Evaluation of the Growth Services Range:

Statistical analysis using firm-based performance data

Prepared by the
Ministry of Economic Development

April 2009

Disclaimer

Acknowledgements

Thanks to the IBULDD team members (past and present) and the Data Laboratory staff at Statistics New Zealand. Thanks to the researchers working with the prototype Longitudinal Business Database for their technical assistance and insightful discussions, particularly Richard Fabling and Lynda Sanderson. Thanks to peer reviewers, David Storey, Dean Hyslop and Hong Tan for constructive comments that considerably improved the quality of this report. All remaining errors are of course the authors' own.

Table of Contents

List of Tables

1 Introduction

1.1 Purpose of this study

The *Growth Services Range* (GSR), which consists of the *Growth Services Fund* (GSF), *Market Development Services* (MkDS) and *Client Management Services* (CMS), was evaluated in 2005[1]. That evaluation provided evidence on outcomes arising from the scheme and the reach and quality of the service provided by NZTE. It used information drawn primarily from surveys of firms that had received GSR grants and/or services. The overall conclusion of the 2005 evaluation was that the GSR seemed effective in improving firm capability in areas that were considered to be important contributors to firm growth.

Questions remained, however, regarding the overall value for money provided by the scheme. As noted in the 2005 evaluation, the ultimate criterion for assessing the success of the GSR is that the total benefits to New Zealand outweigh the total costs. At the time, the information needed for a detailed cost-benefit analysis was not available. However, with the development of the prototype *Longitudinal Business Database* (LBD), we now have a potential means to address this wider question. The LBD is a source of business data held by Statistics New Zealand that addresses several constraints identified by the previous evaluation. It includes a rich set of firm characteristics and performance data, as well as information on participation in business assistance schemes, including GSR. The number of firms contained in the database, and the breadth of variables included, allow the use of econometric techniques to compare outcomes for firms that received GSR with a group of comparable firms that did not. These quantitative estimates of the impact of GSR complement the self-reported assessments of outcomes from the previous evaluation survey of GSR recipients. The approach taken in this report focuses on the direct benefits to firms receiving GSR grants and or services. We focus on the additional value-added in the firms receiving assistance as this is one measure of benefit to society[2]. The relevant costs are the grants and the cost of delivering the service.

It is worth noting that the objectives of GSR and the operation of the scheme have changed since the programme began in 2001[3]. The focus in this report is on the performance of firms receiving GSR assistance between 2001 and 2005.

---

[1] "Evaluation of the Growth Services Range" Research, Evaluation and Monitoring Team, Industry and Regional Development Branch, Ministry of Economic Development, November 2005.

[2] A more comprehensive study would also investigate the wider spillover benefits and displacement effects within the economy; however this is beyond the scope of this study.

[3] The most recent changes occurred in July 2008, with GSF eligibility criteria changing to support a wider range of firm internationalisation activities.

1.2 The Growth Services Range

The *Growth Services Range* consists of a package of grants and services intended to accelerate the development of firms with high growth potential.

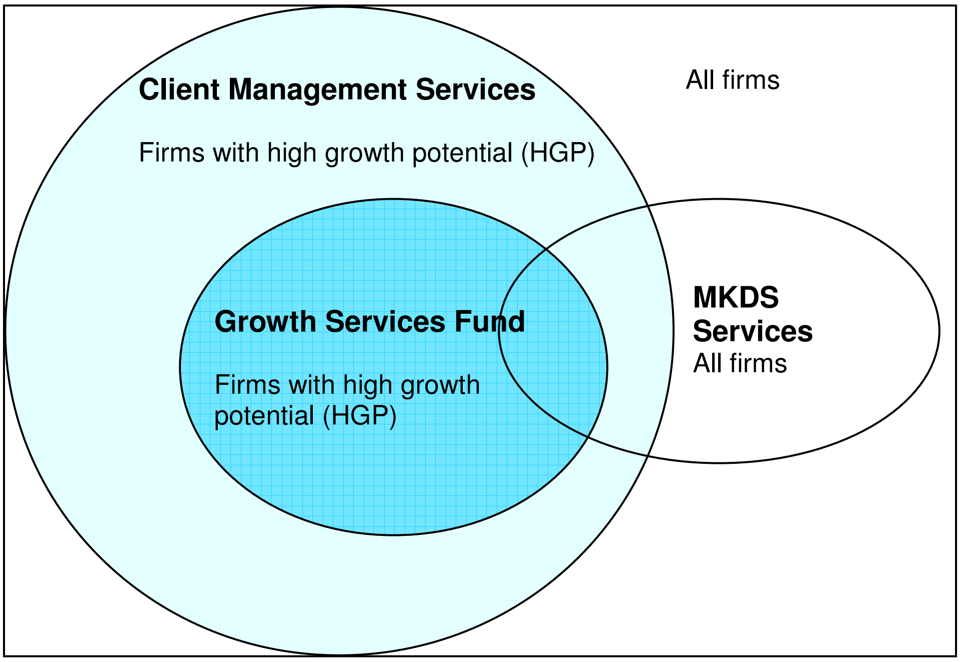
a. *Client Management Services* (CMS), which involves NZTE assigning a client manager to each participating firm ('client') to act as the primary interface between the firm and the services offered by NZTE (allocated $7.4 million in 2007/08);

b. *Growth Services Fund* (GSF), which offers funding assistance for firms to purchase external advice and expertise (allocated $5.9 million in 2007/08); and

c. *Market Development Services* (MkDS), which are provided by NZTE's offshore offices and comprises specialist information, advice and facilitation assistance (allocated $37.0 million in 2007/08[4]).

The relationship of the different components of the scheme is shown in Figure 1.

Figure 1 The Components of GSR



The GSF grant is only available to firms with *high growth potential*. High growth potential *(HGP)* is defined by NZTE as the potential to generate either average 20% per annum revenue growth sustainable for five years, or revenue growth of $5m within five years. All firms that receive a GSF grant also receive CMS advisory services from NZTE. In fact, all

[4] This amount excludes other programmes aimed at identifying and coordinating international market opportunities such as Beachheads ($6.5m); Marketing and Communications ($9.9m) and Generic Market Intelligence ($1.4m).

client firms that are classified as HGP, and some firms that are in the 'pipeline' to becoming HGP, receive the specialist CMS advisory services. By contrast, MKDS services are available to all firms that are deemed capable and willing to pay for the services. HGP clients may receive MKDS services at a reduced cost.

In this evaluation we focus primarily on firms that received GSF grants and CMS advisory services. Specifically, we look at the impact on groups of firms receiving:

a. Client Management Services and a GSF grant (combined GSF)

b. Client Management Services (CMS only)

We also consider the additional impact of receiving MKDS services and of receiving other types of assistance from NZTE and other agencies, but this is not the primary focus. There are some data issues relating to identification of firms that received CMS in the past. These are discussed in further detail in section 4 of this report. Consequently, our most reliable impact estimates relate to the combined impact of GSF and CMS.

1.3 The Evaluation Framework

This evaluation is aimed at assessing the direct impact of receiving GSR grants and/or services on participating firms. We have used a range of micro-econometric techniques to provide quantitative estimates of the impact of GSR. These techniques take into account the characteristics of firms, e.g. firm size and exporting history, and the fact that government assistance is typically not assigned randomly. These techniques have not been used before in New Zealand to evaluate the impact of government assistance to firms (although similar methods are used in the medical and labour policy field, and their use for business assistance is common in other countries). For the purposes of this report, it is necessary to introduce terminology relating to how we measure the impact of business assistance programmes.

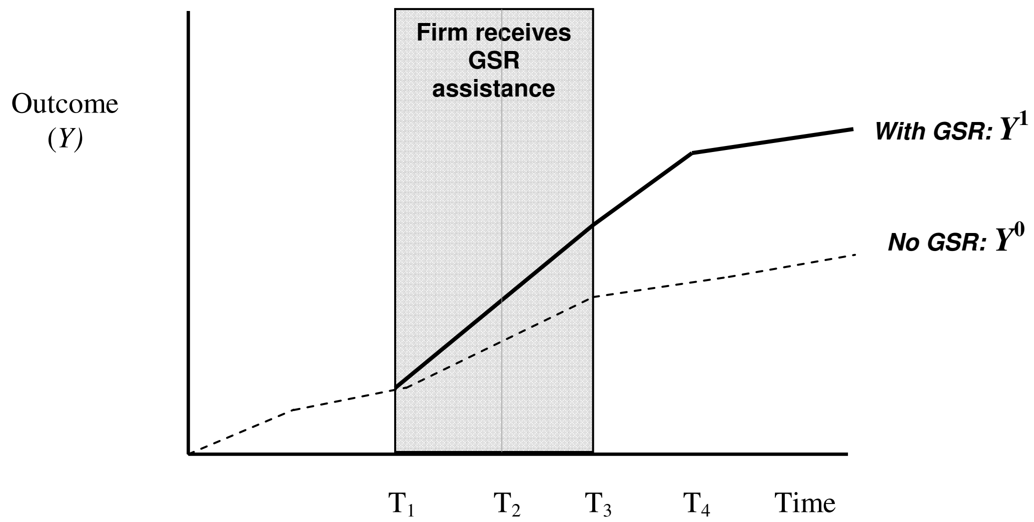
At the heart of evaluation is the following: In seeking to know the impact of a programme on a firm, we wish to compare what happens if they receive government assistance (in the language of the literature: receive the treatment) to what would happen otherwise. If we call the first Y^1 and the second Y^0, then the *treatment effect* for each firm i at any time t is defined as the difference between its potential outcomes:

(1) $$\alpha_{it} = Y_{it}^1 - Y_{it}^0$$

where the outcomes might be a firm's sales (if we are considering just the immediate/intermediate outcomes), value-added or productivity (if we are considering the ultimate intention of the scheme). The fundamental evaluation problem arises because we cannot observe both what would happen if the firm participated in the programme and what would happen if it did not. That is, we do not observe both Y^1 and Y^0. The outcome that we do not observe is called the 'counterfactual'.

The evaluation problem is illustrated in Figure 2[5]. Consider a firm that – in the absence of government assistance – would be expected to follow the outcome profile depicted by the dotted line in the figure. The firm receives assistance in time periods T_1, T_2 and T_3, which results in increases in the outcome levels one year later (as depicted in the bold line, $\boldsymbol{Y^1}$). Note that in the case depicted in the figure, the firm in question would have continued to grow even without the government assistance. This is as one would expect with the 'high growth potential' firms, which are the focus of the GSR. Nevertheless, in receipt of the GSR, the firm grows even faster still. The treatment effect is the difference between the two lines. In our example, there are 3 different periods of assistance. Thus, we can think of the assistance that the firm receives as three separate treatments, or one total treatment spanning three years. In the case described in the figure, there is a permanent impact on the firm as the final outcome level is higher in the assisted case even after the assistance stops. Because there is a lag of one year between each treatment and the impact on the outcome the two lines are identical at T_1 (when the firm first receives treatment), but continue to separate after the final instance of assistance (at T_3) until T_4. From then onwards, the line continues (and thus the assisted and unassisted firm would continue to grow) at a similar rate.

Figure 2 The impact of the GSR



[5] This illustration and discussion are based on similar material in Revesz and Lattimore (2001).

The overall impact (I) for an individual firm i at time t is:

(2) $$I_{it} = \sum_{\tau = T_2}^{T_4} \alpha_{i\tau} R_\tau$$

where T_2 is one year after the initial receipt of assistance, R_τ is a discount factor= $\frac{1}{(1+r)^{\tau - T_2 + 1}}$ and r is a discount rate[6].

Because we do not observe the counterfactual (Y^0), we have to somehow generate it artificially. This is the essence of how we must solve the evaluation problem – how we create a counterfactual. One way to do this is to find a suitable comparison group of firms and compare the outcomes of the firms receiving assistance with those of the control group. However, this is not quite as simple as it first seems. We cannot simply compare a group of firms receiving GSR with another random group selected from the business population. This is because we expect outcomes to vary among firms for all sorts of reasons, regardless of whether they receive assistance. Crucially, they are likely to be influenced by the characteristics of the firm (e.g. big firms produce more than small ones). **In order to make the appropriate comparison, we need to compare like with like.**

One crucial way in which the two sets of firms might differ is in their likelihood of being treated. If only firms with potential for growth receive government assistance, the whole of the business population (some of whom are not expected to grow at all) is not the appropriate comparison. Therefore, the evaluation will tend to overstate the impact of the assistance if comparison is made with the remainder of the population. This is referred to as *selection bias* in the impact estimate. The appropriate comparison group would be firms that also have the potential for high growth, but have not received assistance.

The GSR is very definitely *not* assigned randomly to firms. As we note in section 1.4 below, the GSF grant is only available to firms with high growth potential. Thus, the appropriate control group would be other high growth potential firms. One way to evaluate such a scheme would be to identify a set of high growth potential firms and then randomly assign GSR services to a subset. In such cases, the cost of identifying 'too many' high growth potential firms would be offset by the benefit of being able to compare outcomes. At least two other difficulties emerge with this. First, this requires an evaluation plan to be set out *before* the policy is implemented, which is not always the case. Second, it requires strict

6 See Section 5 for a discussion of appropriate discount rates.

policing to ensure that treatment is indeed random within the group of high growth potential firms.

There are a number of techniques that have been developed to reduce or remove this selection bias and provide robust impact estimates. The techniques that we have considered are outlined in Section 3 and discussed in more detail in an Appendix. There is no golden bullet supplying the perfect estimator for all cases as each estimator provides the correct answer only under certain assumptions. We must consider the institutional nature of the programme – how were firms informed about the programme, what are the eligibility criteria? Another important consideration is what the parameters of interest are – what outcomes are we trying to measure and how will we measure them?

1.4 Outline of this report

Section 2 gives an outline of the GSR programme covering the objectives, desired outcomes and information regarding eligibility for receiving GSR grants and services. Understanding the reasons why firms participate in GSR is vital for reducing selection bias because it helps our selection of an appropriate control group. We also present summary findings from the 2005 evaluation in this section.

Section 3 describes the two broad sets of techniques we use to estimate the impact of GSR. The detailed methodology is outlined in Appendix 1.

Section 4 presents the data used in this study. Appendix 2 presents a detailed table of variables and definitions.

Section 5 presents an overview of our results. Full tables of results are given in Appendix 3 and a detailed technical discussion of results is given in Appendix 4. We include a range of estimates for the average impact of GSR assistance to investigate the sensitivity of our results to different techniques and different underlying assumptions. We also discuss our results in light of previous evaluation evidence.

Section 6 presents conclusions.

2 Programme Information

2.1 Programme objectives

The Cabinet paper[7] that established the GSR stated that its aim is to "accelerate development of firms with high growth potential and enhance their contribution to New Zealand's overall economic growth." The 2005 evaluation has a full discussion of the policy rationale, development of GSR and operational detail. We summarise the main points in this report.

The policy rationale for the programme is that New Zealand firms do not invest sufficiently in expertise and information services that would help them achieve their growth potential. Some of the reasons noted in the Cabinet paper include:

- Firms may not realise benefits of external business advice
- Many firms lack the management expertise and knowledge necessary to grow their businesses;
- Owners may be unwilling to cede control of their businesses or lack confidence to take risks to grow their businesses;
- The pressures of day-to-day management and tight resource constraints may crowd-out a focus on longer term strategic issues;
- Firms may wrongly assume they are too small to export or find the costs and time to set up an offshore network prohibitive
- Some of the benefits may be external to the firm[8]

An intervention logic for the GSR was developed as part of the 2005 evaluation (Figure 3). It was discussed and agreed by MED, NZTE, and MFAT as representing the agencies' joint understanding of what the programme is intended to achieve and how. It includes a summary of the policy problems, activities and intermediate to ultimate outcomes. The focus of this evaluation is on the support offered under the *Growth Services Fund* (GSF) and *Client*

---

[7] from the Integration Cabinet paper EDC (03) 55

[8] Unlike programmes supporting research and development or knowledge transfer within the innovation system, the principal benefits of investing in business advisory services will rest with individual firms. However, there may also be some wider benefits to New Zealand. The skills developed in one firm can be transferred when staff share their knowledge with others or leave to work with other firms. If New Zealand firms increase their presence in offshore markets due to improved market intelligence, there can be flow-on benefits to other domestic firms (e.g., increased recognition of New Zealand products or demonstration benefits where domestic firms see that it is possible to enter new markets profitably (Hausmann and Rodrik, 2005)).

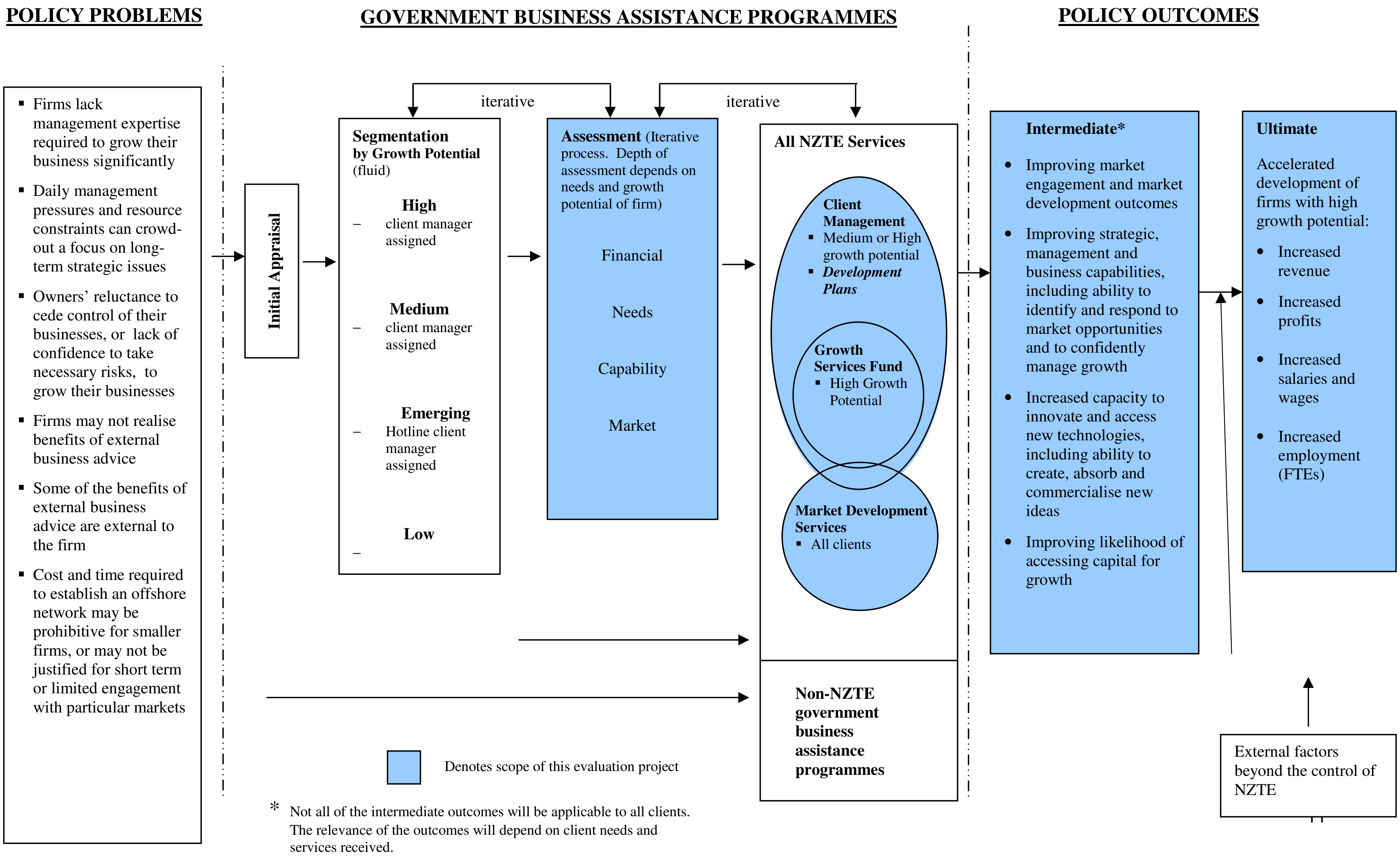
Management Services (CMS) within New Zealand, rather than the offshore networks also funded by the GSR budget.

