



Methods of Evaluating the Impact of Public Policies to Support Small Businesses: the Six Steps to Heaven

David J. Storey

*Centre for Small and Medium Sized Enterprises
University of Warwick*

Acknowledgements. This paper was first formulated as a result of the authors participation in the evaluation group convened by Anders Lundstrom from the Swedish Foundation for Small Business Research (FSF). The paper inevitably reflects the contributions made by other key members of the group such as Dennis De, William Dennis, Lois Stevenson and Jane Wickman as well as Anders himself, Christer Öhman and Hakan Boter. However a key influence on the ideas has been the author's involvement with the Best Practice Working Party on SMEs at the OECD. Early versions of the paper were also presented to the I.G. Conference "Processes of Enterprise Creation and Consolidation" in Rome 12 December 1997, in Mikkeli, Finland on "Growth and Job Creation in SMEs", January 7-9, 1998, and then at the ICSB Conference in Singapore 8-10 June 1998. Helpful feedback was provided in all instances, particularly from Frank Hoy who pointed me towards U.S. literature of which I was unaware. I also appreciated comments received from Marc Cowling.

Abstract. Virtually all developed economies utilize taxpayer's money to provide either free or subsidized assistance to small businesses, the self-employed or to potential small business owners. This paper provides an outline methodology for evaluating the impact of these policies. It identifies six approaches, beginning with the most simple and ending up with the most sophisticated. These are referred to as the 'Six Steps', with step Six being viewed as 'best practice' or 'heaven' in this area.

Keywords: small businesses, public policy, evaluation

1. Introduction

Virtually all developed economies utilise taxpayers money to provide either free or subsidised assistance to small business, the self employed or to potential small business owners. Sometimes this assistance is direct financial payments in the form of subsidies to encourage investment in human or physical capital. In other cases subsidies are in the form of free or subsidised advisory services in starting or developing small business or in specialist areas such as exporting or the use of new technology.

Taxpayers money may also be used to bribe individuals or organisations to behave in a way which is perceived to benefit both small businesses and the economy as a whole. These bribes often take the form of tax relief. For example, wealthy individuals may be given 'tax breaks' to become equity

participants in small or young businesses. Finally, some government procurement programmes focus upon small businesses, and taxpayers money is used to offset any efficiency losses to government by its having to contract with small businesses where these are not optimal suppliers. The wide range of public support programmes to small firms in developed economies and their appraisal is best reviewed in OECD [1995, 1996, 1997].

Given the huge variety of schemes, the diversity of countries in which the schemes are found, and the often inflated claims on the part of those administering the schemes for their effectiveness, it is disappointing that the academic community has been rather slow in seeking to address this area. Perhaps even more seriously, even where the issues have been addressed by small business academics, the methods of evaluation employed have rarely been at the intellectual frontier.

This paper seeks to provide an outline methodology for evaluating the impact of public policies to assist the small business sector. It begins, however, by emphasising the impossibility of conducting an evaluation in the absence of clearly specified objectives for the policy concerned. Ideally, in fact, these objectives should be specified in a quantitative manner in the form of targets.

The paper then moves on to provide a review of the various methodological approaches to evaluation of small business support policies found in developed countries. It does not seek to present a comprehensive review of the area. Instead it provides an analytical framework within which a wide variety of types of analyses can be classified. In total it identifies six approaches, beginning with the most simple and ending up with the most sophisticated. These are referred to as the 'Six Steps' with Step 6 being viewed as 'best practice' or 'Heaven' in this area. The paper also makes a distinction between 'monitoring' and 'evaluation'. Monitoring is viewed as Steps 1-3, with the more sophisticated approaches being classified as 'evaluation' in Steps 4-6.

2. Specification of Objectives

It is a fundamental principle of evaluation that its prerequisite is the specification of the objectives of policy. Unfortunately it appears to be a characteristic of governments in all developed countries to be, at best, opaque about the objectives of small business policy. Many phrases characterise this area. Governments talk about 'creating an enterprising society', or 'maximising SMEs contribution to economic development', or 'enhancing competitiveness', or even 'creating jobs'. So far as this author is aware, however, no developed country produces a clear set of objectives for each component of small business policy. Analysts therefore are required to infer the objectives of policy, rather than having these clearly defined. Only then is

it possible to determine whether or not the target is achieved and hence be able to judge whether or not policy is successful.

Instead, what governments favour are lists of policies. Lists of the various measures which have been introduced to help the small business sector, such as taxation exemptions, late payment, administrative burdens, finance and information provision. Typical of these lists are those presented in ENSR (1997) at a European level or DTI (1998) for a national level.

Analysts such as de Koning and Snijders (1992) have made attempts to compare SME policies in countries. Their work, on EU countries, was only able to compare, using such lists, the number of policy measures focused upon SMEs in policy fields such as Fiscal policies, Export policies, Information and Counselling etc. This is clearly not the same, and indeed is significantly inferior to, specifying objectives.

Not only is there a conspicuous absence of clear objectives for SME policy, but the implied objectives can often be conflicting. The United Kingdom can be taken as illustrative. Table 1 reproduces my earlier (Storey 1994) effort to seek to identify the appropriate objectives for UK small business policy. Note the table only defines the objectives and not the numerical values of the objectives (i.e. targets) themselves.

