Department for Business Innovation & Skills

RESEARCH

BIS RESEARCH PAPER NUMBER 295

Exploring the feasibility of a productivity based approach for evaluating business support interventions

JULY 2016

Note to readers

The Department for Business, Innovation and Skills (BIS) has made significant strides to improve the robustness of its policy evaluations, (an example of which is the Growth Vouchers Randomised Controlled Trial). These evaluations will provide much stronger evidence on the impacts of programmes on those directly supported. This paper explores whether improvements might also be made when taking into account the wider economic impact on growth (displacement). For example, the extent to which those businesses supported grow at the direct expense of rival and more productive businesses. Alternatively, the extent to which growth in supported businesses leads to a more productive use of labour in the economy.

The main focus on this paper is to explore whether a 'productivity decomposition' approach can be used to account for displacement. It does this within a quasi-experimental context, where what would have happened without support is modelled through identifying a counterfactual.

This study finds that this approach can be used to provide further insights about the wider economy impacts from business support interventions. However, this is a first step in seeking to address this problem, rather than the definitive solution, so the following methodological issues should be noted:

- The decomposition is sensitive to how the counterfactual group is identified.

- To establish the reallocative effect the restriction that the treatment group is closely matched to the control group has been lifted, so the conditional independence assumption has been relaxed. This is because the matching may pick the businesses adversely affected by displacement (overstating support impact), or any reallocation towards more productive supported businesses may be missed, (understating impacts).

- The paper is testing the application, benefits, and limitations of this approach. Data from a real policy has been used to conduct this, but this paper is not undertaking a policy evaluation. The application is much simplified than would be the case for an evaluation. For example, the productivity decomposition does not take into account the relative scale of the intervention for each business and the process of identifying a counterfactual is simplified, using fewer variables than would be the case in an evaluation. However, the application does explore the main methodological assumptions and limitations.

Given the innovative and experimental nature of this study, BIS welcomes comments and feedback from analysts and evaluation practitioners on this paper and how it can be taken forward.

Readers wishing to respond to this research paper should send their comments to enterprise.analysis@bis.gsi.gov.uk by 1st October 2016.

Contents

Authors:

Jonas Meldgaard, Carolin Thol and Prabhat Vaze, Belmana

Ben Davies, IFF Research

Dr Michela Vecchi, Middlesex University

Dr Felix Ritchie and Prof Don Weber, University of the West of England

Acknowledgements:

The authors acknowledge helpful comments throughout this study from the BIS project team: James Phipps, Stephen Fernando, Chris Wright and Adam Johnson. There were some very helpful suggestions made at two workshops at BIS. We are also grateful for comments from anonymous referees and the analytical teams at BIS.

Executive Summary

1. The focus of Government economic policy is on improving UK growth and productivity. This often includes providing business support to overcome specific market barriers and so improving the productivity of supported businesses. However, there have been concerns about the extent to which such interventions generate additional indirect, economic impacts, some potentially negative, compared to a situation where there was no intervention. This is because firms that receive such support can benefit at the expense of other firms, for example, as a result of displacement, or the reallocation of productive resources from other firms to the firms supported by the intervention. Yet, it should be noted that there can also be indirect effects that have positive impacts: spillovers from innovation or the clustering of economic activity can provide benefits to firms beyond those directly supported.

2. While the application of current evaluation methodology can lead to accurate estimates of firm level impacts of government intervention (through measuring changes in scale of the firm such as employment or turnover), there has been uncertainty about how to accommodate wider-economy effects such as displacement and spillovers. This raises questions about whether the overall impact of business support interventions is estimated.

3. This paper focuses on the quantitative analysis of displacement. Approaches to evaluating spillovers are also discussed, but only through a qualitative review.

4. The research has sought to address the problem of accurately estimating reallocative effects. The approach taken is innovative: assessing the feasibility of using a productivity decomposition approach in an evaluation context. The research should be considered as a first step in applying a productivity based approach in evaluations and interested academics and analysts are invited to consider the analysis contained within this paper and provide their comments and suggestions to BIS at enterprise.analysis@bis.gsi.gov.uk.

5. Productivity decomposition is an approach that unpacks productivity change by quantitatively assessing displacement in economic activity and whether, through this, there are further productivity changes. Reallocation could see businesses with a productivity advantage capturing market share and resources, further improving overall productivity growth or, more worryingly, resources may move away from the more productive businesses – because of support – implying a negative wider economy impact.

6. Displacement is the term used by evaluators to describe activity increases in supported businesses that arise from activity removed from non-beneficiary businesses. Productivity decomposition would estimate beyond the change in economic activity, particularly how much reallocation contributes to or reduces the overall productivity change.

7. The technique is tested by using a quasi-experimental approach as an illustration. The productivity change in a supported business is compared with a robustly identified control or counterfactual (propensity score matching is used for selection). The decomposition should arrive at the same estimate of net change in productivity across

a treated and a matched control group, the so-called within firm effect as would be estimated using current evaluation methods. Productivity decompositions then additionally shed light on the dynamics accompanying the within-firm effect, as reallocation, including entry/exit, provide further, wider productivity effects.

8. This research is primarily testing feasibility and there are some methodological steps that may be a focus for comment.

9. Overall, the study relies on the Foster-Haltiwanger-Kirzan (FHK) decomposition method; some use is also made of the Griliches-Regev (GR) decomposition method. Both are widely used and understood methods. FHK is relied on because it is most precise in identifying the within firm effects and it analyses any impacts due to entry and exit of businesses. The data analysed in this research suggests the reallocation of resources can be a significant component of overall productivity change. The report also critically evaluates the different productivity decomposition techniques and how to decide on which technique to use in different evaluations.

10. The study then explores how decompositions, combined with careful identification of the counterfactual, can illuminate any wider economy impacts of the support. The innovation is to vary the identification of control groups by the likelihood that wider economy impacts might be material. This is then followed by using the productivity decomposition on the treated and each of the control groups and exploring what the drivers for productivity are in supported businesses and different sets of comparable businesses. The quasi-experimental approach to decomposing productivity changes suggests evidence of reallocation impacts, beyond the within firm effects. Further, these vary in a manner consistent with the likelihood that a business in the control group may be the business affected by displacement. This suggests evaluations can take account of wider economy impacts through displacement.

11. The research is underpinned by a practical application of productivity decomposition, using data on a current business support intervention known as the Enterprise Finance Guarantee (EFG) scheme. Productivity performance is estimated using the administrative data from business returns to government, primarily for business taxes. **It should be noted that the results from this analysis are not an evaluation of the scheme and should not be treated as such, but as an experiment to test the productivity decomposition approach.**

12. This study **represents a first step in using such a productivity decomposition approach** to evaluate impacts of business support. It indicates the feasibility of using the approach in an evaluation, describing the stages involved in implementing it. There are three conditions for the approach to be feasible, largely in the first two preparatory stages, before then operationalising the decomposition.

13. The first condition is with regard to the type of intervention. The study used an intervention which was relatively similar across the supported businesses and offered on a scale sufficient to allow analysis. Where a support measure is customised for individual businesses, then the ability to use this approach satisfactorily is quite low, as it is difficult to measure impacts accurately, primarily because the data available would limit identifying a control group. Interventions targeting specific sectors or types of business usually anticipate economy wide impacts through relatively complex channels, including spillovers. The decomposition cannot estimate this.

14. The second is around the data available for the study. The decomposition requires the construction of a dataset for the businesses supported and a matched control group. The data should provide a time-series of firm-level productivity, measured in a consistent way for both beneficiaries of the support measure and comparable unsupported businesses. In this study, administrative data is used. Such data is becoming increasingly available for analysis and its key advantage is its comprehensiveness. However, the data has a limited number of variables. This research had to focus on labour productivity decompositions, due to the limitations of the available data. Reallocation of other inputs into production, such as intermediate inputs including energy, the use of plant and machinery or other assets, could be significant, but would require considerable investment in data.

15. A significant third factor for this approach is determining the control group and – in particular – the variables available to identify comparable businesses. To some extent, this is a feasibility issue for any evaluation. However, because of the other two conditions, the requirement becomes more constraining. In particular, the need for time series data for both supported businesses and a wider pool of non-beneficiaries may mean the selection of a control group could become infeasible.

16. Hence, because of these limitations, it should be noted that any results from the productivity decomposition in this paper are certainly not any attempt at an evaluation of this government scheme and should not be treated as such.

17. Throughout the study, some limitations - common to many government and wider evaluations - have had to be overcome, such as whether statistical matching can identify a suitable control group. While these issues are not the focus of the study, there are some aspects which are particularly relevant should decomposition be used. A particular area discussed is the extent to which those unsupported businesses that are selected as comparable may be the ones affected by displacement of economic activity, possibly negatively. The report finds this to be the case and – while this is consistent with the presence of wider economy impacts – it means the statistical matching undertaken has to be accompanied by rigorous sensitivity testing.

18. The report makes most use of the FHK decomposition method and some use of GR. These are the most commonly used techniques and so benefit from being relatively well understood. They both seek to model entry and exit and reallocation effects.

19. The modelled estimates of employment impacts in businesses supported by EFG between 2009 and 2014 (noting the caveats outlined in paragraphs 13 – 16) suggest that the supported businesses have a very different change in resources in comparison with the derived control group. This suggests that a productivity decomposition would help to disentangle overall effects on productivity.

20. Furthermore, supported businesses appear to have grown more in employment than the comparison group. In 2014, our estimates suggest that both groups of businesses have similar employment levels, at just over 20 employees. However, businesses that were in existence in 2009 and who received support at some point in 2009-14 had two fewer jobs than the control group in 2009. Overall the additional jobs for beneficiaries were estimated to be approximately 3.1 per firm or a growth of 15%. Secondly, net entry in the supported businesses is also higher. This is primarily because the exit rate in the counterfactual group is three times greater than among the beneficiaries.

Overall, employment growth is around 44% in supported businesses. The figure also suggests entry, which is similar across the two groups.

21. The results of the quasi-experimental decomposition, again noting the caveats, are discussed in the report. They are not evidence of the impacts of EFG. Rather, they demonstrate the analyses that a productivity decomposition produces. The report presents estimates of the within firm impact, entry and exit and reallocation, termed the between effect. The estimates use the FHK method. Results separate between supported and the control group and estimation uses employment for weighting firms together.

22. The decomposition suggests that supported businesses have a productivity growth approximately 2% higher than the control: this is the within effect. The deadweight – that part of the change that would have occurred without treatment – is estimated by the "within" term for the control group, about 4% productivity growth. However, reallocation dominates the within firm productivity growth, reducing productivity by 6% with all but one point being attributable to the between effect in the control group.

23. The between effect shows whether firms capitalise on any productivity advantage they have during the period. The negative terms suggest that firms with low productivity performance increased employment, while better performing firms reduced employment. The between term measures changes in employment as being inversely correlated with productivity growth and/or levels. In these estimates, the results suggest that this was more pronounced for the beneficiaries of support.

24. This is surprising but it is likely to be a mix of measurement problems, alongside actual economic impacts. In economic terms, these estimates may reflect only the medium-term effects. Business support translates into an employment expansion first, reducing labour productivity. These estimates also indicate higher survival in supported businesses, with this outcome generally being considered an early impact of support. Further productivity improvements may then follow in the long-term, especially if the within firm improvement seen in the analysis continues for successive years. Put another way, the dynamics within the supported firms of rising productivity may not have yet fully played out. Reallocation may reflect different levels of labour hoarding, a more short-term reallocative impact.

25. On the measurement issues, the three feasibility constraints described earlier are more likely to affect the estimates of reallocation. While the intervention selected was relatively similar across businesses and of a scale appropriate to the productivity decomposition, the results from this analysis are not an evaluation of the scheme and should not be treated as such. In addition to the earlier constraints, an actual evaluation would analyse the scheme monitoring data to check carefully any patterns in the size, nature and timing of support. It would also look at the logic model to ensure that the impacts are evaluated in a manner consistent with the theory of change for the support. The construction of the control groups, using a propensity score model, may not be selecting firms that are fully aligned or strictly comparable firms to those who received EFG support.

26. The estimates suggest that businesses did see productivity growth, shown by the within effect, but those firms that saw the most growth either did not recruit or shed employment. In order to explore estimating displacement more accurately, the analysis then takes a further set of steps by relaxing the comparability of the control group in a

manner consistent with the likelihood of reallocation. This is because the matching may pick the businesses adversely affected by displacement (overstating support impact) or any reallocation towards more productive supported businesses may be missed (understating impacts). A fundamental question for researchers seeking to improve this approach is the extent to which this is appropriate, given that this has implications for the conditional independence assumption.

27. Some initial findings are presented as the matching assumptions are changed. Broadly, the more comparable the businesses, the higher the treatment impact within the beneficiary businesses. These results have some caveats, but were such findings found to be robust they might reflect displacement – that the most comparable businesses can be detrimentally affected by the support. The logic for this observation would be that, in comparing with businesses that are very similar to the supported businesses, the treatment effect is more likely to include the negative effects on businesses competing with supported businesses.

28. This study is primarily a feasibility study. A key question is whether using decompositions is improving on present measures of displacement: is the productivity decomposition approach better at accounting for displacement than adjusting for market displacement, as is currently done, or than simply assuming full factor (employment) market displacement. The analysis presented in the report serves to illustrate the application of the decomposition to a government support scheme. They indicate the types of analyses that are possible and how they might be interpreted, but further work needs to be done to assure their robustness.

1. Introduction

1. Evaluating the additional impacts on the economy of government support to business has been a focus of researchers and policy-makers for many years. The government's programmes, policies and projects aim to have a net positive economic impact and evaluation is used to establish whether an intervention is successful or whether the supposed effect would have happened anyway or has proven negative (BIS, 2011b). Evaluation usually involves comparing supported businesses to the counterfactual where there was no intervention.

2. Currently, evaluations largely focus on estimating firm level impacts through changes in the scale of the firm (employment or turnover). Support aims to improve the performance of beneficiaries relative to comparable unsupported businesses. There are then methods to measure these direct, additional impacts on productivity growth in the supported businesses. Such analyses provide a good starting point, but may not say much about effects beyond the beneficiary firm as firm growth feeds into aggregate economic growth. BIS (2011b) reviews such wider impacts, identifying that "the main indirect impacts on firms outside of the treatment group come through displacement and spillovers" (p. 59)

3. This research considers these two economy-wide impacts, caused by:

    - **Reallocation of resources**. A first issue is whether the outcome of support is the beneficiary effectively taking resources or market share from competitors; a second is whether those resources originate from more productive firms thereby taking away from any benefits achieved in the supported business. The study looks at whether productivity decompositions can help in estimating these effects.

    - **Spillovers.** Such impacts arise through interventions that promote innovation, co-location and clustering and foreign direct investment. Recent Nesta work has explored the spillovers from patents and the effect of high growth firms to enhance the propensity of other firms to innovate (Sena, Hart and Bonner, 2013). Mason et al. (2009) used a dataset of 45 UK city-regions and found spillovers from the high growth firms on local employment. A key feature of externalities is the richness of methods used to investigate these impacts.

4. This report assesses the feasibility of using an approach to measure the productivity impact of an intervention rather than the change in overall scale of the firm. The study applies productivity decompositions to understand the reallocation dynamics. In particular, the decomposition may begin to answer the questions where reallocation occurs and whether it is productivity improving or not. If not, then this could mean that support has merely displaced economic activity.

5. The report considers externalities due to spillovers, such as from innovation and foreign direct investment and the productivity improvements as spatial interventions deliver agglomeration/co-location and geographical externalities. This study reviews common methods and shows that the approaches are more diverse. The rationale for this diversity is set out and the report offers advice about how an evaluation may navigate through the methodological options.

Evaluation Approaches and Focusing on the Firm-level Impact

6. Much of this report looks at the feasibility of applying productivity decompositions in evaluations. Existing evaluation starts by looking at an intervention's impact on supported businesses in comparison to businesses not receiving support. Evaluation, especially soon after the intervention has occurred, necessarily focuses on what has happened to the businesses targeted by a policy.

7. Business support interventions are quickly followed by "outputs", such as investments in plant or research, specific jobs that are created or safeguarded by the businesses. These are outputs in the intervention's sense, often monitored by the intervention provider. Then over the medium term more profound changes might be observed at the business level. These may be improved sales or entry into a new market. There may be productivity improvements. Evaluators carefully assess whether these are additional, taking account of impacts that would have happened anyway ("deadweight") or when those outside of the targeted businesses have benefitted ("leakage").

8. In focusing on the within firm impact, evaluators may miss "the possibility that assisting one business may have adverse effects on other, non-assisted businesses" (SQW, 2009, p.1). Decision-making expects to take account of all impacts, direct or indirect, negative or positive. But SQW (2009) note that a focus on the direct impact may mean missing significant other indirect effects. Displacement is perhaps the easiest example of why this may bias evaluations. If an intervention raises the output of the supported by taking activity away from other firms, any focus on supported businesses will overstate impact. Some indirect effects are positive, so that missing them would understate impacts, such as spillovers where a technological innovation consequent of support may be adopted in other businesses raising their productivity.

Challenges of Evaluating the Wider Economy Impacts

9. Assessing these wider impacts presents some challenges.

10. A first challenge is to refine the concepts used for wider impacts, particularly displacement. Displacement is about the market context in which businesses are supported. The support may give a business an advantage in markets or cause it to increase demand for inputs. However, market conditions may not be suitable for this. Supply conditions, such as the availability of skilled labour, may need improving. Increased market advantage in itself cannot be considered an additional and beneficial impact if it merely increases the demand for factors of production. The firm then steals market share and displacement occurs.

11. A key refinement to this is that displacement may not always counter the direct beneficial impact a firm receives due to support. If businesses that are supported are more productive than the unsupported businesses from which resources are displaced, then productivity can increase overall through displacement (Box 1).

12. The second challenge in evaluating the economy-wide is estimating these indirect effects. One approach that has been widely used is productivity decomposition. This splits overall productivity growth, separating growth that arises within firms from the reallocation of resources across producers and from entry and exit. There are now a number of UK studies using this technique (Disney et al, 2003; Mason et al, 2014) and some have explored its use as an evaluation method (Harris and Robinson, 2005).

Box 1: Reallocation of Resources

In the economic literature, the movement of factors of production or market share amongst businesses is called reallocation (Syverson, 2011). Evidence about this phenomenon shows that reallocation may sometimes affect productivity positively. In imperfectly competitive settings, overall productivity rises merely by moving resources between firms to the more productive ones (Hall, 1990, and Basu and Fernald, 2002). This suggests that displacement is only part of the story and that, where market frictions exist, there may be further productivity impacts purely due to reallocation of resources between firms with differences in productivity performance.

Syverson covers a range of circumstances where this reallocative effect has been estimated. The estimation distinguishes between productivity growth arising from "within" (productivity growth at a given plant or firm) and "between" (reallocation of resources across existing businesses or entry and exit) sources. Such work has drawn upon plant-level data, which has demonstrated the wide diversity of productivity levels in populations of businesses and, secondly, the dynamics of reallocation. For example, US evidence on retail suggests the primacy of reallocation in contributing to productivity over the within effects suggesting that productivity growth at store-level is modest. It is entry of more productive and exit of the less that drives overall productivity growth.

13. A final set of challenges is around the additional data needed for evaluating wider economy impacts. A key feature of any data is whether it can support analysis beyond the within firm effects. This is often over time: Drews and Hart (2014) stress the importance of taking account of time lags between an intervention and its impact. Further, they note that analysis has to look beyond the beneficiary businesses to the wider population of businesses. Both aspects have meant analysts turning to administrative data. Such data provide a long time series about business outcomes and allow analysis across the whole business population.

Questions Emerging from this Research

14. The report presents a stylised set of 'Stages' to evaluating the economy-wide impacts of an intervention and suggests where current evaluations can incorporate productivity decompositions. It highlights how, especially over the long-term, where the logic model of an intervention predicts significant reallocative impacts due to interventions, there is a role for decomposition. The challenges above have been addressed to some extent, but this raises some questions.

15. Displacement focuses on the moving of resources or market share between firms. While the supported firms can be identified, the firms from which they may displace economic activity is more open. The approach here is to match comparable businesses to the supported businesses. The treatment-control logic focuses on businesses that receive support and those that are identical on average except for the support. The approach then varies the comparison to identify evidence for displacement. It will not identify which employee moved or how a particular product market was captured, but may provide the averages for a group of businesses of interest. A key question is the extent to which this is appropriate.

16. Furthermore, within this general construct, the businesses need to be tracked over time, long enough for displacement and reallocative effects to be analysed. This means a panel, focused on the treated and the matched control, has been used. This decomposes the productivity performance, including entry and exit, of a small part of the universe of businesses. The analysis allows estimation of the within firm effect. A question is whether the observed reallocation in this universe represents the indirect impacts.

17. A final question is whether using decompositions improves on present measures of displacement: is the productivity decomposition approach better at accounting for displacement than adjusting for market displacement as is currently done?

18. The study seeks to cover implementation issues throughout, using an example policy to provide practical applications. In looking at productivity decompositions, the strengths and weaknesses of different decompositions are explored. Then, in the final chapter of the study, wider, indirect impacts that would not usually be estimated through a decomposition, namely the spillovers are considered.

