Research Policy 44 (2015) 1462–1472

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

Assessing an experimental approach to industrial policy evaluation: Applying RCT+ to the case of Creative Credits

Hasan Bakhshi[a], John S. Edwards[b], Stephen Roper[c,*], Judy Scully[b], Duncan Shaw[d], Lorraine Morley[e], Nicola Rathbone[b]

[a] Nesta, 1 Plough Place, London EC4A 1DE, UK
[b] Aston Business School, Aston University, Birmingham B4 7ET, UK
[c] Enterprise Research Centre and Warwick Business School, University of Warwick, Coventry CV4 7AL, UK
[d] Manchester Business School, Booth Street West, Manchester M15 6PB, UK
[e] Warwick Business School, University of Warwick, Coventry CV4 7AL, UK

ARTICLE INFO

Article history:
Received 22 January 2013
Received in revised form 9 April 2015
Accepted 13 April 2015

Keywords:
Evaluation
Experimental
Industrial policy
Innovation
Creative
Qualitative research

ABSTRACT

Experimental methods of policy evaluation are well-established in social policy and development economics but are rare in industrial and innovation policy. In this paper, we consider the arguments for applying experimental methods to industrial policy measures, and propose an experimental policy evaluation approach (which we call RCT+). This approach combines the randomised assignment of firms to treatment and control groups with a longitudinal data collection strategy incorporating quantitative and qualitative data (so-called mixed methods). The RCT+ approach is designed to provide a causative rather than purely summative evaluation, i.e. to assess both 'whether' and 'how' programme outcomes are achieved. In this paper, we assess the RCT+ approach through an evaluation of Creative Credits – a UK business-to-business innovation voucher initiative intended to promote new innovation partnerships between SMEs and creative service providers. The results suggest the potential value of the RCT+ approach to industrial policy evaluation, and the benefits of mixed methods and longitudinal data collection.

1. Introduction

Although unusual in terms of industrial policy, experimental methods of policy evaluation are well established in social policy and development economics. Burtless (1995, p. 63), for example, cites Greenberg and Shroder (1991) who 'identified more than 90 separate field trials involving a wide range of distinctive research areas including health insurance, prisoner rehabilitation, labour supply, worker training and housing subsidies'. Banerjee and Duflo (2008, p. 32) also describe the 'recent surge in experimental work' in development economics. Typically such evaluations involve individual human subjects facing some common socio-economic problem, and random allocation of subjects to a treatment and control group. Differences in outcomes between the treatment and control groups are then attributed to the effect of the policy intervention. In terms of industrial policy, however, experimental policy evaluation approaches remain marginal, with non-experimental, ex post policy evaluations remaining the norm. In the context of small business policy evaluation, Potter and Storey (2007), for example, provide an extensive review of best practice in OECD countries without any mention of either the application or potential for experimental methods. Similarly, UK government guidance on industrial policy evaluation focuses entirely on non-experimental ex post evaluation approaches (BIS, 2009). Related observations might be made in relation to the evaluation of R&D and innovation policies: despite the increasing importance of evaluation as part of the process of development of technology policy, evaluation approaches remain almost universally ex post and non-experimental (Laredo, 1997).

The methodological and practical advantages of experimental and non-experimental evaluation methods have been widely debated in the context of social policy interventions (Burtless, 1995; Heckman and Smith, 1995; Bratberg et al., 2002; Banerjee and Duflo, 2011; Deaton, 2010). Experimental methods based on randomised allocation have – at least in theory – the advantage of transparency and may be more convincing to policy-makers than the results of more complex econometric evaluation approaches (Burtless, 1995). In small samples, however, perhaps less than 300, randomisation may be ineffective at ensuring the homogeneity of

* Corresponding author. Tel.: +44 2476522501.
E-mail addresses: Hasan.Bakhshi@nesta.org.uk (H. Bakhshi), j.s.edwards@aston.ac.uk (J.S. Edwards), stephen.roper@wbs.ac.uk (S. Roper), j.w.scully@aston.ac.uk (J. Scully), duncan.shaw-2@mbs.ac.uk (D. Shaw), lorraine.morley.09@mail.wbs.ac.uk (L. Morley), rathbnj1@aston.ac.uk (N. Rathbone).

http://dx.doi.org/10.1016/j.respol.2015.04.004
0048-7333/

control and treatment groups (Bratberg et al., 2002), though robust experiments can still be run with smaller samples if the intervention is sufficiently powerful (Bloom et al., 2011). Implementing experimental approaches may also lead to other specific biases (Heckman and Smith, 1995), while small-scale experimental studies may fail to replicate the likely macro-impacts of a scheme which is implemented at national level – i.e. they lack external validity (Garfinkel et al., 1990). Perhaps the key advantage of experimental approaches, however, and the central issue with non-experimental evaluation approaches, relates to potential selection biases. In particular, if subjects which are allocated to a treatment group have a higher preponderance of some characteristic which is correlated with outcomes this will lead to bias in the estimation of treatment effects. Such biases may be significant where policy interventions are targeted at particular groups of subjects, where support is allocated on the basis of routinised decision rules, or where there is an element of self-selection into an experiment. One recent study, for example, illustrates how funding allocations in the Norwegian Research Council are based on ex ante project rankings generating a potential selection bias when evaluating the Research Council's funding decisions (Bremnes et al., 2011). In terms of policy for innovation or small and medium enterprises (SMEs), similar selection biases might arise where a policy initiative seeks to back winners or is focussed on firms which have an established track record of growth or innovation. In this sense the receipt of public support may itself need to be treated as endogenous (Garcia and Mohnen, 2010).