Table 1: Intermediate and Final Objective

Intermediate	Final
1. Increase Employment	– Increase Employment – Reduce Unemployment
2. Increase Number of Start Ups	– Increase Number of Start Ups – Increase Stock of Firms
3. Promote Use of Consultants	– Promote Use of Consultants – Faster Growth of Firms
4. Increase Competition	– Increase Competition – Increase Wealth
5. Promote “Efficient” Markets	– Promote “Efficient” Markets – Increase Wealth
6. Promote Technology Diffusion	– Promote Technology Diffusion – Increase Wealth
7. Increase Wealth	– Votes

Source: Storey (1994)

The table distinguishes between intermediate and final objectives. Taking the top line as illustrative, we can identify ‘increasing employment’ as an

objective, with a target being where this objective was given a particular measure – such as increasing employment by 5% over a five year period.

Taking now the objectives, politicians in most developed countries have SME policies because they believe, rightly or wrongly, that SMEs are both currently a major source of employment and likely to be an increasing source of new jobs in the future [Hughes (1997)]. Failure to address/encourage the SME sector may lead to slower rates of job creation and hence unemployment being higher than otherwise.

Whilst political leaders frequently couch their rhetoric in terms of employment creation, their prime concern is, in fact, to seek to reduce unemployment, rather than to increase employment. Increases in employment therefore can be considered as an intermediate objective, with the final objective being that of reducing unemployment.¹ Utilising the vehicle of SMEs to create jobs can however have a mixed effect upon reducing the numbers of individuals registered as unemployed. From the positive side, SMEs are more likely to employ individuals who are comparatively heavily represented amongst the unemployed – unskilled, very young and very old (Brown, Hamilton and Medoff 1990). Yet, in other respects, job creation in SMEs is likely to have only a modest effect upon reducing registered unemployment. This is because SMEs are disproportionately likely to provide part time work and these part time workers (often females) are less likely to be registered as unemployed. There is therefore immediately a question as to whether the real objective of policy in row 1 is the creation of employment or the reduction in unemployment. This is rarely made explicit in policy pronouncements.

A second area of possible conflict between job creation and reduction in unemployment is that the latter can often be reduced by out-migration from a country or region. Policies of job creation, if they are successful, can lead to lower rates of out-migration because workers feel there is a prospect of getting a job in the locality. Success at creating jobs can even, perversely, lead to increased unemployment. Those specifying objectives have to be clear where their priorities lie.

In row 2 of Table 1 an alternative objective of SME policy is articulated. Many countries have policies to encourage individuals to start businesses. This may be related to aspects of objective 1 – such as a view that more people starting businesses leads directly to additional jobs or to reducing the numbers of unemployed. Alternatively, policies to increase the number of start ups may merely reflect (be a result of) a more dynamic economy and one likely to exhibit prosperity in the longer term. However it is widely recognised that policies to assist the start up of new enterprises are most likely to be targeted

1. Here targets are sometimes specified. For example the Swedish government is committed to halving unemployment over a five year period to the year 2002.

upon individuals who are unemployed, since these individuals are the most 'susceptible'. Experience, both in the UK and the United States (Storey and Strange, 1992, Bendick and Egan, 1988), shows these individuals often enter trades with low entry barriers – such as vehicle repairers, window cleaners, taxi drivers, etc. for which there is a finite and highly localised demand. The net effect of such policies is that public money is used to encourage unemployed individuals to start a business in these sectors, but this serves primarily to displace other unsubsidised traders in the locality with no obvious benefit either to the local consumer or to the economy in general. The effect then is to increase the number of start ups i.e. satisfy intermediate objective 2, but also to increase the number of businesses which cease to trade, with little net change in the stock of firms and so not satisfy final objective 2. Even where there is an increase in the stock of firms, there may well be a compensating fall in average firm size without any apparent increase in employment (Storey and Strange 1992). A choice therefore has to be made between Intermediate objective 1 and Intermediate objective 2.

The remainder of table 1 identifies several other objectives which are apparent from observing the characteristics of public support for SMEs in the United Kingdom. The interested reader can consult Storey (1994) for a fuller discussion of other potential conflicts. Perhaps the only objective requiring further comment at this point is that in row 7, where the final objective is 'Votes'. This clearly is a fundamentally different objective, since it is explicitly political, rather than being one of the other more 'economic' objectives specified elsewhere. As noted in Storey (1994), there is nothing undesirable in public policies being focused on the achievement of economic objectives and, as a reward for achieving good economic performance, politicians being re-elected. Indeed such logic is at the cornerstone of democracy. What is more questionable is where policies, using taxpayers money, are couched in terms of economic objectives but are really a mechanism for persuading a numerically significant group (in this case small business owners) to vote for the government through the provision of 'sweeteners'. In many countries there is an overtly political element to small business policies, and failure by analysts to take it into account, would be to underestimate the role which it plays in politician's calculations.

Governments, then, should be required to specify their objective in the provision of small business support. Identifying a wide range of sketchy objectives may serve the government's purpose of being able to point to success if there is an improvement in that objective area, but this is clearly unsatisfactory from the viewpoint of the taxpayer. Paraphrasing Harrison and Leitch (1996) "It is clearly unsatisfactory for the government to claim that the target is anything it happens to hit". Instead governments should set objectives, with an indication of which, if there is more than one, takes priority. Once the

objectives are set, then numerical targets need to be specified. Only then can evaluation take place.

