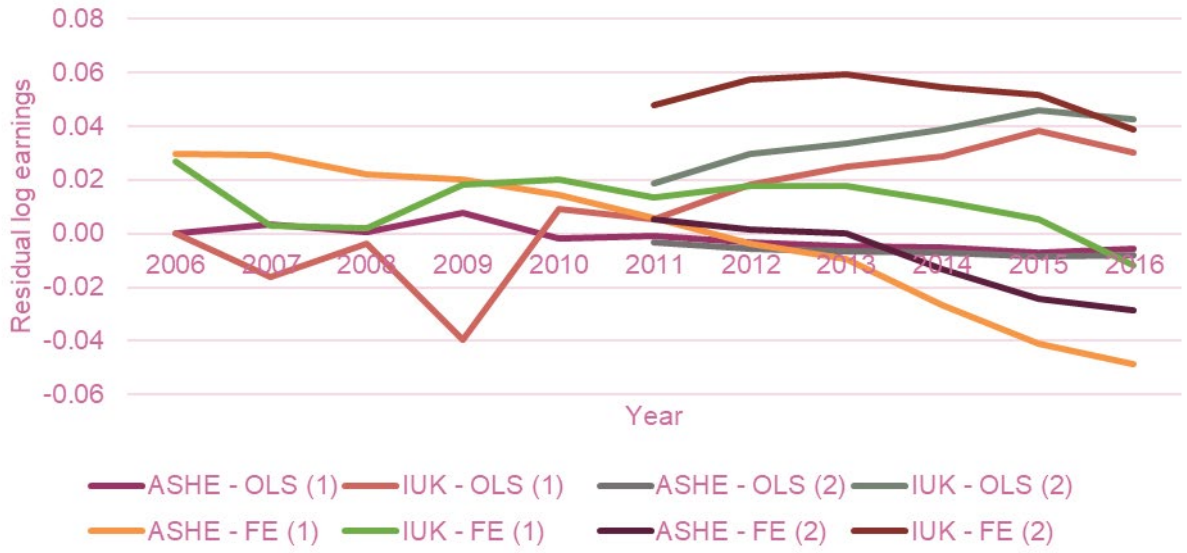
Table E8: Regression analysis of Innovate UK treatment effect

	OLS (1)	OLS (2)	Fixed effects (1)	Fixed effects (2)
IUK treated	0.02***	0.05***	0.00*	0.01**
	(0.00)	(0.00)	(0.00)	(0.00)
Labour force experience	0.01***	0.01***	0.01***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)
Age	0.05***	0.04***	0.07***	0.08***
	(0.00)	(0.00)	(0.00)	(0.00)
Age sq.	-0.00***	-0.00***	-0.00***	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)
Full time	0.80***	0.74***	0.62***	0.55***
	(0.00)	(0.00)	(0.00)	(0.00)
Female	-0.16***	-0.13***	-0.04**	-0.05
	(0.00)	(0.00)	(0.01)	(0.04)
High-skilled	0.62***		0.16***	
	(0.00)		(0.00)	
Medium-skilled	0.17***		0.06***	
	(0.00)		(0.00)	
Employment categories		Sig and +ive	(generally)	
Turnover categories		Sig and +ive	(generally)	
	(0.00)	(0.00)	(0.00)	(0.00)
Low pay	-0.14***	-0.11***	-0.08***	-0.08***
	(0.00)	(0.00)	(0.00)	(0.00)
High tech	0.05***	0.02***	0.04***	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Manufacturing	0.00***	0.01***	0.02***	0.04***
	(0.00)	(0.00)	(0.00)	(0.00)
Scale up	0.02***	0.03***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)
Kn'ledge intensive	0.03***	0.06***	-0.00	0.01*
	(0.00)	(0.00)	(0.00)	(0.01)
Constant	3.86***	4.30***	3.74***	3.62***
	(0.01)	(0.01)	(0.07)	(0.08)
ONET	No	Yes	No	Yes
Region effects	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes
R-sq.	0.67	0.70	0.39	0.37
adj. R-sq.	0.67	0.70	0.39	0.37
within R-sq.			0.39	0.37
overall R-sq.			0.59	0.60
between R-sq.			0.66	0.65
N	1754948	978771	1754948	978771

Note: Dependent variable: Log gross weekly earnings in constant GBP. Treated equal to 1 for all businesses that ever received IUK support from the first year of treatment. Doesn't include businesses that ever only received a voucher. Observations with weekly earnings below £50 and above £10000 excluded. Robust standard errors in parentheses.

In the residual analysis in Figure E2, the picture is more mixed. The graph shows a wage premium only in more recent years, but this has also been falling. One reason behind the inconclusive results could be that the mix of supported businesses changed over the years, and that the effect depends on the type and size of support.

Figure E2: Regression residuals for Innovate UK supported and unsupported businesses



Note: Model numbers corresponding to Table 8. The variable on the y-axis measures the divergence between actual and expected earnings based on the regression results