2. Wider Economy Approach

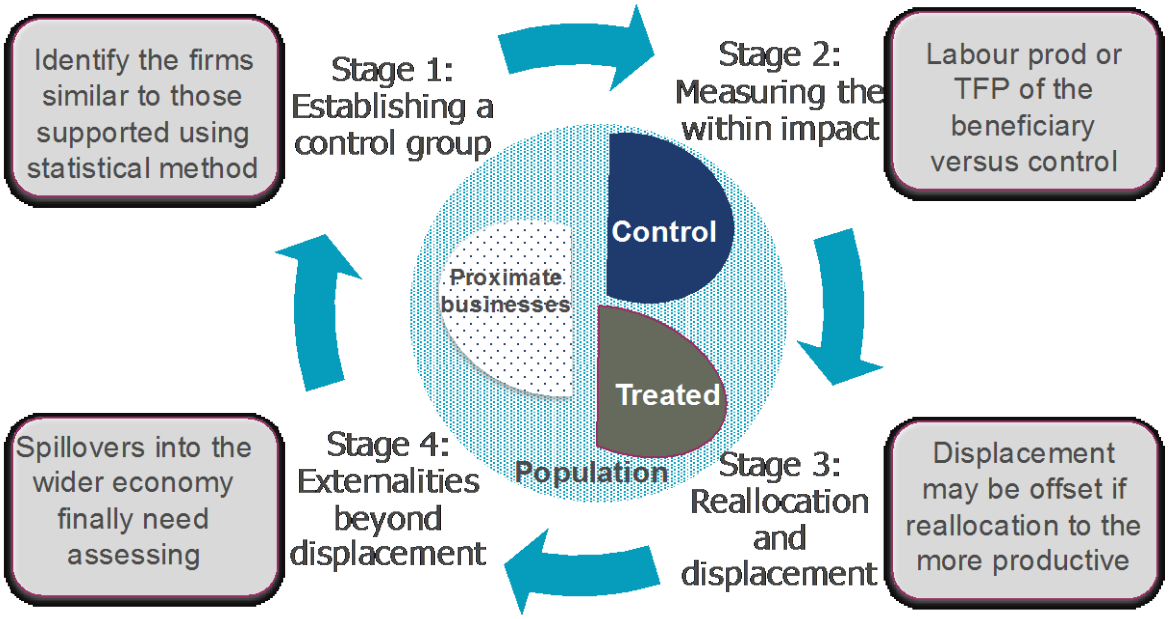
1. This chapter considers how an evaluator might assess the extent and dynamics of wider economy impacts. Steps need to be taken at the right point in an evaluation, and the chapter provides a framework for this. Sometimes wider economy impacts will be significant and measurable; sometimes the scale of an intervention or its likely impacts may mean only a modest wider impact beyond the beneficiary firm. The chapter firstly describes different types of interventions and considers how likely it is that wider economy impacts will occur for each type of support. The chapter then describes a support measure that is used as an example in analyses presented later.

Four Stages when Evaluations should consider Wider Economy Impacts

2. This report divides an evaluation into four stages, with these defined primarily to identify what steps an evaluator might take to incorporate an economy-wide approach. Figure 1 indicates the four stages and the four substantive chapters of the report then focus on each stage:

    - Adapting methods used to identify a counterfactual so that they can consider displacement impacts (in chapter 3). The next chapter considers how the data about the treated and the counterfactuals can be prepared. The control group and treatment group are usually defined by being equally likely to seek support, but with the control group businesses not receiving it.
    - Ensuring the impact measures used in an evaluation are appropriate for wider economy impacts (in chapter 4). The treatment-control logic adjusts impact measures for deadweight: it estimates the impact that would have occurred anyway without treatment (BIS, 2009a). Wider economy impacts have in the past been estimated, but usually less sophisticated measures are used. Businesses may be asked in surveys about proxies for wider economy impacts.
    - Productivity decomposition to measure the wider economy impact reallocation (in chapter 5). There are a variety of productivity decomposition approaches that could be used, replacing the more qualitative measures of chapter 4 with estimates derived using more robust impact measures. Chapter 5 critically assesses the different approaches and considers the different situations in which a specific approach would be appropriate.
    - Measuring spillovers and externality impacts of an intervention (in chapter 6). This chapter looks at how these impacts could be accommodated in an evaluation, not necessarily involving productivity decompositions.

Figure 1: Stages of an Evaluation and the Productivity Approach



3. Much of the work to understand wider impacts has focused on any displacement due to policy interventions. The problem an evaluator faces is whether a control group – by not benefitting from support – is affected by beneficiary businesses stealing either sales or factors of production. Careful selection of the control group, perhaps considering different matched samples, may improve the understanding of any displacement. These alternative control groups are termed “proximate” and identify for the evaluator businesses that may have been affected indirectly.

4. Sometimes the wider economy impacts are large and require some specific analysis to estimate them. The evidence from proxies and multipliers of stage 2 would then be replaced by specific analysis of businesses beyond those supported by an intervention to explore whether and to what extent reallocative impacts have been observed. Economically proximate businesses are those operating in the same markets as supported businesses and wider economy effects may be seen. They may be suppliers or competitors. In the third stage the wider economy perspective is explored. This would use productivity decomposition approaches.

5. In some interventions, there is the potential for externalities. Stage 4, above, would estimate these using approaches based on qualitative and quantitative techniques, as appropriate.

Deciding the Elements of the Wider Economy Approach

6. The process for deciding which elements of the economy-wide approach should be used has three aspects:

- Using the intervention logic model or its appraisal to identify whether wider economy impacts are likely.
- Identify whether impacts are likely to materialise through reallocation or through spillovers or both.
- For reallocative impacts, understand whether entry-exit is likely to be important.

7. A number of studies have presented a typology of interventions. BIS (2009b) consider different interventions in terms of themes, using thematic mapping (Figure 2). The BIS framework differentiates between five broad categories of business support, using the character of the support offered and the stage of growth of the target beneficiary group.

8. To decide whether to evaluate wider economy impacts, a starting point is to understand how and what outcomes/outputs an intervention causes. BIS (2011b) describes logic models. Each intervention will have a distinct theory of change (or logic chain). To decide whether to evaluate wider economy impacts, the key question is whether the particular intervention expects to cause reallocative impacts or spillovers. Where the intervention is targeting a specific business, supporting its productivity improvements, then reallocative impacts are possible. In Figure 2, the interventions in the upper half, consisting of business support that is financial assistance, business advice and human resources, are likely to support productivity improvements in businesses.

Figure 2: Business Support and Growth Stages

Business Support	Start Up	Growth	Maturity
Financial Assistance	• Start up grants and loans • Seed capital	• Working capital finance • Investment incentives • Venture capital	• Bank finance of all types • Access to stock markets • MBO/MBI finance
Business Advice	• Sign-posting • Business planning advice	• Marketing strategy • Organisational change	• Business strategy • Supply chain development • Knowledge networks
Human Resources	• Entrepreneur training	• Management skills • Personal recruitment and training	• Personal management • Specialist skills training
Innovation & Technology	• R&D advice • Product Development	• Quality standards • Technology transfer • Use of ICT	
Physical Infrastructure	• Business Incubator units		

Notes: Table from European Commission (1999) study on SMEs and business support schemes. MBO is management buy-out and MBI is management buy-in.

9. The logic model for an intervention usually starts with the market failure to be addressed. For example, a number of support measures address the issues faced by small businesses. However, such interventions are unlikely to have measurable reallocative impacts, because the scale of the interventions would be modest. Interventions that address credit market failures or information asymmetries meanwhile are targeting resource allocation issues consequent of significant market failures. These types of interventions may be more likely to cause reallocative effects.

10. Secondly, reallocative effects will depend on the types of beneficiary and the markets in which they operate. Product market displacement is likely to be correlated with high levels of market saturation and competition, whereas labour intensity and skill requirements will be correlated with factor market displacement.

11. In chapter 5, different productivity decomposition approaches are discussed. The logic model for an intervention will help identify what approach should be used and there are then two aspects to consider in this choice. Firstly, in evaluating any start-up intervention or one targeting small firms, there is likely to be reallocation through entry and exit, which is likely to be less prevalent for interventions targeting mature firms. The decomposition methods differ between ones where entry and exit are explicitly measured in the decomposition method and ones more suitable for when the policy supports larger businesses so that the likelihood of entry or exit is small.

12. Interventions that anticipate externalities through spillovers should be identifiable through the logic model and, in Figure 2, the lowest two rows are likely to generate such externalities. In UK policy, this is likely to be correlated with the agency providing the business support, for example Innovate UK and UK Trade and Investment interventions generally expecting spillovers due to innovation and trade respectively. ICF GHK (2014) presents an analysis of spillovers from innovation support highlighting how and when interventions can maximise externalities. UKTI has a number of drivers for their interventions, including wider economy impact (BIS, 2011c).

13. Spatial spillovers are promoted in a wider range of interventions than seen in Figure 2. For example, a transport intervention may expect to have spillovers because the reduced travel costs cause markets to function more efficiently. As with economy-wide impacts due to reallocation, the size of the intervention is likely to be important in whether wider economy impacts should be evaluated.

Practical Application of the Productivity Decomposition Approach

14. For this feasibility study, BIS made available data for beneficiaries of an intervention into credit markets – the Enterprise Finance Guarantee (EFG) Scheme – that has been implemented over the past decade. This was used for the empirical work in this study, used to exemplify the approaches being tested. Details about the scheme can be found in annex A.

15. The EFG was a relatively large intervention, with support not being significantly customised to individual businesses. Furthermore, the support was generally to smaller businesses without a particular regional or sectoral target. There were a number of benefits of having access to such data, the most important being that it allowed for empirical testing of the productivity decomposition approach.

16. The study uses this evidence base to test the productivity decomposition method. However, the results presented do not represent an evaluation of the scheme, and should be considered to be experimental in nature. This is because key steps of a full evaluation were not in the scope of this study.

17. The study did not construct a logic model for the intervention or analyse an existing one, an important step in understanding all the intervention's impacts, the timing of the impacts and the likely intermediate outcomes that would be observed. Timing may be a crucial aspect as wider economy impacts are likely to occur with a lag. The study also has not been accompanied by an analysis of the management information of the intervention. The empirical work only uses the fact of a business being given support, not considering the size or nature of individual interventions. Management information can also be used to suggest approaches to identifying a counterfactual, such as using rejected applicants as a pool from which to select comparable businesses that did not receive support. In the annex, details of evaluations of the EFG are cited.

3. Determining the Counterfactual

1. A counterfactual lies at the heart of a robust evaluation. This provides a comparison group of businesses that did not receive any support. The evaluator can estimate the effect of support by estimating any different performance between the supported and the comparison group.

2. This chapter describes the issues for deriving a control group when there are wider economy impacts. The chapter then discusses ways to develop counterfactuals specifically for an evaluation where wider impacts, and reallocative impacts in particular, are to be estimated. A particular challenge is how to draw into the treatment-control approach a means to estimate whether displacement is occurring. Furthermore, the entry and exit that will often characterise the reallocation of resources needs to be included in estimates. The final section then presents data from an application that includes such dynamics in a firm-level dataset.

Counterfactuals, Displacement and Wider Economy Impacts

3. In many recent evaluations (reviewed in annex B), the performance of supported businesses is compared to a counterfactual group. The counterfactual consists of businesses as comparable as possible to those supported but not receiving the intervention. In medical and scientific research, this is achieved by randomising the receipt of the intervention, the randomised control trial (RCT) approach. In many business interventions, the counterfactual businesses are selected using statistical matching on pre-treatment characteristics. So if an intervention generally supports larger businesses in a particular region and industry, non-beneficiaries with these characteristics would be in the control group. Propensity score matching is the statistical technique most commonly used.

4. Using a counterfactual is best practice for evaluating direct economic impacts. After a careful matching, the only difference between the supported business and the counterfactual is the treatment and any difference in performance can be ascribed to the support measure. Such matching can also form the basis for looking at wider economy impacts, but there is a caveat.

5. Both RCTs and statistical matching approaches assume that the treatment does not impact those in the control group. This issue is detailed below, but one innovative aspect of this research has been to use the strengths of administrative data, especially in its coverage of all businesses and over a long time period, to test a strategy to find a counterfactual that might shed light on displacement in particular. This strategy is based on three concerns about measuring wider-economy impacts.

6. A first concern is that wider impacts are likely to be measurable with a considerable lag after a support measure. The first impacts will be directly to beneficiary businesses and this is usually observed in the years after the support. However, reallocative impacts across supported and unsupported businesses and spillovers due to support necessarily will require some time to materialise. This means that the dataset used in an evaluation of wider economy impacts may need to be a repeated set of snapshots over some years after the treatment. Data is used, firstly, to track businesses through time for a period long enough for market stealing or spillovers to be observable.

Secondly, the data will allow an evaluation to consider the effect on market share due to entry and exit of businesses, especially as many interventions materially affect exit levels and the setting up of new businesses is an important part of reallocation.

7. A second concern is that the approaches used to evaluate wider economy impacts are data intensive. This is primarily because the techniques used are highly reliant on consistent measurement across businesses and across time. A further complicating factor is that data is needed for both the supported businesses and those that are used for comparison. The traditional approach to this has been surveys conducted for both groups, asking about displacement impacts or proxies for any such impact. But this has the disadvantage that surveys often have a small sample of non-beneficiaries. This may limit the analysis needed to tackle a third concern.

8. This third concern is that those in the control group are affected by the wider-economy impacts of a treatment. The businesses most similar to the supported businesses may be competitors, and economic activity may move from them to the supported business as a consequence of treatment, i.e. displacement (HMT, 2003). If the beneficiaries are compared to businesses that are adversely affected by the intervention, then overall impact estimates may be overstated.

9. Such effects are often called market displacement. After support, displacement is where the growth of beneficiary firms is at the expense of the market share of competitors. The extent of displacement will be influenced by whether support is freely available. A widely available form of support (e.g. a free website) should have low displacement effects as support is relatively uncontested and should not itself cause one business to have market advantage.

Box 2: Product Market Displacement and Factor Market Displacement

Where supported businesses can steal sales from other, unsupported firms, this is called product market displacement. Such displacement is affected by the nature of the product produced by a firm. The more distinctive the products or services of beneficiaries, the lower product market displacement is likely to be as there is less likelihood that there will be other businesses offering the same products or services.

Factor market displacement occurs where assisted firms take inputs away from non-assisted firms. High displacement may occur for competitors proximate to a supported business in markets where geography matters, such as labour markets. Dahlberg and Forslund (2000) use Swedish data over the period 1987-1996 and find direct displacement effects from programmes that subsidise jobs, but there seems to be no displacement effects from vocational training programs. Crepon's (2012) RCT based study provides evidence on displacement in highly skilled labour.

10. In looking at displacement, a distinction is made between product market and factor market displacement. Box 2 describes this distinction. In selecting businesses to act as a counterfactual to those treated, the product market similarity is often implicitly integrated when statistical matching is employed. In propensity score matching,

industry classification, location and other market relevant characteristics enter the selection model. This means that the beneficiary and counterfactual businesses, on average, operate in the same product market.

11. As with product market displacement, the selection model used to find a counterfactual may provide a counterfactual consistent with factor market displacement. By-and-large, the more accurately the businesses' product market is defined, the more likely the factor markets will be matching. Arguably, the industry of businesses is the best proxy for the input markets that the firm is engaged in. However, matching might also have to include location, average pay levels and size, which may also be an important determinant of competing over production inputs.

Designing a Counterfactual for Wider-Economy Impacts

12. The way an evaluator selects a counterfactual to assess wider economy impacts has to address these issues. For this feasibility study, the data requirements – namely a dataset that provides a good time-series and adequate coverage of both beneficiaries and non-beneficiaries – is largely met through the use of administrative data, described in the next chapter. Using administrative data allows the study access to a rich data about non-beneficiaries. This allows a degree of experimentation in selecting the control group so the study can assess both the likelihood and level of displacement that has occurred. A main outcome of this is that different control groups can be produced and this section discusses the design considerations around this.

Identifying Controls for Reallocative Effects

13. Defining control groups in a way that allows analysts to estimate wider-economy impacts is complicated because there are two divergent effects. To measure what would have happened anyway, an evaluator seeks out businesses that are as comparable as possible to the beneficiaries of support. But these businesses are also likely to be affected by displacement or reallocation. This section outlines an approach, using propensity score matching in a flexible way, to test when there are wider-economy impacts.

14. The possibility that support might affect both the beneficiaries and the businesses in the control group undermines the stable unit treatment value assumption (SUTVA) needed in experimental methods. This assumes that the treatment status of firms does not affect the potential outcomes of other units in the control group.

15. Attanasio (2014) highlights how, where there are indirect wider economy impacts, both experimental and quasi-experimental approaches to identify a counterfactual are biased. In an RCT, the businesses that are randomly selected out of support may be the ones that then benefit (suffer from) any beneficial (detrimental) indirect impacts. This means comparing the supported businesses with the control group would be invalid without some modelling of the bias.

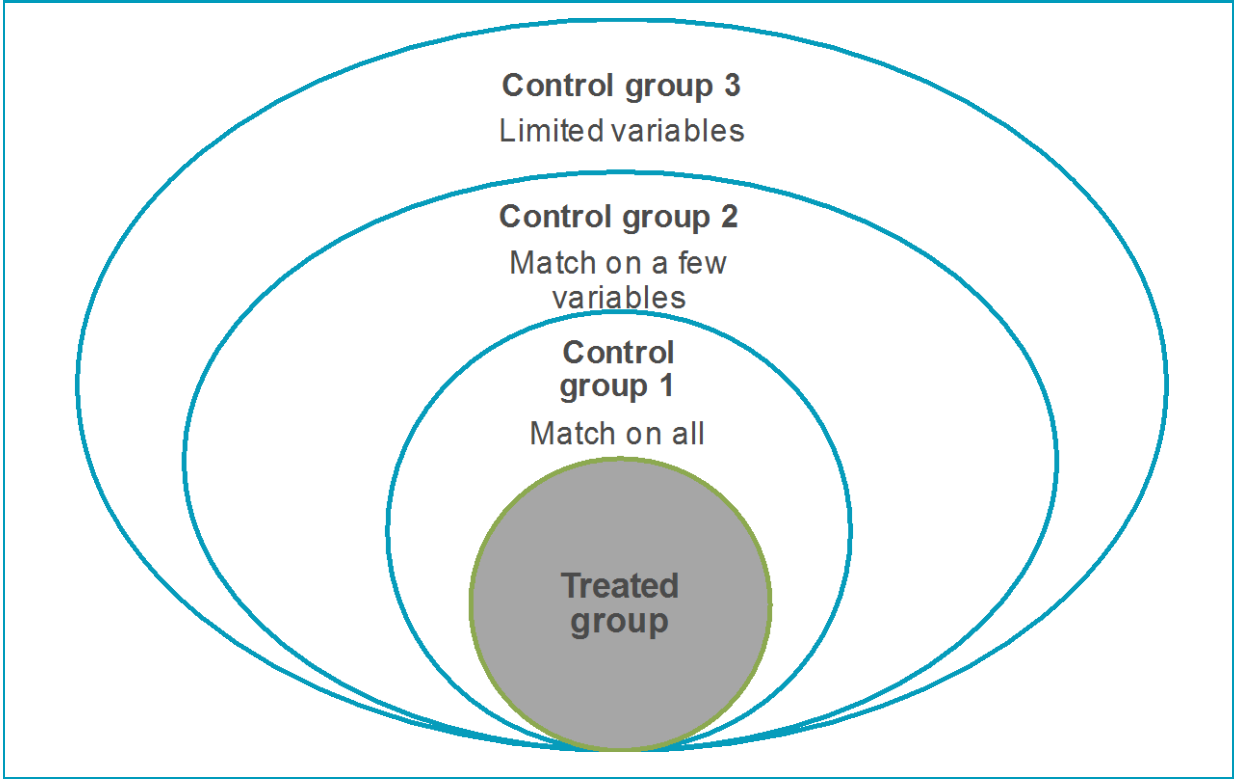
16. Where there is a concern that SUTVA does not hold, such as where there are wider-economy impacts of support, the evaluator's approach is usually to use additional variables which are correlated with the wider impact but unrelated to the treatment impact. These are called instruments. These variables can be used in the estimation stage of the evaluation, ensuring estimates adequately control for the wider impact.

17. In this study, PSM is used to explore this phenomenon. Using a matched control group might not be desirable when looking at any displacement or reallocation effects because SUTVA would not hold. Inadvertently, the matching may pick the businesses adversely affected by displacement (overstating support impact) or any reallocation towards more productive supported businesses may be missed (understating impacts). SUTVA needs to hold for inference from the quasi-experimental PSM approach.

18. However, PSM provides an opportunity to test any effects around whether and to what extent SUTVA holds. This is somewhat akin to the instrument route described above, and Annex C outlines the statistical tests that are used. The annex also includes results for these tests when applied in the statistical work used to provide examples for this study.

19. Figure 3 depicts how PSM is used in this study. The selection of the control group is undertaken in a manner that refines the estimation of any wider-economy reallocative impacts. The models to find a control group are varied in a way that differs by characteristics related to the likelihood of displacement and so indirect impacts, but less likely to be related to the size of the impact of the support.

Figure 3: Multi-Level Controls to understand Reallocation



20. The approach starts by matching to the most similar company, but then allows the matching to worsen. The deterioration in the matching may approximate the effects of being less likely to experience displacement impacts. If a supported business in the same geographical market is selected for the control group as the most similar business, then rerunning the matching but relaxing the similarity on this aspect may correlate with a reduced displacement impact. Such multiple control groups would

indirectly capture the proximity of supported businesses to the businesses selected to be in the control group.

21. While previous studies have not used matching in this manner, studies have experimented with tests of how sensitive measured impacts are to proxies for displacement. For example, Criscuolo et al. (2014) look at the impact of a government support policy and then expand the geographical coverage to see if impact is lowered because the displaced economic activity is reduced.

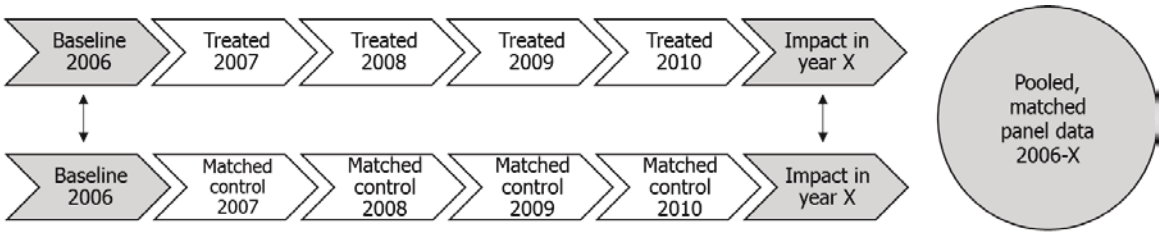
Designing Panel Datasets for an Evaluation

22. Alongside identifying different sets of businesses to act as control groups, a key step is compiling data so that long-term impacts, including firm entry and exit, can be measured. This is done by constructing a panel of both the supported businesses and a control group.