The *intermediate outcomes* comprise both direct and shorter-term effects of the programme, such as changes in attitudes, knowledge, skills, abilities and behaviours of programme participants. The expected policy outcomes of improved management and business capability, increased innovation and adoption of new technologies/ideas, and increased capital for investment aim to increase firm productivity. Firms with more of these things should become more productive.

The *ultimate outcomes* are accelerated development of firms with high growth potential as measured by increased revenue (or sales), profits, and employment. The private benefits to the firm of government assistance are the increases in profits or shareholder value. The measure of importance to government is the benefits to New Zealand in terms of welfare. This is measured by an increase in value-added (defined as sales minus purchases). Thus, we consider value-added to be a better measure than profit for two reasons. First, it measures the impact on overall economic welfare. The second reason is that profits are difficult to measure because there is an obvious incentive to report low taxable profits (Fabling, Grimes, Sanderson and Stevens 2007).

Typically, indicators of ultimate outcomes are influenced by multiple factors beyond the programme, for example, general economic conditions and the state of a firm's own market. Many of these things would be expected to influence the outcomes of similar firms similarly.

Figure 3: NZTE GROWTH SERVICES RANGE **PROGRAMME LOGIC MODEL**



2.2 Programme mechanisms

The Client Management component of GSR evolved from what was originally a pilot called *Fast Forward New Zealand* to determine the benefits of transforming Industry New Zealand's *Business Growth Service* (operating since October 2000) in order to take a more proactive and co-ordinated approach to identifying and selecting high growth businesses and providing more intensive case management for these businesses.

In April 2003, Cabinet agreed to replace Industry New Zealand's *Business Growth Fund* with a more flexible GSF, which offers support for high growth potential firms to purchase external advice and expertise and market intelligence and development services. The GSF was intended to be highly flexible in terms of the level of funding provided. Funding is available for up to 50% of the costs of approved projects and is typically up to $100,000 per company within any 3 year period (with potential to raise this limit for exceptional cases).

2.2.1 Eligibility for Client Management Services

There have been changes in the application process over time. The following describes the processes up to 2005 (the selection time period of interest for this evaluation).

NZTE developed a client engagement model whereby all potential participants should have received an initial appraisal via the Enterprise Hotline (later called Business Evaluation Team). This is a high level assessment of a firm's growth potential and stage of development, which allowed NZTE to categorise firms in terms of their growth potential. If firms were considered to have potential for high growth they were eligible to receive Client Management Services and potentially a Growth Services Fund. These firms were assigned an NZTE Sector or Client Manager to help firms identify the strategies and services to address their needs.

High growth potential is defined as the potential to grow at a rate of 20% and/or $ 5million over the next 5 years. There is no formal checklist for determining whether a firm is classified as high growth potential, however the following factors influence the assessment. Firms should be:

- be high performing and have potential for significant growth;
- have demonstrated commitment to substantial growth;
- have a world-class product, service or intellectual property;
- have determination to be a world class business, typically demonstrated by commitment to a culture of innovation and best practice.

It is important to note that the 2005 GSR Evaluation found that most of the firms classified as high growth potential in 2005 had not come via the Enterprise Hotline. Rather, they were legacy clients from predecessor organisations, Industry New Zealand and Trade New Zealand. This has ramifications for how the impact of new or additional assistance is evaluated, particularly for these firms.

2.2.2 Eligibility for Growth Services Fund[9]

The GSF is only accessible to firms that are receiving Client Management Services. The Sector/Client Managers play a critical role in the application process. A firm's access to a GSF is only by invitation from Sector Managers after a thorough screening process. They assess whether a firm is eligible to receive a grant and assist the firm to compile a development plan. The screening can include background checks, assessments of a firm's capabilities and financials and potential of the proposal to achieve net economic benefit. The NZTE Manager then reviews the firm's eligibility with a NZTE Director and if both agree they proceed with a GSF proposal. The proposal is assessed by an independent external Assessment Panel, The rejection rate by the panel is low due to the prior screening process. In addition to the factors mentioned above as influencing selection for CMS, a firm receiving GSF should also:

- demonstrate a commitment to retain the value of the GSF proposal in New Zealand;
- have 100 or fewer full time equivalent employees and/or annual turnover of less than $NZ50 million.

In summary, the processes for determining whether a firm receives Client Management Services and a Growth Services Fund grant involve a fair degree of subjectivity on the part of NZTE staff. This may be useful and appropriate from an operational perspective, but it is worth pointing out that it makes it very difficult to undertake a robust evaluation due to the difficulty in identifying the most appropriate control group.

2.2.3 Eligibility for Market Development Services

In contrast to the CMS and GSF components of GSR, any firm is able to access Market Development Services if they are prepared to pay for these services. Firms that receive Client Management and/or GSF may receive these services at no or subsidised cost.

[9] This section also describes selection processes up to 2005.

2.3 Programme summary

NZTE and other government agencies provided information to Statistics New Zealand on the number and type of assistance received by individual firms over the past decade. These records were matched where possible to firm records in the prototype Longitudinal Business Database (see Section 4 for more detail). Table 1 shows the number of firms receiving different types of assistance from the Growth Services Range[10].

Table 1 Number of firms receiving Growth Services Range assistance by year

Year ending March	Received GSF	Received CMS only (and never GSF)	Received CMS (and may have received a GSF)	Received MkDS only
2001	15	0	12	762
2002	78	69	156	888
2003	192	165	417	1,242
2004	192	132	474	912
2005	147	150	540	573
2006	93	120	555	468

The numbers of firms receiving GSF grants increased from 15 in 2001 to 192 in 2003 and 2004[11]. The number of recipient firms declined after this period (with a corresponding increase in the size of grants approved). Some firms received multiple grants over different years. All GSF recipients also received CMS services. A large number of firms received CMS (columns 2 and 3) but the majority of those firms had received a GSF grant at some point in time. In 2006, there were 120 firms that were receiving CMS but had never had a GSF grant, out of a total of 555 firms receiving CMS. The total number of firms that had ever received a GSF grant and/or CMS services was 618 by 2006 (not shown in the table). The last column shows that many more firms had accessed MkDS services. For example, in 2003 there were 1242 firms accessing MkDS services, compared to 417 receiving CMS.

Table 2 shows the grant amount paid to GSF recipients by year. Note that this is the amount actually paid out on a reimbursement basis, not the approved dollar value. The amount approved could be considerably higher because the grants are paid on completion of the approved project.

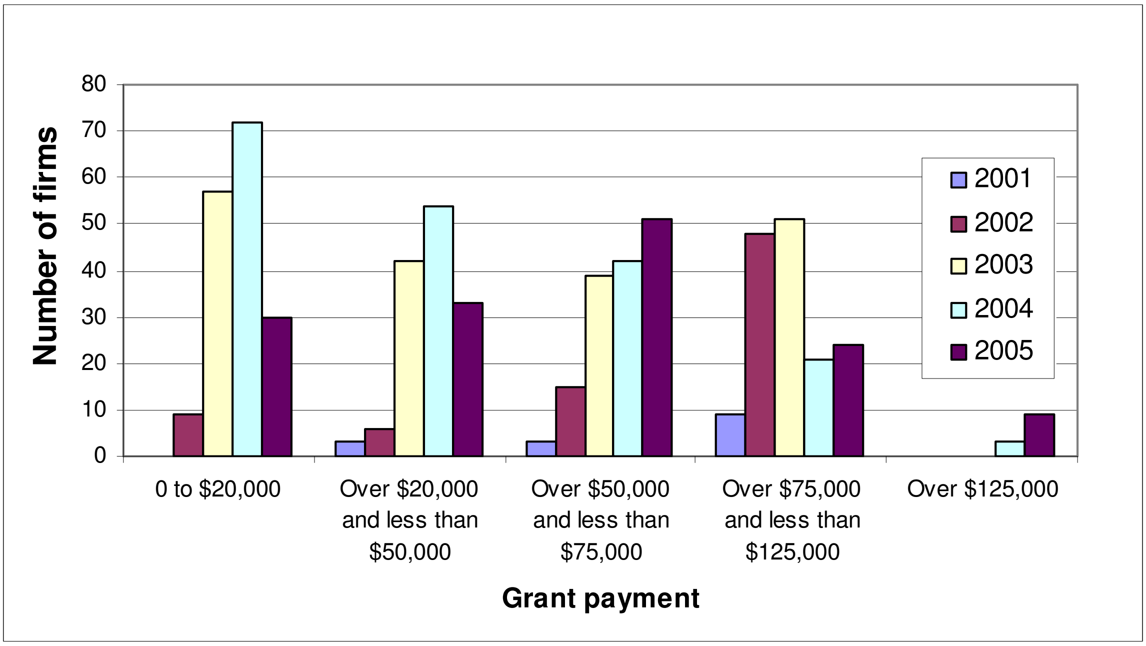
[10] These numbers are lower than the actual numbers because some firms could not be matched to the database. The matching rate is around 85%.
[11] All firm count tables in this report have been randomly rounded to base 3 for confidentiality reasons

Table 2 Amount of grant paid

Year	Total GSF grant paid ($million[12])
2001	1.04
2002	5.75
2003	8.90
2004	7.29
2005	8.89
2006	9.05

The grant payment distribution by year is shown Figure 4. Most grant payments were under $125,000 per year with a reasonably even split between the four different size categories ranging from under $20,000 to $75,000 to $125,000 per year.

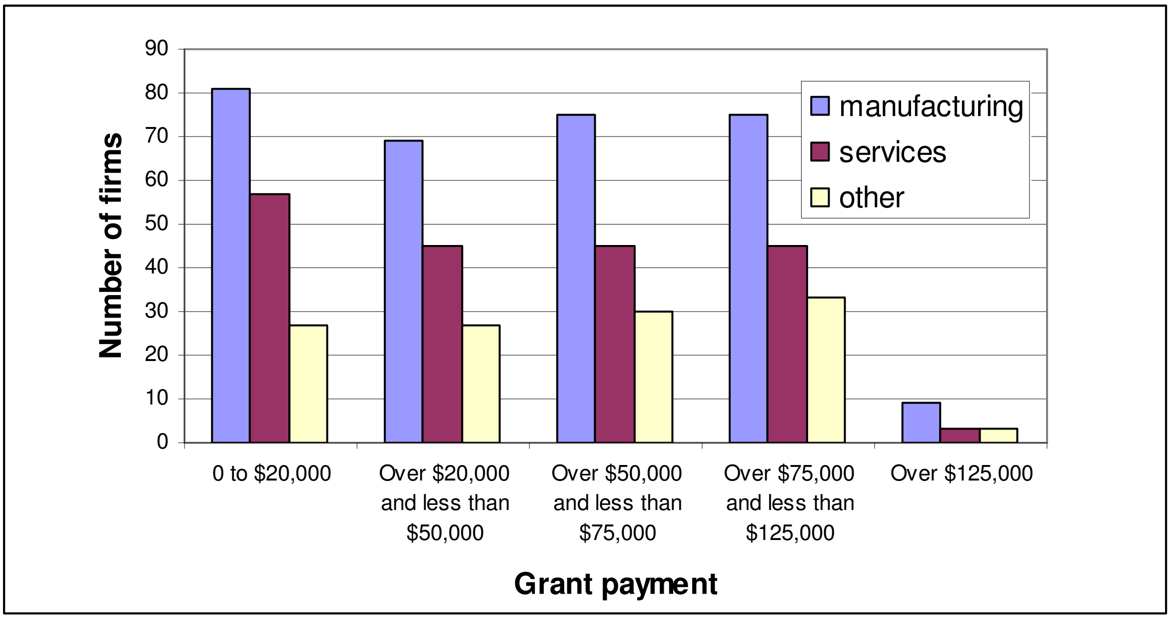
Figure 4 Grant payment distribution by year ending March



The grant payment distribution by industry is shown in Figure 5. The distribution of grant payments by size was similar between the three different industry groups.

[12] Nominal values are shown in this table

Figure 5 Grant payment distribution by industry



2.4 Characteristics of firms receiving GSR assistance

Descriptive statistics of some of our variables are summarised in Table 4 (shown at the end of this section). The data are described in more detail in Section 4. Statistics are shown for two groups: those that received GSR assistance ('Treated') and those that did not ('Untreated'). The assistance in question is whether a firm received the combined GSF grant and CMS services. The total number of treated observations over the full time period is about 1,130 compared to potentially over 2.9 million untreated observations. Each observation corresponds to a firm in a particular year over this time period. In any year, there are around 400,000-500,000 untreated firms. The total number of observations depends on the variable, e.g. the number of firms with employment data drops to 2.1 million (mainly because there are a large number of zero employment firms (Fabling, Grimes, Sanderson and Stevens, 2007).

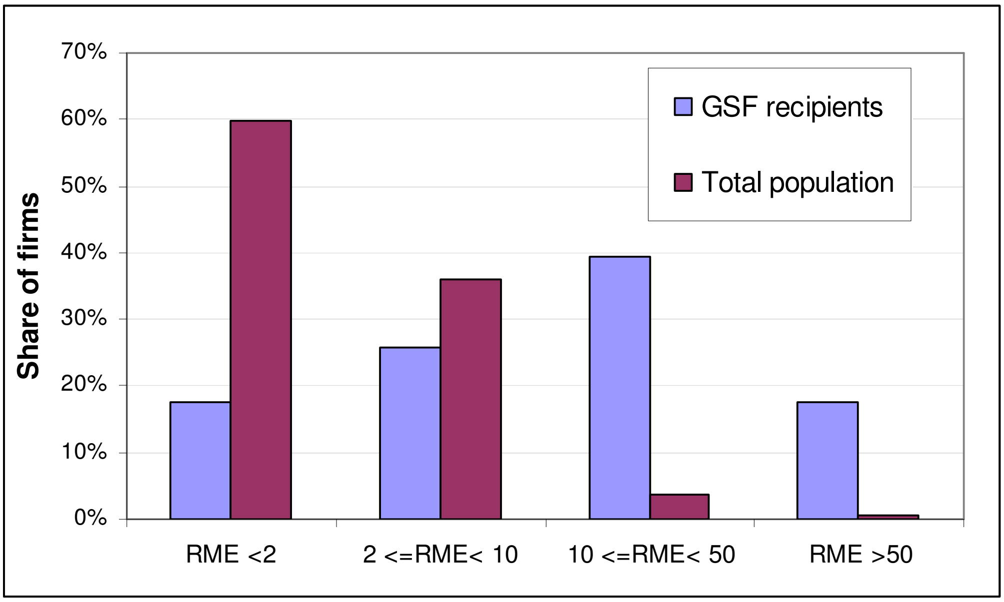
2.4.1 Size and industry of GSR recipients

The table provides some interesting insights into the characteristics of firms receiving GSR assistance. The characteristics of the untreated group reflect the total New Zealand population. It is well-known that most firms in New Zealand are small or medium-sized, e.g., in 2004 around 97% employed fewer than 10 people (SNZ, 2007). Consequently, the average employment count for the untreated total population is much lower than the treated group, at around four people (RME=3.7), between 2000 and 2006[13]. In contrast, firms that

[13] The distribution is highly skewed so care must be taken interpreting the averages; the table displays statistics for the logged values and the non-logged values.

received assistance are significantly larger, employing about 32 people on average. This is illustrated in Figure 6 which shows the distribution of GSF recipients and the total untreated population by RME categories.

Figure 6 Share of firms by employment size



This difference in employment distribution between GSF recipients and the total population could partly reflect industry differences in average firm size. We see from Table 4 that about half of the recipient firms were in the manufacturing industry group and about 30% were in the services industry, compared with about 6.5% and 39% for the total New Zealand population. It is likely that the higher average employment for GSF recipient firms is partly explained by the higher proportion of assisted firms in the manufacturing sector, which has relatively large firm sizes compared with other industries. This is illustrated in Table 3, which shows a breakdown of GSF recipient firms by employment (RME) and industry category. A great proportion of firms in the manufacturing sector (76%) have more than 10 employees compared to services firms (64%) and other industries (58%).

Table 3 Number of GSF recipients by employment size and industry

	Manufacturing	Services	Other industries	Total
RME <2	18	54	36	108 (17%)
2 <=RME< 10	54	69	36	159 (26%)
10 <=RME< 50	141	60	42	243 (39%)
RME >50	90	9	9	108 (17%)
Total	303 (49%)	192 (31%)	123 (20%)	618 (100%)

The higher firm size of the assisted group is also consistent with the difference in average age and levels of capital; GSF recipients tend on average to be four years older and have higher levels of capital than the average New Zealand firm.