As many methodological and implementation issues arise in industrial policy evaluations, which focus on firms, as they do in social policy, where the subjects are likely to be individuals. The selection biases are likely to be as great. However, arguably some of the ethical issues which arise in implementing experimental approaches to social policy may be seen as less significant in industrial policy interventions. It may be more ethically acceptable, for example, to randomly allocate public support amongst firms rather than adopting a similar allocation rule to the distribution of support amongst financially disadvantaged individuals. Nevertheless, the random allocation amongst firms needs to be carefully considered to ensure money is not simply thrown at applicants in the hope that they do something useful with it. Application criteria, an evidenced business need, project governance and, potentially, co-funding can raise confidence in the ethical aspects of investing in companies' innovation potential. This makes it all the more surprising that experimental evaluation approaches in industrial policy are not more common.

In this paper we propose and assess an experimental approach to the evaluation of new industrial policy interventions. Our approach (which we label RCT+ or Randomised Control Trial plus) takes advantage of randomisation, but combines this with a longitudinal and mixed methods data collection strategy to provide causative insights rather than only summative policy assessment. In other words, RCT+ seeks to evaluate the validity of an underlying logic model rather than simply generating point estimates of policy impact (Donaldson and Gooler, 2003), and seeks to explain why these results are observed (Denzin and Lincoln, 2005; Ludwig et al., 2011).

The aim of this paper is to assess the potential value of the RCT+ approach when applied to industrial policy. We make three main contributions. First, we provide an assessment of the value of experimental evaluation approaches to industrial policy initiatives, suggesting an alternative enhanced approach to the development of industrial policy. Second, through an application of the RCT+ approach we extend standard (quantitative) experimental evaluation approaches beyond the summative to provide causal explanations for policy outcomes – i.e. to identify the 'why, how, and at what cost' an intervention may have worked (White, 2008, p. 98). Adopting this type of rigorous qualitative approach provides potentially frame-breaking insights, and may enable the conceptualisation of the context in which an intervention was implemented, facilitating the generalisation of results (White, 2008). Third, our application of RCT+ shows how a longitudinal approach can provide a time profile of policy outcomes, without which policy-makers can make incomplete, and potentially misleading, inference. We apply the RCT+ approach to Creative Credits, a UK-based business-to-business innovation voucher programme designed to foster new innovative partnerships between SMEs and creative service providers.

The remainder of the paper is organised as follows. In Section 2 we discuss the rationale for industrial policy evaluation and review previous debates about the relative merits of experimental and non-experimental evaluation approaches. Subsequent sections of the paper report the application of the RCT+ evaluation approach to the Creative Credits experiment. Section 3 reports the application of the RCT+ approach and its outcomes. Section 4 discusses the implications of the RCT+ experiment and the strengths and limitations of experimental methods in industrial policy evaluation. Parallels with meta-analytical approaches in medicine are also considered. Section 5 concludes the paper.

2. Evaluating industrial policy initiatives

Potter and Storey (2007) identify five reasons why industrial policy evaluation might be undertaken: to establish the impact of industrial policy; to inform the allocation of funding to alternative policy measures; to demonstrate value for money; to stimulate debate about forms of public intervention; and to contribute to improvements in the design and administration of policy interventions. In each case the problem of causal inference is the same, i.e. that the treated and non-treated outcomes for any single firm are never observed (Holland, 1986). The analytical problem this raises is how to estimate the difference between the actual realised outcomes and the potential outcomes if no treatment had been administered. Ideally, the substitute for the unobserved (un-treated) outcome needs to meet two criteria: (i) it should be observable to the researcher; and (ii) it should be an 'internally' valid substitute for the set of un-treated outcomes. Internal validity in this sense requires that 'the only difference between the member of the control group and the member of the treated group corresponds to the fact that the latter is treated and the first one is not' (Reiner, 2011, p. 18).

More comprehensively, Imbens and Wooldridge (2009) outline three situations which describe the allocation of subjects to a control and treatment group. The first, and simplest, is the classical experimental situation of randomised allocation in which allocation is unrelated to outcomes. The second allocation mechanism – 'un-confounded allocation' – occurs where assignment is independent of outcomes but may be related to subject characteristics. Here, where the assignment mechanism is either observable or discoverable, sampling and/or statistical approaches can be used to minimise any systematic differences between the characteristics of the treatment and control groups and provide a valid estimate of treatment effects (Burtless, 1995). In practice, evaluations of SME policy vary in the sophistication of their approach to un-confoundedness. Potter and Storey (2007), for example, cite evaluation studies which compare the performance of treated firms with control groups of 'typical' firms (Lambrecht and Pirnay, 2005), and studies which use 'matched' control groups based on treatment and control groups with similar baseline characteristics (Lerner, 1999). However, despite careful matching or selection of control groups, the potential remains for bias in terms of the background characteristics of the two groups (Bratberg et al., 2002).