23. The figure indicates how this may be operationalised, especially where the evaluation has administrative data so that entry can be integrated into a treatment by starting the panel in the years before support is provided. Matching is undertaken for each year of support separately using firms in the year before support. If the intervention spreads across multiple years, this continuous matching will capture entry effects of both treated and non-treated companies. This allows estimation of the resource reallocation effects of entry prior to treatment, such as testing if more or less productive entrants are supported compared to the comparison group.

24. As illustrated in Figure 4, taking an intervention stretching its support from 2007-2010 as an example, the analyst will have to generate four separate matched control groups corresponding to each year of treatment. For these control groups, the analyst will have to match on pre-treatment characteristics (i.e. matching treated firms in 2008 on their observable characteristics in 2007). After this, the analyst should construct a panel from the initial treatment year to the year where impacts are expected, for all treated and matched firms. Following this procedure will allow for entry, as some firms treated in 2010 would have been incorporated into the panel after the baseline year (2007).

Figure 4: Continuous Matching to Construct Dynamic Dataset



25. Constructing the dataset this way also poses a second problem. Using one baseline year for all cohorts of treated firms means there is a lack of precision, as the analyst will not know whether productivity gains happened before or after the actual intervention. For example, a firm treated in 2010 might show a productivity increase from 2006-2014. In this case, the analyst would have to ascribe this productivity

increase to the intervention, even though there is the possibility that the increase happened before. On the whole, since the matched controls are derived on a year-on-year basis, one would expect such conclusions to be fairly robust.

Designing a Panel for a Business Support Intervention

26. Implementing Figures 3 and 4 can be undertaken using a quasi-experimental approach on administrative data to select a counterfactual. In this study, the ONS Business Structures Database is used, linked to the beneficiaries of a recent BIS support measure as given in management information. Analysis uses data about businesses at enterprise level, which aggregates over plants and outlets when a business has multiple units.

Table 1: Beneficiaries and Counterfactual Businesses in a Panel Designed to Estimate Economy-wide Impacts

Year of support	Businesses supported that could be linked	2009	2010	2011	2012	2013	2014
2009 Cohort	6,586 businesses received support of which 3,535 could be linked to the administrative data.	3,535 x 2 = 7,070	6,999	6,266	5,957	5,394	4,797
2010 Cohort	5,324 businesses received support of which 2,422 could be linked to the administrative data.	4,063	2,422 x 2 = 4,844	4,405	4,237	4,000	3,664
2011 Cohort	3,311 businesses received support of which 1,427 could be linked to the administrative data.	2,089	2,362	1,427 x 2 = 2,854	2,710	2,606	2,474
2012 Cohort	2,716 businesses received support of which 1,016 could be linked to the administrative data.	1,304	1,491	1,710	1,016 x 2 = 2,032	1,959	1,901
2013 Cohort	3,153 businesses received support of which 1,075 could be linked to the administrative data.	1,263	1,432	1,609	1,836	1,075 x 2 = 2,150	2,056
2014 Cohort	102 businesses received support in 2014 of which 34 could be linked to the administrative data	19	23	28	36	44	34x 2 = 68
Total Obs.	Overall panel size as beneficiary & control businesses enter and exit	15,808	17,151	16,872	16,808	16,153	14,960

27. In Table 1, an example of applying Figure 3 is presented for a business support intervention. A control set of businesses is identified for each of the years using propensity score matching. In matching, all businesses were used and all variables used to identify comparable businesses. Later analysis will use panels constructed in a similar manner but using different matching models.

28. The support measure has operated over a number of years and the matching to a control is undertaken for six cohorts of treatment occurring over the period, 2009-14. The table indicates the number of observations in the panel for a particular year and by the year of treatment. Altogether about 19,000 businesses are tracked in the panel, calculated by summing the number given in the left column.

29. As the support is given to small and medium sized firms, there is a degree of churn in businesses, with firms entering, then receiving support and sometimes exiting. The presence of entry/exit means each year has fewer businesses than the total tracked. Of the 4,844 businesses that were selected for the 2010 cohort, half beneficiaries and half the matched control for the beneficiaries, only 4,063 existed a year earlier. In the subsequent years after 2010, more businesses exit so that 3,664 firms remain by 2014.

30. Table 2 indicates how employment growth, entry and exit is different between those in the beneficiary group compared to those in the control group.

31. There is some evidence that resource reallocation differs between supported businesses when compared to the wider business population and the businesses considered comparable to the beneficiaries. In 2009, the employment in supported businesses was about the same as that in the businesses selected for the control group. However, Table 2 then indicates how dynamics of the two sets of businesses differ.

32. There are three observations from the table.

Table 2: Supported and Counterfactual Businesses Employment Shares and Levels

	Average employment		Businesses in 2009: Shares of employment			Businesses in 2014: Shares of employment		
	2009	2014	Total	Survive 2009-14	Exit	Total	Entrants 2009-14	Survive 2009-14
Counterfactual	18	21	53.1%	36.5%	16.6%	43.2%	10.0%	33.2%
Supported	16	22	46.9%	37.7%	9.2%	56.8%	13.5%	43.3%
TOTAL	17	21	100%	74.2%	25.8%	100%	23.5%	76.5%

33. Firstly, the supported businesses grow in employment more than the comparison group. In 2014, both groups of businesses have similar employment levels, at just over 20. However, businesses that were in existence in 2009 and who received support at some point in 2009-14 have increased employment by two posts more than the counterfactual during the period. Overall, the additional jobs for beneficiaries were 3.1 or a growth of 15%. Secondly, net entry in the supported businesses is higher. Entry is more common amongst firms that subsequently receive support than the businesses that are comparable. However, this is primarily because the exit in the counterfactual

group is almost double the beneficiaries: about 17% of the employment in 2009 is in control group businesses that do not survive to 2014. For supported businesses, only 9% by employment of businesses close.

34. Finally, it can be seen that the share of employment in the continuing firms – those that are present in 2009 and 2014 – increases more for the firms given the support than the comparison group which reduce in share. In the right hand column, the two sets of businesses are comparable using employment shares in 2014. Whereas the supported businesses that survived the period had 38% of employment in 2009, they had an employment share of 43% in 2014; for the comparison group the shares are 36% in 2009 falling to 33% by 2014. The supported businesses' employment growth was higher amongst the survivors than the comparable businesses that also survived the period.

Concluding Comments

35. Resource movements drive allocative efficiency, with efficiency raised when the movements are towards more productive uses. The share changes in Table 2 measure the relative movement of the labour resource between the supported businesses and the businesses in the counterfactual group. Later chapters will investigate whether these resources move to more productive businesses, which would be an economy-wide impact.

36. By using a counterfactual group, the tracking of these changes can be compared with businesses that are similar. However, that similarity in size, industry or location could correlate with economic proximity which means the direct impacts of the support on firms may also be at the expense of the businesses in the control group. This is termed displacement and means that estimation blurs the impact of support with the wider economy impacts due to reallocation.

37. The focus of this chapter has been the matching techniques used to identify a control group for an evaluation. These techniques are a key part of evaluations, specifically to find comparable businesses to provide evidence about what would have happened without treatment. The chapter considers how matching may be undertaken where research seeks to understand reallocative impacts. In particular, the chapter described how a panel of businesses can be constructed, with firms that have received support and matched, comparable non-beneficiaries. Comparing the dynamics of the two sets of businesses indicates differences in growth, entry and exit between the supported and those firms in the counterfactual.

38. Chapter 5 will use panel datasets to look at the productivity effects of such reallocation. First, however, the next chapter considers approaches which do not specifically measure wider economy impacts. Rather, they make adjustments approximating such impacts, a simpler and quicker means to evaluate wider impacts.

4. Direct Impact Measurement

1. Current evaluations seek to measure accurately the firm-level impact of an intervention, looking at the change in scale of a firm (turnover or productivity). This chapter looks at such measures of "within" impact.

2. The within impacts being evaluated may differ across evaluations and over time. In the early years after an intervention, the focus may be on changes in a business's activities or inputs. An innovation-focused intervention might look at R&D expenditures. Over time, the focus may move to outcome measures, such as productivity. To measure the so-called "within firm" impacts, both survey and administrative data is used, measuring the impacts for supported businesses compared to the control group.

3. This chapter covers methods that seek to evaluate economy-wide impacts without necessitating the collection or compilation of data from businesses not supported by an intervention. One common approach has been to use multipliers of the impacts on the supported firms to quantify the wider effects. A second approach has been to survey beneficiary businesses about their expectations of economy-wide impacts, particularly displacement.

Measuring Impacts of an Intervention

4. Within firm effects of an intervention may be measured by gathering information on business performance before and after the intervention and determining how much of this change can be attributed to the intervention. Often, this is estimated through welfare measures such as gross value added (GVA) or proxies, such as business turnover. Also, soon after the intervention, the focus may be intermediate impacts of the intervention: the jobs created, research and innovation. Guidance such as BIS (2009) provides more detail, especially consideration of the strengths and weaknesses of different sources.

5. Surveying the beneficiaries of an intervention has been a key data source for evaluations, offering both impact evidence and helping the analyst to understand any further direct or indirect impacts. One of the key concerns is the extent to which the evaluation will be based on a self-reported impact. This self-assessment of, for example, a firm's satisfaction, perceived difference and additionality of a programme or intervention is likely to be highly subjective and may lead to both over- and underestimations of impact. BIS (2009b) specifically guides survey designers about more objective questions to understand impacts.

6. The impacts that are measured can be diverse, especially when considering long-term impacts. With interventions aimed at skills, the impact may be best measured via the effect on the individual. Approaches centre on wage premiums, which are then used to assess the value added from such improved qualifications based on CEE (2003).

7. Total factor productivity measures the growth in output after accounting for that part of growth that is attributable to key inputs. The usual inputs covered extend beyond labour to capital and intermediate inputs. The remaining output growth is then viewed as due to the improved use of all the resources used in production. Evaluators

recognise the importance of providing this more complete account of the impacts of an intervention but the data requirement becomes quite high.

8. For evaluating economy-wide impacts, the measure used need not differ from that used in estimating direct business impacts. However, once selected, the evaluator will begin to consider measurement of business performance beyond the beneficiary firm. This may require consistent data about the impact not only in the firms directly affected by the intervention but beyond that in businesses in the counterfactual group. One source for such information is administrative datasets.

Using Administrative Sources for Employment and Turnover Impacts

9. Defining employment consistent with the key administrative datasets means some of the rich firm-level data sources can be used. ONS defines employment as "... full and part time employees on the payroll plus the number of working proprietors employed" (ONS, 2001; ONS, 2010, page 6). The ONS Business Structures Database (BSD) reports estimates of this measure.

Box 3: Business Structures Database

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.

Employment may be analysed at reporting unit and local unit. The local unit is the ONS plant level unit (for example a workshop, factory, warehouse, office, mine or depot) situated in a geographically identified place. The reporting unit usually represents an enterprise, 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. For many smaller businesses, matters are often greatly simplified because a single local unit represents the entire enterprise and so is the reporting unit.

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.

A key issue is whether an entry on the IDBR is deemed active, with various statuses possible. Active firms are identified as enterprises with at least one local unit for which live data is available, with live data becoming unavailable if a firm stops trading or falls below the VAT threshold (Evans and Ritchie, 2009).

10. Once data has been collected about beneficiaries, estimating the "within" impact of the intervention is the focus of the evaluator (Bryson et al. 2002). A common approach is the difference-in-differences estimator, which uses the statistical advantages of having samples of treated and control units before and after the treatment. Such data is used to yield two levels of differences. The first level relates to the before/after difference for both groups of beneficiary and non-beneficiary firms. The second level difference is found by subtracting the first level difference in the non-beneficiary group from that of the beneficiary group. This double differencing removes two biases in comparisons between beneficiary and non-beneficiary firms as it allows the researcher to control for permanent differences between the two groups of firms and to control for time trends unrelated to the treatment (Imbens & Wooldridge 2009).

Assessing Wider-Economy Impacts using Within Firm Evidence

11. There are two areas where the within firm impacts of an intervention can be used to estimate the wider economy impacts. The first – using a multiplier – seeks to measure the indirect impacts as increased demand following support causes further economic activity. Secondly, the displacement caused by support can be assessed indirectly by asking beneficiary businesses about the markets they operate in and deriving proxies for displacement.

Using Multiplier Estimates to Indirect Impacts

12. BIS (2009a) defines the multiplier effects of an intervention in terms of the further economic activity stimulated by the direct benefits of an intervention. It splits these impacts into a portion attributable to the income of the additional employment by the business supported. The supply multiplier then captures the additional activity of suppliers to the supported businesses. The key analytical framework underpinning such analysis is the input-output tables produced by ONS and Scottish Government. BIS (2009a) presents estimates of multipliers by the different types of interventions and at regional and sub-regional levels. These are based on almost 300 programmes and projects.

13. Multiplier analysis is a partial one. It explores what happens after an intervention, assuming there are no supply constraints in the economy. After support, businesses are assumed to be able to increase production without then exploring the market effects of this increase. This clearly has limitations for estimating indirect effects. A criticism has been that such analysis – especially for modelling wider economy impacts – will ignore whether and to what extent a support measure bids up overall prices and wage levels for inputs and decreases the prices of outputs. This would be a main channel for indirect effects.

14. The general caveat to the results has traditionally been the relative age of the underlying input-output tables for the UK and the lack of a set of region-specific tables. As the tables represent the technology of the economy, the age is particularly problematic in industries that are innovative and changing, often the areas of the economy where public support is targeted. Further, there is a degree of regional specialisation in the UK, implying that a more nuanced local picture would improve estimates.

Asking Beneficiaries about Displacement

15. Displacement is where support for a business causes economic activity to move from unsupported businesses, reducing the additional impact. In order to assess the displacement effect, beneficiaries are surveyed and asked what proportion of their business would be taken by local area competitors and competitors elsewhere in the UK, if they were to cease trading (i.e. by implication, what business they were taking from their competitors). A different size to the market asked about (e.g. within 20 miles or whole of UK) then allows estimates at different geographies. SQW (2009) summarises the regional level displacement estimates from a collection of studies on business support schemes. It demonstrates the variability in the scale of displacement.

16. The differences in the estimates are driven by a series of reasons. There are likely to be different methods of estimation, different policies (e.g. access to finance, skills, innovation), a range of target groups for the policies (e.g. size of business, sector), variation in the size of the region concerned and specific contexts that will all affect the final estimate.

Box 4: Estimates of Displacement

English Partnerships (2008) estimates displacement for business support by various sizes of area. At district level it estimates 31% displacement, at county level 49%, and 75% for both regional and UK level displacement. It also suggests that displacement differs with the type of intervention. For example, displacement is estimated to be higher for generic business support (49%) compared to access to finance (19%), based on such interventions only for the North East.

Evaluations of business support interventions also have widely varying estimates. Though, for RSA, Arup (2001) estimates job displacement at 24% (for local and other assisted areas), which is similar to that of the earlier evaluation of SFIE.

More recent evaluations have measured displacement at a national level. A study of the Business Link Helpline by BIS (Ecorys, 2012) suggested almost 90% of firms faced strong competition from other UK firms, indicating high displacement. Using a more detailed survey method to calculate displacement estimated for pre-starts at 46% and for established businesses at 66%.

The data also supports that displacement increases with size of area. For example, the evaluation of Business Links finds local displacement (within 20 mile radius) is 54% compared to national level (UK) displacement of 93% (PACEC, 1988). Those firms which sell internationally and have international competitors will also cause less displacement at a national level as they are likely to displace economic activity overseas.

Displacement tends to be higher in the tourism and services sectors (PA Consulting and SQW, 2005; Optimal Economics, 2012), and lower for “high-tech” firms and firms in niche markets (Lenihan and Hart, 2002). This is likely to be the case because high-tech firms have more unique products and fewer competitors, so their sales cannot so easily be taken by other firms compared to the services sector with high levels of competition.

Box 4: Estimates of Displacement

International evidence from Finland which considers the factors underlying displacement finds that displacement of output in non-assisted areas is more likely if the firm is young, large, has higher production technology and the investments are larger than usual (Tervo, 1990).

17. Using datasets that reflect plant level information, such as the Business Structures Database, there is potential to look at the intervention impact at plant level and at enterprise level. There is then the possibility to explore displacement within businesses (as undertaken by Criscuolo et al., 2012) by looking at the extent to which plant level employment impacts may be associated with employment falls elsewhere in an enterprise. Specific action has been taken to assure the quality of the local unit data, but the stage 1 analysis benefits from this early analysis of additionality.

Concluding Remarks

18. This chapter has discussed options for making adjustments for displacement. This chapter also describes a data source to estimate the direct impacts of interventions in terms of the scale of a business, as measured by employment. This is often the prime effect of business support.

19. If robust adjustments for displacement can be made, an evaluation may be able to include economy-wide impacts with relatively little additional analysis. Adjustments for displacement provide a reasonable first approximation and, in some circumstances, such as if there is little difference in productivity across businesses, the reallocation effect would be small and the average adjustment for displacement may be quite accurate.

20. However, research shows how, in imperfectly competitive settings, the reallocation across plants is a significant aspect of the productivity story (Hall, 1990, and Basu and Fernald, 2002). The main insight is that overall productivity rises merely by moving resources between firms to the more productive ones, an effect not considered in displacement adjustments. Whereas displacement seeks to quantify whether resources transfer from one firm to another, the allocative efficiency terms in productivity decomposition refine this by measuring the productivity change due to the reallocation of resources. It begins to answer the question about why the productivity decomposition approach is better at accounting for displacement than adjusting only for product market displacement – as is currently done – or simply assuming full factor (employment) market displacement.

5. Wider Economy Impacts due to Reallocation

1. Productivity decomposition, pioneered by Olley and Pakes (1996), explores how much of productivity growth results from the reallocation of resources. This chapter assesses whether productivity decomposition can be used in an evaluation to estimate an intervention's wider-economy impacts and especially whether, after support, resources move towards the most productive businesses, enhancing any within firm effects.

2. The previous chapters covered approaches that should be seen as necessary stepping-stones to the application of decompositions. Productivity decompositions seek to nest any wider economy impacts alongside the firm impacts covered in the previous discussion. It then quantifies the reallocation impacts of markets shares changing and entry and exit.

3. In the next section productivity decompositions are introduced. The chapter then turns to evidence from an intervention. It uses an intervention targeting the smallest businesses, so a relatively limited set of impacts are considered. The work has also been undertaken using the different control groups of the evaluation.

Decomposing Productivity Changes

4. The theoretical base for productivity decomposition goes back to the contribution of Schumpeter and to the idea of creative destruction, formalised more recently by Caballero and Hammour (1994). Firms can differ on many dimensions, such as product differentiation, technologies, experimentation or new product and/or processes, managerial abilities, age and size. Aggregate productivity measures can hide these different sources of productivity gains. Understanding the determinants of heterogeneity across firms and how they are affected by policy interventions may contribute to the understanding of how the aggregate productivity evolves over time (Ahn, 2001).

5. Productivity is characterised by high heterogeneity and persistence even within narrowly defined sectors, although persistence is decreasing over time (Baily et al., 1992). This means that highly productive firms at one point in time tend to be very productive in subsequent years. The reverse is also true, i.e. low productivity plants tend to keep underperforming with respect to average productivity and they are more likely to exit. This is a common result across several studies, controlling for other factors such as establishment size and age (Olley and Pakes (OP), 1996; Baily et al., 1992).

6. The allocative efficiency/creative destruction can be affected by product and labour market regulations. For example, OP's study suggests that deregulation in the telecommunication equipment industry led to an increase in productivity generated by a reallocation of capital and a shift in production towards more productive plants. Bartelsman et al. (2013), in a cross-country study of productivity decomposition, find that allocative efficiency has increased over time, particularly in Eastern Europe following the transition towards a market based system.

7. Due to the increasing application of productivity decomposition, the literature provides several reviews (Davis and Haltiwanger 1999, Foster et al. 2001). In this study, six decompositions are considered and annex B describes the theory. All six are presented in Table 3, which gives the strengths and weaknesses of each of the different approaches. Whereas OP focuses on reallocation among continuing firms, the other decompositions also consider the effects of entry and exit.

8. In later parts of this study, the focus is on the Foster, Haltiwanger and Kirzan (FHK, 2001) approach, with Griliches-Regev (GR, 1995) also presented. FHK estimate the within firm effect more precisely, something which is of prime interest in policy evaluation. It separates reallocative effects into other terms of the decomposition. GR estimates some of the reallocation effect separately but includes some in the "within-firm" term, lessening this decomposition's usefulness for evaluation but making the estimates more robust to economic changes, a feature significant where a study looks at a long period of time. In this study, focusing on only three to four years, the FHK method is preferred, because this advantage is less important.

9. Both FHK and GR look at the impact of entry and exit. The exit effect is of interest in evaluations, primarily because supported firms often have higher survival rates than non-beneficiaries and this is a reallocative impact. Entry is also included in this study, because decompositions look at productivity change during a particular period. It is possible to start the analysis window before support, so that both supported businesses and unsupported comparable businesses may have been set up during the period being analysed.

10. Table 4 gives evidence from the UK. Disney et al. (2003) assess the importance of external and internal restructuring over the period 1980-1992, using both labour productivity and TFP. The analysis compares the performance of three methods: BHC, FHK and GR. This study distinguishes between productivity performance in single establishment firms and in establishments that are part of a multi-business. Interestingly, the latter group is the most dynamic and it contributes greatly to overall productivity growth via the net entry effect, while the contribution of single establishments is very small, particularly when considering the within and the net-entry effects. This means that restructuring within single establishments affects productivity much less than restructuring within establishments that belong to multi-establishment firms.