2.4.2 Other government assistance of GSR recipients

Another interesting point relates to the high likelihood of GSF recipients receiving other types of assistance from NZTE and other agencies. About 20% of firms that had received a GSF grant had also received another type of NZTE grant, although not necessarily in the same year (*L_other_gsp*=0.207). Roughly, half of GSF recipients had also accessed MkDS prior to or at the same time as receiving GSF(*L_inv_job*=0.476). This is not surprising as GSF recipients receive intensive Client Management Services from NZTE staff and are therefore also likely to get advice regarding other types of business assistance provided by NZTE. More surprising perhaps is the fact that roughly half of the group (45%) had also received assistance from other government agencies, primarily FRST *(L_non_nzte*=0.446). There are several possible reasons for this: government agencies may be looking for firms to support from a small pool of firms; the schemes are well aligned; or some firms are serial consumers of government assistance.

2.4.3 Exporting behaviour of GSR recipients

Assisted firms are more likely to export, or be in a group that exports. About 44% of GSF recipients had exported during or prior to their receiving a grant, compared with less than 3% in the total population *(export_ind*=0.379 and *group_export_ind*=0.059). Clearly, there is a link between firms deemed to have the potential to grow their business and hence receive GSF assistance and exporting experience. GSF firms are also slightly more likely than unassisted firms to be foreign owned (*foreign owned*=0.019).

2.4.4 Outcomes of GSR recipients

Finally, we look at firm outcomes and see that these are all significantly higher than the unassisted firms. The average sales for a GSF recipient was around $6.88 million between 2000 and 2006; much higher than the average sales for an unassisted firm for the same time period ($783,606). Similarly, the average value added and productivity of GSF recipients was about $2.6 million and $82,000 per employee, respectively, whilst the unassisted averages were about $298,000 and $62,000 per employee. The group of firms receiving assistance are clearly higher performing than average. The question remains: are they higher performing solely due to their size, higher levels of capital or other characteristics, or has GSR assistance contributed to improving their performance? The econometric techniques used in this evaluation attempt to resolve whether these differences can be attributed to a firm receiving a GSF grant and NZTE advisory services.

Table 4 Descriptive statistics 2000 - 2006

Variable	Definition	Group	Number of Observations	Mean	Std. Dev	*t* statistic[14]
L_grant_gsf	Cumulative $ from GSF grants	T U	1,130 2,933,664	$63,190 n/a	$46,866 n/a	n/a
ln_sales	Log of sales	T U	1,130 2,933,664	14.578 11.305	1.862 1.892	59.1
sales	Sales	T U	1,130 2,933,664	$6,880,000 $783,606	$14,400,000 $16,700,000	14.2
age	Firm age	T U	1,130 2,933,664	13.988 9.141	12.38 8.567	13.2
ln_rme	Log of employment	T U	1041 2,172,302	2.828 0.743	1.311 0.927	51.3
rme	Employment	T U	1,130 2,172,302	31.513 3.729	45.546 48.894	20.5
ln_capital	Log of depreciation	T U	666 2,158,152	10.793 8.409	1.768 1.511	34.8
capital	Depreciation	T U	666 2,158,152	$171,679 $28,075	$343,797 $1,065,204	10.8
ln va	Log of value added	T U	976 2,389,906	13.743 10.679	1.633 1.734	58.6
va	Value added	T U	976 2,389,906	$2,628,830 $298,008	$8,474,066 $5,782,713	8.6
ln_prod	Log of labour productivity	T U	918 1,832,595	10.926 10.097	0.86 1.296	29.2
prod	Labour productivity	T U	918 1,832,595	$82,178 $62,437	$136,651 $1,271,325	4.3
L2_grantgsf	Cumulative $ from GSF grants	T U	1,130 2,933,664	$36,624 n/a	$42,688 n/a	n/a
L_grant_other	Cumulative $ from other NZTE grants	T U	1,130 2,933,664	$5,825 $31	$72,610 $5,412	2.7
L2_grant_other	Cumulative $ from other NZTE grants	T U	1,130 2,933,664	$2,440 $17	$14,068 $4,212	5.8
L_grant_non	Cumulative $ from non NZTE grants	T U	1,130 2,933,664	$54,974 $136	$151,099 $7,580	12.2
L2_grant_non	Cumulative $ from non NZTE grants	T U	1,130 2,933,664	$43,483 $84	$129,617 $5,014	11.3
manufacturing	Manufacturing sector dummy	T U	1,130 2,933,664	0.523 0.064	0.499 0.245	30.9
services	Services industry dummy	T U	1,130 2,933,664	0.302 0.391	0.459 0.488	-6.5
foreign owned	Foreign owned indicator	T U	1,130 2,933,664	0.019 0.006	0.138 0.078	3.2
export_ind	Exporting dummy for firm	T U	1130 2933664	0.379 0.018	0.486 0.136	25.0
group_export_ind	Exporting dummy for other firms in the group	T U	1123 2932739	0.059 0.008	0.236 0.092	7.2

[14] *t* statistic for the difference in means between treated and untreated observations

Variable	Definition	Group	Number of Observations	Mean	Std. Dev	*t* statistic[14]
L.other_gsp	Dummy for whether firm received other NZTE grants	T U	1130 2933664	0.207 0.001	0.405 0.031	17.1
L.inv_job	Dummy for whether firm accessed MKDS services	T U	1130 2933664	0.476 0.003	0.499 0.054	31.9
L.non_nzte	Dummy for whether firm received other non-NZTE grants	T U	1130 2933664	0.446 0.003	0.497 0.054	30.0

T = Treated, U = Untreated
Variables prefixed with 'L_" are lagged by one year; "L2_" are lagged by two year

2.5 Previous evaluation findings

2.5.1 Scope and methodology

The 2005 evaluation examined the effectiveness of GSR in achieving the desired outcomes outlined in the logic model. It focused on firms that had received GSR assistance between 2000 and 2005. The evaluation used a combination of quantitative and qualitative methods including: interviews with 35 firms and their NZTE Sector (Client) Managers; an online survey with responses from over 400 firms; analysis of data from NZTE's database and interviews with other NZTE staff and a number of Economic Development Agencies and business associations. This provided rich qualitative data on outcomes as well as the factors which affect the programme's impact. However, it was not possible at that stage to assess the performance of firms receiving GSR against a control group due to lack of suitable data.

2.5.2 Achievement of Intermediate Outcomes

The evaluation found that GSR appeared to be effective at improving firm capability in areas that are thought to be important for driving firm growth. The methodology focused on asking firms to identify changes in firm performance following receipt of assistance. Responses from GSR recipients showed the following benefits:

- A majority of firms believed that their market knowledge had improved since receiving GSR and most felt that this was partly due to receiving the assistance. In some cases, this resulted in cost savings as firms found out about a lack of market opportunities;
- Over half the firms also thought that they had improved connections and networking with international markets as a result of GSR assistance.
- Over half the firms reported improvement in a range of strategic, management and business capabilities since receiving GSR support. Improvement rates were highest for business knowledge and practices; strategic planning; sales and marketing practices and quality or process improvement practices.
- Over 2/3 of firms had improved innovation performance, including an increase in R&D; development and/or implementation of new products and processes and improved knowledge of new products, processes or technologies.

Firms that received GSF grants were significantly more likely to experience improvements compared to firms that only received MKDS (data for firms that received no GSR assistance were unavailable). Firms that received both GSF and MKDS had higher improvement rates than those that only received GSF. The impact of CMS advisory services was not able to be measured conclusively. The evaluation noted potential selection bias issues affecting their results – firms with stronger capabilities in the intermediate outcome areas could have been more likely to receive GSF. An attempt was made to mitigate this by phrasing questions to focus on improvements in performance after receipt of assistance.

2.5.3 Achievement of Ultimate Outcomes

The 2005 evaluation was unable to draw firm conclusions on the ultimate outcome of increasing firm growth. The main issue encountered by the evaluators was the short time lag between the intervention and the measurement of impact and the lack of control group data. However, a small group of firms that received GSF grants (or predecessors) between 2000 and 2001 were asked to estimate their growth rates since receiving the grant. About 40% (25%) of these firms reported that they had experienced over 20% average annual growth in sales (profit) since receiving the grant funding. We are now in a position to examine growth rates based on actual administrative data, instead of relying on survey responses.

Table 5 shows the share of firms with greater than 20% annual growth in sales for different groups receiving different types of GSR assistance and those that did not receive assistance (non-GSR). The first column shows shares calculated from all observations pooled over the time period 2000 to 2006. The data show that a larger share of firms receiving GSF or CMS have high growth rates compared with those that accessed MKDS or those that had no GSR assistance. This includes growth rates both before and after receiving assistance. In the next two columns we break this down into years preceding assistance and year of receipt and afterwards (post-receipt). The pre-receipt shares, at 40% and over, are higher than the pooled estimates and are much higher than the average for the rest of the population (30%). This is consistent with NZTE selecting fast growing firms to receive their Client Management Services and GSF grants.

However, the share of firms reporting high growth drops markedly after receipt of GSR assistance. Why is this? One simple interpretation is that assistance causes a decline in the growth rate of firms. However, there are many factors that are not taken into account in these comparisons, such as macroeconomic changes between 2000 and 2006 resulting in population-wide declines in growth. (Most of the post receipt observations occur after 2004, whereas pre receipt observations are more numerous in the first part of the time period). Moreover, one might expect it to be difficult to sustain such high levels of growth over several years (Hull and Arnold, 2008). The methodology used in this report addresses these issues in order to provide more reliable estimates of impact of GSR. We discuss this in the next section.

Table 5 Share of firms with more than 20% annual growth rates in sales

Year ending March	Pooled 2000 to 2006	Pre receipt	Post receipt
GSF	37%	44%	32%
CMS only	40%	43%	36%
MkDS only	32%	40%	29%
non GSR	30%	N/A	N/A

3 Method

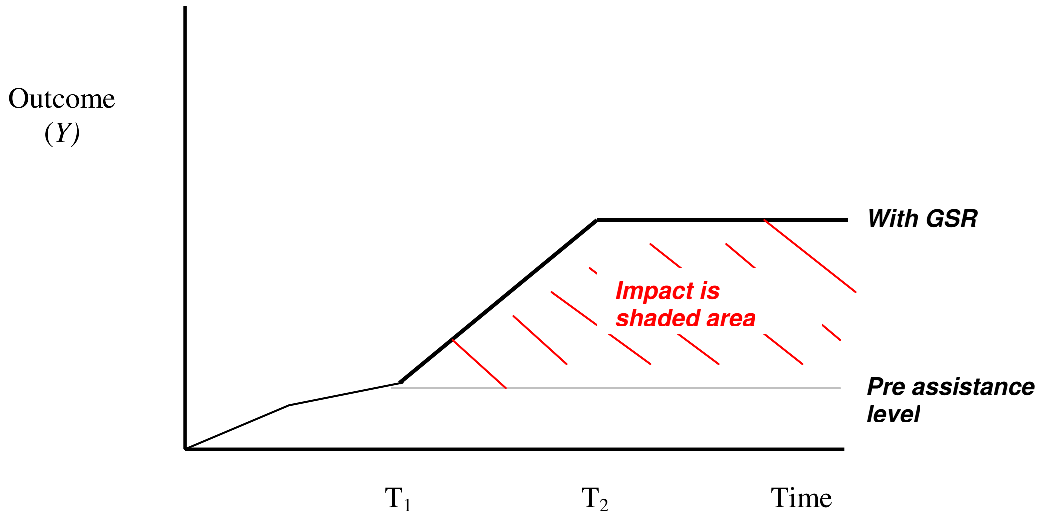
This section presents a brief outline of the techniques used in this study. The technical detail is discussed further in Appendix 1. As we stated in the introduction, all econometric methods attempt to reduce selection bias in order to assess the treatment effect. We focus on two sets of econometric techniques to estimate the direct benefits to firms of GSR.

3.1 Panel models

The first set of techniques takes advantage of the fact that we have several years of data to assess changes in firm performance after introduction of business assistance. We present results from a range of models that differ in terms of complexity but which all attempt to remove effects of factors which are common to all firms – whether assisted or not.

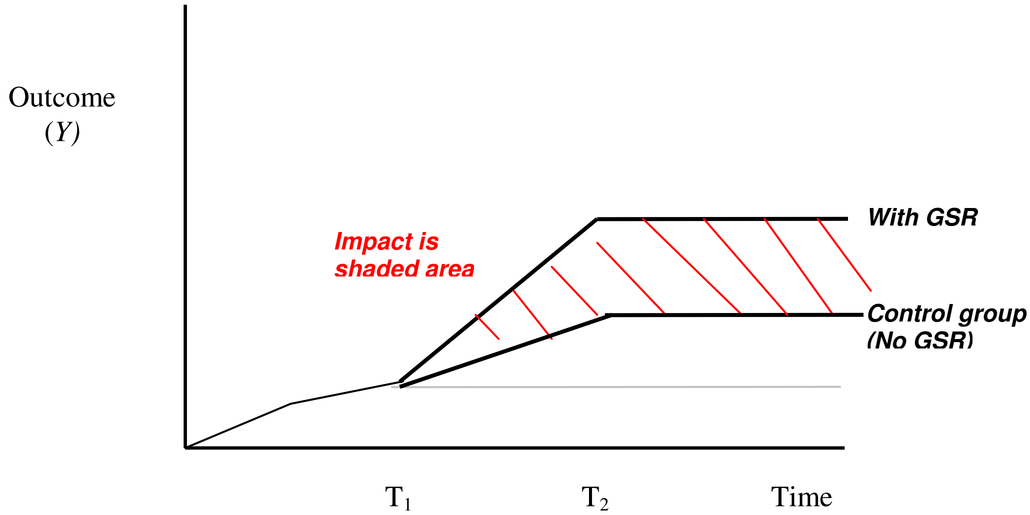
The simplest estimator one could calculate (*before-after*) compares the outcomes of firms before and after they received help. However, there are many reasons why a firm's performance might improve or decline from one period to the next. This calculation would attribute any change in performance wholly to the assistance (see Figure 7 and compare to Figure 2).

Figure 7 Before-after comparison



The *difference-in-difference* estimator looks at changes in time before and after assistance for two groups: the group of firms receiving assistance and a control group. The impact of the assistance equals the difference in the changes, or 'difference-in-differences'. In comparing firms before and after treatment, it assumes that the treated and control groups would grow and perform in the same way in the absence of assistance. Any remaining changes are attributed to the treatment (see Figure 8)

Figure 8 Difference in difference



The control group is chosen on the basis of similarity of firm characteristics such as size, industry, exporting activity and any other factors that enable us to assume similar performance to the treated group without assistance. This method relies on having comprehensive and high quality firm data over time.

This approach assumes that there are no differences in the way treated and untreated firms respond to external factors over time. There are at least two reasons why this may not be valid:

- First, treated and untreated firms may respond differently to changes in widespread influences, such as macroeconomic conditions. If, for example, treated firms were in one region and the controls in another, it is entirely possible that the two could experience different economic cycles.
- Second, the difference-in-differences estimator is sensitive to choices made by treated firms. For example, firms may just be in a bad year, and bad years tend on average to be followed by a return to normal. It may only be when cash or resources are short that firms look for government assistance.

If this is the case, we would expect treated firms to grow more than untreated ones merely because the only way is up[15].

We have used two ways of describing firms receiving support under the GSR programmes: a dummy variable model where a firm either receives assistance or not, and an intensity model where the treatment depends on the duration and amount of assistance received.

In this report, we focus on the average treatment effect across a group of firms receiving GSR assistance. This measures the additional impact on performance due to receiving assistance over and above what might be expected for a firm of similar size, exporting behaviour and other characteristics.

The outcomes we consider are sales, value-added and labour productivity. The explanatory variables relate to sales, employment, capital, exporting behaviour, industry and previous government assistance from NZTE and other agencies. We include only those we believe will have a significant effect on performance.

A common cause of bias occurs when there is a two-way relationship between the outcome and a factor influencing the outcome, such as between sales and employment. Not only do sales respond to changes in employment (as when a firm expands), but employment also responds to changes in sales (as in response to a demand shock). In these cases, the employment variable is considered to be *endogenous* and the resulting estimate of the treatment effect may be biased. One way of reducing this bias is to use lagged versions of the endogenous variable. These will be predetermined and therefore unable to be influenced by the current outcome. We examine the sensitivity of our treatment effect estimates by using lagged versions of any potentially problematic variables, such as employment, exporting activity and capital.

The most sophisticated models we have used are dynamic, allowing for the possibility that current outcomes are influenced by previous outcomes. As we shall see in the results section, these models provide our best estimates of the treatment effect. We note however that these techniques do not appear to be used widely in policy evaluation, perhaps due to data limitations[16].

---

[15] This may be further affected by differences in drop out rates (i.e. firm failure) between the two groups, since we can only compare changes in firms that are in both years.

[16] Although they are common in the analysis of firm behaviour. For examples relating to labour demand, see Arellano and Bond (1991) and Stevens (2005).

3.2 Matching Models

Another set of techniques matches firms in receipt of assistance with firms or groups of firms that do not receive it on the basis of a set of firm characteristics.

The first step is to identify the characteristics that predict whether a firm receives assistance. The *propensity score* is the conditional probability of receiving a treatment, in our case GSR assistance, given pre-treatment characteristics. We need to make two assumptions: (a) that the exposure to treatment is random once we account for firm characteristics and (b) that potential outcomes are independent of treatment assignment once we account for the firm characteristics. The balancing hypothesis means that firms with the same propensity score must look the same (have the same mean characteristics), regardless of whether or not they receive assistance.

We estimate the propensity score with variables related to previous sales, previous sales growth, employment, capital, exporting behaviour, industry and previous government assistance from NZTE and other agencies. Testing the balancing hypothesis plays an important part of model specification in estimating the propensity score. Once we have groups of firms made up of treated and untreated firms with a similar likelihood of receiving assistance, we can test the balancing assumption.

Once we have balanced the treated and control firms in each group, we move on to estimate the effect of receiving GSR assistance. There is a range of possible matching techniques. Essentially, there is a balancing act to be made between using too few matches and too many. If we use too few matches, we (a) do not take account of the heterogeneity of outcome in our large population of firms and/or (b) potentially increase the 'cost' of using a 'bad' match. If we use too many, we run the risk of matching the treated firms with firms that are unlike them and thus no use as comparators (Morris and Stevens, 2007).