It is interesting to note that, whilst this paper was in preparation, the issue received ‘heavyweight’ support in the UK from the House of Commons Select Committee (1998). It said:

The Government has yet to state clearly what its policy objectives are with regard to SME policy; how the achievement of these broad objectives can be assessed, or how existing policy measures fit within a broader context . . . the means by which competitiveness can be measured and the reasons for targeting competitiveness, in terms of its impact on employment, unemployment, GDP and other indicators can only be guessed at . . . we are not convinced that the Government’s SME policy is characterised by sufficient structure and focus. We recommend that, as a matter of urgency, the government define the objectives of SME policy. The objectives chosen must be accompanied by measurable targets, with a timetable for their attainment. Clearly such a development would be highly desirable and ought to be implemented with all speed.

3. The Six Steps

This section makes the unrealistic assumption that objectives, either of small business policy as a whole, or of the particular programme under consideration, are specified. The remainder of the paper seeks to review how, in practice, appraisal is undertaken. Table 2 identifies the six steps and these are ranked in terms of sophistication, with Step 1 being the least sophisticated and Step 6 being the most sophisticated.

Table 2: The Six Steps

Monitoring	
STEP I	* Take up of schemes
STEP II	* Recipient’s opinions
STEP III	* Recipient’s views of the difference made by the assistance
Evaluation	
STEP IV	* Comparison of the performance of ‘Assisted’ with ‘Typical’ firms
STEP V	* Comparison with ‘Match’ firms
STEP VI	* Taking account of selection bias

Although all six steps are often referred to as ‘evaluations’ in the literature, Steps 1-3 can be considered as merely monitoring, with only Steps 4-6 being evaluation.² The difference between monitoring and evaluation is that the latter are attempts, demonstrating analytical rigour, to determine the impact of the policy initiatives. Monitoring, on the other hand, merely either documents activity under the programme or reports participant’s perception of the value of the scheme. In short, the difference between monitoring and evaluation is that monitoring relies exclusively upon the views of the recipients of the policy. Evaluation however seeks, by some means, to contrast these with non-recipients, in order to present a ‘counter-factual’. The difference between actual changes and the ‘counter-factual’ is viewed as the impact of the policy – or its ‘additionality’.

3.1. Monitoring

Step 1: Take up of Schemes

Table 3 describes Step 1. This monitoring procedure identifies the characteristics and nature of the take up of the scheme. For example, it might

Table 3: Take Up of Schemes

Questions			
* How many firms participated?			
* What sectors were they in?			
* What locations were they in?			
* How big were these firms?			
* How much money was spent?			
Problems			
* Tells you almost nothing about policy effectiveness			
* Tells you almost nothing about satisfying objectives			
Examples			
Author:	Date:	Topic:	Country:
USA Delegation to OECD	1997	Small Business Investment Co.	USA

2. Monitoring has narrower objectives than evaluation. It is limited to observing and recording practical indicators of inputs and outputs Evaluation has two prime aims: – An improving and learning aim, and A proving aim” Bridge et al (1998)

quantify the number of firms which participated in a particular scheme, their sectoral distribution, the size of such enterprises and possibly their regional distribution. Step 1 reviews also frequently include public expenditure on the schemes, so that it is possible, for example, to identify expenditure by firm size or the proportion of expenditure 'consumed' by particular regions. What is much less frequently available is information on the money received by individual firms, since this is thought to contravene a confidentiality relationship between government and the enterprise.

The data used in Step 1 are primarily collected by the public sector for accounting purposes. In many instances they appear as Appendices to government documents but, because they are collected simply for accounting purposes, they do not even seek to evaluate whether the monies have been effectively spent. Their sole concern is to document expenditure, making it clear that expenditure is compatible with the purpose for which it was intended. In short, Step 1 serves an accounting and legal function, but plays no economic role.

The second section of Table 3 makes it clear that, whilst Step 1 appraisals are the most frequently conducted, the results obtained provide no insight whatsoever into policy effectiveness. They do not even seek to answer the question 'To what extent did the policy achieve the types of objectives outlined in Table 1?'

Despite their ubiquity, Step 1 appraisals can only be considered as the 'building blocks' for evaluation. Whilst they provide data on the numbers of firms participating, and on expenditure, these items are not linked.

Step 2: Recipient's Opinions

In Step 2 those firms who participated in the schemes are asked for their opinions. For example, those participating in subsidised training activities are asked about whether they felt there was value in the training provided; firms in receipt of subsidised loans are asked about whether they thought the loan to be valuable; those who participated in export counselling services are asked whether they felt the advice was helpful and whether it led to new orders. Firms participating in Loan Guarantee Schemes are asked about whether they would have received funding for a project without the availability of the scheme.

Firms are also normally asked about the application procedures to participate in the Scheme to determine whether these can be streamlined. For example, firms are asked about how they became aware of the service, about the complexity of the application procedure and whether the application was speedily and fairly handled by the bureaucrat. Step 1 data is therefore 'objective' financial accounting data, whereas Step 2 seeks to obtain the

viewpoint of the firms both on the effectiveness of the scheme and on its accessibility.

The 'Problems' section of Table 4 however shows that, despite the frequency of such studies, Step 2 information does not help determine whether objectives are achieved. Take for example participants upon training courses: here participants are often asked to express an opinion as to whether they felt the training to be of use to them and whether it was professionally delivered - the so-called 'happy sheets'. It is however a strong leap of faith to believe that satisfaction with the course delivered relates to enhanced firm performance; yet it is only enhanced firm performance which will be related to the objectives of policy³.