11. In Mason et al. (2014) the implementation of the dynamic decomposition shows that much of the reallocation takes place within and between continuing firms, rather than as the result of entry and exit. This is consistent with results based on the OP static decomposition. Net entry shows a negative contribution to productivity growth because of new entrants with below average productivity levels and some exiting firms with above average productivity. This could be the consequence of anti-competitive practices or market imperfections.

12. Riley et al. (2014) analyse labour productivity dynamics before and after the financial crisis with the objective of investigating whether the shock to credit supply led to a substantial drag on aggregate productivity. They compare results across four different methods (MP, GR, FHK and their own hybrid decomposition) and find that restructuring within firms mainly drives productivity. The within-firm restructuring is driving productivity growth in the pre-crisis period (2003-2007) and it also determines the

decrease in productivity after the crisis. Contrary to their working hypothesis, external restructuring does not have a large impact on productivity growth.

Table 3: Productivity Decomposition Approach, Compare and Contrast

	Method of decomposition	Brief	Strengths and weaknesses	Recommendation for evaluation studies
Static	**OP:** Olley, S. and Pakes, A. (1996) "The Dynamics of Productivity in the Telecommunications Industry." Econometrica, 64(6): 1263-1298.	Decomposes productivity into an unweighted average of plant level productivity and a covariance term.	This provides a way of decomposing productivity changes into a component capturing shifts in the productivity distribution and another component capturing market share reallocations via the change in covariance. Does not account for dynamic components, i.e. entry and exit. However, changes in reallocation over time can easily be computed using first differences. Does not allow for any inferences about reallocation between groups of treated and untreated companies, but gives valid within-group estimates. The main strength is its simplicity.	OP decomposition should be run as a preliminary estimation and used as a benchmark for consistency across other decomposition methods. Can easily be updated on a yearly basis and check the evolution of the within and the reallocation component.
Dynamic	**BHC:** Baily, M., Hulten, C., and Campbell, D. (1992). Productivity dynamics in manufacturing plants. Economic Activity: Microeconomics, (4): 187-267.	Within and between firm effect. No reference average productivity level for entry and exit firms.	Simplest dynamic decomposition. Divide firms into survivors, entrants and exitors. In the BHC decomposition, the contribution of entry is always positive and the contribution of exit is always negative, regardless of the aggregate productivity of the groups, which is an unbiased picture of the contribution.	Could be used in cases where entry and exit is uncommon (i.e. in cases where treated companies are large and consolidated firms and where exit is more likely to be due to M&As than actual deaths).
	GR: Griliches, Z. and Regev, H. (1995). Firm productivity in Israeli Industry: 1979-1988. Journal of Econometrics, 65, pp. 175-203.	Within and between firms effect. Uses the average aggregate productivity level between the two periods, as the reference productivity level for assessing entry/exit.	Lowers the bias produced by not having a reference average productivity level for entry and exit. Bias relating to entry/exit is still present though for two reasons: reference productivity does not change over time while exit is assessed at time (t) and entry at time (t-1); reference productivity is computed without distinguishing between survivors, entrants and exitors. When there is productivity growth driven by productivity improvements in surviving firms, then the productivity of surviving firms is higher than the reference productivity level used to benchmark the entrants/exiters. This implies an upward bias to the estimation of the contribution of entry, and hence a downward bias of the contribution for the remaining two groups of firms. The component reflecting market share reallocations among surviving firms is most severely biased downwards. Bias more prevalent when looking at periods of large productivity changes.	GR estimates of the within effect weights using period average share, more reflective of economic change; means reallocative impacts included in within term. Could be used in cases where entry and exit is uncommon. If entry/exit is prevalent in the analysed data and the productivity changes are large over the period measured, the GR will produce biased estimates.
	FHK: Foster, L., Haltiwanger, J.C., and Krizan, C.J. (2001). Aggregate Productivity Growth: Lessons from Microeconomic Evidence. in New Developments in Productivity Analysis.	Within and between and cross firms effect. Uses industry average at time t-1 as a benchmark for entry and exit.	Produces a biased estimate of the contribution of entry and exit for reasons similar to the GR. The bias might be more serious here because reference productivity is evaluated at (t-1) rather than averaged across two periods. As for the GR, bias will be more prevalent when looking at periods of large productivity changes or long periods of time. Better suited than MP for assessment of relative importance of internal (within-firm) restructuring and external restructuring, as it includes a covariance term.	Gives the most precise estimation of within, between and cross firm effects. Index uses base year weights so may become biased as windows of analysis is lengthened, especially if period of large changes in growth rate or considerable entry/exit.
	MP: Melitz, M. and Polanec, S. 2012. "Dynamic Olley-Pakes Productivity Decomposition with Entry and Exit" *NBER Working Paper*.	Dynamic version of the static OP decomposition. Productivity levels of entrants are compared to those of continuing firms in year of entry; for exiters the relevant comparison is continuing firms in year of exit.	Unbiased estimation of entry/exit, which feeds into a better estimation of the remaining terms, which is crucial when running analysis on groups where the business demography and productivity growth rates are volatile. Unbiased when applied to longer time-periods as well, where productivity growth is expected to be higher. Lacks the full breakdown of productivity changes in the surviving/continuing firms, which is provided by FHK (no cross-term). The within component corresponds to an unweighted rather than a share-weighted mean of productivity for the continuing firms, and this can generate bias and higher volatility in the different components. MP is the most complicated estimation process.	Preferred decomposition when entry/exit is common and if the growth in productivity is more volatile. Relatively new method with fewer applications means may be better to use in conjunction with other methods.
	R: Riley et al.(2014)	It combines the treatment of entry and exit MP with the treatment of surviving firms in GR to better capture the contribution of surviving firms	It is an extension of the MP and it is considered the best decomposition method. Provides a more precise computation of the contribution of the within component compared to MP. It is more complicated to implement and it is used in only one study so comparison of results is constrained.	Given the novelty of this decomposition, it might be better to use it in conjunction with other methods.

13. The implementation of productivity decomposition requires a set of choices not only about the decomposition method but also in terms of data definitions. Studies have estimated TFP and labour productivity, in some instances comparing the performance of both (Harris and Robinson 2005, Disney et al. 2003, Ahn 2001). Although more demanding in terms of data requirement, measures based on TFP are sometimes preferred because they take into account different factors of production.

14. When studies undertake a comparison between the two measures, results tend to differ. For example, Harris and Robinson (2005), using UK data, find that restructuring within plants is important when considering labour productivity but much less relevant when TFP is used. A similar conclusion concerns the between effect, i.e. there is not a shift of resources towards better performing plants that positively contributed to TFP growth, even though there is some contribution to labour productivity growth. These conclusions are consistent with Mason et al. (2014).

Table 4: Recent UK Studies using Productivity Decompositions

Study	Period	Sector	Within	Between	Cov	Net Entry
Disney et al 2003	1980-92	Manu	0.48	-.04	-0.01	0.49
Harris & Robinson 2005	1990-98	Manu	0.08	-0.22	0.62	0.52
Mason et al 2014	2003-07	Man & Serv	1.25	0.92	-1.08	0.00
Riley et al 2014	2001-05	Man & Serv	1.29	0.88	-1.45	0.28

Note: Decomposition based on FHK. Figures are computed from table 7 in Disney et al. (2003), table 5 in Harris and Robinson (2005), table 6.1 in Mason et al. (2014) and table 1 in Riley et al. (2004).

15. Another choice regards the computation of the shares used to derive weighted productivity, output shares or employment shares. Results seem to be less sensitive to the decision regarding the shares. Mason et al. (2014) carry out the decomposition using both. They find that output shares produce higher productivity but the general trends are very similar across the two methodologies. This study also shows that results are quite sensitive to the method of aggregation used, a bottom-up approach or a top-down aggregation. Comparing the two, Mason et al. (2014) find that, although the general trends are similar, different sectors appear to drive productivity changes across the two methods.

16. In Table 4, a summary of results is presented from four UK based studies, based on the same data set (Annual Respondents Database - ARD) and the same decomposition method (FHK). Looking at the first two rows of the table, it can be seen that the contribution of the within component in Disney et al. (2003) is much larger than in Harris and Robinson (2005). Net entry is similar in the two studies, while the covariance term differs substantially. Results in the last two rows show that the two most recent studies produce comparable figures for the within effect only, while values for the other three components vary sensibly.

Productivity Decomposition for Policy Evaluation

17. Harris and Robinson (2005) use productivity decomposition to evaluate the impact of Regional Selective Assistance (RSA), a capital subsidy linked to the creation and safeguarding of employment. This intervention operates in assisted areas only and it is mainly directed towards manufacturing companies, including foreign owned ones. The policy is of interest because its impact has also been assessed using instrumental variable regression methods (Harris and Robinson 2004, Criscuolo et al. 2012) so it allows a comparison across alternative evaluation tools.

18. The authors undertake the analysis using two definitions of productivity, labour and total factor productivity (TFP). Revenue shares to compute TFP are derived from the separate estimation of a production function. The productivity decomposition, based on the method described in Foster et al. (2001), is undertaken for the whole economy, assisted plants and non-assisted plants, as well as by region and by two digit industrial sectors, all within manufacturing. The starting period of the analysis is 1990, the final year of a period of strong growth in the UK. This was followed by a severe recession (up to 1994) and then a period of flat growth in manufacturing. This detail is important, because the Foster et al. (2001) evaluation of the contribution of entering firms is particularly sensitive to the cycle and tends to underestimate their contribution in periods of recessions.

19. The results of the productivity decomposition show that for the overall sample, total productivity increased by 29.6%, with the largest contribution coming from the covariance term, which captures the increasing market shares of high productivity plants (18.5%). The second largest component is from entering firms (11.7%) while restructuring within firms is positive but it only accounts for 2.3% of the total variation[1]. The relative importance of the components is different from later studies, such as Mason et al. (2014) and Riley et al. (2014).

20. Results also show that RSA assisted plants are characterised by smaller productivity growth in the same period (8.5%) with negative contributions from both entering and exiting firms. When looking at the results based on TFP, growth in RSA assisted plants is negative (-1.4%), with a negative between effect (-2.3) and a negative contribution from entering firms (-0.03%). Hence, RSA assisted plants are particularly underperforming in terms of TFP. However, lower productivity for RSA recipients is expected as these plants were underperforming at the outset, and were granted support for this reason. In terms of regional analysis, results reveal that the RSA assisted plants do worse in those regions where RSA assistance is concentrated. Compared to the overall economy, entry and exit play a significantly smaller role in RSA assisted plants.

21. The authors state that the decomposition approach provides detailed information of how a scheme such as the RSA affects productivity growth. The fact that the assisted plants are less productive suggests that RSA, by limiting entry and exit, might be protecting inefficient plants that contribute negatively to the overall productivity

[1] The balancing terms are the within component (2.35%) and the between component (-6.59%).

performance. However, the use of productivity decomposition might not be the best tool to evaluate a policy with a clear employment target.

22. A problematic issue with this study is that the number of plants affected by the RSA is quite small compared to the total, particularly when plants are classified by region and industrial sectors. This means that the presence of outliers can severely affect the results. Moreover, the counterfactual used for the evaluation only partially accounts for similar characteristics in assisted and non-assisted plants, hence it does not allow to compare 'like for like'. For example, assisted plants in Scotland perform worse than non-assisted plants in the same region, but it is not possible to check whether the difference in performance is also caused by size/age of the plant and/or the type of industry where the firm operates. Therefore, it is difficult to assess how the same or similar plants would have performed without the implementation of the policy.

23. Finally, results are very sensitive, in this instance, to the measure of productivity used, with negative TFP growth and positive labour productivity growth in all plants. This implies that policy implications can be very different, hence the methodology does not provide a clear guidance.

24. In a parallel analysis of RSA assisted plants[2], Harris and Robinson (2004) evaluate the impact of the policy on productivity using panel data production function estimation (GMM). Here, the authors can account for several firms' characteristics, including industry and regional dummies, age of the firm, foreign ownership and business cycle effects. Results for the overall sample partially confirm those based on productivity decomposition, i.e. the RSA assisted plants have below average levels of productivity, but productivity performance increases following assistance. When the analysis is confined to plants located in Assisted Areas only, the study does not find statistically significant differences across assisted and non-assisted plants, with the exception for Scotland which experiences a post-assistance increase in productivity. There are also some positive productivity effects in sectors that were most dependent on RSA, in the regions where RSA is mostly concentrated, such as the Chemical, Metal goods, electrical engineering, food and clothing.

25. The related study by Criscuolo et al. (2014) assesses the impact of RSA assistance on employment, investment, productivity and number of firms, using instrumental variable estimation to account for the endogeneity of the treatment (receiving RSA). Their results show that a 10% investment subsidy causes a 7% increase in employment, with about half of this arising from incumbent firms and the other half due to greater net entry. The size of the firm matters as the positive effect of the treatment is confined to smaller firms. The increase in employment mainly comes from large reductions in unemployment, hence the policy does not generate job displacement. Similarly to the analysis by Harris and Robinson (2004 and 2005) this study does not find additional effects on productivity. Since less productive plants receive more subsidies, the programme lowers measured aggregate productivity because it increases the employment shares of low productivity firms.

[2] Harris and Robinson (2004) also assess the impact of the SMART/SPUR policy, whose aim is to encourage innovative activity in small and medium sized enterprises (SME).

Productivity Decomposition: An Illustration

26. In this section, some of the decomposition methods are applied to the panels created earlier. The support measure, the Enterprise Finance Guarantee (EFG), is primarily an intervention to ameliorate imperfections in financial markets, particularly as small and medium sized enterprise access capital.

27. We present these results in order to demonstrate the analysis a productivity decomposition produces but it is important to note they are not evidence of the impacts of EFG. This is due to several reasons. First, this analysis does not take into account the scale of EFG support but is instead binary to whether a firm received EFG or not. We also only focus on labour productivity which is one aspect of factor market displacement, and do not consider product market displacement in our analysis. Finally, the derivation of the control group could be improved if data allowed, in particular around the variables available to identify comparable businesses. To some extent, this is a feasibility issue for any evaluation. However, because of the other two conditions, the requirement becomes more constraining. Further analysis of how an actual evaluation would be conducted using a productivity decomposition is outlined in paragraph 45.

28. Tables 5-8 indicate decomposition results using a control-treatment framework for the support measure. Initially, two productivity decompositions, Griliches-Regev (GR) and Foster-Haltiwanger-Kirzan (FHK) are used; though later results focus on FHK. Each decomposition seeks to disaggregate the total growth in productivity into the components attributable to different types of business and sources of productivity growth.

29. Both decompositions focus on three sources of productivity growth. The first is the productivity effects within a firm. These are internal improvements that raise the productivity of the businesses. For example, access to credit will allow a business to invest in productivity improving investments. In the evaluation, these would be direct impacts of the support.

30. 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.

31. The entry and exit terms are the final component. They are important in this analysis because, as indicated earlier, the support measure used in the study appears to have an impact on exit of businesses, with much higher rates of exit in the businesses that did not receive support when compared to those that did. 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 recession, a focus of the analysis has been the latter half of the panel, as the UK emerged from recession.

Using Productivity Decomposition to Explore Reallocation

32. Tables 5 and 6 provides our estimates from decomposing labour productivity. The first results are for the whole economy in the periods 2011-14 and 2008-14, focusing on the supported businesses and a sample of comparable non-beneficiaries. The tables then unpack some estimates of productivity performance since the recession for these firms, whereas Table 6 focuses on just those businesses supported and the counterfactual firms.

33. Table 5 analyses productivity performance 2011-14, when the total growth in labour productivity in approximately 2.8 million businesses was 3.6%. The Table shows this overall productivity and then considers how much of this is accounted for by productivity growth in the businesses that were supported and from the control group. This represents 6,200 firms evenly split between the supported and the control. The second half of the table focuses on the change in labour shares and employment levels.

34. The decompositions "add up", so that the two different disaggregations both total to the overall productivity growth. The table does not have any evaluation analysis: it is merely deriving approximations from what is happening in the businesses that are the target for the support measure with the productivity performance of wider economy, in order to demonstrate what a productivity decomposition would look like.

35. The first result is unsurprising. The productivity impacts of supported businesses form a very small portion of overall productivity. The lower panel to Table 5 indicates the modest share of total employment that the analysed businesses represent. The businesses supported, plus the matched control group, represent around 0.6-0.7% of employment. Even the largest government intervention will support too few businesses for the effects not to be lost in aggregate productivity.

Table 5: Whole Economy Decomposition vs Treated & Control, 2011-14

Design	Within	Reallocation	Entry/Exit	Total
All Business and those Supported with Control, 2011-14				
All firms excl analysed	5.02%	-0.95%	-0.50%	3.57%
Supported & Control	0.07%	-0.04%	-0.01%	0.02%
Total	**5.1%**	**-1.0%**	**-0.5%**	**3.6%**

Employment Levels and Shares: All vs Analysed Firms, 2011-14								
	Average employment		**Businesses in 2011: Shares of employment**			**Businesses in 2014: Shares of employment**		
	2011	**2014**	**Total**	**Survive 2011-14**	**Exit**	**Total**	**Entrants 2011-14**	**Survive 2011-14**
All ex Analysed	11	11	99.4%	87.7%	11.7%	99.3%	87.6%	11.7%
Analysed	15	19	0.6%	0.5%	0.1%	0.7%	0.7%	0.1%

	Average employment		Businesses in 2011: Shares of employment			Businesses in 2014: Shares of employment		
	2011	2014	Total	Survive 2011-14	Exit	Total	Entrants 2011-14	Survive 2011-14
TOTAL	11	11	100%	88.2%	11.8%	100%	88.2%	11.8%

Note: Table reports log productivity growth over period, not annualised. The productivity decomposition total covers all businesses, with the splitting between the beneficiaries and matched control ("Analysed") and All businesses excluding Analysed

36. The reallocation column points to whether employment changes are correlated with productivity improvements at a firm level. One noticeable feature in these approximations is that the businesses targeted by support have a greater proportion of their productivity change attributable to reallocative effects than is the case in the wider economy. However, the term is negative, suggesting that in this productivity decomposition, the businesses have employment growth which is not correlated with productivity change, with employment moving away from the relatively more productive businesses.

    The changes in employment share also suggest that supported businesses and those in the control group are growing faster in terms of employment during the period. In these businesses, employment stood at 15 in 2011 but had risen to 19 by 2014. No such rise was seen in the wider business population.

37. Table 6 disaggregates the total productivity seen in the 16,750 analysed businesses, distinguishing between supported businesses and the comparable set of non-beneficiaries. The measure used in the productivity decomposition is employment-weighted labour productivity. The weights used in the decomposition reflect employment shares across ***both the supported and control*** group, as with Table 5. This means the decomposition focuses on the relative performance of the supported and counterfactual businesses, for within, reallocative and entry/exit impacts.

Table 6: Labour Productivity Decomposition, 2011-14 versus 2008-14

Design	GR/FHK	Within	B/w	Entry	Exit	Total
Employment weighted, 2011-14						
Counterfactual	**GR**	0.1%	2.2%	-0.2%	0.2%	2.3%
Supported	**GR**	0.2%	1.2%	-0.7%	-0.4%	0.3%
Total		**0.3%**	**3.4%**	**-0.9%**	**-0.2%**	**2.6%**
Counterfactual	**FHK**	3.6%	-1.3%	-0.2%	0.1%	2.2%
Supported	**FHK**	6.1%	-4.6%	-0.7%	-0.4%	0.3%
Total		**9.7%**	**-6.0%**	**-0.9%**	**-0.3%**	**2.6%**

Employment weighted, 2008-14						
Counterfactual	**GR**	-5.8%	0.3%	-2.5%	0.5%	-7.5%
Supported	**GR**	-3.1%	-0.7%	-5.8%	0.7%	-9.0%
Total		**-8.9%**	**-0.4%**	**-8.3%**	**1.2%**	**-16.5%**
Counterfactual	**FHK**	-3.5%	-1.6%	-3.1%	1.4%	-6.7%
Supported	**FHK**	3.1%	-7.3%	-6.6%	1.0%	-9.8%
Total		**-0.4%**	**-8.8%**	**-9.7%**	**2.4%**	**-16.5%**

38. The 2011-14 decompositions points to potential estimates of within firm additional growth in productivity; our decomposition illustrates approximate increases of 2-3% due to the support. This is using the FHK decomposition, which measures the within effect more precisely than GR. FHK uses the base year employment shares to weight firm level productivity and the 3.6% growth seen in the counterfactual is compared with the 6.1% growth in the supported businesses. In contrast when GR is used the within effect is considerably more modest. GR uses an average of employment shares at the start and end of the period, thus somewhat incorporating the reallocation during the period of employment.

39. Even though both indices are general approximations subject to the caveats noted earlier and not estimates of the impact of EFG, they indicate the significance that reallocation can have. The decomposition suggests that supported businesses in 2011-14 have a 6.1% productivity growth according to FHK, but our estimates suggest the reallocation of employment towards firms with weaker productivity performance, there is reduction of productivity of 4.6%. The 2011-14 results indicate that very little of this approximate productivity change occurs due to entry. This is primarily because the majority of businesses receiving EFG would have been supported prior to 2011. This makes the top half of Table 5 more like a conventional evaluation, tracking supported businesses and the counterfactual only for the period after treatment. Reallocation effects due to entry and exit is more ambiguous than the within firm and between firm impacts, as a closure (entrant) may reduce (increase) product varieties or have effects on the asset base in the economy.