4 Data

4.1 The Longitudinal Business Database

The firm data used in our analysis comes from the prototype Longitudinal Business Database (LBD), which contains data for financial years 2000 to 2006 from a number of sources. The data are provided at the enterprise level, defined as a business or service entity operating in New Zealand. The spine of the LBD consists of the Longitudinal Business Frame (LBF), to which are attached Goods and Services Tax (GST), financial returns (IR10) and aggregated Pay-As-You-Earn (PAYE) returns provided by the Inland Revenue Department (IRD). All data are annualised to firms' actual balance date, and then assigned to the closest year ending 31st March. The data are described in more detail in Fabling, Grimes, Sanderson and Stevens (2007).

The LBD records information about ownership, including the relationship of an enterprise to any parent or subsidiary enterprises. Most enterprises operate independently. However, a small proportion of the population are part of a group structure sharing a common group-top enterprise with other enterprises. These groups, whilst few in number, account for a substantial proportion of total employment and value-added (Fabling, Grimes and Sanderson, 2008). We include variables that take into account whether a firm is part of a group, such as when a firm belongs to a group that exports or one that receives government assistance.

4.1.1 Government assistance data in LBD

The database includes participation data for business assistance schemes administered by New Zealand Trade & Enterprise (NZT&E), Foundation for Research, Science and Technology (FRST) and Te Puni Kōkiri (TPK). These agencies provided lists of firms that had received assistance with information on the duration and amount of assistance. These details were probabilistically linked (on contact details) to the LBD. This results in a matching rate of about 85% of the firms. All schemes that provide direct assistance to businesses are included. Some of the government programmes matched to the LBD are current and others have been terminated or superseded by other programmes. The matched scheme data extend back to the 1990s while our analysis is restricted to 2000-2006 because of the coverage of the LBD.

Whilst, the information held in LBD regarding grant programmes is of good quality, we have less reliable data relating to the type and intensity of services received by firms from NZTE staff. NZTE have started to collect detailed information regarding services in the last couple of years so it is straightforward to identify current HGP firms. However, historically these data were not collected so we have only sketchy information regarding past services received. For the purposes of this evaluation, we used two snapshots of HGP clients and a best guess for their status preceding and between those two points in time to construct a time series. One snapshot is the current list (as of July 2008). The other is a list of HGP firms compiled for the 2005 evaluation. If a firm was on both lists then it was assumed to be an HGP client for the intervening years. If the firm was on the current list but not in 2005 then it was assumed to have started receiving CMS in the year ending March 2008. If it was not on the current list but appeared in 2005 then we assumed the services stopped in year ending March, 2005. If a firm was an HGP client in 2005, then we assumed that the start date occurred on the first recorded date of engagement with NZTE.

It is important to realise that there could be significant correlations between different types of assistance. Many firms access more than one type of government assistance. We include variables relating to other NZTE programmes and assistance from other agencies as explanatory variables.

4.1.2 Business population

We restrict our population to include "private for profit" firms and exclude households, ANZSIC division M (Government Administration and Defence) and firms located offshore. We only include firms that are considered to be economically active[17] and have at one time been included in Statistics New Zealand's Business Frame and have therefore been assigned an ANZSIC industry classification. The number of distinct firms in our population is around 700,000 (about 440 to 500 thousand firms in any one year between 2000 and 2005); a large number of these had zero employment over the entire period they were active (SNZ, 2007).

4.1.3 Model variables

The outcome variables and explanatory variables are described in Appendix 2. The Growth Services Range aims to accelerate the development of firms with high growth potential and stimulate their economic growth. We test the scheme impact on the

[17] Defined by a number of criteria, including non-zero GST sales/purchases; RME and selected IR10 variables

following growth outcomes: sales[18], value-added (defined as sales – purchases)[19] and labour productivity (value-added per employment). The last variable could be considered to be an ultimate outcome for any business assistance programme regardless of its intermediate objective, whether that be increased growth in sales, exports or R&D activity. Consequently, it is likely to be the most difficult to observe on the relatively short time scales associated with the study.

Although, the LBD contains a wide array of financial and other information we focus only on a subset of firm characteristics –those thought to be most important in order to approximate random selection into treatment.

[18] The sales and purchases derive from GST returns. GST returns have the advantage of much higher coverage rates across the population compared to IR10 or Annual Enterprise Survey data. However, there is one issue that affect our study: the sales and purchases variables should include government grants and subsidies from the Crown (with some exceptions). The inclusion of GSF grant amounts will cloud interpretation of the treatment effect on sales. For example, if the grant is recorded in the same year as it was approved then we may underestimate the impact on sales in the following year. Similarly, if the grant is recorded in the year following approval, then the treatment effect on sales may be overestimated. Unfortunately, we do not have payment date information in LBD so it cannot be e removed from the data – although it will be included in future. We do know from external data that roughly half the GSF recipients receive payment in the same year as approval and the other half receive payment one year following approval (although a small number receive payment at longer lags). The amounts dispersed between the two groups are roughly equal. Therefore, it is likely that the two different biases in our treatment effect may be cancelled out. This problem also appears to affect other sources of sales such as those from IR10 and Annual Enterprise Survey.

[19] More correctly we would also include changes in stocks as a consumption variable. However, Fabling *et al.*, 2007 have shown that the relative contribution of stock adjustment to value-added is small, so we exclude it from our analysis in order to retain a higher number of observations.

5 Results

In this section we outline our results. We consider four methods: two panel models and two propensity score matching models. These take different approaches to the evaluation problem – in particular how they consider selection bias. The two panel models consider the impact of business assistance in two ways. The first measures the average impact of GSR assistance on firms (what we call the 'dummy variable' model), the second considers how the impact of treatment depends upon the intensity of treatment, measured by the dollar value of assistance (what we call the 'intensity model'). The two matching methods consider the impact of GSR assistance on the level of outcomes and on the change (or 'first difference') in outcomes.

One of our most effective tools for removing selection bias is to use panel methods to remove factors that we cannot observe but are fixed in time. For example, consider a firm that has excellent management practices, a culture of innovation and an above average skilled workforce. Say, this firm receives a GSF grant and is included in the treated group. One might argue that it is these less tangible characteristics of the firm that have driven higher sales growth, which has been falsely attributed to the grant. However, because our panel methods remove any constant unmeasured factors, they are only a concern if they have also changed over the differencing period. We think the panel model results that remove the fixed effects are closer to the true treatment impact than those that do not[20].

In summary, although there is some variation in the estimates from the different techniques, as one would expect, the estimates from two quite different techniques are broadly similar. This gives us some confidence in our results. Because no method can completely purge the results of all sample selection bias, we therefore believe that the estimates represent an upper bound. That is, they are likely to represent a best possible scenario for treatment impact. We present a number of impact estimates to illustrate the range of estimates.

A summary of our results is presented in Table 6. The summary table shows the range of estimated impacts due to GSR assistance (excluding MkDS) on sales, value-added and productivity for our preferred model specifications. The average

[20] The methods that remove fixed effects include all panel models and the propensity score matching with first differencing.

impact of GSR assistance is measured by the *Average Treatment Effect on the Treated* (ATT) (defined in Appendix 1). The full set of results tables is included in Appendix 3. A comprehensive technical discussion of these results can be found Appendix 4.

Table 6 Summary Results for Firms receiving GSR assistance

Method	Outcome variable	Average ATT estimate[1]	Additional outcome per firm per year[2]
Panel : dummy variable	**Sales**	9%	$215,000
Panel: intensity model		4%[3]	$102,000
Matching: levels[4]		17%	$410,500
Matching: differences		14%	$370,000
Panel : dummy variable	**Value added**	(4%)[5]	$36,800
Panel: intensity model		4%[3]	$34,100
Matching: levels		18%[6]	$147,000
Matching: differences		11% [6]	$86,000
Panel : dummy variable	**Productivity**	9%[3]	$4,900 per worker
Panel: intensity model		6%	$3,400 per worker
Matching: levels		insignificant	N/A
Matching: differences		insignificant	N/A

Notes:

1) For the panel models, the average ATT estimate is 100*(exp(α_0 -1) where α_0 is the coefficient for the treatment variable. For the intensity model, the treatment coefficient is also multiplied by 1000 times the average grant size. See the appendix for discussion.
2) The additional outcome per firm per year is calculated by multiplying the ATT by the mean value of the outcome for unassisted firms. Mean values are taken from Table 7.
3) However, ATT is not statistically significant for all of the preferred model specifications
4) We present average estimates for manufacturing and services industries for all matching results unless stated (levels and differences). We exclude 'other' industry estimates from these averages because we have reservations about their validity because the numbers of observations are smaller and paired firms could be from quite different industries.
5) However, ATT is not statistically significant for any of the preferred model specifications
6) However, ATT only significant for manufacturing industry, not services.

5.1 Impact on Sales

The results show that the combined impact of receiving a GSF grant and associated Client Management Services had, on average, a positive impact on firm sales. The size of the average impact depends on the method employed and is estimated to be between 4-17%[21]. The additional sales due to assistance are over and above the levels of sales that would have been achieved due to other factors, such as, firm size, levels of capital, industry type, and levels of sales prior to receiving GSR assistance. They are also over and above the impact of receiving other types of grants and services from NZTE and other government agencies because we controlled for those separately.

We found that the impact due to GSR assistance was best modelled as a mean effect (the levels of sales increase after receiving assistance and remain high but further growth (or decline) cannot be attributed to assistance). Interestingly, the impact was strongly dependent on the year of treatment and was significantly more positive for firms receiving assistance in earlier years (2002 and 2003). The impact was insignificant and possibly negative for firms receiving treatment in 2005[22]. This may be linked to changes in selection processes over time. Recall that firms are selected to receive assistance if they are assessed by NZTE staff as having potential for high growth. One of the ways in which earlier treated firms were different from those receiving assistance in later years is that many of them had established histories with NZTE or its predecessor organisation. It is possible that those that NZTE thought would benefit the most were the firms first selected for assistance. Hence, the more positive treatment impacts for the earlier years.

How does the treatment effect of 4-9% translate into dollar values of additional sales? The additional sales are calculated by multiplying the treatment effect by the mean sales of firms that are similar to those that received GSR in every way except that they did not receive any assistance. We can use the output of the matching techniques to identify these 'like' firms and their mean outcome levels. We use these values to calculate additional outcomes due to treatment that are shown in the last column of Table 6. Using the weighted overall average sales for untreated firms values and a treatment effect of 4-9% gives additional sales of $102,000 -$215,000 per firm one year following assistance.

5.2 Impact on Value-added

The value for money received from GSR assistance is calculated from the additional value-added generated compared to the cost of delivering the programme. Our econometric results for value-added were less conclusive than for sales. Impact estimates ranged from 4-18% depending on technique; however, many estimates

[21] We think the true estimate is closer to the lower end of this range, as we discuss in the text below.
[22] Recall, *Lgsf2006* refers to firms receiving assistance prior to 2006, i.e., in 2005.

were not statistically significant. Similar to the sales estimates, we believe that the higher end of the range is upwards biased due to selection bias. Even the lower values of 4% impact may overestimate the true impact.

An impact of 4% corresponds to additional value-added of around $34,100 to per firm per year following assistance. This is a one off impact and does not appear to change over time. (We did not measure any additional impact in value-added in subsequent years although we allowed for this possibility with our model.)

The total impact of GSF on firms receiving assistance between 2001 and 2005 can be estimated by summing the value-added contributions from each firm for all the years over which they experienced heightened levels of value-added post GSF funding (see equation 2). We need to specify a realistic duration of the impact, i.e. the average number of years that a firm might experience higher value-added. To illustrate, we assume that the average impact persists for 3 years. Without discounting, this translates to an additional $57million or 149% of the cost associated with the grant between 2001 and 2005[23]. (We do not have outcome data to assess the impact of firms receiving assistance in 2006). Choosing a discount rate of 10.5%, this drops to 134% of the grant paid.

Of course, the choice of 3 years for the duration of the impact is arbitrary. We estimated the sensitivity of our results to differing durations of impact and differing discount rates. Results are shown in Table 7 where we display the total impact in dollars divided by the total cost (grant paid to firms plus overhead) for our low and high impact estimates. As can be seen from that table, the results are sensitive to both assumed discount rate and duration of impact.

Table 7 Projected Total Impact versus Costs

Nominal Discount rate[24]	1 Year	3 Years	5 Years
0%	50%	149%	249%
6.50%	49%	139%	219%
10.50%	49%	134%	203%
12.50%	49%	132%	197%

[23] Without discounting, the total grant paid between 2001 and 2005 is $31.9million. The 2005 evaluation estimated the cost of delivering the programme to be around 20% of the grant value. Thus the total cost of delivering the grant between 2001 and 2005 is $38.2million.

[24] We are making a rough adjustment for inflation of 2.5% and examining the sensitivity to real discount rates of 4,8 and 10%. The 10% figure is that recommended by the NZ Treasury until July 2008, when it lowered this to 8% (http://www.treasury.govt.nz/publications/guidance/costbenefitanalysis). The 4% figure is closer to that used in other jurisdictions, e.g. HM Treasury in the UK (http://www.hm-treasury.gov.uk/data_greenbook_index.htm).

Assuming reasonable values for the duration of impact (3 - 5 years) and discount rates (10.5%) gives value for money estimate of 134 -203%. This is an encouraging result because the benefits of the programme exceed the costs. Of course, the value for money calculation is approximate and does not take into account other effects which might decrease the net economic benefit, such as displacement effects. However, it is difficult to believe that inclusion of displacement effects would significantly reduce the net benefit. However, it is likely that despite our best attempts the treatment effect still includes selection bias resulting in impact estimates that are too high.

5.3 Impact on Productivity

Our results for productivity are less conclusive. There is evidence for significant impact on productivity of firms from the panel models. The size of the impact is between 6% and 9% which translates to an additional $3,400 to $4,900 per worker per year. However, matching methods show no significant impact associated with GSR assistance.

5.4 Comparison with previous evaluation evidence

How does this compare to previous evaluation evidence? In 2005, firms reported improvements in a number of areas important to firm growth as a result of receiving GSR assistance. These ranged from improvements in market knowledge, international connectedness to management and other business capabilities. These improvements should be reflected in improvements to firm's sales, value-added and productivity. Our results show significant positive impacts on these outcomes due to GSR, which corroborates the firms' perceptions of improvements in their business activities and capabilities.

The previous evaluation was unable to provide conclusive evidence of the programme's impact on firm growth due to a lack of a control group. Therefore, we have no direct measures of impact to compare with our econometric results. However, surveyed firms were asked to provide estimates of growth between 2000 and 2001. Results from early recipients of GSF grants were analysed separately to provide an estimate of post intervention growth. About 40% (25%) of those firms reported strong growth in sales (profit) since receiving GSR assistance. Most attributed their profit growth at least partly to GSR assistance. Our preliminary analysis of growth rates of all firms in the 2000 to 2006 time period (see Table 5) showed that roughly a third of all firms that had received a GSF had annual growth over 20% in sales in the years following assistance. This is consistent with the survey findings. However, our analysis also showed that the growth preceding the GSR assistance was significantly higher than the period following assistance. This highlights the issues with interpreting trend information on firms without good data before and after intervention and without information on other factors that could influence growth.

The econometric techniques provide the most accurate estimate of impact by teasing out the component of firm growth that is directly attributable to the assistance. While a crude comparison of growth data pre and post GSR assistance could be interpreted as GSR assistance having a negative impact on firm growth, our econometric techniques show that there is in fact a significant positive impact. This is consistent with the surveyed firms reporting that at least part of their growth in profits was attributable to GSR assistance. Thus while the growth rate of GSR firms declined following receipt of assistance (possibly due to macroeconomic changes), our results show that the decline in growth was reduced due to the fact they received GSR.

5.5 Comparison of different components of GSR

The focus in this evaluation has been on the GSF and Client Management Services. The initial comparison of sales growth rates (see Table 5) suggested that firms receiving CMS alone performed better (i.e. had a higher share of firms with strong growth) than firms receiving a combination of GSF grants and services. However, we were unable to differentiate between the impact due to GSF grant and the impact due to the associated services using econometric techniques. We attempted to do so by assessing the impact of the services for those firms that did not receive GSF grants, but did receive services. Unfortunately the impact estimates were variable and the uncertainties were higher. This may reflect in part the poorer quality data relating to services received compared to grant information. Similar, to the previous evaluation, we can draw no conclusions about the relative impact of the GSF and CMS components of GSR.

Our results do show evidence that accessing MkDS services results in an additional impact on firms, over and above that due to receiving the other components of GSR assistance. The additional impact for firms accessing MkDS is about 6% compared to those that don't use these services and this impact appears in turnover, value-added and productivity. This is consistent with previous evaluation evidence which found that firms that received all components of GSR performed better on average (had larger shares of firms reporting strong growth) than those that received GSF and CMS alone. Interestingly, the impact of government assistance received from other agencies (primarily FRST) is also positive and significant (2-10%). This assistance is mainly related to technology development and might be expected to have longer lags than the GSR assistance. We detect a short-term impact although we are not using the best model to measure longer term impact.

5.6 Comparison with international evidence

Our results are comparable to other international studies aimed at evaluating the impact of business assistance on small to medium enterprises (e.g.,OECD 2007 and World Bank 2007). As those reviews show, the more sophisticated impact assessments that use econometric techniques to isolate selection bias (i.e., approaches that could be classified as Step V1 within Storey's 'Six Steps to Heaven' framework[25]) are still the exception rather than the rule. It is typical to find little or no significant impact on the ultimate outcomes of value-added or productivity using more sophisticated approaches, although some studies have detected impact at the intermediate outcome level.

For example, a review of support for small businesses in Northern Ireland and the Republic of Ireland (Roper and Hewitt-Dundas, 2001) used selection models to determine the impact of firm performance. The study found that grant support had no impact on turnover growth or profitability of firms although they did detect a positive impact on employment growth. Another study (World Bank 2007) presented results from a range of techniques, including propensity score matching and first differences, to assess the impact of SME business support in Mexico. They found evidence to suggest that assistance improved intermediate outcomes, such as training and adoption of technology, but none to suggest any improvement in productivity or other final outcomes for firms. Similar to our study, the results were sensitive to the methodology adopted to determine impact.