Table 4: Recipient's Opinions

Questions			
* Course Participants: Did they like it?			
* Firms : Were there problems in applying?			
: Were procedures too slow?			
: Cumbersome?			
* How much money was spent?			
Problems			
* Even if they like it, it does not tell you if it is effective			
* All it can do is offer insights into policy delivery			
- but that is not they key question			
Examples			
Author:	Date:	Topic:	Country:
Moint	1998	Export Assistance	USA
Rogoff and M-S Lee	1996	Small business support services, in general	USA
Ernst & Young	1996	Business Links	USA

In short, whilst such assistance may make the recipients happier - and conceivably more likely to vote for the politicians - it does not necessarily

3. Despite this, the link is frequently made. For example, the Barclays 1998 small business review on training reports high levels of satisfaction reported by owner-managers on training courses (91% felt that quality was good or very good), but no attempt was made to link this to formal performance measures. Despite this Barclays asserted that they believe it to be a critical element to a successful small business. However where such links have been sought through careful work [Hughes et al (1998)] associations are very weak or non-existent.

relate to the economic objectives of the policy, such as increasing the competitiveness of the firm or job creation.

If the objective of the investigation is, in part, to identify the problems with accessing aid then only addressing these questions to those firms which were successful in overcoming any barriers leads to biases. In particular it is likely that those who have surmounted the barriers will have a more 'positive' view than those who were discouraged. Questioning only participant firms fails to estimate the extent to which firms are discouraged from participating in a scheme by the real or imagined barriers which exist. It is therefore of paramount importance that the views are sought of all relevant businesses - whether or not they applied. From this, a list of applicants, but who did not access the aid, must also be drawn. Only in this way is it possible to obtain an accurate measure of the extent of any application barriers.

Overall, Step 2 appraisals can offer some insight into policy delivery (especially when combined with the views of non recipients), but they remain almost irrelevant to determining the effectiveness of policy. This is because there may be no link between the views of the firm on the value of the policy and the ability of the policy to achieve the objectives specified in Table 1. For example, the privately rationed firms will prefer public subsidies with high dead-weight elements and might be tempted to speak positively about such policies if they felt this was likely to influence government provision of such subsidies. On the other hand some firms may be more truthful, yet the evaluator has no means of distinguishing the truthful from the selfish firms.

Step 3: Recipients views of the difference made by the assistance

In Step 3 recipients of policy are asked, not simply whether they liked the policy – the happy sheets – but also whether they thought this made any difference to the performance of their firm. Normally quantitative estimates are sought, to determine whether the initiative provided additionality in terms of additional jobs, sales, or profits.

Table 5 shows that, in the more 'sophisticated' Step 3 appraisals, firms may also be asked questions as to what would have happened to them if they had not been in receipt of the policy initiative. Perhaps, most difficult of all, firms may be asked to estimate the extent to which, if there is any enhanced performance on their part, this is at the expense of other firms. Such questions are designed to estimate the extent of any 'displacement'.

The 'Problems' section of Table 5 shows there are several fundamental problems with this approach in addition to those referred to in Step 2. The most important of these is the extent to which businesses are capable, even if they choose to be truthful, of conducting the mental gymnastics required to answer such questions. To ask a small manufacturer to estimate the extent to

Table 5: Recipient's view of the difference made by the assistance

Questions			
* Did firms think it provided 'additionality'?			
* Would firms have done it anyway?			
* Does it cause 'displacement'?			
Problems			
* Even if they like it, it does not tell you if it is effective			
* All it can do is offer insights into policy delivery – but that is not they key question			
Examples			
Author:	Date:	Topic:	Country:
Moint	1998	Export Assistance	USA
Rogoff and M-S Lee	1996	Small business support services, in general	USA
Ernst & Young	1996	Business Links	USA

which the provision of a loan or subsidised advisory service received two or three years previously influenced the subsequent profitability of his/her firms merely encourages guessing. There are so many influences upon the performance of small enterprises that being able to attribute precisely a number, or even a range, is an unreasonable question.

In many instances it is a perfectly understandable reaction of businesses to provide answers which they think the questioner wishes to hear in order to be able to continue untroubled with the running of their business. If they do adopt this response there is, yet again, no way of checking.

Whilst some entrepreneurs will provide the answer which they think the questioner wishes to hear in order to get them out of the door - and by implication therefore overestimate the impact of the initiative - others may adopt the reverse strategy. Many entrepreneurs are fiercely proud of their business and are very reluctant to admit to receiving any assistance whatsoever. Such individuals are therefore likely to underestimate the contribution of policy by claiming that any improvements in their business reflected their entrepreneurial skills, rather than public money. Faced with these extreme groups the analyst has no basis for judging which of the two are numerically dominant in any group.

There is also the issue of when such questions should be asked, and of whom. Clearly they cannot be asked at the time of the loan since any effects (on profitability/sales etc.) will not have had an effect. On the other hand a period of more than three years after the loan will mean that too many other

influences will have affected firm performance. A balance therefore has to be struck between not waiting long enough for effects to appear and waiting so long that recall deteriorates.

Finally, it is the case for both Step 2 and Step 3 appraisals that interviews can only be conducted for firms which continue to trade. It is very difficult to contact enterprises which are no longer trading and yet all firms are the target for policy. To have responses only from surviving firms will clearly bias the interpretations placed upon the effectiveness of the policy, serving to make the outcomes more positive than would be the case by the inclusion of both survivors and non-survivors..