40. Table 6 cover all industries for the 2008-14 period and for the post-recession period, 2011-14. Looking across the two periods allows distinguishing the recession impacts when credit markets were particularly affected and cyclical impacts were pronounced. The results covering the recession indicate that the supported businesses are growing in terms of productivity faster than the matched control group. The decomposition compares productivity growth between the two groups firstly, for example, within firm productivity growth contributes 3.1% to productivity in the supported firms while the control groups within firm productivity declines by 3.5%, both using FHK. The GR estimates are more similar.

41. There was considerable reallocation in this period. For example, using the FHK method for estimating the reallocative effects to supported firms during 2008-14, a negative

7.3% between effect is observed in the decomposition for firms that received EFG support. In particular, the treated firms generally survived in larger numbers than the unsupported. One reason may be that support was focused on businesses during 2009-10 as firms emerged from the recession and the UK economy delivered generally low productivity growth. This meant a significant transfer of employment to the supported businesses, but productivity growth was not correlated with employment movement. This may point to labour hoarding, with firms accepting a productivity fall rather than laying off employees, and such a driver may have been more pronounced in the supported businesses, because of the additional support they received raising survival rates.

42. This was further exacerbated by many supported businesses being recent entrants. The businesses in the counterfactual also include a number of new entrants. Both FHK and GR compare the productivity of entrants at the end of the period with some average level of firm productivity and entrants – both supported and those in the control – appear to have lower than average productivity. There are alternative ways to look at the effect of entry, especially as the period of considered is quite long, and this may improve the estimation of the entry effect.

43. There are caveats to such an analysis. The results are primarily for illustrative purpose, so key steps of an evaluation have not been undertaken. A key aspect in an evaluation would be to look at the intervention's logic model to understand the expected impacts in terms of timing, type and extent. There has been a recent evaluation study of the EFG support measure. This careful evaluation estimates that a quarter of the benefits to assisted firms would displace other local firms (Allinson et al., 2013). This study also concluded that 39% of businesses' improved performance were not additional (deadweight) and displacing.

44. The results also depend crucially on the treatment-control logic, and its quality depends on the quality of the propensity score matching. Allinson et al. (2013) make some important improvements on the comparison group by using information about whether non-beneficiaries secured finance from alternatives to the support measure. Where this is taken into account in analysing the treatment, it reduces the overall treatment effect. Such data is not available for the current study.

45. Overall, the tables suggest that reallocative effects are high though we cannot be certain of this due to the above data limitations. Reallocation of resources towards or away from firms that experience productivity differences is a driver for overall productivity. The comparison of the supported businesses with a counterfactual then explores what happens if these effects are measured in a sub-population of businesses that are statistically similar, except for the benefit of support.

46. An approach to improve the PSM is to focus the matching on pools which are inherently more comparable. In addition, there is value in considering only a few years' performance for supported businesses and then tracking their performance in comparison with a control group. The next two sections consider results using these refinements.

Defining Control Groups where Reallocative Effects are High

47. Table 7 indicates results when matching and analysis is refined to focus on sub-samples of firms supported by EFG. The subsamples explore the effects of matching

refined by variables related to economic proximity. The examples used here are relatively simple and so are limited in their accuracy, starting from a decomposition where all beneficiaries are matched to a counterfactual using relatively imprecise, in economic proximity terms, variables about the businesses. For example, the location of a business is initially approximated using the Government Office Region. Replacing this with its travel to work area allows matching to take into account which labour market a business finds itself in.

48. The decompositions and matching is then repeated but with the counterfactual and supported businesses used being only those in the service sector. This restriction of the pool approximates economic proximity in a number of dimensions. Generally, the service sector is less capital intensive and so this refinement of the decomposition may mean more similar production processes. Further, the outputs of services may have more in common than all industries, suggesting product market displacement due to treatment is more likely to be captured.

49. Table 7 indicates the overall growth in productivity, producing illustrative estimates of within, reallocation, and entry/exit components of the total as in previous tables. Moving down the table, the economic proximity increases. It should be noted that – unlike in Table 6 – the difference in effect is estimated rather than the total. The decomposition disaggregates the total effects, but because the discussion focuses on the differences observed between the counterfactual and the supported businesses, Table 7 highlights this.

50. The difference in the within firm effect gradually rises as proximity increases. This may reflect the increasing effect of displacement, with the supported businesses being compared to businesses that are more likely to be competing for resources or sales. The reallocation effects are generally negative and these effects are more pronounced in the supported businesses. The scale of reallocation is quite similar.

51. Entry and exit effects have been added together to provide an overall estimate as this has generally not been a large impact. It is noticeable that the net effect is slightly more significant for all industries, reducing as the decomposition is applied to service sector businesses.

52. A feature of the results is a transition from an overall impact which is negative to one that is positive. This may be a consequence of the increased precision of the matching but is also related to the focus shifting from all industries towards services. The labour decomposition measure does not take account of capital input and the post-recession period has been one of reducing capital intensity (paralleling labour hoarding). This is noticeable in the manufacturing productivity estimates in particular, which show negative growth. Removing production industries probably improves the appropriateness of using a labour productivity measure and then removes this downward pressure.

Table 7: Labour Productivity Decomposition and Economic Proximity				
FHK Decomposition	**Within**	**Reallocation**	**Entry/Exit**	**Total**
All Industries matched using Government Office Region, 2011-14				
Counterfactual	3.6%	-1.3%	-0.1%	2.2%
Supported	6.1%	-4.7%	-1.1%	0.3%
Difference/Effect	**2.5%**	**-3.4%**	**-1.1%**	**-1.9%**
All Industries matched using Travel to Work Areas, 2011-14				
Counterfactual	3.2%	-1.8%	-0.3%	1.1%
Supported	6.3%	-4.9%	-1.1%	0.3%
Difference/Effect	**3.1%**	**-3.1%**	**-0.8%**	**-0.8%**
Services matched using Government Office Region, 2011-14				
Counterfactual	2.4%	-1.3%	-1.2%	0.1%
Supported	7.0%	-5.1%	-0.8%	1.0%
Difference/Effect	**4.6%**	**-3.8%**	**-0.3%**	**-1.2%**
Services matched using Travel to Work Area, 2011-14				
Counterfactual	1.8%	-0.4%	-0.6%	0.9%
Supported	7.0%	-4.4%	-1.1%	1.5%
Difference/Effect	**5.1%**	**-4.0%**	**-0.5%**	**-0.7%**

53. One general observation is that the reallocative effects are negative. This has been a feature of FHK more generally, an index measures which distinguished the within effect from the reallocation effects well. In the types of businesses that are being supported, it is clear that employment growth is not highly correlated with productivity performance. As noted above, the scale of this needs some further work to ensure that the productivity measure is correct. However, this is consistent with a view that displacement is a problem, and that the reallocative effects dampen any beneficial within-firm impacts

54. There are caveats to these results. While they show a pattern consistent with expectations, this relies on a fairly blunt instrument to estimate proximity. Because of the relatively few variables available in the administrative data about the markets in which a business operates, with sector and location being the key variables, the models used to match are very basic. They prove unsatisfactory, for example, when modelling the manufacturing businesses.

Concluding Comments

55. Displacement focuses on the moving of resources or market share between firms. While the supported firms are defined, the pool of firms from which they may displace economic activity is more open. The approach here is to match to the supported businesses comparable businesses derived as a control group. The treatment-control logic is seeking to focus on businesses that receive support and those that are identical on average except for the support. The logic for this is that within that pool of businesses, any reallocative effect, while not identifying which employee moved or how a particular product market was captured, does provide the averages for a group of businesses of interest. A key question is the extent to which this is appropriate.

56. Secondly, within this general construct, the way in which matching is conducted can be tailored to the support given. Further, it requires the businesses to be tracked over time, long enough for displacement and reallocative effects to be analysed. Therefore, a panel, focused on the treated and the matched control has been used. Decompositions are then estimated on the productivity performance, including entry and exit, of a small part of the universe of businesses. The analysis allows estimation of the within firm effect, but a question is whether the observed reallocation in this universe represents the indirect impacts.

57. A final question is whether using decompositions is improving on present measures of displacement: is the productivity decomposition approach better at accounting for displacement than adjusting for product market displacement as is currently done or simply assuming full factor (employment) market displacement. The results presented in the chapter give some initial estimates for the Enterprise Finance Guarantee. They indicate the types of analyses that are possible and how they might be interpreted, but considerable further work needs to be done to assure their robustness.

6. Evaluating Spillovers

1. Whereas the productivity decomposition accounts for dynamics of reallocation, it does not assess externalities in the form of spillovers. The presence of such wider impacts means that businesses may fail to account for the wider societal benefits from their activities and decisions. If this is the case, there is a rationale for cost effective government intervention to provide incentives to encourage entrepreneurs to undertake the activities that lead to spillovers.

2. The externalities that are derived from spillovers are diverse and evaluation has usually used specific data collection to understand the extent of such impacts. These may be business surveys or qualitative evidence. There are also some relevant administrative datasets that have been employed. In the more data intensive work, the effort has often been in deriving appropriate variables to capture both the drivers of and the impact of the spillovers. Often, for assessing spillovers, a mixed method approach may be appropriate.

3. This research does not quantify any positive spillovers resulting from the examples of government interventions used. This chapter provides a review of case studies where spillovers, such as innovation leakages and spatial spillovers have been estimated. For an assessment of wider economy impacts they should be included where data allows.

Types of Spillovers

4. An intervention may have economic effects beyond those directly supported through spillovers. This could be the intention of a support, as these indirect impacts can be substantial. Sometimes, the spillovers are unintentional or even negative and policies would seek to reduce these harmful effects. Sometimes the spillovers are modest in scale and so the evaluation would have to be proportionate in its use of tools to capture the effects.

5. In three areas, spillovers have been significant and studies have developed tools to measure these impacts. Firstly, the chapter focuses on innovation. The business input that is targeted for such spillovers typically is research and development. Here the spillovers occur as businesses beyond the ones that conduct research and development can benefit from or improve on innovations. These provide wider economy impacts.

6. Secondly, there are geographical spillovers. Economic geography has modelled the effects of clustering and co-location both on product markets and on factor markets, exploring how this affects productivity levels. Here, the input driving any spillovers is sometimes an investment in infrastructure or co-ordinated industry-wide focus on skills or advice. Finally, a significant impact of international trade has been the spillovers associated with inward investment. Inward investment brings the intellectual property of multinationals into the production processes; it also opens up overseas markets. These permeate through UK supply chains and businesses associated with the investment.

Knowledge Spillovers and Evaluation Frameworks

7. One important source for spillovers is innovation activity. BIS (2014) highlights how innovation results in market spillovers as businesses commercialise the products of research and development. There has been a drive 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. The input of research and development by businesses firstly results in commercially viable technologies. In an open innovation system, these are improved through sharing and collaboration.

8. The UK's innovation support is quite advanced, working across businesses, universities and government to promote research leadership in knowledge-based industries. As policy has focused on capturing and maximising economy-wide impacts, evaluators are developing tools to measure the spillovers from the different types of support. Broadly, the support offered is categorised as 'investments' supporting businesses to invest in innovation, as 'infrastructure' (such as Catapult centres) and interventions to support 'collaboration'. The final two – especially when targeted to increase the openness of innovation – target wider impacts.

9. For assessing wider economy impacts, BIS (2014) first notes indicators that would guide the evaluator regarding the likelihood of spillovers. Many indicators are related to the technology and institutional set-up in which the innovations are being supported. Where a technology has wide applications, or is embedded in a relatively open architecture, then this increases the likelihood of spillovers. 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.

10. Where spillovers due to innovation are likely, any measurement strategy is likely to employ a diversity of methods. BIS (2014) suggest that there is a need for qualitative assessment of knowledge spillovers. Any qualitative assessment would then be drawn into developing and applying a multi- criteria analysis.

11. Some studies have sought to quantify spillovers from the perspective of supply chain and market impacts (which tend to assume that the business users of the knowledge are close to identical to the producers of the innovation and/or that existing inter-industry linkages are stable and provide a reasonable proxy). Using these assumptions, these studies have used econometric methods to quantify the value of spillovers for particular industrial sectors.

12. The assessment of knowledge spillovers is especially important when the intervention consists of investment in R&D projects. The rationale for such investments is usually that R&D tends to be underprovided by the private sector due to positive externalities. Hence, the assessment of spillovers already plays an important role in policy appraisal. Assessment frameworks distinguish between private and social returns to R&D. Roper et al. (2004) propose a framework to assess regional spillovers of R&D projects. First, they provide a "checklist" of potential private and social benefits. Second, they describe how the profile of the R&D facility itself and the character of the innovation system in which the R&D project is located influence the share of the spillovers that is accruing to the host region. They propose using these two factors as "filters" through which to assess the spatial distribution of benefits.

Quantitative Approaches to Innovation Spillovers

13. Understanding the inputs to innovation has relied on business surveys about expenditure in research and development and the business skills, processes and management around that R&D. The most common sources of data are two surveys, Business Expenditure on Research and Development and the EU-wide Community Innovation Survey. Variables to understand business innovation activity then cover R&D expenditure, staff that are research focused and the extent to which businesses have delivered changes to products or production processes through their innovation. These latter variables are used to understand the transition from inputs into innovation to outputs, especially to ones that drive firm level productivity.

14. Bloom et al. (2013) point out that spillovers can cut two ways: technological spillovers can benefit everyone, but there can also be market-stealing effects on the product market side. There are two distinct types of spillovers. The first are technology (or knowledge) spillovers, which may increase the productivity of other firms that operate in similar technology areas. The second type of spillovers is the product market rivalry effect of R&D. Whereas technology spillovers are beneficial to other firms, R&D by product market rivals has a negative effect on a firm's value due to business stealing.

15. Analysing this involves two steps. The first uses a general analytical framework to develop indicators correlating specific technology innovation performance indicators (market value, citation-weighted patents, productivity, and R&D) with business performance. The predictions differ across performance indicators, thus providing identification for the technology and product market spillover effects. Second, the study empirically distinguishes a firm's position in technology space and product market space using information on the distribution of its patenting across technology fields, and its sales activity across different four-digit industries. This allows construction of distinct measures of the distance between firms in the technology and product market dimensions. The study develops a methodology for deriving the social and private rates of return to R&D, measured in terms of the output gains generated by a marginal increase in R&D over heterogeneous firms.

Case Studies and Spillovers

16. A case study approach has often been applied early on in the life of an innovation intervention as part of a process and early impact evaluation. This is used when quantitative data is not yet available, or to complement an on-going quantitative impact evaluation. These often explore the scope for spillovers in an intervention.

17. London Economics (2010) employed a comparative case study approach involving the selection of 11 cases of companies that had relocated against a comparison group of 8 cases where relocation had not taken place. None of the reviewed studies were explicit with regards to the wider theoretical population to whom their findings were generalizable.

18. Methods employed within the reviewed studies included document review, analysis of management information, surveys, and both face to face and telephone interviews. Managers, delivery staff, project participants or beneficiaries, and other local stakeholders were the focus of the surveys and interviews. The use of multiple methods enabled researchers to triangulate the evidence and use data gathered using

one method (for example, interviews) to help to interpret and explain data collected using other methods (for example, surveys).

19. In the reviewed case studies that attempted to cover additionality, the approach tended to focus upon directly asking managers, delivery staff and beneficiaries about what they thought would have happened in the absence of the intervention. This included whether it would have been possible to raise funding from elsewhere, to what extent and in what form a project would have been able to continue functioning, and whether the intervention helped collect additional funding from other sources.

20. As London Economics (2010) emphasised, this approach to understanding the counterfactual is 'necessarily hypothetical and so may not correctly indicate the actions that would have been undertaken in reality'. However, qualitative interviews did enable the identification of additional non-financial and less tangible benefits of the intervention, including increased staff expertise, locally tailored provision, and the spillover effects of improved coordination and linkages with both local and national support. The survey indicated that firms were forming links with UK organisations, suggesting that there is potential for future knowledge-based spillovers. Around 60% of respondents stated that they use UK suppliers, while a similar proportion held collaboration agreements with UK organisations.

Spillovers through Trade and Foreign Investment

21. There is considerable evidence on the importance of export orientation as a determinant of business productivity. For a small firm, exporting, which requires sinking some fixed economic costs to orientate itself for overseas trade, is a viable option only for the more productive, innovative firms (Helpman et al. 2003). Those fixed costs have been modelled in terms of distance to overseas markets (Bernard et al., 2007), information asymmetries (Iacavone and Javocik, 2008), etc. And some evidence is emerging about the selection effect versus more dynamic aspects, such as how productivity increases through the business preparing to export (Harris and Li, 2007). This research indicates the "inputs" to firms participating in international trade.

22. The evidence on inward investment shows that high quality foreign direct investment (FDI) confers productivity enhancing spillover benefits. Such FDI can benefit UK exporters, because they have the "absorptive capacity" to do so. Girma and Gorg (2005) present evidence on the heterogeneity of firms in this capability.

23. Evaluations of the interventions that target trade and investment (such as Driffield, 2010) have generally used a mixed method approach. They often analyse quantitatively both management information about the intervention and the effects on businesses in key variables, such as exports. Many interventions will target a particular business activity, such as research and development and the effect on business use of inputs will then be related to outcomes including trade related ones. Spillovers are then analysed using both qualitative and quantitative approaches. For the former, surveys will indicate the trade and investment relationships of businesses and whether spillovers are anticipated. Quantitative approaches to productivity enhancing spillovers will depend on the transmission mechanisms expected.

Spatial Spillovers

24. There is a developed literature concerning the potential for spatial spillovers. These are thought to be important contributory factors behind regional growth dynamics. Attempts to estimate the presence and significance of spatial spillovers have begun by understanding the theoretical basis for the spillovers before focusing on measurement.

25. Although there are many different theories regarding spatial spillovers, three core themes have emerged (Capello, 2009). First, knowledge spillovers occur between firms within a geographical area. These occur when one firm creates knowledge and that knowledge dissipates to other firms (Fischer, 2006). Given that Tobler's (1970) first law of geography highlights that "Everything is related to everything else, but near things are more related than distant things," it logically follows that knowledge, which is generally assumed to be non-excludable and non-rival, is more likely to transfer between individuals who interact more frequently and in greater depth, and hence such knowledge can be more quickly embedded in the production process by economic agents in close proximity.

26. Second, a market, and in particular a monopolistically competitive market, is spatially defined. When a large firm in an industry expands its production, improves its productive efficiency and/or introduces a change in its productive process then it can influence other firms that are horizontally or vertically linked across the supply chain; these effects are often termed industrial spillovers. When firms along that supply chain co-locate within a geographical area, patterns of spatial spillovers could be identifiable and could reflect positive externalities brought about through productivity enhancing effects. These industrial spillover effects can be brought about through technological advances and good managerial practices that are implemented into the production process (Grilliches, 1992) and may stimulate increases in labour force training (Capello, 2007). Thus businesses who achieve clear benefits from government support could enhance the productivity of firms along their supply chain, and if these firms co-locate then they could be identified using measures of spatial spillovers.

27. Third, when a local economy develops and grows, either through strong development of one industry or the evolutionary development of a cross-section of industries, then it is often expected that firms located in neighbouring areas benefit. These growth spillovers have a spatial pattern in line with Tobler's first law of geography. Although this is the most general perspective of geographically related production spillovers, it should not be forgotten that these effects are mediated by trade and transport costs.

Agglomeration Economies

28. These three categories of spatial spillovers have obvious synergies with agglomeration economies, which describe the costs and benefits that firms realise from locating relatively close to one another. Typically, agglomeration economies relate to the ability of firms to benefit from economies of scale (which are internal to the firm) and network effects (which exist between firms); agglomeration diseconomies can also occur which are associated with the costs incurred of co-location, such as the adverse effects of greater competition, a lack of superior technologies relative to competitors, a comparative lack of economies of scale and/or scope, and congestion.

29. Agglomeration economies are associated with at least two main important areas of analysis: urbanisation economies and localisation economies. Urbanisation economies

(Marshall, 1920) arise in larger conurbations due to a relatively large pool of workers who have a variety of skills, combining with increasing returns to scale (within firms) in intermediate inputs. The latter may be facilitated by the relative ease of communication (between firms) and obtaining supplies, workers and innovative ideas due to the proximity among firms.

30. Localization economies (Jacob, 1969) arise in a conurbation from the close proximity of many firms in the same industry. They stem from more firms in the same industry attracting more workers (who exchange knowledge, ideas and information) who have the skills to work in that industry (though a scarcity or inflexibility of such workers can negate these potential benefits). In addition, the co-location of firms can collectively push down the price of intermediate good purchases. Greater competition within the sector may also stimulate firms to increase innovative capacity to stay ahead (or at least with) the rest of the market.

31. Theory suggests that when firms in related industries cluster, the costs of production may decline significantly, due to competing multiple suppliers, greater specialisation and division of labour. Even when multiple firms in the same sector (i.e. competitors) cluster, there may still be advantages because that cluster attracts more suppliers and customers than a single firm could have done alone. This concentration of economic activity in cities is the reason for their existence, persistence and growth, while the accumulation of skilled workers over time has led to the urban wage premium (Glaesar and Mare, 2001).

32. It is theoretically possible to identify whether urbanization and/or localization agglomeration economies exist and whether they are stronger in one location relative to another. Comparison should be made with a range of other areas or conurbations. Differing industrial structures may mean that although urbanization economies may be present across most areas, the underlying localization economies may be sector specific and therefore will be distinctly different between conurbations.

Evaluating the Importance of Policy in Stimulating Spatial Spillovers

33. Empirical research has sought to identify the presence of spatial spillovers. The economics literature provides empirical evidence which supports the presence of spatial spillovers but it is rare for a study to differentiate whether their sources are knowledge, industrial or simply growth spillovers. Empirical difficulties arise in measuring individual local units accurately, as data collection is often at a level that aggregates across plants. This counters approaches to measure spatial spillovers, missing business responses to distinctly different policy initiatives in different areas.