We have primarily focused on the average impact across all firms, i.e., the impact averaged across all GSR recipients. However, international evidence suggests that the impact can vary significantly between different types of firms. For example, Wren and Story (2004) used selection models to gauge the impact of advisory services and found that the impact was most effective for firms with 10-80

[25] Storey (2000) presents a framework for classifying impact assessments as a gradation from monitoring activities to evaluation. Step I is the least sophisticated procedure (e.g. monitoring the take-up of a programme) and Step 6 the most sophisticated (comparison of treated firms with a control group, taking into account the selection bias).

employees and insignificant for smaller firms. This sort of information is particularly useful for policy makers who are interested in refining the objectives or implementation of a particular scheme. We have examined some aspects, e.g. differences between different sectors using matching techniques. When considering how these policies and their delivery might be improved it will be useful to explore these further. A useful starting point will be to investigate whether the GSR impact is more effective for different firm sizes.

6 Conclusions

The aim of this report was to provide quantitative estimates of the direct benefits to firms receiving GSR assistance in comparison to similar firms that did not receive assistance. Our main conclusions are:

- GSR recipients differ from the average New Zealand firm. Compared to the average: they have higher levels of employment, sales, value added and productivity; are more likely to export and be in the manufacturing sector; and are more likely to receive other types of government funded business assistance.
- GSR assistance has a significant positive impact on the sales of firms receiving the assistance[26]. The impact on value-added and productivity due to GSR assistance is less conclusive. We have measured this additionality in firm performance due to GSR using econometric techniques. These new results are consistent with the 2005 evaluation findings based on self-reports of assisted firms[27].
- We found that the impact was best modelled as a mean effect - the levels of sales increase after receiving assistance and remain high but neither continue to grow further nor decline. We found no significant additional impact at longer lags, although we allowed for this in our models.
- When presenting econometric results, it is important to present ranges of estimates due to the sensitivity of results to methodology. Although, the broad similarity of impact estimates using two different techniques gives us some confidence in our results, we have found it very difficult to remove selection bias from our estimates. While we are fortunate to have a rich dataset in order to exploit these techniques to their full potential, we are hampered by the fact that the GSR programme is targeted at firms that are likely to do well, regardless of whether they receive any help from government.
- We believe that the lower end of the ranges presented are the closest to the true impact, and that even those estimates should be considered an upper

[26] See footnote 18 for a discussion of potential bias in our treatment effect from the use of GST data. The direction of the bias is unknown and there is nothing we can do to mitigate it until we have linked improved grant payment information into the LBD. This linking is planned for 2009.

[27] The previous evaluation was unable to provide conclusive evidence of the programme's impact on firm growth due to a lack of a control group. However, firms reported improvements in a number of areas important for firm growth which they attributed to GSR assistance.

bound. This corresponds to a treatment effect of 4% for sales and value-added and 6% on productivity. (These estimates are accurate provided that the main unobserved differences between GSR firms and the control group influencing outcomes are fixed in time)

- The treatment effects translate to $102,000 higher sales, $34,1000 higher value-added and $3,400 per worker higher productivity for GSR recipients compared to similar firms that did not receive assistance. This result applies to the average impact for firms that received GSR between 2001 and 2005. The impact was strongest for firms receiving assistance prior to 2004. It was insignificant and possibly negative for firms receiving assistance in 2005. This may be due to changes in selection processes over time. It will be interesting to see whether this trend in reduced impact persists.
- We have estimated the value for money of GSR assistance over this period. This takes into account the total impact of firm performance compared to the costs of the programme. With reasonable choices of parameters, we find that total additional value-added compared to the costs is about 134 -203%. For government interventions to offer value for money, we would expect that the total additional benefits should exceed the total costs. This is, therefore, an encouraging result however we believe it represents an upper bound.
- Our analysis is based on surviving firms for the period of analysis from 2001 to 2006. There may be systematic differences in the probability of firm survival between assisted and unassisted firms. For example, part of the benefit of business assistance could be a reduction in the probability of firm failure. Alternatively, firms that receive assistance may be more likely to survive than those that do not. This is another type of selection bias. Our methods compare the sales, value-added and productivity outcomes for treated and untreated firms only where information on these are available, i.e. only for firms that survive. Excluding failed firms from our analysis may bias the results. However, as the two examples above show, the direction of such a bias is uncertain.
- This report focuses on the impact of the policy averaged across all assisted firms. While beyond the scope of this study, it would be useful to ascertain which factors in combination with GSR assistance provide the largest improvements in firm performance. For example. GSR assistance may be

more effective if targeted at particular firm sizes. This can be accommodated within an econometric methodology but is beyond the scope of this study.

7 Bibliography

Arellano, M. and Bond, S. (1991), 'Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations,' *Review of Economic Studies*, 58, pp. 277-97.

Arellano, M. and Bover, O. (1995), 'Another look at the instrumental variables estimation of error-components models.', *Journal of Econometrics*, 68,pp. 29-51.

Baum, C., Schaffer, M. and Stillman, S. (2003), 'Instrumental variables and GMM: Estimation and testing', Boston College, Dept. of Economics, Working Paper N9. 545, February 2008.

Becker, Sascha. O., and Ichino, Andrea, (2002), 'Estimation of average treatment effects based on propensity scores', *Stata Journal*, 7.2, pp. 358-77.

Blundell, R. and Bond, S. (1998),'Initial conditions and moment restrictions in dynamic panel data models.', *Journal of Econometrics*, 87, 11-143.

Fabling, Richard, Arthur Grimes and Lynda Sanderson (2008), 'Export Market Choices for New Zealand Firms', Presentation to CAED'08, Budapest

Fabling, Richard, Arthur Grimes, Lynda Sanderson and Philip Stevens, (2007), 'Some Rise by Vice and Others by Virtue Fall: Firm Dynamics, Market Structure and Performance', Presentation to New Zealand Association of Economists Annual Conference, Christchurch.

Hausmann, R and D. Rodrik (2003) Economic Development as Self-Discovery, *Journal of Development Economics*, 72.

Hull L., and Arnold, R, (2008), 'New Zealand firm growth as change in turnover', mimeo, Ministry of Economic Development.

Morris, M., and Stevens, P., (2007), 'Evaluating Government Business Assistance Programmes: An exploratory analysis using a sample NZT&E scheme', mimeo, Ministry of Economic Development.

OECD (2007), 'OECD Framework for the Evaluation of SME and Entrepreneurship Policies and Programmes' – ISBN 978-92-64-04008-3

Revesz, J. and Lattimore, R. (2001), "Statistical Analysis of the Use and Impact of Government Business Programs", Productivity Commission Staff Research Paper, Ausinfo, Canberra, November.

Roodman, D. (2006),'How to Do xtabond2: An Introduction to "Difference" and "System" GMM in Stata', Working Paper, Number 103, Centre for Global Development.

Roper, S. and Hewitt-Dundas, N. (2001), 'Grant Assistance and Small Firm Developments in Northern Ireland and the Republic of Ireland', Scottish *Journal of Political Economy* ,48 (1) , p.p 99-117

Rubin, D., (1974), 'Estimating Causal Effects to Treatments in Randomised and Nonrandomised Studies, *Journal of Educational Psychology*, 66, pp. 688-701.

Smith, J., (2000), 'A Critical Survey of Empirical Methods for Evaluating Active Labor Market Policies', *Zeitschrift für Volkswirtschaft und Statistik*, 136.3, pp. 1-22.

SNZ (2007), Potential Outputs from the Longitudinal Business Database, Statistics New Zealand, http://www.stats.govt.nz/NR/rdonlyres/7390A3FB-F43D-4ADB-A957-11432BB6A1B0/0/PotentialOutputsfromtheLongitudinalBusinessDatabase.pdf

Stevens, P.A., (2007). "Skill shortages and firms' employment behaviour," *Labour Economics*, vol. 14(2), pp. 231-249

Storey, D., (2000), 'Six Steps to Heaven: evaluating the impact of public policies to support small businesses in developed economies', in Landstrom H and Secton D.L.(eds) *Handbook of Entrepreneurship,* Blackwells, Oxford, pp 176-194..

Wooldridge (2006). *Introductory Econometrics A Modern Approach*, 3rd ed. USA, Thomson South-Western.

World Bank (2007), 'How well do Small and Medium Enterprise programs work? Evaluating Mexico's Small and Medium Enterprise Programs using panel firm data', Mexico Country Management Unit, Poverty Reduction and Economic Management Division, Latin America and Caribbean Region, World Bank.

Wren, C. and Storey, D., (2002), 'Evaluating the effect of soft business support upon small firm performance', *Oxford Economic Papers*, 54, p.p. 334-365.

Appendix 1: Detailed Method

This section presents a brief outline of the techniques used in this study. We focus on two sets of econometric techniques to estimate the direct benefits to firms of GSR. One set takes advantage of the panel nature of the data to remove factors that are fixed in time. This can sometimes eliminate or reduce sources of estimation error, such as omitted variables or relationships between the treatment variable and other independent regressors. The other method uses firm characteristics to match firms that receive assistance with those that do not. Note that the focus of these techniques is the direct benefits of business assistance for the firms in receipt of assistance. Indirect benefits of GSR, such as knowledge spillovers to other firms, are not estimated using these techniques.

A1.1 General model

We start with a general model linking a firm's outcome (such as sales of productivity) to a number of firm characteristics (such as the firm size or industry grouping), and to whether or not the firm received GSR assistance:

(3) $$Y_{it} = \boldsymbol{\beta}\mathbf{X}_{it} + \alpha_t D_{it} + \varepsilon_{it}$$

where

(4) $$\varepsilon_{it} = u_i + v_{it}$$

(5) $$E(u_i v_{jt}) = 0 \, , E(v_{it}) = 0$$

$\mathbf{X}_{it}$ is a vector of terms representing the firm characteristics or confounding variables for firm i at time t. D_{it} is a treatment variable indicating whether a firm received GSR assistance prior to that time period. The residual ε_{it} is split into a time invariant component u_i (the fixed effect) and an idiosyncratic term v_{it}. The simplest model assumes the impact of the treatment occurs at one time period (i.e., one year) after receiving GSR assistance. We will also include time dependent treatment effects to test whether longer term impacts are significant.

The α_t coefficient is the average value of the treatment effect across a group of firms receiving GSR assistance. This is known as the average treatment effect of the treated, ATT. It is the expected value of the difference between the assisted and unassisted outcomes for firms that have been selected into the scheme conditional on the characteristics X of the firms.

We consider two conceptual models for the treatment variable D_{it} .

Model i) *Dummy treatment variable*:

Let $d_{ik} = 1$ once a firm has received their first assistance, then

$$D_{it} = \begin{cases} 1, & \text{if } d_{it} = 1, t \geq t_0 \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

Model ii) *Intensity model:*

Let $Investment_{ik}$ represent the actual dollar value received from a grant[28], then the treatment variable is the cumulative dollars received prior to that time:

$$D_{it} = \sum_{k=1}^{t-1} Investment_{ik} \tag{7}$$

The first model uses a dummy variable for treatment. We assume that that the effect of receiving a grant has two components: a one-off permanent effect on outcomes and a term that depends linearly on time, i.e., $\alpha_t = \alpha_0 + \alpha_1(t - t_0)$. Thus, we interact D_{it} with a time dummy variable (years since treatment) to estimate the latter term. We also allow the one-off component α_0 to depend on time.

The second model assumes a grant's impact depends on the amount of grant assistance over time. We examined the time varying component of impact in two different ways: by including a linear term $\alpha_1(t - t_0)$ and by including additional terms corresponding to different lags since treatment.

Unfortunately, we are unable to provide an intensity measurement associated with CMS advisory services because this type of data has not been recorded. We only know whether the services were received or not so we use a dummy variable to indicate whether a firm received CMS services.

[28] We use nominal dollar values as we will be transforming the above equation by first differencing.

A1.2 Panel Methods

First difference and Fixed Effects

The standard panel data approach is to transform the data to remove the fixed effects u_i. This removes any differences between the treated and control firms that are fixed in time. Two common methods are difference-in-difference estimation and fixed effects (or within) estimation; one transforms the data by first-differencing; the other by de-meaning. If there are endogenous regressors or omitted variables that are fixed in time, then these transformations remove any associated bias.

Mathematically, the first difference estimator (FD) is obtained by first differencing the outcome equation (3) to give:

$$\text{(8)} \quad \begin{aligned} \Delta Y_{it} &= Y_{it} - Y_{i,t-1} \\ &= \boldsymbol{\beta}\Delta\mathbf{X}_{it} + \alpha\Delta D_{it} + \Delta v_{it} \end{aligned}$$

where Δ is the lag operator, e.g., $\Delta X_t = X_t - X_{t-1}$ and we assume α is constant for clarity.

The fixed effect transformation (FE) takes deviations from the firm means:

$$\text{(9)} \quad Y_{it} - Y_i = \boldsymbol{\beta}(\mathbf{X}_{it} - \mathbf{X}_i) + \alpha(D_{it} - D_i) + (v_{it} - v_i)$$

where firm means are denoted by single subscript *i*. Constant regressors, such as, e.g. industry dummies, are removed by both transformations.

The first difference-in-difference and fixed effect methods are identical if there are only two time periods. Both methods assume that the error term v_{it} is uncorrelated with the regressors X and selection into treatment D, and also that v_{it} is serially uncorrelated. If these assumptions hold then both estimates are consistent but the FE estimator is more efficient. (e.g. Wooldridge 2006). The requirement that treatment selection is uncorrelated with the idiosyncratic error means that there can be no pre-treatment dips in performance[29]. The fixed effects are removed by differencing and so we need not be concerned about any correlation between the fixed effect and treatment selection or other control variables.

[29] For more, see the example given in section 3.1 of the main text.

Dynamic Model extensions and Instrumental Variables

We next consider a dynamic model extension of the general model, where current outcomes are also influenced by past outcomes, i.e.

(10) $Y_{it} = \gamma Y_{it-1} + \boldsymbol{\beta}\mathbf{X}_{it} + \alpha_t D_{it} + \varepsilon_{it}$

The standard first-difference or fixed effects approach has problems because the transformed lagged dependent variable is correlated with the transformed error term which violates our assumptions for unbiased estimates. To illustrate using the first-difference approach: $\Delta Y_{it-1} = Y_{it-1} - Y_{it-2}$ will be correlated with the transformed error $\Delta v_{it} = v_{it} - v_{it-1}$ through the v_{it-1} term.

One way of overcoming this issue is to use instruments for troublesome variables and apply the Two-Stage Least Squares (2SLS) approach. This requires the presence of an instrument which is related to the endogenous variable of interest, in this case ΔY_{it-1}, but is itself uncorrelated with the error term. Lagged versions of the endogenous variables are potential candidates for instruments because these variables are predetermined and cannot be influenced by current outcomes. Potential instruments for ΔY_{it-1} are ΔY_{it-2} or Y_{it-2} because both those variables are correlated to ΔY_{it-1} but they are not mathematically related to the differenced error Δv_{it}. (However they may be still be correlated with Δv_{it} if the idiosyncratic errors are serially correlated). Note ΔY_{it-1} cannot be instrumented away with a double lag Y_{it-2} in the case of the fixed effects transform, because the transformed error term contains all error terms from all time periods.

One of the problems with using lagged variables as instruments is the loss of observations in a short time series – in the case above, we need to start our estimation at *t=3* because we need a lagged instrument at *t=2.* Panel methods such as difference and system General Methods of Moments (GMM) (Arellano-Bond 1991, Arellano-Bover(1995), Blundell-Bond (1998)) are newer, more general techniques that use a similar rationale to cope with endogenous regressors whilst minimising data loss. The techniques take full advantage of the availability of panel data and make fewer assumptions about the underlying data generating process. They are designed for panels with many observations (large N) and few points in time (small T). There is a question around suitability of these types of techniques when T is too small, however, as may be the case with our dataset. A good review paper on dynamic panel estimators with practical guidance for implementation is given by Roodman (2006). We present difference GMM results in this report. We note however that the GMM techniques do not appear to be used widely in policy evaluation, perhaps due to data limitations.

Summary of approach for panel models

All our results are based on a first-order dynamic model (see Equation 10) and use a first difference transform to remove the fixed effects. We examined the sensitivity of our results to different instrument sets for the lagged dependent variable. We also suspected that some of our control variables could be endogenous, particularly those related to RME, capital and exporting. For example, there could be a two way relationship between sales and RME where higher employment might lead to higher sales, which in turn might lead to more employment. Therefore, we also instrumented some of our control variables. Unfortunately, we were unable to find any suitable instruments for the treatment variables, although we tried various combinations of the dependent and control variables.

The specifications are summarised in the following table:

Table 8 Panel Model Descriptions

Model	Stata Code	Description
1	*Ivreg2*	*Instrumented*: ΔY_{it-1} *Instruments*: Y_{it-2}
2	*Ivreg2*	*Instrumented*: ΔY_{it-1} *Instruments*: Y_{it-2}, X_{it-1}, X_{it-2}
3	*Ivreg2*	*Instrumented*: $\Delta Y_{it-1}, \Delta Y_{it-2}$ *Instruments*: Y_{it-3}, $X_{it-1}, X_{it-2}, X_{it-3}$
4	*Ivreg2*	*Instrumented*: $\Delta Y_{it-1}, \Delta X_{it}$ *Instruments*: Y_{it-2}, X_{it-1}, X_{it-2}
5	*Ivreg2*	*Instrumented*: $\Delta Y_{it-1}, \Delta Y_{it-2}, \Delta X_{it}$ *Instruments*: Y_{it-3}, $X_{it-1}, X_{it-2}, X_{it-3}$
6	*Xtabond2*	*Instrumented*: $\Delta Y_{it-1}, \Delta X_{it}$ *GMM style instruments*: Y_{it-4}, X_{it-3}, X_{it-4}
7	*Xtabond2*	*Instrumented*: $\Delta Y_{it-1}, \Delta Y_{it-2}, \Delta X_{it}$ *Instruments*: Y_{it-4}, Y_{it-5} $X_{it-1}, X_{it-3}, X_{it-4}$
Y refers to the outcome variable; X refers to the control variables ln_rme, ln_rme2, ln_capital,ln_capital2 and export_ind		

All our estimates produce standard errors that are robust to heteroskedasticity and arbitrary patterns of autocorrelation within individuals. We used *ivreg2* for Models 1-

5 (Baum et al, 2003). Models 6 and 7 aim to improve the efficiency of the previous estimates by using GMM style instruments which incorporate deeper lags and replace missing values in the instrument set with zeros so the sample size is not reduced. The Stata code is *xtabond2* and we chose options 'noleveleq' and 'robust' (Roodman, 2006).