Overall, therefore, monitoring alone is incapable of offering policy relevant insights into policy effectiveness, where the objective of policy is to enhance the performance of SMEs. This is because the effect of policy cannot be estimated simply by seeking the views of recipient firms, even if these views were honestly provided. It is only capable of soliciting views from operational businesses so, if one objective of policy is to raise survival rates of firms, then this procedure is precluded. To overcome these problems it is necessary to compare the assisted firms with groups of firms not assisted by the policy. This is defined as evaluation. Its challenge is to isolate the appropriate group of firms with which to make the comparison, and to hold constant all other influences.

3.2. Evaluation

Step 4: Comparison of the performance of 'Assisted' and 'Typical' firms.

The earlier discussion of Table 2 emphasised that a key distinction between monitoring and evaluation was that monitoring focused exclusively upon firms which have been assisted by policy. Yet to evaluate the impact of the policy it is necessary to decide what would have happened to businesses in the absence of policy - the so-called 'counter-factual'. The effect of policy is therefore defined to be the difference between what actually happened and what would have happened in the absence of policy.

Step 4 estimates this impact by comparing the performance in firms assisted by the policy with those which have not been assisted. The inference is that any difference in the performance of the two groups can be attributed to the impact of the policy.

In Table 6 assisted firms are compared with typical firms in the population. For example, employment or sales growth in assisted firms is compared with typical firms; alternatively the differences in survival rate of assisted firms may be compared with the survival rates of firms more generally in the economy.

Table 6: Comparison of the performances of 'Assisted' with 'Typical' firms

Approach			
* Employment/Sales growth of assisted firms compared with 'typical' firms?			
* Survival of assisted firms compared with 'typical firms'			
Problems			
* Assisted firms are not typical			
Examples			
Author:	Date:	Topic:	Country:
Chrisman et al	1985	Subsidised Consultancy	USA
Deschoolmeester 1 et al	1998	Management Training	Belgium

The advantage of this approach is that, for the first time, a 'control' group of enterprises is identified. This enables comparisons between the 'assisted' and the 'control' group to be made; it also enables comparisons, in principle, to be made between the survival and non-survival of firms in both groups.

The problem, as noted in the second half of Table 6, is that firms in receipt of assistance may not be typical of firms in the economy as a whole. For example those firms where the entrepreneur seeks training, even where this is subsidised by the state, may be more likely to be growth orientated than firms more generally throughout the economy. Those seeking training from a premier University Business School are more likely to have graduates in the business than 'typical' firms. As Deschoolmeester et al (1998) show, those seeking training are generally younger and have significantly better educated owners than the population of firms as a whole. Given this, they may also be starting businesses in different sectors. Equally, firms seeking advisory services may be more 'aware' businesses and therefore more likely to be better performing businesses. Thirdly there may be sectoral or geographical characteristics of recipients, which distinguish them from the population of firms overall. These effects can be either positive or negative.

For example some SME policies are focused upon the unemployed or 'at risk' groups. A classic example is Law 44 in Italy, described in detail in OECD (1995, 1997). This is a scheme which targets young unemployed individuals in Southern Italy; it provides financial and mentoring support to these people in starting up and developing, during their early years, their businesses. It would clearly be inappropriate to compare these businesses with typical Italian small firms for at least two reasons. The first is that these businesses are founded by young people, the survival rate of whose businesses is known to be markedly lower than those of other age groups [Cressy and Storey (1994)]. The second difference is that the economic and trading environment of

Southern Italy is significantly more difficult than in other parts of that country, making it more difficult for new businesses to flourish. For these two reasons, to compare Law 44 firms directly with 'typical' Italian firms and attribute any difference in performance to the Law would be to risk seriously underestimating impact.

The study by Deschoolmeester et al (1998), comparing start-ups of businesses from 'graduates' of the Vlerick school finds marked differences in age, sector and education between the graduates and the population of firms. All these factors will influence the subsequent performance of the firm. To attribute performance differences to the provision of the training requires explicit account to be taken of these factors.

It is therefore necessary to more explicitly take into account the factors likely to influence the performance of the assisted and non-assisted firms and to seek to hold these constant. This process is called matching.

Step 5: Matching

In Step 5 researchers identify a specific 'control group' with which to compare the assisted businesses. For example, if a policy were implemented to enhance the survival rates of new businesses then it would clearly be inappropriate to compare survival rates of assisted new businesses with that of typical small firms because it has been consistently shown that young businesses have lower survival rates than longer-established businesses (Storey 1994). It is also consistently shown that larger firms have higher survival rates than smaller firms (Storey 1994). Failure to take account of these elements would clearly bias the picture. Equally, if the scheme were focused upon high-tech businesses, then these types of businesses generally have faster growth rates than the SME population as a whole. Hence it would be unreasonable to compare the performance of the two groups of firms and infer that the difference in performance is attributable to the policy.

For these reasons Table 7 shows that Step 5 appraisals formally identify a 'control group' of firms. These are called 'match' firms, and matching generally takes place on four factors known to influence, to different extents, the performance of firms. In principle the 'assisted' and the 'match' firms would be expected to be identical on the basis of age, sector, ownership and geography.

Given such controls it is then possible to compare the performance of both groups over the same time period. The inference drawn is that any differences in performance between the two groups are attributable to the policy.