34. Capello (2009, p. 640) correctly emphasises that the empirical identification of spatial interactions among observations has pushed the literature on spatial spillovers in a specific direction. This highlights the role played by the mere physical proximity in the complex processes of local interactions instead of improving understanding of the relevance and contributory effects of distinct underpinning theoretically derived mechanisms. For instance, the location of foreign firms in close proximity to domestic firms can generate co-agglomeration effects including employment generation (Barrios, Bertinelli and Strobl, 2003).

35. Co-location can be the result of policy incentives to stimulate business parks, enterprise zones or similar developments. Unfortunately, the empirical literature on this

front is inconclusive as it is known that firms typically relocate to such enterprise zones in order to benefit from location subsidies and the potential supporting facets of co-location; consequently, output drops in the firm's previous location. Although the empirical evidence is mixed, the displacement effect can have no beneficial effect on net employment but may have some, albeit small, beneficial effect on productivity (Hanson and Rohlin, 2013) although industrial spillovers may still be beneficial if policy incentives enhance productivity along a supply chain (Markusen and Venables, 1999).

36. Conceptually there are many difficulties in measuring spatial spillovers. First, the quality and frequency of interactions between institutions and/or local actors can be enhanced by stronger networking which is often unobservable and/or difficult to measure in strength. Although Romer (1986, 1990) argues that the nature of knowledge is non-excludable and non-rivalrous, it may be the case that networking activities break down the excludability and rivalry characteristics of knowledge associated with a lack of familiarity of the true gains that can be accrued from implementing and embedding the ideas into the production process.

37. Second, the influence of these interactions between institutions and/or local actors on decisions depends crucially and critically on human behaviour, and not least habit. Capello (2009, p. 640) argues that "the capabilities of economic agents to absorb and utilise spillovers – implicitly assumed by spatial spillover theory to be invariant with respect to geographic space – depend on strongly territorialised and unimitatable assets, such as local trusts, sense of belonging and creativity."

38. Third, the difficulty in identifying empirical support for a particular policy initiative is due to the lack of effective proper economic evaluations, where a clear and proper control group is established. Firms in any control group should be affected by the same global, regional and local market mechanisms as those firms that receive the treatment. This strong limitation, especially when monopolistically competitive local economies feature in the landscape. This has often meant that the effect of policy initiatives can only be properly assessed on a case-by-case basis and can only be evaluated ex post, which makes ex ante predictions extremely difficult. Moreover, if Camagni and Capello's (2009) perspective that a market area is a 'territory' that is characterised by both geographical proximity (agglomeration economics, district economies) and a cognitive proximity (share behavioural codes, common culture, mutual trust and sense of belonging, identical managerial objectives) is adopted, then it makes accurate implementation of propensity score matching practically impossible.

39. Fourth, an ever-present concern in economic geography problems is the spatial scale for analysis, commonly called the Modifiable Areal Unit Problem (MAUP, see Gehlke and Biehl (1934), Openshaw (1983) and Arbia (1989). Openshaw famously bemoaned that "areal units used in many geographical studies are arbitrary, modifiable, and subject to the whims and fancies of whoever is doing, or did, the aggregating." The MAUP elucidates that what is appropriate as a market area for a firm working in one sector and/or local economy may be completely different for a firm working in another sector and/or local economy.

40. The importance of MAUP issues have been developed and analysed by various researchers. They show that spatially identifiable patterns could reflect an area selection problem (Holt et al., 1996) that could be due to spatial aggregation biases (Okabe and Tagashira,1996; Tobler, 1989), susceptible to spatial autocorrelation relationships that vary at different spatial scales (Arbia et al., 1996), the differential

effects of distinct zoning patterns (Openshaw, 1996; Green and Flowerdew, 1996) and/or scale effects on parameters of spatial models (Amrhein and Reynolds, 1996; Wrigley et al., 1996).

41. To summarize, the ability to accurately assess the presence and strengths of spatial spillovers is difficult, challenging and currently more of an art than a science. There is discernible scepticism over the accuracy and reliability of estimates generated using spatial econometric methods due to the inability to attribute spillover effects to specific theoretical causes or policy initiatives. Similarly, application of propensity score matching techniques to identifiable productivity effects of policy initiatives are sensitive to firm-level characteristics, such as unobservable managerial objectives.

Concluding Remarks

42. A case study approach has often been decided upon early on in the evaluation of an intervention where significant spillovers are anticipated. The scope of the case study is defined around the likely economy-wide impacts as the evaluator would, through an analysis of the logic model of the intervention, have been able to articulate the mechanics of how externalities might arise in businesses beyond the treated firms.

43. The case studies are implemented with a mix of methods. London Economics (2010) employ a comparative case study approach involving the selection of 11 cases of companies that had relocated against a comparison group of 8 cases where relocation had not taken place. None of the reviewed studies are explicit with regards to the wider theoretical population to whom their findings were generalizable.

44. Methods used include document review, analysis of management information, surveys, and both face to face and telephone interviews. Managers, delivery staff, project participants or beneficiaries, and other local stakeholders were the focus of the surveys and interviews. The use of multiple methods enables researchers to triangulate the evidence and use data gathered using one method (for example, interviews) helps to interpret and explain data collected using other methods (for example, surveys).

45. Trade spillovers have been quantified in a similar manner. There is considerable evidence on the importance of export orientation as a determinant of business productivity. For a small firm, exporting, which requires sinking some fixed economic costs to orientate itself for overseas trade, is a viable option only for the more productive, innovative firms (Helpman et al. 2003). Evaluations of the interventions that target trade and investment (such as Driffield, 2010) have generally used a mixed method approach. They often analyse quantitatively both management information about the intervention and the effects on businesses in key variables, such as exports. Many interventions will target a particular business activity, such as research and development and the effect on business use of inputs will then be related to outcomes including trade-related ones. Spillovers are then analysed using both qualitative and quantitative approaches. For the former, surveys will indicate the trade and investment relationships of businesses and whether spillovers are anticipated. Quantitative approaches to productivity enhancing spillovers will depend on the transmission mechanisms expected.

46. A developing area for evaluation is the increased use of econometric methods to estimate spillovers. Bloom et al. (2013) is a recent example where the spread of an innovation is separately modelled using patent citation data, which correlate very well

with the dissemination of ideas as they are adopted by businesses. Relating such indirect productivity enhancement from an innovation to any support given by government may provide an evaluation method.

7. Conclusions

1. This feasibility study suggests that productivity decompositions should not be seen as a stand-alone method for policy evaluations. However, conducting a productivity decomposition using the approach outlined can add real value towards understanding the impacts of a government intervention.

2. The technique work effectively when embedded within standard quasi-experimental evaluation methods. There are some caveats, because the approach is data intensive and – so far – the key useable datasets have relatively limited numbers of variables. However, the study recommends PSM using administrative data because it allows refining the matching process and statistical tests at various stages to ensure using a decomposition is feasible. The paper sets out four stages for an evaluation where a decomposition might be considered. In the first two, which focus on identifying a control group and establishing that there is the data needed to understand reallocative effects, the feasibility of using the approach for a particular intervention can be explored.

3. The decomposition itself is a complex but manageable form of analysis. After arriving at a robust estimate for the productivity effects of a given intervention through standard techniques, analysts can use productivity decompositions to shed light on the underlying dynamics of these effects; such as the re-allocation effects on productivity, as well as the entry and exit of firms. This builds on standard evaluation methods to estimate primarily the within effect alone.

4. One of the benefits of the technique is that it can be used to explore the indirect impacts, by testing with different control groups. The control groups should be based on theoretical expectations from the intervention logic model. Testing the technique on multiple control-groups will provide evidence on how effects permeate from the treated firm to the sector or region or other aspect associated with the intervention's theory of change through reallocation.

5. Using the decomposition method is thus a sequential one: 1) estimate treatment effects using standard techniques, 2) decompose this estimate using the decomposition techniques, 3) change control-group specifications. Productivity decomposition applied this way adds value by opening the "black box" of a standard within-firm estimation. Robustness is ensured both through triangulation of findings between methods and through testing different control-group specifications.

6. There are challenges for implementing the productivity perspective. The decompositions are data intensive, though any improvements in data can quickly be integrated into the method, such as capital stock or firm specific price data. However, the method requires the developing of an unbalanced panel where the supported businesses are matched with comparable businesses that did not benefit taking account of entry and exit. For this to be achieved, some recommendations with regard to the approach are:

    - Use the logic model of the intervention to understand whether wider economy impacts are likely to be significant and help construct datasets that reflect this;

    - The analysis puts a high premium on long time series measuring both supported and unsupported business and this suggests using administrative data;

- With regard to the design of the counterfactual firms, the report applied propensity score matching on the data as means of demonstrating the productivity decomposition; other econometric techniques may also be valid. As stated previously, the advantage of the decomposition approach demonstrated in this reports is that the control group's specification can be changed and the resulting analysis compared.

- There is some choice on productivity decomposition method. The report makes most use of the Foster-Haltiwanger-Kirzan (FHK) methods. This is commonly used techniques and so benefit from being relatively well understood. It has two advantages over other methods, important for the particular intervention. Firstly, it models entry and exit, something shown to be material in this case. Secondly, it separates the "within" firm and reallocative impacts in an intuitive manner.

7. Some initial findings are presented as the matching assumptions are changed. Broadly, the more comparable the businesses, the higher the treatment impact within the beneficiary businesses. These results have some caveats, but were such findings found to be robust they may reflect displacement – that the most comparable businesses are detrimentally affected by the support.

8. This study is primarily a feasibility study. A key question is whether using decompositions is improving on present measures of displacement: is the productivity decomposition approach better at accounting for displacement than adjusting for market displacement as is currently done or simply assuming full factor (employment) market displacement. The results presented in the report give some initial estimates for a support scheme. They indicate the types of analyses that are possible and how they might be interpreted, but considerable further work needs to be done to assure their robustness.

References

Ahn, S. (2001) "Firm Dynamics and Productivity Growth", OECD Economic Department Working Papers, 297.

Allinson, Gordon, Paul Robson and Ian Stone (2013) “Economic Evaluation of the Enterprise Finance Guarantee (EFG) Scheme”, Report to BIS, February.

Arbia, G. (1989) *Spatial Data Configuration in Statistical Analysis of Regional Economic and Related Problems.* Kluwer.

Arup (2001) “Evaluation of Regional Selective Assistance, 1991-1995”, Report to DTI, National Assembly of Wales and Scottish Executive from Arup Economics and Planning.

Attanasio, O. P. (2014) “Evidence on Public Policy: Methodological Issues, Political Issues and Example” Scandinavian Journal of Public Health, 42(Supp 13): 28-40

Baily, M.N., Hulten, C. and Campbell, D. (1992) "Productivity Dynamics in Manufacturing Plants" Brookings Papers on Economic Activity (Microeconomics): 187-247.

Barrios, Salvador, Bertinelli, Luisito and Strobl, Eric, 2003. "Coagglomeration and Growth," CEPR Discussion Papers 396, Centre for Economic Policy Research.

Bartelsman, E., Haltiwanger, J. and S. Scarpetta (2013) "Cross-Country Differences in Productivity: The Role of Allocation and Selection". American Economic Review, 103(1): 305-334.

Basu, Susanto and Fernald, John G. (2002) “Aggregate Productivity and Aggregate Technology.” European Economic Review, June, 46(6), pp. 963-91.

Bernard, A., J. Jensen, S. Redding and P. Schott (2007) “Firms in International Trade” Journal of Economic Perspectives. 21(2): 105-130

Bloom, Nicholas, Mark Schankerman and John Van Reenen (2013) “Identifying Technology Spillovers and Product Market Rivalry”, Econometrica, 81(4): 1347-93.

BIS, (2009a) “Research to improve the assessment of Additionality”, Department for Business, Innovation and Skills

BIS (2009b) “RDA Evaluation: Practical Guidance on Implementing the Impact Evaluation Framework”, December.

BIS (2011a) “International Trade and Investment – the Economic Rationale for Government Support.” BIS Economics Paper No 13

BIS (2011b) “Guidance on Evaluating the Impact of Interventions on Business”. BIS Working Paper, August, https://www.gov.uk/government/publications/impact-evaluation-guidance-for-business

Blundell, R. and MaCurdy, T. 1999. “Labour supply: a review of alternative approaches” in Ashenfelter, O. and Card, D. (ed), *Handbook of Labor Economics: Vol III*, Amsterdam.

Bryson, A. Richard Dorsett, Susan Purdon (2002) “Use of propensity score matching in evaluation of active labour market policies”. Working Paper for Department of Work and Pensions

Breinlich, H., G. Mion, P. Nolen, and D. Novy. (2012). “Intellectual Property, Overseas Sales, and the Impact of UKTI Assistance in Entering New Overseas Markets”. Report to UKTI.

Caballero, Ricardo, and Mohamad Hammour (1994) “The Cleansing Effects of Recessions,” American Economic Review, 84(5): 1356-68.

Camagni R. and Capello R. (2009), “Knowledge-based Economy and Knowledge Creation: The Role of Space”, in Fratesi U. and Senn L. (eds), *Growth and Competitiveness in Innovative Regions: Balancing Internal and External Connections*, Springer Verlag, Berlin, pp. 145–166

Capello R. (2009), “Spatial Spillovers and Regional Growth: A Cognitive Approach”, European Planning Studies, vol. 17, n. 5, pp. 639-658

CEE (2003) “Returns to Education: A Non-Technical Summary”, Centre of the Economics of Education CEE Work and Policy Discussion.

Coad, Alex, Marc Cowling and Joshua Siepel (2012) “Growth Processes of High-Growth Firms in the UK”, Nesta Working Paper 12/10

Collins, D and Balarajan, M (with A Bryson) (2011) “Survey Questions for Impact Evaluations which rely on beneficiaries self-assessment: evidence and guidance”. Report for BIS.

Cowling, Marc and Nick Isles (2005) “Sent to Coventry? The re-employment of the Longbridge 5,000”, Report by the Work Foundation. http://www.theworkfoundation.com/downloadpublication/report/137_137_sent%20to%20coventry.pdf

Crepon, Bruno, Esther Duflo, Marc Gurgand, Roland Rathelot and Philippe Zamora (2012) “Do Labour Market Policies have Displacement Effects? Evidence from a Clustered Randomized Experiment.” NBER Working Paper 18597, December.

Criscuolo, Chiara, Ralf Martin, Henry Overman and John Van Reenen (2012) “The Causal Effects of an Industrial Policy”, NBER Working Paper 17842

Dahlberg, Matz and Anders Forslund (1999) “Direct Displacement Effects of Labour Market Programmes: The Case of Sweden”. Uppsala University Economics Working Paper 1999-22.

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

Driffield, N. Jun Du, Mark Hart, Jim Love and Stathis Tapinos (2010) “A Comparative Evaluation of the Impact of UK Trade & Investment’s R&D Programme and Other UKTI Support that Impacts R&D”. Report to UKTI.

Drews, Cord-Christian and Mark Hart (2014) “Feasibility Study: Exploring the Long-Term Impact of Business Improvement Services”. ERC Research Paper, University of Aston.

Ecorys (2012) “Assessment of the Business Link Helpline”. Report to BIS, December.

English Partnerships (2008) “Additionality guide”, October

European Commission (1999) “Thematic Evaluation of Structural Fund Impacts on SMEs”, Synthesis Report to the European Commission, July. http://ec.europa.eu/regional_policy/sources/docgener/evaluation/pdf/eval_sme.pdf

Foster, L., Haltiwanger, J. and Krizan, C.J. (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.

Gehlke, C. E.; Biehl, Katherine (1934). "Certain effects of grouping upon the size of the correlation coefficient in census tract material". Journal of the American Statistical Association 29 (185A): 169–170

Gerber and Green (2012). *Field experiments: design, analysis and interpretation***.** New York: W W Norton and Company

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.

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

Hall, Robert (1992) “Invariance Properties of Solow’s Productivity Residual” In Peter Diamond (ed.) *Growth, Productivity, Employment*. Cambridge: MIT Press.

Hanson, Andrew and Shawn M. Rohlin, 2013. "Do Spatially Targeted Redevelopment Programs Spill-Over". Regional Science and Urban Economics, 43 issue 1, pp. 86-100

Harris, R. and Q C Li (2013) “Study of the Relationship between Innovation, Exporting and the use of E-commerce (CIS5)”. January, Report to UKTI

Harris, R., and C. Robinson (2004) “Industrial Policy in Great Britain and its Effects on Total Factor Productivity in Manufacturing Plants, 1990-98”, Scottish Journal of Political Economy, 51(4), September.

Harris, R. and Robinson, K. 2005. "Impact of Regional Selective Assistance on Sources of Productivity Growth: Plant-Level Evidence from UK Manufacturing, 1990-98". Regional Studies, 39(6): 751-765.

Helpman, E., Melitz, M.J., Yeaple, S. (2003) “Export versus FDI”. Harvard Institute of Economic Research, Discussion Paper 1998.

HMT (2003) *The Green Book – Appraisal and Evaluation in Central Government***,** London: HM Treasury

Holt, David, David Steel, Mark Tranmer, Neil Wrigley (1996) "Aggregation and ecological effects in geographically based data". Geographical Analysis 28 (3): 244–261.

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

Iacovone, Leonardo & Beata S. Javorcik (2008) "Multi-product exporters: diversification and micro-level dynamics," Policy Research Working Paper Series 4723, The World Bank.

Imbens, Guido W. and Jeffrey M. Wooldridge (2009) “Recent developments in the econometrics of program evaluation”. Journal of Economic Literature 47(1): 5-86.

Lenihan and Hart (2002) “Evaluating the Impact of Enterprise Ireland Assistance: methodological considerations when estimating deadweight and displacement.”

London Economics (2012) “A review of the evidence base for UK Trade & Investment’s support for firms in High Growth Markets”. Final Report to UK Trade & Investment

London Economics (2012a) “Evaluation of the impact and cost effectiveness of UKTI’s regional network support”. Report to UKTI.

London Economics (2012b) “Evaluation of the Fiscal Stimulus Initiative”. Report to UKTI.

Markusen, James, and Anthony J. Venables (1999) “Foreign Direct Investment as a Catalyst for Industrial Development”. NBER Working Paper No. 6241; European Economic Review, Vol. 43: 335-356

Mason G., Bishop K. and Robinson C. (2009) “Business Growth and Innovation: The wider impact of rapidly-growing firms in UK city-regions”, National Endowment for Science, Technology and the Arts, Nesta.

Melitz, M. and Polanec, S. 2012. "Dynamic Olley-Pakes Productivity Decomposition with Entry and Exit". NBER Working Paper 18182.

Mole, Kevin, Mark Hart, Stephen Roper, and David Saal (2008) “Assessing the Effectiveness of Business Support Services in England: Evidence from a Theory Based Evaluation”. Small and Medium Sized Enterprise Research Centre Working Paper 93, Warwick Business School.

Olley, G. S. and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry". Econometrica, 64(6): 1263-1297.

ONS (2001). “Review of the Inter-Departmental Business Register”. National Statistics Quality Review Series Report No. 2, Office for National Statistics.

ONS (2010). “Annual Business Inquiry: Background Information”. Technical report, Office for National Statistics, Newport.

Openshaw, Stan (1983). *The modifiable areal unit problem*. Norwick: Geo Books

Optimal Economics (2012) “Tourism Marketing Return on Investment: The Impact of Displacement”. Report to DCMS.

PA Consulting, and SQW Ltd. (2006). “Evaluating the Impact of England’s Regional Development Agencies: Developing a Methodology and Evaluation Framework”. DTI Occasional Paper 2

Public and Corporate Economic Consultants (2011) “Evaluation of the Collaborative Research and Development Programme”, Technology Strategy Board

PA Consulting and SQW (2005) “Developing a Methodology and Evaluation Framework”. Report to BIS.

Riley, R. Chiara Rosazza-Bondibene and Garry Young (2014) “The Financial Crisis, Bank Lending and UK Productivity: Sectoral and Firm-Level Evidence”, National Institute Economic Review 228.

Roper, S., N. Hewitt-Dundas, and J.H. Love (2004). “An ex ante evaluation framework for the regional benefits of publicly supported R&D projects”, *Research Policy*, vol. 33, pp. 487-509.

Sena, V., M. Rogers, and N. Smith. (2011) “The Role of Foreign Direct Investment in the UK IP and the Impact of UKTI Support”. Report to UKTI.

Sena, Vania, Mark Hart and Karen Bonner (2013) “Innovation and UK high-growth firms’, Nesta Working Paper 13/12, November.

SQW Consulting (2009) “Pushing the Boundaries of Impact Evaluation.” SQW Limited.

Syverson, Chad (2011) “What Determines Productivity”, Journal of Economic Literature, 49(2): 326-365

Tervo, Hanno (1990) “Factors Underlying Displacement: an analysis of Finnish Regional Incentive Policy Using Survey Data on Assisted Firms”. Applied Economics, 22(5): 617-28.

Torgerson, D. J. and Torgerson, C. J. (2008) *Designing randomised trials in health, education and the social sciences: An introduction*. London: Palgrave Macmillan

Annex A: Intervention Example

Practical Application for the Feasibility Study

1. Data from a recent business support policy – the British Business Bank's Enterprise Finance Guarantee (EFG) – is used in this feasibility study to demonstrate some of the empirical approaches used to evaluate economy-wide impacts. The support was to improve access to finance particularly for small and medium enterprises. This report does not represent an evaluation of the EFG scheme, but uses it to demonstrate the approaches outlined.

2. The management information about the intervention has been linked to ONS administrative datasets. Where the links have been successful, the administrative data can be used to track employment and turnover for at least a decade and to identify when the business started trading or when it ceased to trade. Matching has been high. A key feature of the ONS source is that UK data on employment and turnover is comprehensive for all businesses. This means that the key performance variable of labour productivity is available for beneficiaries and non-beneficiaries of business support. (Data on other outcomes, such as investment, research and development is collected but the coverage is more limited especially for small and medium sized businesses.)