A1.3 Matching Models

Another set of techniques takes a 'structuralist' approach and models all the determinants of the outcome. Matching estimators work by matching firms in receipt of assistance with firms or groups of firms that do not receive it. This match is made on the basis of a set of firm characteristics. Conditioning on all relevant characteristics is problematic when there are many of them. This leads us to matching via an index, e.g. propensity score matching (see below). The pros of matching are that it overcomes many of the problems with regression. It is less restrictive than regression because it doesn't specify a linear or log-linear functional relationship between the outcome and factors that influence it. It does not difference away factors that are fixed in time that may be important for influencing the outcome. The cons are that it is sometimes difficult to satisfy the required assumptions, particularly with a very large sample[30].

Propensity score matching

The first step is identify the characteristics that predict whether a firm receives assistance. The *propensity score* is the conditional probability of receiving a treatment, in our case GSR assistance, given pre-treatment characteristics. Rosenbaum and Rubin (1983) show that if exposure to treatment is random within cells defined by X, it is also random within cells defined by the values of the following variable:

$$p(X) \equiv \Pr(D = 1|\mathbf{X}) = E(D|\mathbf{X})$$

Thus we can write the average treatment effect on the treated as

$$\begin{aligned}\alpha_{ATT} &\equiv E(Y^1 - Y^0|D = 1)\\ &= E\{E(Y^1 - Y^0|D = 1, p(\mathbf{X}))\}\\ &= E\{E(Y^1|D = 1, p(\mathbf{X})) - E(Y^0|D = 0, p(\mathbf{X}))|D = 1\}\end{aligned}$$

We need to make two assumptions to derive the above equations:

(a) that the exposure to treatment is random once we account for firm characteristics, i.e. $D \perp X|p(\mathbf{X})$

(b) that potential outcomes are independent of treatment assignment once we account for the firm characteristics, i.e. $Y^0, Y^1 \perp D|p(\mathbf{X})$. Assumption

30 In particular, the balancing hypothesis-see below.

(a) – the *balancing hypothesis* – means that firms with the same propensity score should look the same (have the same value of the vector **X**), regardless of whether or not they are treated. Testing the balancing hypothesis is an important part of model specification in estimating the propensity score.

We estimate the propensity score using a standard probit model with variables related to previous sales, previous sales growth, employment, capital, exporting behaviour, industry and previous government assistance from NZTE and other agencies. We use Becker and Ichino's (2002) *pscore* routine for this task.

After estimating a probit model of the probability of treatment, we split the sample into equally-spaced strata according to their propensity score. For each stratum, we test whether the propensity scores for treated and untreated firms are not different. We continue to divide the respective stratum until the propensity scores are not different for all strata. Once we have strata of firms made up of treated and untreated firms with a similar likelihood of receiving assistance, we test the appropriateness of the balancing hypothesis. That is we perform a *t*-test of the means of the explanatory variables for the two types of firms. If these are all insignificant, we consider our data to be balanced and appropriate for estimation of the treatment effect. Exploratory analysis showed that many variables which we thought would be important for explaining participation in GSR failed the balancing hypothesis and were therefore excluded[31].

Once the balancing hypothesis is satisfied we move on to the second stage of computing the treatment effect. In what follows, we denote the set of control units that are matched to an assisted firm i as $C(i)$ and the outcomes of treated and untreated (control) firms as Y^T and Y^C respectively[32]. We also drop the time subscripts because there are essentially only two periods, pre- and post–treatment. Next we discuss the different types of matching estimators.

[31] Since the balancing hypothesis is testing many comparisons of means, it is simple to see how a very large sample will allow one to state whether two sub-populations are *statistically* distinguishable, even if they are not necessarily *economically* different. For example because of the low standard errors, we may be able to say with a high degree of certainty that the population of treated firms has on average 1c more sales than the untreated one – however, it is a moot point as to whether we would consider such a small different as qualitatively important. Moreover, with ten strata and ten variables, we would fully expect any comparison of means to find at least one statistically significant difference at the 99% confidence level, *even if the populations were drawn from the same population*.

[32] The discussion that follows and the programming of the Stata code to calculate the estimators come from Becker and Ichino (2002).

Nearest neighbour matching

The nearest neighbour is selected by choosing the untreated firm with the nearest propensity score. That is:

$$C(i) = \min_{j} \left\| p_i - p_j \right\|, \; i \neq j$$

Thus, nearest neighbour matching is likely to match one firm in receipt of assistance to the most similar firm that is not in receipt of assistance (i.e. $C(i)$ is a singleton set). In practice there may be more than one firm that satisfies this condition. It may be the case that there are two firms the same distance away, but one with a value of p_j that is higher and one with a value that is lower than p_i (what Becker and Ichino, 2002, call *forward* and *backward* matches). In these cases, we use two methods to calculate the average treatment effect, ATT. The first is to randomly select one of these (possibly groups of) observations and the other to give them equal weight. The ATT is computed by taking the average of the treatment effects calculated by comparing the outcomes of the assisted firms with their untreated nearest neighbours from the control group. In most samples, these are likely to yield similar results. However, since we have a huge pool of potential controls, the number of firms with the same propensity score will increase. In general, the likelihood of ties increases with the number of controls and decreases with the number of variables (particularly continuous ones). These two methods do not differentiate between firms that are the same distance in the same direction from the treated firm (i.e. those for whom p_i-p_j=p_i-p_k $i \neq j \neq k$). Since the random method is essentially arbitrary, our *a priori* preference is for the weighted method.

Stratified matching

Stratified matching uses the groups (strata) used to test the balancing restriction when calculating the propensity score. Because the covariates within each stratum, or block (b) are balanced, receipt of government assistance can be considered random within each stratum. Thus, we can compute the within-stratum average treatment effect as the average of the outcome for the treated minus the average treatment for the untreated for each stratum b:

$$\alpha_b^S = \frac{\sum_{i \in I(b)} Y_j^T}{N_b^T} - \frac{\sum_{i \in I(b)} Y_j^C}{N_b^C}$$

where $I(b)$ is the set of firms in block b. In order to calculate the overall ATT for all B blocks, we weight each block by the fraction of treated units it represents:

$$\alpha^S = \sum_{b=1}^{B} \alpha_b^S \frac{\sum_{i \in I(b)} D_i}{\sum_{\forall i} D_i}$$

If we assume independence of outcomes across units, we can obtain standard errors analytically using the following formula for the variance:

$$\mathrm{Var}(\alpha^S) = \frac{1}{N^T}\left[\mathrm{Var}(Y_i^T) + \sum_{b=1}^{B} \frac{N_b^T}{N^T}\frac{N_b^T}{N_b^C}\mathrm{Var}(Y_j^C)\right]$$

Note that if there is exactly one treated and/or one control in one or more of the blocks, the ATT in that block can be computed, but the standard error cannot. However, we do also calculate standard errors using bootstrapping techniques.

Which matching method is best?

When considering these different variations on the propensity score matching method, the question naturally arises: which is the best method? Asymptotically, all of them should provide the same results. As the sample size increases, they all become closer to the case where we compare only exact matches (Smith 2000). Ultimately, the choice one makes depends on the trade off between bias and variance. Some of the estimators, or methods of implementation (e.g. nearest neighbour with or without replacement) are better in small samples. In slightly larger datasets (i.e. those with a larger number of potential matches) some form of over-sampling, such as using the kernel estimator, might be useful[33].

With a larger sample, it seems a waste to throw away all the potential extra information to be provided by our large group of untreated firms. The question that arises is: how useful is this extra information? In New Zealand, the distribution of firms is significantly different to other advanced economies in the sense that there are a number of relatively large firms and many small (by international standards) firms but very few medium-sized firms. Since there appears to be something that prevents small firms becoming medium sized, it must be questioned as to how much information the large firms provide about small firms, who are in the majority. These issues are complex and are the subject of policy debate. Nevertheless, they do raise

[33] In an earlier exploratory piece using the LBD, Morris and Stevens (2007), we considered a larger number of estimators, including the kernel and radius estimators, but concluded that they were inappropriate for such a dataset as ours.

questions about the appropriateness of using the whole of the population of untreated firms, even when common support has been imposed.

One method that does help with this is the stratified matching method. The strata are chosen to ensure that both the propensity score and the characteristics of the firms (the X variables) of the treated and untreated firms therein are not different. Similarity is in this case is tested statistically. Because the two sets of firms (treated and untreated) cannot be distinguished from each other, assignment to treatment is essentially random. In such cases, the ATT is simply the difference in outcomes of the treated and untreated firms. Such balanced samples are often difficult to achieve when there is a large set of potential controls, because of the increase in statistical power of our tests[34]. Thus, if one can achieve a set of balanced strata in a large dataset such as ours, the stratified matching method becomes very attractive.

Summary of approach for matching models

We present results for 3 different types of models to estimate ATT. We use Becker and Ichino's (2002) routines to estimate the ATT for each technique. These are run after *pscore.*

1. Nearest neighbour matching I
2. Nearest neighbour matching II
3. Stratified matching

1) Models 1 (and 2) uses nearest neighbour methods. The techniques match each firm in receipt of GSR assistance to a single (or occasionally multiple nearest neighbours in the event of a tie) unassisted firm. The match is based on propensity values. The treatment effect is then the difference between the outcome of the firm receiving GSR assistance and its single control firm. In the event of a tie, the type 1 method randomly selects just one of these observations. We use the *Stata* routine *attnr* to estimate ATT, restricting the computation to the region of common support and calculating bootstrapped standard errors.

2) Model 2 is the same as 1 except that in the case of a tie it uses all observations with equal weight. We use the *Stata* routine *attnw* to estimate ATT, restricting the computation to the region of common support.

3) Model 3 uses stratified matching based on the groups (strata) used to test the balancing hypothesis. If the covariates within each group are balanced, the assisted

[34] See discussion in fn 31

and unassisted firms cannot be distinguished from each other and receipt of assistance can be considered to be random within each stratum. So we can compute a treatment effect for each stratum by comparing the average outcome of assisted firms with unassisted firms within each stratum. The overall ATT is calculated using a weighted average of all the individual stratum estimates, with the weighting based on the number of observations in each group. We use the *Stata* routine *atts* to estimate ATT, restricting the computation to the region of common support.

We present standard errors with all of our estimates. In many cases, these can be computed analytically. However if there is exactly one treated and/or one control in one or more of the strata, the standard error cannot be computed. In these cases, we calculate standard errors using bootstrapping technique.

Appendix 2: Description of variables

Table 9 Outcome variables

Outcome Variable	*Description*	*Source*
ln sales	log of sales	The source data on sales and purchases come from the Business Activity Indicator database, which is based on GST data from the Inland Revenue Department. All businesses which conduct taxable activity are required to register for GST if their annual turnover was greater than $40,000 (previously $30,000). This constitutes total sales and income S_I for the period (including GST). It also includes zero rated supplies Z. This is adjusted using data on zero-rated sales as follows $Sales = \frac{8}{9}(S_I - Z) + Z$. GST data should also include government grants and subsidies except where these are intended for overseas use for international development. Ideally, we would exclude government grants from the sales variable used to estimate a treatment effect. We examined alternate sources of the sales variable (from IR10 and Annual Enterprise Survey data) to see whether we could find a measure that excluded government grants. Both these alternate sources have categories where government revenue should be reported separately. However, it was not clear that the grant information was recorded separately so we retain GST data as our measure because of the greater coverage of New Zealand businesses. Ideally, we would subtract the known government grant funding from the GST sales measures. However, we do not have accurate payment date information yet so we are unable to do this. We plan to include this additional information in 2009.
ln va	Log of value added	Value added is the difference between sales and purchases. The purchases data in the BAI also come from the GST data. They relate to ‘Total purchases and expenses (including GST) for which tax invoicing requirements have been met’ and include an estimate for imported goods and the use of private goods and services in taxable activity. Our purchases data are adjusted to exclude GST.
ln prod	Log of labour productivity	Value-added divided by rolling mean employment

Table 10 Control variables

Firm characteristics		
Variable	***Description***	***Source***
ln rme *(ln rme2)*	log of employment (log squared employment)	Employment data are from aggregated Pay-As-You-Earn (PAYE) data from Inland Revenue Department (IRD). Employment is measured using an average of twelve monthly PAYE employee counts in the year. This is known as Rolling Mean Employment (RME). It includes an annual count of working proprietors.
ln capital *(ln capital2)*	Log of depreciation (log of squared depreciation)	Capital services are approximated by depreciation. We use depreciation data from the Annual Enterprise Survey imputed IR10 data.
age	Years since first recorded activity	We use the birth date of a firm from the LBF to calculate an initial estimate of age of firm. Because some firms have sales or other information records prior to the official birth date, we check the BAI and government scheme databases to see whether we have any information in years prior to the official birth date in LBF and replace the birth date with the earliest record year if necessary.
in group	Dummy variable for group status	This variable is one if the firm is independent, i.e. is a single enterprise and not in a parent/subsidiary relationship with other firms between 2000 & 2006. This is the period covered by LBF which holds the ownership information. It is zero if first is involved in a group for year within the period covered by LBF.
manufacturing *services*	Dummy variables for manufacturing or services industries.	Firms are assigned to three broad industry groups: Service industries, that is ANZSIC codes J (communications services), K (finance and insurance), L (property and business services), N (education), O (health and community services) or P (cultural and recreational services), and two manufacturing groups (ANZSIC divisions C20-C49 and C50-C99). Our initial analysis showed no significant differences between the 2 separate manufacturing dummies so we merged them for this paper
export ind *group_export_ind*	Dummy variable for whether a firm exports or is in a group that exports	The export indicator is one if the value of exporting revenue in any year is greater than zero. It is zero otherwise. Similarly, the group export indicator is one when any other firm in the group has exporting revenue greater than one. The exporting revenue is based on Customs data. It is collected from forms completed by firms whenever they import or export any physical goods.
foreign	A dummy variable for foreign-owned firms	The foreign indicator is one if a firm is owned or controlled by a non-resident. This variable is based on an IR4 response.

Government assistance variables[35]		
Variable	***Description***	***Source***
gsf *grant gsf* *slopegsf*	A variable to indicate whether firm received treatmen (GSF + advisory services)	*gsf* equals one if the firm if a firm has received a GSF grant and advisory services (received treatment) and 0 otherwise. *gsf* is interacted with year dummies to give time dependent estimates of α_0. *slopegsf* is the number of years since first treatment and is 0 if firm is not treated. Both variables are 0 before treatment. *Grant gsf* is the cumulative grant dollars received by the firm prior to that point.
cms *slopecms*	A variable to indicate whether firm received treatment (advisory services)	*cms* equals one if the firm if a firm has received a advisory services and 0 otherwise. *cms* is interacted with year dummies to give time dependent estimates of α_0 *slopecms* is the number of years since first treatment and is 0 if firm is not treated. Both variables are 0 before treatment.
other *group other* *grant other*	A dummy variable for whether a firm received any grant from NZTE other than the scheme being evaluated	*Other gsp* equals one if the firm has ever accessed any other grants from NZTE or 0 otherwise. Similarly, *group other gsp* is one if any other member of the group has accessed another grant from NZTE. Both variables are zero before receipt of other grants and one afterwards. *Grant other* is the cumulative grant dollars received by the firm up to that point. We include the following: In-scope grant programmes include: Australia New Zealand Bio-technology Partnership Fund, Business Growth Fund, Cluster Development Programme, Better by Design, Enterprise Awards Scheme, Enterprise Development, Enterprise Development Grants - Capability Building component, Enterprise Development Grants- Market Development component, Enterprise Network Funding, GIF Sector Projects, Growth Services Funds, Major Investment Fund, Regional Partnerships Programme and Strategic Investment Fund.
non nzte *group non nzte* *grant non*	A dummy variable for whether a firm accessed any non-NZTE grant	*Non nzte* equals one if the firm has accessed any grants from FRST or TPK in a given year. Similarly, *group non nzte* equals one if another group member has accessed FRST/TPK grants. Both variables are zero before receipt of grants and one afterwards. FRST schemes included in the GAP database are: Grants for Private Sector Research and Development, Technology for Business Growth programmes, Technology for Industry Fellowships, SmartStart and other TechLink schemes. The TPK administered schemes that have been matched are the Māori Business Facilitation Service and Māori Tourism Facilitation Service.

[35] These variables only appear as lagged variable in our models: lags of one year are denoted by L_ before the variable name; lags of 2 years are L2_ before the variable names

Government assistance variables[35]		
Variable	***Description***	***Source***
inv job *group inv job*	A dummy variable for whether a firm accessed NZTE's Market Development Services	*Inv job* equals one if the firm has accessed any of NZTE's Market Development Services (MKDS) in a given year. Similarly, *group inv job* equals one if another group member has accessed MKDS. Both variables are zero before any services are accessed and one afterwards. MKDS relate to services provided for firms under some kind of full or partial cost recovery.
Gap history	A dummy variable for whether a firm has received any other government assistance that is not the treatment scheme	*Gap history* equals one if the firm has accessed any grants or services from FRST or TPK or any grants or services from NZTE excluding the treatment scheme, from 2000 to the year in question. The variable is zero before receipt of any grants and one afterwards.