However Table 7 shows that, even here, there are both technical and inferential problems. The technical problem is that perfect matching upon all four criteria simultaneously can be difficult. Ideally such matching should take

Table 7: Comparison with 'Match' firms

Approach			
* Compare assisted with 'Match' firms on the basis of:			
– age			
– sector			
– ownership			
– geography			
* Compare performance of both groups over same time period			
Problems			
* Perfect matching on four criteri can be very difficult			
* Sample selection bias			
– More 'motivated' firms apply			
– Attribute differential performance to scheme and not to motivation			
Examples			
Author:	Year:	Topic:	Country:
Storey & Westhead	1994	Science Park Evaluation	UK
Lerner	1997	Small Business Investment Companies	USA
Hart & Scott	1994	Financial assistance	UK

place immediately prior to the time at which the policy is implemented so that the performance of the two cohorts can be monitored over time. In practice this rarely happens. Instead information may be available for the assisted firms over a period of time, but then the control group is constructed as part of the evaluation procedure after the policy has been in operation. This means that it can be difficult to accurately estimate the survival/non-survival impact of policy - and yet this is a crucial element of SME policy initiatives in most countries.⁴

Whilst there are technical problems in constructing the sample, there are also major inferential problems. In particular, even if the four matching characteristics are held constant there may be other factors, which are not, where the two groups differ. In the terminology of labour economics, whilst it is possible to take account of 'observables', it is much more difficult to take

4. A classic example of this are the attempts which began to be made in 1997 to evaluate the impact of Business Links in the UK. These seek to provide 'soft' assistance to small firms but Business Links had been in operation for three years before any Step 5 type evaluations were contemplated under the Inter-Departmental Working Group on Impact Assessment of Business Support.

account of ‘unobservables’ (Lalonde 1986, O’Higgins 1994). The ‘observables’ can be considered to be age, sector, etc., as discussed above. The key ‘unobservables’ in this context can be considered to be the possibly linked issues of motivation and selection.

Taking motivation first; it may be that although firms do not differ in terms of ‘observables’, those who seek assistance are more dynamic and growth orientated. They may be run by individuals who are more aware, better networked and more open to new ideas. If we compare the performance of assisted and non-assisted firms and find the former outperform the latter, it may be tempting to infer the difference is attributable to the policy. But, if the two groups also differ in terms of motivation, any performance differences may reflect motivation rather than policy impact. In technical terms the motivated firms are self-selecting and this has to be taken into account. This is subsequently referred to as ‘self selection’.

A second source of selection bias occurs where the scheme providers choose some applicants and not others; this is called administrative selection. Illustrations of administrative selection include the SBIR programme in the United States (Lerner 1997), Law 44 in Italy or the Prince’s Youth Business Trust in the United Kingdom. In all three schemes an individual or a business applies to participate. A judgement is made as to whether that individual is suitable (we assume all individuals are eligible but that resources are deemed insufficient to fully satisfy all eligible applicants). Under this selection procedure it is reasonably assumed that the selectors will seek to identify the ‘best’ cases, or at least seek to avoid the ‘worst’ cases. Otherwise there would be no value in a selection procedure.

We have to assume the selectors are capable of making informed judgements – otherwise there would be no point in having selectors. In this case the performance of the selected group will, even if the policy yielded no benefits whatever to the firms, be superior to that of the ‘match’ group since the better cases are being selected. It therefore cannot be inferred that the whole of the observed difference between the assisted group and the non-assisted group in terms of performance is attributable exclusively to the policy.

Two factors are likely to enhance this bias. The first is the extent of competition for the funds. If 99 out of 100 applicants are successful, sample selection bias is likely to be less than where only 10 applicants in every 100 are successful. Secondly the ability of the selectors to make good decisions is also of considerable importance. Our judgement is that, since so many small business support policies are selective, and substantial resources are devoted to the selection procedure, it must be believed, at least by policy makers, that selection makes a difference. Quite simply, the bigger the difference which the selection makes, the bigger the deflation component required from the use of control groups, which only take account of ‘observables’.

Step 6: Taking Account of Selection Bias

How then do we seek to overcome these problems? Table 7 shows that Step 6 procedures seek to compare assisted with matched firms, taking account of sample selection. Two procedures can be employed. The first is the use of statistical techniques which seek to explicitly take account of sample selection bias. These have become standard practice within the labour economics literature dealing with assessing the impact of training upon subsequent employment prospects of individuals. (Dolton et al 1989, O'Higgins 1994). The analysis utilises the technique originally formulated by Heckman (1977). In non-technical terms the Heckman 2-step adjustment procedure formulates a single equation to explain the selection procedure and then, taking of the selection procedure factors, formulates a second equation to explain performance change, taking account of factors included in the selection equation.

The value of the procedure is that the extent, if any, of selection can be taken into account. Thus the selection equation generates a coefficient (inverse Mills ratio) which is significant where selection is present. Where it is not, then a Stage 5 procedure is perfectly valid.

Where selection is shown to be present the impact deflation can be considerable. For example, the Wren and Storey (1998) analysis of the impact of the United Kingdom's subsidised consultancy services showed that, taking no account of selection, the policy appeared to raise the survival rate of firms by up to 16% over an eight year period and raised it up to 3% over a two year period. However, when account was taken of selection, these fell to 5% over the long term and 2% in the short run. Failure to take account of selection can therefore lead to serious overestimates of the impact of policy and whilst this can be favoured by some policy makers, it clearly is not in the public interest.