Enterprise Finance Guarantee

3. EFG was introduced in January 2009 in response to the credit crunch, as a comprehensive and wide-ranging replacement for the long established Small Firms Loan Guarantee (SFLG) scheme. The rationale for both schemes has been to address the market failure in the provision of debt finance. Viable businesses may fail to obtain normal commercial loans, because they lack adequate security or a proven financial track record.

4. Allinson et al. (2013) is an evaluation of the Guarantee, using a survey designed to collect information on additionality by including questions on deadweight and market displacement amongst EFG supported businesses. This evaluation provides evidence confirming the rationale for the EFG scheme, namely that it is being targeted at small businesses affected by lack of security against which to borrow. Such businesses would otherwise not be able to access conventional bank loans and this is especially important when the supply of credit is constrained in the economy. The supported businesses are not found to be ailing or weaker than other businesses in terms of their growth performance. Their subsequent business performance is at least as good as other borrowing businesses, but EFG acts as an enabling catalyst for these businesses affected by a lack of collateral or track record to help them fulfil their growth potential.

5. The logic of the intervention suggests that some reallocative impacts are possible. In Allinson et al., displacement is estimated using two questions in the telephone survey. Firstly, businesses were asked whether they competed with local and national firms and secondly they were asked whether, if they ceased trading immediately, all of their sales would be taken up by a UK based company within one year. The number of businesses that indicated all their sales would be taken up by local competitors was 24.5 per cent. EFG supports small and medium enterprises so wider impacts are also

likely if entry or exit amongst supported businesses is correlated with productivity differences.

6. Evaluation of the EFG and, indeed, the original case for the Small Loans Guarantee, makes no mention of spillovers or externalities due to the support. The objective of the intervention is not related to innovation or clustering and the use of loan guarantees is not material to foreign direct investment decisions. So, in assessing any wider economy impacts, the approach need not consider externalities, as these appear unlikely to occur.

Direct Impacts in this Study's Illustration

7. The EFG beneficiaries were linked to the Business Structures Database and a panel of borrowers was matched to comparable businesses. The treatment appeared to create about three jobs per borrower. This quantifies the scale effect of the treatment, adding about 15% to the size of treated businesses over and above comparable groups.

8. These results have not undergone the same level of scrutiny as a full evaluation. However, they are consistent with the careful evaluation undertaken by Allinson et al. (2013). That evaluation used businesses self-reported assessment of performance and scheme impact. The analysis used econometric techniques to control for any differences in the EFG recipient businesses compared to the general population. As the intervention was largely to make credit available to businesses that had constrained access to finance, the evaluation also looked at businesses that reported no problems accessing finance by surveying businesses that had borrowed money without any government support.

9. Overall, they find that EFG beneficiaries appear to have fared well in terms of employment and sales growth compared to non-borrowers, but grew at a lower rate than other borrowers. The analysis takes account of deadweight and displacement and estimates benefits to the economy over a 2-3 year time duration. For the 6,700 participants drawing down an EFG loan in 2009, 0.96 jobs created per recipient business and 1.84 jobs saved per recipient business, results similar to those in Table 4 that use administrative data in a simpler treatment-control framework.

Using EFG to Illustrate Decompositions

10. The analysis in this study is not an exhaustive evaluation of EFG in three important dimensions. Firstly, an evaluation should begin with the interventions logic model or theory of change to direct what evaluation approach should be taken and the methods suitable for the evidence gathering. This grounds the evaluation in an understanding of the different impacts envisaged, both positive and negative, the timing of these impacts and then helps decide the correct evaluation approach to be used. Overall, this means a balanced picture of impacts can be derived.

11. The present work has used the management information available about EFG but in a limited manner. A second departure from a full evaluation is this study has not used key variables about the support, such as the size of support or its precise structure. Interventions will have various processes associated with the disbursement of support as well, which provide vital information to assess impact.

12. Finally, there are a number of key improvements to the matching to a counterfactual that could be envisaged. Recent evaluations have sought to use information about rejected applicants, or the past behaviour of businesses seeking support to improve the selection of a control group of businesses.

Annex B: Determining a Control

Methods to determine a counterfactual

Randomised Control Trials (RCTs)

1. In essence, RCTs are comparative studies in which participants are allocated at random to either receive or not receive an intervention. Such trials have long been a staple of medical research but are increasingly being employed in social and economic studies. RCTs are seen as the gold standard in establishing impact. As Collins and Balarajan (2011) state, if executed properly, RCTs ensure that there are no systematic differences between participants and non-participants on either observed or unobserved characteristics.

2. As such, RCTs can offer policymakers the prospect of learning about the effectiveness of programmes that avoid biases which plague other evaluation approaches. However, by necessity RCTs place a heavy emphasis on avoiding bias at the design phase. Unlike other approaches to understanding policy impact (e.g. quasi-experimental approach), the focus is on designing-out bias at the outset, rather than necessarily adjusting for them statistically at the end of a study.

3. There is a large literature on some of the challenges and biases that can inadvertently frustrate the objective of achieving a robust and credible study. There is also a growing body of work setting out some of the approaches that can be taken should certain challenges materialise. It is important that these challenges are considered and addressed in order for data collection and analysis to be rigorous and robust.

4. Attansio (2014) reviews some of the issues for evaluation of attributing impacts to the intervention using RCTs. A first key issue is that in many circumstances an RCT is impractical. Where an intervention cannot exclude a randomly selected sub-sample, then the RCT is not feasible. Attansio also considers the situation where a treatment indirectly affects those in the control group. A training programme may mean the individuals randomised out of treatment face much stiffer competition for employment than had the treatment not occurred. The difference in outcomes between the supported and control group would overstate treatment effect if this affects the employment outcome.

5. Torgerson and Torgerson (2008) provide an accessible account of some of the practical difficulties often faced in ensuring an RCT maintains credibility. These include:

    - Subversion – where stakeholders who are relied upon to cooperate with the trial deliberately sabotage the design – i.e. allocation can end-up being non-random.
    - Attrition – particularly where the patterns of sample loss over time differ between treatment and control samples.
    - Recruitment – it is important to consider whether the study sample covers the population within which estimates of causal effects are being assessed. This is an issue of external validity rather than internal validity but also whether one can estimate the correct parameter of interest.

- Non-compliance – this can be single-sided where the treatment group do not take up the support on offer, or two-sided, where controls access treatments (Gerber and Green, 2012).
- Resentful demoralisation and compensatory rivalry - where human subjects adjust their behaviour as a result of knowledge of their status with respect to randomisation.

6. The final set of difficulties, with regard to rivalry, is similar to indirect impacts. They highlight that, despite the strong preference for using RCTs in evaluations and a growing number of studies using this design, the worry that the control group may be adversely affected by the treatment has to be explored. The rest of this study focuses more on the quasi experimental approaches to identifying a control group, but many of the methods used will have cross over to the RCT design.

Quasi-Experimental Methods

1. Bryson et al. (2002) describe the evaluation problem quasi-experimental approaches seek to resolve. In an ideal world, the evaluator would like to establish the outcome for a firm with and without the intervention. After an intervention, it is possible to observe outcomes for the firm. But, of course, the performance of a firm without that intervention is unknown. Simply using non-users of the intervention has problems, in particular, because users and non-users are likely to differ in economic performance.

2. Propensity score matching addresses this problem by constructing a comparison group of non-users that are similar to the users along a range of characteristics not directly influenced by the intervention. So, if the firms being supported are generally large, the matching process will seek out firms that are not treated but are a similar size. Propensity score matching is a quasi-experimental approach in trying to find a distribution of firm characteristics that would be independent of whether a firm used an intervention or not.

3. Table B1 indicates studies where a propensity score matching is used to provide a control group from an administrative dataset. Breinlich et al (2012) evaluates UK Trade and Industry's Overseas Market Introduction in Service (OMIS) using the approach. They use large-scale business accounts datasets, FAME, which provides a population of all businesses that did not receive support. OMIS beneficiaries are larger than the average business and therefore a propensity score match is used to generate comparators. The differences in size may influence the performance of businesses over and above that attributable to the intervention. Breinlich et al (2012) note that the technique provides a matched sample to control for these differences but only using the observable characteristics in the dataset.

4. Table B2 indicates studies where the control group was determined using statistical matching but where some of the limitations of using administrative data only were overcome. It is often the case that some specific data is needed about non-beneficiaries so that a survey covering both beneficiaries and non-beneficiaries is conducted. Mole et al (2008) use responses to the Business Link Operator (BLO) Economic Impact Survey in 2005 and derive a control group after the survey. A group of non-assisted firms was surveyed, the sampling frame drawn from the Dun & Bradstreet UK database. Then corrections to this random sample needed to provide a suitable control are undertaken using ex post matching.

5. An alternative route used in the London Economics study into UKTI's international trade advice is to use propensity score matching ex ante to adjust a survey sample to be as similar to the treated group as possible. Surveys were then conducted to collect the data from both beneficiaries and the businesses in the control group. This relies on having a register or some other sampling frame that provides appropriate data on non-beneficiaries.

6. A problem for the propensity score technique is its reliance on observed evidence to perform matching. If there are unobserved variables that determine whether a business participates in an intervention and those aspects of the business correlate with business performance, then the evaluation will be biased. This may justify ex post matching as in the BLO study. Data is collected to both understand the non-beneficiaries and also match a control group from the surveyed non-beneficiaries. The modelling firstly analyses the decision to participate in the intervention using a probit model and in a second stage then uses that modelling to adjust the performance outcome for the evidence on selection that the first stage provides, the Heckman approach.

Table B1: Studies using Quasi-Experimental Methods

Study	Breinlich et al (2012)	Sena et al (2011)	Criscuolo et al. (2012)	London Economics (2012)
Intervention	UKTI OMIS	UKTI Inward Investment Assist	Regional Support Assistance	High Growth Markets Programme
Period	2006-8	2005-7	1986-2004	2007-8
Primary data source	ORBIS, a database derived from Companies House	FAME, a Companies House derived database	ONS Annual Respondents Database	FAME, a Companies House derived database
Other data	Oxford Firm-Level Intellectual Property DB	Oxford Firm-Level Intellectual Property DB	ONS surveys	Internationalisation Survey 2008
Businesses in sample	14,477 of 20,730	1,110 of 1,800	41,828 of 54,322	403 of 412

7. An important issue in operationalising the Heckman type model is the avoidance of too much overlap between the selection and performance models. In the probit models, the focus is on external characteristics of the firm which may have been visible ex ante, and which may have provided the basis for administrative criteria for the targeting of assistance.

8. A related approach of identifying causal effects of regional policy is conducted in Criscuolo et al. (2012). They use micro level data to construct a quasi-experimental framework to identify the causal effects of the UK's Regional Selective Assistance programme on firm performance. They generate an instrument for recipient status of state aid by exploiting changes in the area-specific eligibility criteria. The eligibility

criteria in the UK are determined by the European Commission's guidelines for regional development policies that also underlie the Structural Funds Programme. The revision of regional eligibility for structural funds before each programming period also determines the provision of Regional Selective Assistance to firms in the UK and may therefore be used as an exogenous instrument. The authors find a significant positive effect of state aid on investment as well as on employment. The majority of regional aid also ends up going to larger firms because they tend to be more effective at obtaining subsidies (Criscuolo et al., 2012).

Table B2: Studies using Score Matching with Beneficiaries Survey

Study	Mole et al (2008)	London Economics (2012a)	London Economics (2012b)	Driffield et al. (2012)
Intervention	Business Link	UKTI International Trade Advice	UKTI Fiscal Stimulus Initiative advice	UKTI Inward investment advice and R&D
Period	2006-8	2010-12	2009-2012	2007
Primary data source	Beneficiary survey and Dun and Bradstreet non-beneficiary sample	User and non-user surveys, 300 of each	Beneficiary survey (150) and Dun and Bradstreet non-beneficiary (500) sample	User and non-user surveys, 400 of each. PIMS
Key collected data	Employment, turnover, management actions	Employment, turnover	Employment, turnover, views on access to advice on FSI markets	R&D
Matching	Ex post Heckman match	Ex ante propensity score match		Ex post Heckman match

9. Methodologically, the approach has several layers all using sensible instruments to correct for evaluation problems. A first issue is that the intervention is likely to be correlated with firm specific, unobserved shocks. The RSA is likely to target firms facing difficulty so under-estimating the effect of the treatment on the treated (TT). The instrument used to estimate a consistent ATE is the maximum amount of intervention available, recognising that in firms outside the selected areas this would be zero. The study then notes that many businesses would be multi-plant and a second model is then estimated where the instrument is applied recognising this and using plant level data. This estimates the LATE.

10. As instruments, organisational factors are chosen, which may initially have been unobservable but which may nonetheless have contributed to performance. Driffield et al. select variables for their first stage modelling and explain the logic. The survey asks about whether businesses have a business plan – this is typically linked to the ability to obtain funding but independent of performance. What firms were seeking to do R&D for – in terms of sales, innovation, exports etc. These appear correlated with UKTI support, but uncorrelated with R&D spend.

Annex C: Propensity Score Matching

The PSM Method

1. While RCTs are often considered the 'gold standard' of policy evaluation, they are often not feasible on ethical grounds, or because considerations for evaluation are only given after implementation. While there are other options to construct a control group, this study uses propensity score matching (PSM). The following section will outline the formal PSM model and then present some results regarding the quality of the PSM models used in this study.

2. The treatment variable Tr_{it} is equal to 1 if an enterprise i receives treatment in period t. Let $\Delta y^{1}_{i,t+1}$ be the productivity of enterprise i at time t+1 after receiving treatment in period t and let $\Delta y^{0}_{i,t+1}$ be the hypothetical performance of the same enterprise i at the same time t+1 had it not received treatment in period t. The productivity effects of the support in enterprise i, called the average treatment effect on the treated, can then be expressed as follows:

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

3. The second term in equation (2), $- E\left(\Delta y^{0}_{i,t+1} \middle| Tr_{it} = 1\right)$ is the counterfactual mean or the hypothetical productivity change in enterprises treated, 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 similar to the treatment group in all respects except the exposure to the treatment (Ravallion, 2003). The estimation of the causal effect in this case becomes:

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

where $E\left(\Delta y^{1}_{i,t+1} \middle| Tr_{it} = 1, X_{i,t-1}\right)$ is the mean productivity change at time t+1 of the enterprises receiving support at time t; $E\left(\Delta y^{0}_{i,t+1} \middle| Tr_{it} = 0, X_{i,t-1}\right)$ is the mean productivity change of the control group at time t+1; and $X_{i,t-1}$ 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 the support on productivity. A practical constraint arising from the application of such matching techniques is that exact matching across multiple covariate indexes 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 treatment becomes:

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

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

4. For consistent estimates of the productivity effects, two key assumptions must hold: the conditional independence assumption and the common support assumption. Conditional independence means that there are no unobservable differences between treated and non-treated enterprises after conditioning for $X_{i,t-1}$, 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:

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

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 productivity 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).

5. 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:

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

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

7. 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 treatment is estimated, i.e.:

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

where Tr_{it} is the dummy for first treatment; $X_{i,t-1}$ includes all enterprise level variables that affect the probability of receiving treatment, $Z_{i,t-1}$ controls for external factors that are likely to affect the probability of receiving treatment and Tr_{it} 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.

8. 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:

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

9. A common problem in PSM is the occurrence of bad matches (i.e. the nearest neighbour is not very near). Given the size of the matching pool this should not be a problem for the beneficiaries in the case of EFG. 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 treatment (i.e. matches between 2012 beneficiaries and 2013 beneficiaries are not possible). Further, the common support assumption, discussed earlier, rules out the perfect predictability of treatment given the observed covariates. This ensures the existence of potential matches in the pool of non-treated enterprises from the BSD.

10. Another assumption that is required for matching and all of the other partial equilibrium estimation strategies is the so-called SUTVA assumption, where SUTVA means stable unit treatment value assumption. This assumption says that the impact of the support on one case does not depend on whom else, or on how many others, are in the support programme. As Sianesi notes (2001) SUTVA is in fact the assumption that the model's representation of outcomes is adequate, that is that the observed outcome for an individual exposed to treatment depends only on the individual and not on what treatments other individuals receive nor on the mechanism assigning treatment to individuals and that whether the individual participates only depends on the individual.

11. Propensity score matching provides an estimate of the effect of a "treatment" variable on an outcome variable that is largely free of bias arising from an association between treatment status and observable variables. However, matching methods are not robust against "hidden bias" arising from unobserved variables that simultaneously affect assignment to treatment and the outcome variable. One strategy for addressing this problem is the Rosenbaum bounds approach, which allows the analyst to determine how strongly an unmeasured confounding variable must affect selection into treatment in order to undermine the conclusions about causal effects from a matching analysis. Instrumental variables (IV) estimation provides an alternative strategy for the estimation of causal effects, but the method typically reduces the precision of the estimate and has an additional source of uncertainty that derives from the untestable nature of the assumptions of the IV approach. A method of assessing this additional uncertainty is proposed so that the total uncertainty of the IV approach can be compared with the Rosenbaum bounds approach to uncertainty using matching methods. Because the approaches rely on different information and different

assumptions, they provide complementary information about causal relationships. (DiPrete and Gangl, 2004).

Alternative matching algorithms

12. 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 observations are weighted by the distance of their propensity scores. 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 variable. Because variables have been simplified, there is a finite number of bins – the permutations that the different coarsened variables take – and matching becomes possible.

13. 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). For the current study, nearest neighbour matching has been chosen for a number of reasons. First, this is the most straight-forward method and requires no arbitrary choices (for example the size of a radius, or the level of coarsening). It performs reasonably well where the number of potential matches is high, so that propensity scores will be very close most of the time. This is the case with the data used here, with the BSD 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 & Kopeinig, 2005).

14. The fact that matching is performed on a single metric and needs searching only for the closest score also reduces the need for computational power, which is an important consideration when dealing with a large database. Finally, and perhaps most importantly, nearest neighbour matching performed well in reducing the bias between treatment and control group, which is the goal of PSM. To assure that this was the case, several tests of matching quality have been performed, which are discussed next.

Probit selection models

15. The factors leading to selection into support can be described by a probit model. These selection models are presented in the different panels of Table C1. Reassuringly, the estimation results are consistent across sectors of the economy, years, and model specifications.

Table C1.1: Probit model for selection into EFG in 2009

	Total				Services			
	β	SE	β	SE	β	SE	β	SE
Turnover (ln)	0.18	0.01	0.18	0.01	0.18	0.01	0.18	0.01
Employment (ln)	0.16	0.01	0.16	0.01	0.12	0.01	0.12	0.01
Employment share	-42836.02	7169.21	-42439.38	7143.47	-29502.86	5846.13	-29238.98	5825.24
Turnover share	-52593.69	6205.84	-52880.69	6205.82	-23741.89	4436.75	-23855.77	4436.75
Age	-0.04	0.00	-0.04	0.00	-0.04	0.00	-0.04	0.00
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
H-index	0.05	0.10	0.04	0.10	-0.17	0.14	-0.19	0.14
Low pay	-0.07	0.02	-0.07	0.02	-0.04	0.02	-0.04	0.02
High-tech	0.15	0.02	0.15	0.02	0.15	0.02	0.15	0.02
Constant	-3.79	0.05	-3.83	0.04	-3.73	0.06	-3.79	0.04
Regional controls	GOR		TTWA		GOR		TTWA	
Number of obs.	1651406		1651426		1191574		1191591	
Log likelihood	-18086.511		-18108.206		-13420.108		-13429.565	
Pseudo R2	0.10		0.10		0.08		0.08	

Table C1.2: Probit model for selection into EFG in 2011

	Total				Services			
	β	SE	β	SE	β	SE	β	SE
Turnover (ln)	0.19	0.01	0.19	0.01	0.19	0.01	0.19	0.01
Employment (ln)	0.12	0.01	0.12	0.01	0.10	0.01	0.10	0.01
Employment share	-18262.88	7725.06	-18232.85	7706.82	-11338.20	6111.17	-11172.23	6079.39
Turnover share	-94478.33	12454.65	-94772.41	12450.34	-53583.52	9361.26	-53646.74	9348.49
Age	-0.04	0.00	-0.04	0.00	-0.04	0.00	-0.04	0.00
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
H-index	0.38	0.11	0.37	0.11	0.12	0.15	0.10	0.15
Low pay	0.03	0.02	0.03	0.02	0.06	0.02	0.06	0.02
High-tech	0.10	0.02	0.10	0.02	0.08	0.03	0.09	0.03
Constant	-4.13	0.07	-4.14	0.05	-4.08	0.08	-4.08	0.06
Regional controls	GOR		TTWA		GOR		TTWA	
Number of obs.	1949935		1949952		1438333		1438348	
Log likelihood	-10366.664		-10383.353		-8034.2695		-8041.3215	
Pseudo R2	0.09		0.09		0.08		0.08	

Table C1.3: Probit model for selection into EFG in 2013								
	Total				Services			
	β	SE	β	SE	β	SE	β	SE
Turnover (ln)	0.18	0.01	0.18	0.01	0.21	0.02	0.20	0.02
Employment (ln)	0.17	0.02	0.17	0.02	0.13	0.02	0.13	0.02
Employment share	-71498.62	16977.76	-71172.65	16944.19	-47431.38	14701.09	-47302.27	14686.54
Turnover share	-154431.80	23368.10	-155033.90	23394.95	-121599.80	20641.38	-122130.10	20672.09
Age	-0.03	0.00	-0.03	0.00	-0.03	0.00	-0.03	0.00
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
H-index	0.38	0.12	0.37	0.12	0.24	0.15	0.24	0.15
Low pay	0.05	0.02	0.05	0.02	0.08	0.02	0.08	0.02
High-tech	0.04	0.03	0.03	0.03	0.01	0.03	0.01	0.03
Constant	-4.19	0.08	-4.29	0.06	-4.28	0.09	-4.36	0.07
Regional controls	GOR		TTWA		GOR		TTWA	
Number of obs.	2505893		2505975		1880512		1880580	
Log likelihood	-8568.599		-8566.6419		-6730.6289		-6726.9763	
Pseudo R2	0.08		0.08		0.08		0.08	

Note: Columns for "Total" and "Services" refer to the total economy and only the services sector included in the estimation, respectively. Different model specifications included either Government Office Regions (GOR) or Travel to Work Areas (TTWA) as regional controls.