Appendix 3: Tables of Results

A3.1 Impact of combined GSF using panel methods

The following tables present panel model results for assessing the impact on firms of receiving a GSF grant and associated NZTE intensive client management services. Results are presented for Models 1, 5, 6 and 7 – these are described in the Detailed Methods section. The following apply to these tables:

- Year dummies and constants have been omitted from the tables for clarity.
- All cumulative dollars (in the intensity models) have been divided by 1000.
- The L and L2 operators preceding variable names refer to lags of one or two years of that variable.
- Standard errors are shown in parentheses.
- * significant at 10%;** significant at 5%; *** significant at 1%

Table 11 Impact on Sales for firms receiving combined GSF/services (dummy variable)

Specification #	**(1)**	**(5)**	**(6)**	**(7)**
Lln_sales	0.206***	0.506***	0.645***	0.233***
	(0.004)	(0.014)	(0.060)	(0.023)
L2ln_sales		0.002	0.009	
		(0.005)	(0.009)	
Lgsf2003	**0.052**	**0.149****	**0.190****	**0.136****
	(0.055)	(0.068)	(0.081)	(0.067)
Lgsf2004	**0.084**	**0.125****	**0.154*****	**0.155*****
	(0.060)	(0.052)	(0.050)	(0.045)
Lgsf2005	**0.068**	**0.140***	**-0.015**	**0.044**
	(0.065)	(0.075)	(0.070)	(0.051)
Lgsf2006	**-0.215****	**-0.177**	**-0.101**	**-0.020**
	(0.102)	(0.129)	(0.119)	(0.086)
slopegsf	**0.024**	**-0.008**	**0.090**	**0.095****
	(0.020)	(0.020)	(0.060)	(0.048)
Lothergsp	0.043**	0.050***	0.050**	0.058***
	(0.018)	(0.019)	(0.024)	(0.021)
Linvjob	0.039**	0.021	0.016	0.077***
	(0.018)	(0.019)	(0.026)	(0.020)
Lnon	0.039**	0.052**	0.064**	0.094***
	(0.018)	(0.023)	(0.029)	(0.020)
ln_rme	0.332***	0.194***	0.077	0.149***
	(0.003)	(0.015)	(0.049)	(0.023)
ln_rme2	0.080***	-0.013*	-0.002	
	(0.001)	(0.007)	(0.015)	
Lcapital	-0.038***	-0.022***	-0.946***	-0.162***
	(0.002)	(0.006)	(0.342)	(0.033)
export_ind	0.043***	-0.005	-0.753***	-0.577***
	(0.005)	(0.011)	(0.198)	(0.190)
Observations	1053690	699210	754080	1053690
Number of clusters=	331175	245830	268435	331175
Anderson statistic (identification/IV relevance test)	78567	7099		
Hansen J statistic		392		
Arrelano-Bond test for AR(2),z=			4.159	2.127
Prob > z			0.000	0.033
Sargan test of overid. restrictions			84.606	70.758
Chi2:			29.000	18.000
Hansen test of overid. restrictions			103.917	81.478
Chi2:			29.000	18.000

Table 12 Impact on Sales for firms receiving combined GSF/services (intensity model)

Specification #	(1)	(5)	(6)	(7)
Lln_sales	0.20551***	0.50686***	0.64400***	0.23288***
	(0.00409)	(0.01376)	(0.06013)	(0.02317)
L2ln_sales		0.00129	0.00829	
		(0.00510)	(0.00907)	
Lgrant_gsf	**0.00047**	**0.00054**	**0.00030**	**0.00110****
	(0.00040)	(0.00042)	(0.00061)	(0.00046)
Lgrant_other	-0.00002	-0.00001	0.00000	0.00001
	(0.00009)	(0.00010)	(0.00012)	(0.00011)
Lgrant_non	0.00029**	0.00030**	0.00032**	0.00047***
	(0.00012)	(0.00015)	(0.00015)	(0.00015)
L2grant_gsf	**0.00065**	**-0.00019**	**-0.00013**	**0.00064**
	(0.00047)	(0.00048)	(0.00067)	(0.00057)
L2grant_other	0.00056	-0.00004	0.00008	0.00077
	(0.00055)	(0.00005)	(0.00018)	(0.00058)
L2grant_non	-0.00020	-0.00015	-0.00035**	-0.00025*
	(0.00012)	(0.00012)	(0.00016)	(0.00013)
ln_rme	0.33199***	0.19466***	0.08344*	0.13540***
	(0.00326)	(0.01540)	(0.04943)	(0.02392)
ln_rme2	0.08037***	-0.01423*	-0.00986	
	(0.00147)	(0.00735)	(0.01580)	
Lcapital	-0.0379***	-0.0223***	-0.8655**	-0.1616***
	(0.00179)	(0.00615)	(0.34122)	(0.03274)
export_ind	0.0430***	-0.0081	-0.9080***	-0.7782***
	(0.00501)	(0.01142)	(0.21833)	(0.21006)
Observations	1053690	699210	754080	1053690
Number of clusters=	331175	245830	268435	331175
Anderson statistic (identification/IV relevance test)	78092.264	7047.386		
P-val (null:eqn is underidentified)	0.000	0.000		
Hansen J statistic		397.147		
p-val		0.000		
Arrelano-Bond test for AR(2),z=			4.147	2.117
Prob > z			0.000	0.034
Sargan test of overid. restrictions			82.355	71.178
Chi2:			29.000	18.000
Hansen test of overid. restrictions			101.889	79.540
Chi2:			29.000	18.000

Table 13 Impact on Value-added for firms receiving combined GSF/services (dummy variable)

Specification #	(1)	(5)	(6)	(7)
Lln_va	0.104***	0.373***	0.43511***	0.29162***
	(0.003)	(0.015)	(0.07932)	(0.03058)
L2ln_va		0.003	0.01793	
		(0.005)	(0.01615)	
Lgsf2003	**0.074**	**0.085**	**0.04050**	**0.00399**
	(0.103)	(0.123)	(0.12441)	(0.11304)
Lgsf2004	**-0.023**	**0.062**	**0.09015**	**0.08312**
	(0.080)	(0.087)	(0.07320)	(0.06556)
Lgsf2005	**-0.072**	**-0.064**	**0.00267**	**0.04422**
	(0.105)	(0.134)	(0.08375)	(0.07355)
Lgsf2006	**-0.054**	**-0.087**	**-0.03837**	**-0.00900**
	(0.098)	(0.116)	(0.10291)	(0.09364)
slopegsf	**0.018**	**0.017**	**0.05004**	**0.04261**
	(0.025)	(0.028)	(0.06177)	(0.05899)
Lothergsp	0.063**	0.079**	0.07851**	0.07993**
	(0.027)	(0.033)	(0.03816)	(0.03503)
Linvjob	0.020	0.039	0.07811**	0.08349***
	(0.022)	(0.029)	(0.03671)	(0.02812)
Lnon	0.038	0.049	0.11992***	0.14442***
	(0.026)	(0.035)	(0.04157)	(0.03295)
ln_rme	0.323***	0.169***	0.09650	-0.01049
	(0.005)	(0.025)	(0.06372)	(0.04158)
ln_rme2	0.080***	-0.001	-0.04830*	
	(0.002)	(0.013)	(0.02557)	
Lcapital	0.029***	0.042***	-0.41057	-0.3053***
	(0.002)	(0.007)	(0.46682)	(0.04236)
export_ind	0.011	-0.045**	-1.1315***	-1.1581***
	(0.008)	(0.019)	(0.33124)	(0.32457)
Observations	781300	499765	531385	781300
Number of clusters=	267095	187820	201620	267095
Anderson statistic (identification/IV relevance test)	93714	3798		
P-val (null:eqn is underidentified)	0.000	0.000		
Hansen J statistic		137		
p-val (null: instruments valid)		0.000		
Arrelano-Bond test for AR(2),z=			8.158	6.900
Prob > z			0.000	0.000
Sargan test of overid. restrictions			68.546	41.340
Chi2:			22.000	18.000
Hansen test of overid. restrictions			68.073	45.970
Chi2:			22.000	18.000

Table 14 Impact on Value-added for firms receiving combined GSF/services (intensity model)

Specification #	**(1)**	**(5)**	**(6)**	**(7)**
Lln_va	0.10382***	0.37385***	0.44163***	0.29001***
	(0.00334)	(0.01535)	(0.08157)	(0.03070)
L2ln_va		0.00227	0.01700	
		(0.00474)	(0.01665)	
Lgrant_gsf	**0.00021**	**0.00057**	**0.00088**	**0.00094***
	(0.00047)	(0.00052)	(0.00064)	(0.00057)
Lgrant_other	0.00010	0.00008	0.00018	0.00015
	(0.00007)	(0.00013)	(0.00017)	(0.00013)
Lgrant_non	0.00010	0.00012	0.00028	0.00039*
	(0.00015)	(0.00017)	(0.00024)	(0.00023)
L2grant_gsf	**0.00036**	**0.00007**	**-0.00004**	**0.00046**
	(0.00056)	(0.00064)	(0.00080)	(0.00079)
L2grant_other	-0.00011	-0.00008	0.00018	0.00021
	(0.00009)	(0.00010)	(0.00030)	(0.00031)
L2grant_non	0.00010	0.00017	0.00001	0.00003
	(0.00011)	(0.00013)	(0.00017)	(0.00015)
ln_rme	0.32267***	0.17029***	0.10575	-0.02789
	(0.00481)	(0.02457)	(0.06524)	(0.04319)
ln_rme2	0.07999***	-0.00315	-0.06069**	
	(0.00198)	(0.01260)	(0.02685)	
Lcapital	0.02906***	0.04157***	-0.34219	-0.3024***
	(0.00184)	(0.00721)	(0.47719)	(0.04252)
export_ind	0.01153	-0.04763**	-1.3401***	-1.3752***
	(0.00791)	(0.01921)	(0.37174)	(0.36049)
Observations	781300	499765	531385	781300
Number of clusters=	267095	187820	201620	267095
Anderson statistic (identification/IV relevance test)	93105.388	3774.705		
P-val (null:eqn is underidentified)	0.000	0.000		
Hansen J statistic		142.466		
p-val		0.000		
Arrelano-Bond test for AR(2),z=			2.371	6.789
Prob > z			0.018	0.000
Sargan test of overid. restrictions			54.217	45.202
Chi2:			29.000	18.000
Hansen test of overid. restrictions			61.154	46.809
Chi2:			29.000	18.000

Table 15 Impact on Productivity for firms receiving combined GSF/services (dummy variable)

Specification #	**(1)**	**(5)**	**(6)**	**(7)**
Lln_prod	0.131***	-0.194***	0.48574***	0.47763***
	(0.003)	(0.010)	(0.09921)	(0.05735)
L2ln_prod		0.078***	0.04355**	
		(0.004)	(0.02160)	
Lgsf2003	**0.075**	**0.184***	**0.05850**	**0.01052**
	(0.104)	(0.109)	(0.13866)	(0.12097)
Lgsf2004	**0.022**	**0.133****	**0.16458****	**0.10863**
	(0.070)	(0.066)	(0.07893)	(0.06927)
Lgsf2005	**-0.058**	**-0.021**	**0.06651**	**0.05197**
	(0.107)	(0.116)	(0.08729)	(0.07781)
Lgsf2006	**-0.050**	**-0.039**	**-0.00970**	**0.00479**
	(0.097)	(0.096)	(0.11221)	(0.09926)
slopegsf	**0.021**	**0.029**	**0.07905**	**0.05017**
	(0.025)	(0.026)	(0.07387)	(0.06147)
Lothergsp	0.066**	0.064**	0.06965*	0.09180**
	(0.028)	(0.028)	(0.04134)	(0.03731)
Linvjob	0.024	0.053**	0.09027**	0.09373***
	(0.022)	(0.023)	(0.03832)	(0.03075)
Lnon	0.043	0.067**	0.14774***	0.16634***
	(0.026)	(0.028)	(0.04389)	(0.03648)
ln_rme	-0.689***	-0.779***	-0.4287***	-0.7723***
	(0.005)	(0.020)	(0.07695)	(0.04004)
ln_rme2	0.083***	-0.076***	-0.1188***	
	(0.002)	(0.012)	(0.02904)	
Lcapital	0.034***	0.059***	0.04454	-0.3431***
	(0.002)	(0.006)	(0.50954)	(0.05161)
export_ind	0.014*	-0.074***	-1.1224***	-1.0563***
	(0.008)	(0.016)	(0.35249)	(0.33158)
Observations	764810	499765	512575	764810
Number of clusters=	260675	187820	193300	260675
Anderson statistic (identification/IV relevance test)	168085.557	3396.644		
P-val (null:eqn is underidentified)	0.000	0.000		
Hansen J statistic		702.358		
p-val (null: instruments valid)		0.000		
Arrelano-Bond test for AR(2),z=			1.696	7.185
Prob > z			0.090	0.000
Sargan test of overid. restrictions			75.527	43.541
Chi2:			29.000	18.000
Hansen test of overid. restrictions			75.514	51.385
Chi2:			29.000	18.000

Table 16 Impact on Productivity for firms receiving combined GSF/services (intensity model)

Specification #	(1)	(5)	(6)	(7)
Lln_prod	0.13080***	-0.1476***	0.49357***	0.4750***
	(0.00318)	(0.01086)	(0.10219)	(0.05764)
L2ln_prod		0.07802***	0.04262*	
		(0.00352)	(0.02217)	
Lgrant_gsf	**0.00043**	**0.00108****	**0.00129***	**0.00108***
	(0.00046)	(0.00045)	(0.00068)	(0.00062)
Lgrant_other	0.00012*	0.00009	0.00022	0.00020*
	(0.00007)	(0.00009)	(0.00020)	(0.00011)
Lgrant_non	0.00012	0.00003	0.00032	0.00045*
	(0.00015)	(0.00015)	(0.00027)	(0.00025)
L2grant_gsf	**0.00036**	**0.00031**	**0.00030**	**0.00061**
	(0.00056)	(0.00051)	(0.00092)	(0.00084)
L2grant_other	-0.00013	-0.00003	0.00015	0.00012
	(0.00009)	(0.00010)	(0.00033)	(0.00031)
L2grant_non	0.00011	0.00018	0.00014	0.00011
	(0.00011)	(0.00012)	(0.00017)	(0.00015)
ln_rme	-0.68937***	-0.7939***	-0.4160***	- 0.7947***
	(0.00515)	(0.02224)	(0.07915)	(0.04158)
ln_rme2	0.08339***	0.00475	-0.1304***	
	(0.00209)	(0.01343)	(0.03059)	
Lcapital	0.03438***	0.07983***	0.08329	- 0.3401***
	(0.00185)	(0.00642)	(0.52253)	(0.05183)
export_ind	0.01398*	-0.0441***	-1.3255***	-1.300***
	(0.00805)	(0.01585)	(0.39340)	(0.36933)
Observations	764810	499765	512575	764810
Number of clusters=	260675	187820	193300	260675
Anderson statistic (identification/IV relevance test)	167984.764	3387.845		
P-vall (null:eqn is underidentified)	0.000	0.000		
Hansen J statistic		712.210		
p-val		0.000		
Arrelano-Bond test for AR(2),z=			1.695	7.088
Prob > z			0.090	0.000
Sargan test of overid. restrictions			77.933	48.036
Chi2:			29.000	18.000
Hansen test of overid. restrictions			75.567	53.240
Chi2:			29.000	18.000

A3.2 Impact of Client Management Services using panel methods

The following tables present panel model results for assessing the impact on firms of receiving NZTE intensive client management services (CMS only). Results are presented for Models 1 to 7 – these are described in the methodology section. The following apply to these tables:

- Year dummies and constant terms are omitted from the tables for clarity.
- The L and L2 operators preceding variable names refer to lags of one or two years of that variable.
- Standard errors are shown in parentheses.
- * significant at 10%;** significant at 5%; *** significant at 1%

Table 17 Impact on Sales for firms receiving CMS only (dummy variable)

Specification #	(1)	(5)	(6)	(7)
Lln_sales	0.207***	0.505***	0.64208***	0.23320***
	(0.004)	(0.014)	(0.05947)	(0.02314)
L2ln_sales		0.002	0.00910	
		(0.005)	(0.00903)	
Lcms2003	**0.192**	**-0.025**	**0.20724**	**0.35636***
	(0.210)	(0.098)	(0.18028)	(0.18663)
Lcms2004	**0.143****	**0.145***	**0.05949**	**0.14726****
	(0.070)	(0.079)	(0.07630)	(0.06460)
Lcms2005	**0.559****	**0.553****	**0.16962***	**0.21961****
	(0.250)	(0.259)	(0.09252)	(0.08869)
Lcms2006	**0.077**	**0.037**	**0.01252**	**0.10719**
	(0.099)	(0.099)	(0.09122)	(0.08160)
slopecms	**-0.024**	**-0.033**	**-0.00906**	**0.04718**
	(0.042)	(0.054)	(0.07377)	(0.06156)
Lothergsp	0.042**	0.054***	0.05266**	0.05376**
	(0.019)	(0.021)	(0.02516)	(0.02192)
Linvjob	0.040**	0.020	0.00739	0.07400***
	(0.019)	(0.020)	(0.02677)	(0.02053)
Lnon	0.047**	0.057**	0.07137**	0.09647***
	(0.019)	(0.023)	(0.02942)	(0.02134)
ln_rme	0.332***	0.193***	0.07986	0.15260***
	(0.003)	(0.015)	(0.04898)	(0.02306)
ln_rme2	0.080***	-0.011	-0.00186	
	(0.001)	(0.007)	(0.01531)	
Lcapital	-0.038***	-0.021***	-0.986***	-0.163***
	(0.002)	(0.006)	(0.34685)	(0.03272)
export_ind	0.042***	-0.005	-0.7469***	-0.5247***
	(0.005)	(0.011)	(0.19437)	(0.19024)
Observations	1052075	698015	752830	1052075
Number of clusters=	330780	245475	268065	330780
Anderson statistic (identification/IV relevance test)	78582.029	7219.286		
P-val (null:eqn is underidentified)	0.000	0.000		
Hansen J statistic		390.207		
p-val (null: instruments valid)		0.000		
Arrelano-Bond test for AR(2),z=			4.094	2.089
Prob > z			0.000	0.037
Sargan test of overid. restrictions			85.639	70.184
Chi2:			29.000	18.000
Hansen test of overid. restrictions			104.128	80.872
Chi2:			29.000	18.000

Table 18 Impact on Value-added for firms receiving CMS only (dummy variable)