Many policy makers, however, are not happy with these statistical methods because the procedures are so complicated and technical that they feel uncomfortable. Their discomfort is supported by the findings of LaLonde (1986) who compared the use of random panels with the econometric analysis and found the former to yield superior results. Random panels are particularly valuable if the object is to take account of 'committee selection' but they are of only limited value when taking account of 'self selection'. If we take, as an example, individuals or entrepreneurs who seek to obtain an award either of finance or advice - an example might be SBIR or Law 44 - then it would be appropriate to make a selection in the normal way, but, as a control, to allow a random sample of applicants access to the award without selection. The performance of the random access group would then be monitored over the same time as that of the selected applicants. If the selected applicants differ significantly in terms of observables from the random applicants then this would also have to be taken into account in the analysis. Nevertheless the

Table 8: Compare assisted with 'Match' firms taking account of sample selection**Approach**

- * Use of Statistical Techniques: Heckman 2 Step Estimator for testing and adjustment
- * Use of Random Panels

Problems

- * Policy makers (and some academics) feel uneasy about statistical 'adjustment'
- * Use of random panels could mean public money is given to firms/people who we know will not benefit

Examples

Author:	Year:	Topic:	Country:
Wren & Storey	1998	Subsidised Marketing Consultancy	UK
Westhead & Storey	1998	Undergraduate placement programme (STEP)	UK

prime purpose of the random access group is to seek to take account of the 'administrative selection' influence.

The effect of the impact of the policy would be the difference in performance between the assisted and the control group, after also eliminating the influence of selection.

The second part of Table 8 however shows problems remain even with these two approaches. Many analysts feel that the fairly complex statistical analysis in the Heckman 2-step procedure⁵ is difficult to communicate in simple language. Even if they understand it themselves, politicians, faced with having to explain the Heckman 2-step to taxpayers and the small business community, would risk being branded as indulging in 'statistical hocus pocus'. It is therefore unattractive on these grounds.

In principle, the use of random panels is more attractive because it is more easily understood. But, if it is known with some degree of certainty that only a small proportion of firms will significantly benefit from the scheme, and it is also known, in advance, the characteristics of those who will benefit - i.e. selection is accurate - then public money is being wasted in providing assistance to businesses which are unlikely to succeed. The business community itself could therefore justifiably complain that money which otherwise could be usefully used on the scheme is being wasted upon businesses with little prospects, in order merely to evaluate the impact of the scheme.

5. Despite the fact that Heckman is now a standard procedure in mid-range statistical analysis packages such as STATA or LIMDEP. It is not available on basic packages such as SPSS.

Overall, however, the key message is that selection, both in the form of 'self-selection; and 'administrative selection' is an important issue. Failure to take it into account seriously risks overestimating the impact of policy. Where administrative selection, in particular, is clearly prevalent there is a strong case for the limited use of random panels. Where self selection is likely then the more sophisticated statistical analysis has to be conducted - even if explaining the outcome to politicians could be tricky!

4. Conclusions

If public money is spent on SME support then it is vital that evaluation of the impact of these initiatives takes place. Unfortunately evaluation is not possible unless objectives, which are clear and, in principle, measurable are specified. Too often objectives are either not specified, or specified in a way which is overly vague and incapable of being used as the basis for deciding whether or not the policies are successful. In our judgement these objectives should be quantified and become explicit targets.

This paper has also argued that evaluation and monitoring are not identical. We view monitoring as collecting information about the firms in receipt of the scheme, together with financial information of monies expended. We also view monitoring as seeking only the opinions of recipients of the scheme. On the other hand evaluation seeks to compare performance of recipients with other groups of individuals or enterprises. Unfortunately most policy initiatives in OECD countries currently are merely monitored, rather than evaluated. In the terminology of this paper, such appraisals as are conducted rarely pass beyond Step 3, and in many instances do not pass beyond Step 1.

There are problems with all stages in the evaluation procedure, but currently best practice is Step 6. In our judgement new SME policies should ensure that, prior to their implementation, a budget is set aside to ensure that an evaluation plan is established to achieve at least a stage 5 level of evaluation. Governments are failing in their responsibilities to their taxpayers if they continue to finance 'evaluations' which are below those of Stage 5.

From the viewpoint of the research community it is important for the most sophisticated analysis possible to be undertaken. Almost all small business policies involve an element of selection - either 'administrative selection' or 'self selection'. The challenge to researchers is to seek to address the issue of selection. The problem, however, is the payoff to researchers is likely to be negative. This is because the experience chronicled in this paper suggests that, the more sophisticated and careful the analysis, the weaker the apparent impact of 'policy'. This is because the sophisticated analyst does not attribute to the policy, effects which are actually attributable to other influences such as selection, or firm characteristics.

Unfortunately the realpolitik of the situation is that policy makers generally (but not always) wish to demonstrate the effectiveness of their policies. They are therefore likely to favour sloppy analysts who are capable of “demonstrating” major policy impacts and disfavour careful analysts.

This has potentially serious consequences for the serious research community; it means we risk exclusion from the policy arena because we do not have access to data. Even where access is granted, the data will almost certainly have not been collected in an ideal way, i.e. not collected prior to the policy being implemented; no data on ‘control’ firms included; data on ‘administrative selection’ not collected.