Testing the quality of the matching results

16. 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 balance tests – look at whether the two groups are similar in terms of their characteristics after matching. Table C2 presents results regarding how similar the two groups are in terms of the average for key characteristics. A second set of tests focuses on what may be missed in the selection modelling. In particular, they look at how sensitive the average treatment effect is to different assumptions about unobserved variables when matching to generate a control group. These so-called Rosenbaum bounds tests are illustrated in Table C3.

Balancing tests

17. 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.

18. Running this test, called PS-test, before and after matching, it can be assessed what reduction in bias was brought about by PSM. Results are presented in tables C3. Ideally, the bias in the matched sample should be close to zero. Indeed, only very few variables still show a significant bias after matching. Only in the total economy sample in 2009 is there more than one variable that still has a bias after matching; in all other samples, there was at most one variables with such problems.

Table C2: PS-tests on matching balance

Sector	Variable	2009 Means of matched samples Treated	2009 Means of matched samples Control	2009 Bias in matched sample (%)	2011 Means of matched samples Treated	2011 Means of matched samples Control	2011 Bias in matched sample (%)	2013 Means of matched samples Treated	2013 Means of matched samples Control	2013 Bias in matched sample (%)
Total	Turnover (ln)	6.31	6.41	-6.8**	6.13	6.22	-6	6.11	6.17	-4
	Employment (ln)	1.92	2.00	-6.3*	1.78	1.84	-5.7	1.82	1.83	-1.1
	Total emp. share	0.00	0.00	-1.5*	0.00	0.00	-1.4	0.00	0.00	0.2
	Total turnover share	0.00	0.00	-0.6	0.00	0.00	-0.8*	0.00	0.00	-0.1
	Relative productivity	-0.01	0.07	-9.1***	-0.07	-0.03	-4.4	-0.12	-0.04	-9.1**
	Age	9.08	9.52	-5	6.92	7.41	-5.7	8.36	8.92	-6.1
	Age^2	151.11	169.07	-5.9**	107.91	129.41	-8**	140.73	165.55	-7.8
	H-index	0.02	0.03	-4.3	0.03	0.03	-5.3	0.03	0.04	-4.8
	Low pay	0.32	0.37	-11.5***	0.41	0.45	-7.4*	0.44	0.44	-0.2
	High-tech	0.17	0.15	6.9**	0.14	0.11	7.6*	0.11	0.10	1.1
	Employment	18.72	24.76	-1.5*	15.62	19.94	-1.4	14.34	13.84	0.2
	Turnover	1783.60	2312.90	-0.6	1239.60	1840.40	-0.8*	1073.40	1149.00	-0.1
Services	Turnover (ln)	6.31	6.34	-2.4	6.13	6.16	-1.8	6.11	6.08	2
	Employment (ln)	1.92	1.96	-2.8	1.78	1.80	-2.1	1.82	1.81	0.8
	Total emp. share	0.00	0.00	-1	0.00	0.00	-0.9	0.00	0.00	0.4
	Total turnover share	0.00	0.00	-0.5	0.00	0.00	-0.5	0.00	0.00	0.2
	Services emp. share	0.00	0.00	-1	0.00	0.00	-0.9	0.00	0.00	0.4
	Services turnover share	0.00	0.00	-0.5	0.00	0.00	-0.5	0.00	0.00	0.2
	Relative productivity	-0.01	0.05	-6.2**	-0.07	-0.02	-5.3	-0.12	-0.07	-5.3
	Age	9.08	9.02	0.7	6.92	6.95	-0.4	8.36	8.72	-4
	Age^2	151.11	153.52	-0.8	107.91	112.07	-1.5	140.73	153.54	-4
	H-index	0.02	0.03	-6.2*	0.03	0.03	-3.1	0.03	0.04	-5.8
	Low pay	0.32	0.36	-7.7**	0.41	0.47	-12.5***	0.44	0.46	-3
	High-tech	0.17	0.16	3.8	0.14	0.12	4.5	0.11	0.09	5.7
	Employment	18.72	22.88	-1	15.62	18.45	-0.9	14.34	13.05	0.4
	Turnover	1783.60	2237.50	-0.5	1239.60	1610.10	-0.5	1073.40	957.89	0.2

Note: Significance levels of bias in matched sample: ***<0.01, **<0.05, *<0.1.

Sensitivity tests

19. 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.

20. However, the analysis may be biased if there are unobservable variables that drive selection into treatment as well as the outcome, employment growth. Typically, one such variable is managerial capability and ambition. If more ambitious managers are more likely to apply for support, the supported businesses would differ in an important respect from untreated businesses, and it may be the manager's ambition rather than the treatment that causes higher productivity growth.

21. 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).

22. 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.

23. The panels of Table A5 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.

24. The results for 2009 indicate the level of bias would have to quite high before the estimated firm level productivity growth, used as the outcome variable, would be significantly different from that estimated were no bias assumed. This suggests the 2009 results are robust. The results for 2011 and 2013 are less clear cut. At modest levels of bias, the PSM matches would give different results with regard to productivity comparisons between the control and treated groups.

25. Apart from the underlying matching being problematic, there are two other explanations for this. The first is that the period being considered becomes quite short, especially matching in 2013 and then exploring for significant differences between 2013-14. The problem is that the outcome may yet be very robust. A second issue is that the method employed in this study – productivity decomposition – is not a conventional outcome

variable, in being firm specific tracked over time. The study looks at reallocation across businesses. The study has not found approaches to using this test on such outcomes focusing on the correlation across firms in productivity change and employment change.

Table C3.1: Rosenbaum bounds test for 2009

	Total Economy								Services							
	Controlling for GOR				Controlling for TTWA				Controlling for GOR				Controlling for TTWA			
	Significance of selection effect		Point estimate		Significance of selection effect		Point estimate		Significance of selection effect		Point estimate		Significance of selection effect		Point estimate	
Γ	Positive	Negative	Upper	Lower	Positive	Negative	Upper	Lower	Positive	Negative	Upper	Lower	Positive	Negative	Upper	Lower
1.00	0.00	0.00	0.06	0.06	0.00	0.00	0.09	0.09	0.00	0.00	0.07	0.07	0.00	0.00	0.10	0.10
1.02	0.00	0.00	0.05	0.07	0.00	0.00	0.08	0.10	0.01	0.00	0.06	0.08	0.00	0.00	0.09	0.11
1.04	0.01	0.00	0.04	0.08	0.00	0.00	0.07	0.11	0.02	0.00	0.05	0.09	0.00	0.00	0.09	0.12
1.06	0.04	0.00	0.04	0.09	0.00	0.00	0.06	0.11	0.04	0.00	0.04	0.10	0.00	0.00	0.08	0.13
1.08	0.09	0.00	0.03	0.10	0.00	0.00	0.05	0.12	0.09	0.00	0.03	0.11	0.00	0.00	0.07	0.14
1.10	0.17	0.00	0.02	0.11	0.01	0.00	0.05	0.13	0.16	0.00	0.02	0.11	0.00	0.00	0.06	0.15
1.12	0.30	0.00	0.01	0.11	0.03	0.00	0.04	0.14	0.26	0.00	0.02	0.12	0.01	0.00	0.05	0.16
1.14	0.44	0.00	0.00	0.12	0.07	0.00	0.03	0.15	0.38	0.00	0.01	0.13	0.03	0.00	0.04	0.16
1.16	0.60	0.00	-0.01	0.13	0.15	0.00	0.02	0.16	0.51	0.00	0.00	0.14	0.06	0.00	0.04	0.17
1.18	0.74	0.00	-0.01	0.14	0.25	0.00	0.01	0.16	0.63	0.00	-0.01	0.15	0.11	0.00	0.03	0.18
1.20	0.84	0.00	-0.02	0.15	0.38	0.00	0.01	0.17	0.75	0.00	-0.02	0.16	0.18	0.00	0.02	0.19
1.50	1.00	0.00	-0.12	0.25	1.00	0.00	-0.09	0.27	1.00	0.00	-0.12	0.27	1.00	0.00	-0.08	0.29

Table C3.2: Rosenbaum bounds test for 2011

	Total Economy								Services							
	Controlling for GOR				Controlling for TTWA				Controlling for GOR				Controlling for TTWA			
	Significance of selection effect		Point estimate		Significance of selection effect		Point estimate		Significance of selection effect		Point estimate		Significance of selection effect		Point estimate	
Γ	Positive	Negative	Upper	Lower	Positive	Negative	Upper	Lower	Positive	Negative	Upper	Lower	Positive	Negative	Upper	Lower
1.00	0.46	0.46	0.00	0.00	0.08	0.08	0.04	0.04	0.36	0.36	0.01	0.01	0.24	0.24	0.02	0.02
1.02	0.59	0.34	-0.01	0.01	0.13	0.04	0.03	0.05	0.47	0.26	0.00	0.02	0.34	0.16	0.01	0.03
1.04	0.70	0.23	-0.01	0.02	0.21	0.02	0.02	0.06	0.57	0.18	-0.01	0.03	0.44	0.11	0.00	0.04
1.06	0.80	0.15	-0.02	0.03	0.31	0.01	0.01	0.07	0.68	0.12	-0.01	0.04	0.55	0.06	0.00	0.05
1.08	0.87	0.09	-0.03	0.04	0.43	0.00	0.01	0.07	0.76	0.07	-0.02	0.04	0.65	0.04	-0.01	0.06
1.10	0.92	0.05	-0.04	0.05	0.54	0.00	0.00	0.08	0.84	0.04	-0.03	0.05	0.74	0.02	-0.02	0.07
1.12	0.96	0.03	-0.05	0.05	0.66	0.00	-0.01	0.09	0.89	0.02	-0.04	0.06	0.82	0.01	-0.03	0.07
1.14	0.98	0.01	-0.05	0.06	0.75	0.00	-0.02	0.10	0.93	0.01	-0.04	0.07	0.88	0.01	-0.04	0.08
1.16	0.99	0.01	-0.06	0.07	0.83	0.00	-0.03	0.11	0.96	0.01	-0.05	0.07	0.92	0.00	-0.04	0.09
1.18	0.99	0.00	-0.07	0.08	0.89	0.00	-0.03	0.11	0.98	0.00	-0.06	0.08	0.95	0.00	-0.05	0.10
1.20	1.00	0.00	-0.08	0.08	0.93	0.00	-0.04	0.12	0.99	0.00	-0.07	0.09	0.97	0.00	-0.06	0.11
1.50	1.00	0.00	-0.18	0.18	1.00	0.00	-0.14	0.22	1.00	0.00	-0.16	0.19	1.00	0.00	-0.16	0.21

Table C3.3: Rosenbaum bounds test for 2013

	Total Economy								Services							
	Controlling for GOR				Controlling for TTWA				Controlling for GOR				Controlling for TTWA			
	Significance of selection effect		Point estimate		Significance of selection effect		Point estimate		Significance of selection effect		Point estimate		Significance of selection effect		Point estimate	
Γ	Positive	Negative	Upper	Lower	Positive	Negative	Upper	Lower	Positive	Negative	Upper	Lower	Positive	Negative	Upper	Lower
1.00	0.00	0.00	-0.05	-0.05	0.13	0.13	-0.02	-0.02	0.01	0.01	-0.05	-0.05	0.02	0.02	-0.04	-0.04
1.02	0.00	0.01	-0.05	-0.04	0.08	0.20	-0.03	-0.02	0.00	0.01	-0.05	-0.04	0.01	0.04	-0.04	-0.04
1.04	0.00	0.01	-0.06	-0.04	0.05	0.28	-0.03	-0.01	0.00	0.02	-0.06	-0.04	0.01	0.06	-0.05	-0.03
1.06	0.00	0.03	-0.06	-0.03	0.03	0.38	-0.04	-0.01	0.00	0.04	-0.06	-0.03	0.00	0.10	-0.05	-0.03
1.08	0.00	0.05	-0.07	-0.03	0.01	0.48	-0.04	0.00	0.00	0.07	-0.07	-0.03	0.00	0.15	-0.06	-0.02
1.10	0.00	0.08	-0.07	-0.02	0.01	0.59	-0.05	0.00	0.00	0.10	-0.07	-0.02	0.00	0.20	-0.06	-0.02
1.12	0.00	0.12	-0.08	-0.02	0.00	0.68	-0.05	0.01	0.00	0.15	-0.08	-0.02	0.00	0.27	-0.07	-0.01
1.14	0.00	0.18	-0.08	-0.02	0.00	0.77	-0.06	0.01	0.00	0.20	-0.08	-0.02	0.00	0.35	-0.08	-0.01
1.16	0.00	0.25	-0.09	-0.01	0.00	0.83	-0.06	0.02	0.00	0.27	-0.09	-0.01	0.00	0.43	-0.08	0.00
1.18	0.00	0.33	-0.09	-0.01	0.00	0.89	-0.07	0.02	0.00	0.34	-0.09	-0.01	0.00	0.52	-0.08	0.00
1.20	0.00	0.42	-0.10	0.00	0.00	0.93	-0.07	0.03	0.00	0.42	-0.10	0.00	0.00	0.60	-0.09	0.00
1.50	0.00	1.00	-0.16	0.05	0.00	1.00	-0.14	0.09	0.00	1.00	-0.16	0.05	0.00	1.00	-0.15	0.06

Annex D: Productivity Decompositions Methods

1. This annex expands on the introduction to productivity decompositions provided in chapter 5 and the table

Different decomposition methods - static decomposition

2. The seminal work on productivity decomposition is Olley Pakes (OP) analysis of productivity following the deregulation of the US telecommunication industry. In this study the authors first estimate a production function and then derive the coefficients to compute their measure of multi factor productivity (TFP) at the firm level. More specifically, productivity is given by:

$$p_{it} = \exp(y_{it} - b_l l_{it} - b_k k_{it} - b_a a_{it})$$

where the parameters 'bs' are derived using an estimation method that accounts for selectivity bias and attrition bias. TFP is a residual term which captures technical and efficiency advantage of firms over and above measured inputs. The decomposition allows the analysis of how the market process distributes these advantages. If changes in technology and efficiency are primarily due to new establishments, then TFP should be dominated by external factors. If changes occur within firms (for example via learning by doing) then the bulk of TFP growth should be explained by internal factors (Disney et al. 2003).

3. Average productivity is calculated annually using the weighted sum of productivity in individual plants, where the weights are represented by the output share of each plant. This aggregate measure is given by:

$$p_t = \sum_{i=1}^{N_t} s_{it} p_{it}$$

where s_{it} is plant i 's share of output at time t. The share is computed for industrial sectors, at the 2-4 digit level. The aggregate measure is decomposed into two parts:

1) unweighted average of plant level productivities, $\overline{p_t}$; and
2) sample covariance between productivity and output shares. A positive covariance implies that more productive firms have largest output shares, i.e. it signal the presence of allocative efficiency, whereby resources are allocated to the most productive firms (reallocation of resources within surviving firms), $\sum_{i=1}^{N_t}(s_{it} - \overline{s_t})(p_{it} - \overline{p_t})$.

4. Increases in aggregate productivity can be due to increases in average productivity (within productivity), increases in allocative efficiency or both. In OP this second component is particularly important as the deregulation of the US telecommunication industry caused large changes in the number of firms in the sectors, allowing the entry

of new firms and the increasing competitive environment for the incumbents. It seems quite reasonable that, in this situation, allocative efficiency would play an important role, i.e. liberalisation allowed a reallocation of output to more productive plants. In fact their study shows that, while the contribution of unweighted average productivity did not change very much between 1975 and 1987, allocative efficiency increased throughout the period.

5. The OP decomposition has also been used to describe differences in productivity across eight countries (US, UK, France, Germany, Netherlands, Hungary, Romania, Slovenia) in Bartelsman et al. (2013). The study carries out the decomposition for each country and then compares the covariance terms across countries. Their results reveal that the covariance between output/employment shares and productivity is always positive, with the exception of Romania, implying that improvements in allocative efficiency is affecting several countries, although levels are higher in the United States compared to the rest of the sample. When looking at changes in the covariance measure, they found that allocative efficiency has increased over time particularly in Eastern European countries since the early 1990s, following the transition towards a market based system.

6. Mason et al. (2014) apply the OP method to analyse UK productivity performance in the years leading to the financial crisis, looking at both manufacturing and services. This study finds that the within component is the main driver of productivity. Looking at differences across sectors the authors show that the contribution of allocative efficiency is stronger in the service sector at the beginning of the period, indicating a shift of resources towards more efficient use in this industry. However, over time the contribution of allocative efficiency decreases. Compared to the two studies discussed above, these results suggest that external restructuring is particularly important following changes in market regulations that affect competitiveness. In periods of relative institutional stability, the within component is the main driver of productivity growth.

7. The application of this decomposition method is straightforward as it is simply the difference between share weighted average sector productivity and the un-weighted average sectoral productivity (Mason et al. 2014). However, the static nature of this decomposition has been considered a potential shortcoming as it does not allow to analyse the contribution of firms entering and exiting the market on aggregate productivity hence it misses out what it is considered, both theoretically and empirically, an important determinant of average productivity growth[3].

Dynamic decomposition

8. The importance of accounting for the contribution of entry and exit to aggregate productivity has led to the development of different methodologies which extend the OP method. These dynamic decompositions distinguish between three types of firms, surviving or continuing firms (C), entering (N) and exiting (X) firms, and the contribution to aggregate productivity is modelled for each type. Entry and exit will raise overall productivity so long as exitors are less productive than entrants.

[3] According to Foster et al. (2001) the static nature of the OP decomposition can also be advantageous because alternative methods that account for entry and exit produce very unstable results.

9. These decompositions are generally more complex than the OP and they can include up to five components. These usually include a within effect, which measures changes in productivity due to reallocation of resources within firms, a between effects, which implies a reallocation across firms, an entry and an exit effect. Within productivity can increase because of the introduction of new technology and/or organizational changes. The decomposition of Foster et al. (2001) - hereafter FHK - includes an additional term, a covariance term between market shares and productivity. This considers whether increases in productivity correspond with increases in market shares, similarly to the covariance term in OP.

10. One of the distinguishing features of the different extensions is the computation of the contribution of the different types of firms and particularly of the exiting and entering firms. Most models compare firms' productivity to a reference value, which is the industry average at time t in Griliches and Regev (1995) - hereafter GR:

$$\Delta P_{it} = \sum_{e\in C} \bar{s}_e \Delta p_{et} + \sum_{e\in C} (\overline{p_{et}} - \bar{P}_i)\Delta s_{et} + \sum_{e\in N} s_{et}(p_{et} - \bar{P}_i) - \sum_{e\in X} s_{et-1}(p_{et-1} - \overline{P_i}),$$

11. The first term is the within firm productivity growth, the second identifies gains in productivity that come from high-productivity firms' expanding market shares and the last two terms account for productivity growth due to entry and exit.

12. The reference value is the industry average at time t-1 in FHK:

$$P_{it} = \sum_{e\in C} s_{et-1}\Delta p_{et} + \sum_{e\in C} (p_{et} - P_{it-1})\Delta s_{et} + \sum_{e\in C} \Delta p_{et}\Delta s_{et} + \sum_{e\in N} s_{et}(p_{et} - P_{it-1}) - \sum_{e\in X} s_{et-1}(p_{et-1} - P_{it-1})$$

13. It is a different reference productivity for each type of firms in Melitz and Polanec (2012)[4] - hereafter MP. The only method that does not rely on a reference category is Baily et al. (1992) - hereafter BHC. This method has met some criticisms in the literature as it does not allow an easy interpretation of the restructuring effect (Disney et al. 2003)[5].

14. Reference to productivity at time t-1, as in FHK, has been criticised on the basis that it overestimates the contribution of entrants in periods of high productivity growth and, conversely, underestimate it in periods of decreasing productivity. For example, the contribution of entrants in FHK is measured as: $\sum_{e\in N} s_{et}(p_{et} - P_{it-1})$. Such contribution will be positive if the productivity of entrants is higher than average productivity in the previous period. For a growing economy this is likely to be the case, which leads to an overvaluation of the contribution of entrants on aggregate productivity. To avoid this distortion, MP use different reference categories: for surviving firms this is the level of productivity in the previous period ($p_{Ct} - p_{Ct-1}$), for entering firms it is the average level

[4] $\Delta P_{it} = \Delta p_C + \Delta cov_C + s_{Nt}(s_{it}p_{Nt} - s_{it}p_t) + s_{Xt-1}(s_{it-1}p_{Nt-1} - s_{it-1}p_{Xt-1})$, Melitz and Polanec (2012).

[5] The Baily et al. 1992 decomposition can seriously bias the net entry effect downwards. If the market shares of the entrants are low and exitors are high, the impact of net entry might be negative even if entrants were more productive.

of productivity of surviving firms in the same period, $s_{Nt}(p_{Nt} - p_{st})$, and for exiting firms it is the level of productivity of surviving firms at time t-1, $s_{xt}(p_{xt-1} - p_{st-1})$.

15. Although the specification of entry and exit in MP is more sophisticated than in previous models, the contribution of surviving firms results in a disconnection between the measurement of the contribution of external restructuring between existing firms and entrant/exiting firms. In addition, and differently from GR and FHK, the within component corresponds to an unweighted rather than a share-weighted mean of productivity for the continuing firms, and this can generate bias and higher volatility in the different components. To address this issue, Riley et al. (2014) develop a hybrid decomposition which combine the treatment for continuing firms in GR (1995) with the treatment of entry and exit of MP.

This publication available from www.gov.uk/bis

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

Department for Business, Innovation and Skills
1 Victoria Street
London SW1H 0ET
Tel: 020 7215 5000
Email: enquiries@bis.gsi.gov.uk

BIS/16/295