The challenge then for the research community is to persuade policy makers that it is in their long term interests to carefully appraise policy, and to be involved with that appraisal prior to policies being introduced. Unfortunately such ideas may be somewhat naive since policy makers with apparently often very limited budgets prefer ‘cheap and cheerful’ research which will yield them ‘positive’ findings, rather than accurate and careful research where policy impact is likely to be less. The emphasis which most governments have upon competitive tendering for research contracts only serves to reinforce these competitive advantages of the ‘cheap and cheerful’ brigade.

References

- Barclays Bank (1998), *Small Business Review*, London.
- Bendick, M. Jr., and Egan, M.L. (1987), "Transfer Payment Diversion for Small Business Development: British and French Experience". *Industrial and Labor Relations Review*, 40, pp.528-42
- Bridge, S. O'Neill, K. and Cromie, S. (1998), *Understanding Enterprise, Entrepreneurship and Small Business*. MacMillan, London
- Brown, C., Hamilton, J., and Medoff, J. (1990), *Employers Large and Small*. Harvard University Press, Cambridge, Mass.
- Christman, J.J., Nelson, R.R. Hoy, F. and Robinson, R.B. (1985), "The Impact of SBDC Consulting Activities". *Journal of Small Business Management*, July, pp.1-11.
- Cosh, A. Duncan, J. and Hughes, A. (1998), "Investment in Training and Small Firm Growth and Survival: An Empirical Analysis for the UK 1987-95". *DfEE Report No.36*
- Deschoolmeester, D. Schamp, T. and Vandenbroucke, A.M. (1998), *The Influence of Management Training on Entrepreneurial Attitudes and Managerial Techniques of SMEs*. deVlerick School of Management, Ghent.
- Dolton, P.J., Makepeace, G.H. and van der Klaaw, W. (1989), *Occupational Choice and Earnings Determination: The Role of Sample Selection and Non-Pecuniary Factors*. Oxford Economic Papers 41: 573-94.
- DTI (1998), *Small Business Action Update*, Department of Trade and Industry, London, June.
- Elstrot, J.B. (1987), "Procedures for Improving the Evaluation of SBDC Consulting Activities". *Journal of Small Business Management*, January, pp.67-75.
- ENSR (1997). *The European Observatory for SMEs*, EIM, Zoetermeer, Netherlands.
- Ernst & Young (1996), *Evaluation of Business Links*. Department of Trade and Industry, London.
- Harrison, R.T. and Leitch, C.M. (1996), "Whatever You Hit Call the Target: An Alternative Approach to Small Business Policy", in *Small Firm Formation and Regional Economic Development*, M.W. Danson (ed), Routledge, London and New York.
- Hart, M. and Scott, R. (1994), "Measuring the Effectiveness of Small Firm Policy: Some Lessons from Northern Ireland". *Regional Studies*, Vol.28, No.8, pp.849-858.
- House of Commons, Trade and Industry Committee (1998). "Small and Medium Sized Enterprises", *Sixth Report*, The Stationary Office, London.
- Hughes, A. (1997), *Small Firms and Employment*, ESRC Centre for Business Research WP 71, September, University of Cambridge.
- de Koning, A. and Snijders, J. (1992). "Policy on Small and Medium Sized Enterprises in Countries of the European Community", *International Small Business Journal*, Vol.10, No.3, pp.25-39.
- LaLonde, R.J. (1986), "Evaluating the Econometric Evaluations of Training Programs with Experimental Data". *American Economic Review*, Vol.76, No.4, pp.604-620.
- Lerner, J. (1997), *The Government as Venture Capitalist: The Long Run Impact of the SBIR Programme*. Harvard Business School.
- Moint, A.H. (1998), "Small Firms Exporting: How Effective are Government Export Assistance Programs". *Journal of Small Business Management*, January, pp.1-15.
- O.E.C.D. (1995, 1996, 1997), "Best Practice Policies for Small and Medium Sized Enterprises" Paris.
- O'Higgins, N. (1994), "YTS, Employment and Sample Selection Bias". *Oxford Economic Papers* 46: 605-28.
- Rogoff, E.G. and Lee, M.S. (1996), "Putting Government's Role in Perspective: The Impact of Government Programmes on Entrepreneurs and Small Business Owners". *Journal of Developmental Entrepreneurship*, Vol.1, No.1, Spring, pp.57-73.
- Sardar, J.H., Ghosh, D. and Rosa, P. (1997), "The Importance of Support Services to Small Enterprise in Bangladesh". *Journal of Small Business Management*, Vol.35, No.2, April, pp.26-36.

- Segal Quince Wicksteed (1991b), *Evaluation of the Consultancy Initiatives - Third Stage*. Department of Trade and Industry, HMSO, London.
- Storey, D.J. (1994), *Understanding the Small Business Sector*. Routledge/TTP, London.
- Storey, D.J. and Strange, A. (1992), "Entrepreneurship in Cleveland 1979-1989: A Study of the Effects of the Enterprise Culture". Employment Department, *Research Series* No.3.
- Westhead, P. and Storey, D.J. (1994), *An Assessment of Firms Located on and off Science Parks in the UK*. London, HMSO.
- Wren, C. and Storey, D.J. (1998), "Estimating the Impact of Publicly Subsidised Advisory Services upon Small Firm Performance: The Case of the DTI Marketing Initiative". *SME Working Paper* No.58.