Specification #	**(1)**	**(5)**	**(6)**	**(7)**
Lln_va	0.104***	0.373***	0.41263***	0.29494***
	(0.003)	(0.015)	(0.07722)	(0.03071)
L2ln_va		0.003	0.02240	
		(0.005)	(0.01598)	
Lcms2003	**-0.059**	**-0.134**	**-0.12083**	**0.00063**
	(0.143)	(0.141)	(0.15386)	(0.16643)
Lcms2004	**0.044**	**0.081**	**0.04162**	**0.09058**
	(0.128)	(0.143)	(0.10116)	(0.10171)
Lcms2005	**0.623**	**0.671**	**0.24174***	**0.26720****
	(0.387)	(0.436)	(0.14086)	(0.13529)
Lcms2006	**0.087**	**0.130**	**0.03481**	**0.06003**
	(0.149)	(0.132)	(0.13495)	(0.13948)
slopecms	**-0.040**	**-0.088**	**-0.02557**	**0.02430**
	(0.060)	(0.063)	(0.10296)	(0.09938)
Lothergsp	0.057*	0.061*	0.04967	0.06722*
	(0.029)	(0.034)	(0.03987)	(0.03662)
Linvjob	0.026	0.042	0.07781**	0.08821***
	(0.023)	(0.030)	(0.03808)	(0.02930)
Lnon	0.044	0.048	0.11371***	0.14141***
	(0.027)	(0.036)	(0.04128)	(0.03407)
ln_rme	0.323***	0.169***	0.10418*	-0.00411
	(0.005)	(0.024)	(0.06282)	(0.04119)
ln_rme2	0.080***	-0.001	-0.04645*	
	(0.002)	(0.012)	(0.02469)	
Lcapital	0.029***	0.042***	-0.487	-0.310***
	(0.002)	(0.007)	(0.46407)	(0.04257)
export_ind	0.012	-0.042**	-1.0911***	-1.0404***
	(0.008)	(0.019)	(0.31626)	(0.32189)
Observations	780005	498820	530425	780005
Number of clusters=	266755	187525	201315	266755
Anderson statistic (identification/IV relevance test)	93827.980	3877.203		
P-val (null:eqn is underidentified)	0.000	0.000		
Hansen J statistic		134.651		
p-val (null: instruments valid)		0.000		
Arrelano-Bond test for AR(2),z=			2.147	6.976
Prob > z			0.032	0.000
Sargan test of overid. restrictions			55.296	41.524
Chi2:			29.000	18.000
Hansen test of overid. restrictions			66.296	48.160
Chi2:			29.000	18.000

Table 19 Impact on Productivity for firms receiving CMS only (dummy variable)

Specification #	**(1)**	**(5)**	**(6)**	**(7)**
Lln_prod	0.131***	-0.149***	0.45959***	0.47899***
	(0.003)	(0.011)	(0.09492)	(0.05731)
L2ln_prod		0.078***	0.04845**	
		(0.004)	(0.02110)	
Lcms2003	**-0.020**	**0.046**	**-0.18584**	**0.07377**
	(0.172)	(0.116)	(0.13866)	(0.26460)
Lcms2004	**0.053**	**0.051**	**0.07533**	**0.12840**
	(0.133)	(0.121)	(0.11635)	(0.12362)
Lcms2005	**0.622**	**0.710****	**0.25453***	**0.25236***
	(0.390)	(0.351)	(0.15201)	(0.14590)
Lcms2006	**0.085**	**0.201**	**0.02457**	**0.02090**
	(0.150)	(0.127)	(0.14970)	(0.14580)
slopecms	**-0.040**	**-0.053**	**0.02424**	**0.04802**
	(0.059)	(0.064)	(0.11383)	(0.10899)
Lothergsp	0.059**	0.045	0.03584	0.08163**
	(0.029)	(0.029)	(0.04288)	(0.03903)
Linvjob	0.029	0.052**	0.09304**	0.09745***
	(0.024)	(0.024)	(0.03930)	(0.03204)
Lnon	0.047*	0.044	0.13816***	0.16213***
	(0.028)	(0.029)	(0.04306)	(0.03756)
ln_rme	-0.689***	-0.791***	-0.4317***	-0.7652***
	(0.005)	(0.022)	(0.07571)	(0.03985)
ln_rme2	0.084***	-0.000	-0.1166***	
	(0.002)	(0.013)	(0.02785)	
Lcapital	0.034***	0.079***	-0.03051	-0.3445***
	(0.002)	(0.006)	(0.49786)	(0.05163)
export_ind	0.014*	-0.043***	-1.0450***	-0.9532***
	(0.008)	(0.016)	(0.33265)	(0.33020)
Observations	763520	498820	511625	763520
Number of clusters=	260335	187525	193000	260335
Anderson statistic (identification/IV relevance test)	167814.288	3467.050		
P-val (null:eqn is underidentified)	0.000	0.000		
Hansen J statistic		694.754		
p-val (null: instruments valid)		0.000		
Arrelano-Bond test for AR(2),z=			1.496	7.210
Prob > z			0.135	0.000
Sargan test of overid. restrictions			77.923	43.513
Chi2:			29.000	18.000
Hansen test of overid. restrictions			79.333	52.776
Chi2:			29.000	18.000

A3.3 Impact of combined GSF using matching method

This section presents matching model results for assessing the impact on firms of receiving a GSF grant and associated NZTE intensive client management services. We considered two versions of each treatment: one using levels and the other using first differences. The matching is performed separately for three different industry groupings: manufacturing; services and other industries. We only consider the impact of receiving combined GSF and CMS services because we were unable to obtain robust estimates for firms that only received CMS services using this method.

Table 20 Estimation of propensity scores

	Manufacturing *All grants* ***D***	**Manufacturing** *New Grants* **Δ *D***	**Services** *All grants* ***D***	**Services** *New Grants* **Δ *D***
ln_sales	0.126***	0.137***	0.103***	0.111***
	(0.013)	(0.013)	(0.015)	(0.016)
gap_history	1.240***	1.121***	1.517***	1.400***
	(0.056)	(0.059)	(0.064)	(0.070)
Yr2002	-0.684***	-0.661***	-0.948***	-0.842***
	(0.145)	(0.141)	(0.303)	(0.298)
Yr2004	0.164***			
	(0.059)			
Yr2005	0.099*	-0.034	0.097	0.097
	(0.060)	(0.061)	(0.062)	(0.068)
Observations	106045	113360	487010	438350
Pseudo R^2	0.32	0.29	0.32	0.29

- *Dependent variable is the probability of receiving a GSF grant (D) in Columns (1,3) or new GSF grants (ΔD) in columns (2,4).*
- *Sales and government assistance history variables are calculated in the year of receipt of grant*
- *Standard errors in parentheses*
- ** significant at 10%; ** significant at 5%; *** significant at 1%*
- *Observations have been rounded to nearest 5*
- *Constant terms were included in both regressions*

Table 21 Final blocks for propensity score: Manufacturing

Block number	Mean of Propensity Score	*Number N*		
		Controls		
1	9.12E-05	58,010	5	58,015
2	0.000781	15,390	15	15,400
3	0.001563	4,735	30	4,765
4	0.003125	955	5	960
5	0.00625	480	0	485
6	0.0125	915	15	935
7	0.025	1,260	40	1,295
8	0.0375	1,040	70	1,110
9	0.05	1,380	105	1,485
Block Total		84,165	285	84,450

- *Observations have been rounded to nearest 5*
- *Estimated on region of common support*

Table 22 Final blocks for propensity score: Services

Block number	Mean of Propensity Score	*Number N*		
		Controls		
1	4.87E-05	290,365	15	290,380
2	0.000195	36,985	25	37,010
3	0.000391	5,955	10	5,960
4	0.000781	975	0	975
5	0.001563	195	0	195
6	0.003125	85	0	90
7	0.00625	425	0	425
8	0.0125	1,775	45	1,820
9	0.025	1,235	45	1,280
10	0.05	25	0	25
Block Total		338,020	140	338,160

- *Observations have been rounded to nearest 5*
- *Estimated on region of common support*

The following tables present the results of the matching models using the blocks and propensity scores described above and in the main text. For all tables:

- Observations have been rounded to the nearest 5
- Standard errors are shown in parentheses.
- * significant at 10%;** significant at 5%; *** significant at 1%

Table 23 Impact on sales for firms receiving GSF - Manufacturing

Dependent variable = ln sales ; Treatment = GSF	*(1) Stratification*	*(2) Nearest neighbour (nw)*	*(3) Nearest neighbour (nd)*
N Controls	82,915	280	280
N Treated	290	290	290
ATT	**0.140****	0.104	0.104
Bootstrapped Std. Err.	0.062	0.089	0.087
Bootstrapped *t*-statistic	2.253	1.159	1.189

Dependent variable = Δ(ln sales); **Treatment =ΔGSF**	*(1) Stratification*	*(2) Nearest neighbour (nw)*	*(3) Nearest neighbour (nd)*
N Controls	74,515	225	225
N Treated	225	225	225
ATT	**0.148*****	**0.160*****	**0.160*****
Bootstrapped Std. Err.	0.022	0.058	0.058
Bootstrapped *t*-statistic	6.641	2.747	2.737

Table 24 Impact on sales for firms receiving GSF – Services

Dependent variable = ln sales ; **Treatment = GSF**	*(1) Stratification*	*(2) Nearest neighbour (nw)*	*(3) Nearest neighbour (nd)*
N Controls	331,235	135	135
N Treated	135	135	135
ATT	**0.180****	0.050	0.033
Bootstrapped Std. Err.	0.082	0.138	0.154
Bootstrapped *t*-statistic	2.199	0.364	0.217

Dependent variable = Δ(ln sales); **Treatment =ΔGSF**	*(1) Stratification*	*(2) Nearest neighbour (nw)*	*(3) Nearest neighbour (nd)*
N Controls	303,065	110	110
N Treated	110	110	110
ATT	**0.113***	-0.005	-0.001
Bootstrapped Std. Err.	0.069	0.132	0.096
Bootstrapped *t*-statistic	1.648	-0.041	-0.012

Table 25 Impact on value-added for firms receiving GSF -Manufacturing

Dependent variable = ln value-added; **Treatment = GSF**	*(1) Stratification*	*(2) Nearest neighbour (nw)*	*(3) Nearest neighbour (nd)*
N Controls	76,790	255	255
N Treated	265	265	265
ATT	**0.165*****	0.114	0.114
Bootstrapped Std. Err.	0.059	0.089	0.103
Bootstrapped t-statistic	2.781	1.282	1.099

Dependent variable = Δ(ln value-added); Treatment =ΔGSF	*(1) Stratification*	*(2) Nearest neighbour (nw)*	*(3) Nearest neighbour (nd)*
N Controls	67410	200	200
N Treated	200	200	200
ATT	**0.091***	**0.113***	0.113
Bootstrapped Std. Err.	0.053	0.064	0.072
Bootstrapped *t*-statistic	1.731	1.774	1.565

Table 26 Impact on value-added for firms receiving GSF –Services

Dependent variable = ln value-added; Treatment = GSF	*(1) Stratification*	*(2) Nearest neighbour (nw)*	*(3) Nearest neighbour (nd)*
N Controls	120,690	115	115
N Treated	115	115	115
ATT	0.165	0.155	0.155
Bootstrapped Std. Err.	0.124	0.152	0.166
Bootstrapped t-statistic	1.332	1.021	0.935

Dependent variable = Δ(ln value-added); Treatment =ΔGSF	*(1) Stratification*	*(2) Nearest neighbour (nw)*	*(3) Nearest neighbour (nd)*
N Controls	114,765	85	85
N Treated	85	85	85
ATT	0.030	-0.059	-0.059
Analytical Std. Err.	0.093	0.194	0.118
t-statistic	0.321	-0.306	-0.501

Table 27 Impact on productivity for firms receiving GSF -Manufacturing

Dependent variable = ln productivity; Treatment = GSF	*(1) Stratification*	*(2) Nearest neighbour (nw)*	*(3) Nearest neighbour (nd)*
N Controls	72013	255	255
N Treated	265.	265	265
ATT	-0.046	-0.082	-0.082
Bootstrapped Std. Err.	0.033	0.093	0.070
Bootstrapped t-statistic	-1.386	-0.887	-1.177

Dependent variable = Δ(ln productivity); **Treatment =ΔGSF**	*(1) Stratification*	*(2) Nearest neighbour (nw)*	*(3) Nearest neighbour (nd)*
N Controls	63570	200	200
N Treated	200	200	200
ATT	-0.028	0.025	0.025
Bootstrapped Std. Err.	0.054	0.061	0.062
Bootstrapped *t*-statistic	-0.519	0.416	0.409

Table 28 Impact on productivity for firms receiving GSF –Services

Dependent variable = ln productivity; **Treatment = GSF**	*(1) Stratification*	*(2) Nearest neighbour (nw)*	*(3) Nearest neighbour (nd)*
N Controls	112510	110	110
N Treated	115	115	115
ATT	-0.032	-0.039	-0.039
Bootstrapped Std. Err.	0.078	0.149	0.107
Bootstrapped t-statistic	-0.405	-0.260	-0.363

Dependent variable = Δ(ln productivity); **Treatment =ΔGSF**	*(1) Stratification*	*(2) Nearest neighbour (nw)*	*(3) Nearest neighbour (nd)*
N Controls	107025	80	80
N Treated	80	80	80
ATT	-0.017	-0.090	-0.090
Analytical Std. Err.	0.108	0.152	0.187
t-statistic	-0.162	-0.590	-0.481

Appendix 4: Discussion of Results

A4.1 Panel methods

We present the results for this set of techniques in Table 11 to Table 16. The treatment in these cases is receipt of a GSF grant (or predecessor) and the associated advisory services. The even-numbered tables (Table 11, Table 13, Table 15) refer to what we have called the ‘dummy variable’ model, where we use a binary treatment indicator for the receipt of assistance. In order to asses how the impact of receiving GSF and attendant CMS might vary from year to year, we have included a treatment indicator for each year separately. In order to examine the possibility of an additional effect, beyond the new trajectory signified by the dummies, we include a slope dummy (*slopegsf*). When interpreting the results shown in these tables, bear in mind there are these two components to the treatment effect: a component associated with the mean impact of treatment and a component that depends on time since treatment. To illustrate: a significant coefficient of 0.10 for *Lgsf2003* implies that the mean impact is 10% for firms first receiving a grant prior to 2003 (i.e. in 2002). If the *slopegsf* was also significant and equal to 0.01, then the total impact would increase by an additional 1% every year after treatment. This would mean that those firms that received a GSF grant in 2002 had on average 11% higher levels of outcomes in 2003 than similar firms that did not receive assistance. In 2004, the average outcome would be 12% higher than similar untreated firms. If we have successfully removed all causes of bias, this increase would be solely due to the fact the firm received GSR assistance. Note that we only have data on four years of treatment, so this effect is estimated on a maximum of four years. This is not to be taken as suggesting that this additional 1% growth per year is permanent effect, causing firms to grow at that this additional rate indefinitely.

The odd-numbered tables (Table 12, Table 14 and Table 16) refer to the ‘intensity model’, where we examine how the impact of assistance varies with its dollar value. In order to account for more complex dynamics in the impact of assistance on outcomes using this model, we include both a first and second lag of the intensity of treatment variable (*Lgrant_gsf* and *L2grant_gsf*, respectively).

The estimated impact of treatment on outcomes is over and above the level of that outcome that is explained by firm size, exporting assistance, other types of government assistance and other control characteristics. The contributions due to these other factors are measured by each of the coefficients associated with the control variables. As we are interested in the impact following receipt of assistance, all the government assistance variables and the treatment effect are lagged by at least one year. All treatment-related coefficients have been highlighted in bold for clarity.

Each table presents results from a range of model specifications (see Appendix 1). The reason we present a range of results is to show the sensitivity of ATT estimates to various techniques. Part of this variation is to be expected, as we move from our model where only lagged dependent variable is instrumented (model (1)) to more appropriate specification that accounts for more general endogeneity. As we shall see the estimated impact of GSR assistance does vary according to which method is used and we discuss possible causes as we present the results. We aim to present a feasible range of impact estimates from our preferred model specifications (these are shaded grey in the tables). We omit the results from Specifications 2 to 4 from Table 12 for brevity and focus on estimates from Specifications 5 to 7.

Impact on sales for firms receiving combined GSF/services

Dummy treatment variable

Table 11 presents results using four different specifications to estimate the impact of assistance on the sales of firms. Recall that the specifications differ primarily in the way in which endogenous variables are instrumented. Column (1) presents results for the simplest specification, which only instruments the lagged dependent variable (*Lln_sales*). It is instructive to consider this case to illustrate the issue of endogeneity, even though it is not our preferred model for reasons discussed below.

The treatment effects presented in column (1) are mainly small and positive, but not significant. The exception is firms that received GSF grants in 2005 (*Lgsf2006*) - these appear to have significantly *lower* sales than comparable firms following treatment. All of the other firm characteristics are significant. Current levels of sales are strongly influenced by levels of sales in the preceding year, suggesting that shocks to sales are persistent. Sales are higher in firms with more employees and lower levels of capital, *ceteris paribus*. Exporting is also associated with higher sales. There is a discernable impact from other types of government assistance with this model. For example, the coefficient relating to whether a firm accessed MkDS services (*L_inv job*) is 0.039. This implies that a firm accessing MkDS services has about 4% higher sales the following year, over and above the additional impact due to the other factors. Firms that had received other grants from NZTE (*Lothergsp*) or grants from other agencies (*Lnon*) also had slightly higher sales than those that received no assistance. However, as we discussed above we have reason to suspect that a number of these variables are endogenous.

In columns (5) to (7) of Table 11, we examine the sensitivity of our estimates to a broader instrument set, as outlined in Table 8 in the previous appendix. Our sensitivity analysis confirms that we did need to instrument the suspect control variables (those relating to RME, capital and exporting). We obtained the most stable estimates when we included a wider set of variables and deeper lags in the instrument set. We tested to see whether our instruments were valid. All our instrument sets passed the relevance test confirming that the instruments we used were correlated with the endogenous regressors. However, none of our instrument sets passed the overidentifying restrictions tests implying that our instruments were either not truly exogenous or they had been incorrectly excluded from the regression. (We think the former is likely because inclusion of suspect instruments, e.g. *L2ln_sales*, did not help matters much). If we had a longer time series we could try even deeper lags in the GMM style instruments but we are limited by our relatively short dataset. We sought the best possible instrument set which showed the least evidence of serial correlation in the residuals (small values of z in the Arrelano-Bond test).

The results suggest that if we account for slow adjustment in sales, firm size, capital, exporting behaviour and previous government assistance then we can see a statistically significant impact for treated firms in some years. The mean impact clearly depends on the year of treatment. This can be seen in both our preferred specifications. The impact ranges from about 13-15% for firms treated between 2002 and 2004, to negative and insignificant for firms treated in 2005. These estimates suggest that either: (a) there is something different about the firms that were treated in the earlier years that allowed them to achieve more benefit from the assistance; or (b) there is something different in the way in which firms are selected and treated. This may reflect changes in selection processes over time. It may be that prior to 2005, most GSF recipients were firms with an established history with NZTE and its predecessor organisations (this is consistent with the 2005 evaluation findings that most firms classified as high growth potential in 2005 were legacy clients from predecessor organisations). It is possible that these firms were primed for growth. From 2005 onwards, NZTE may have needed to extend its net wider to select new clients (e.g. through referrals from the Enterprise Hotline) who may not have been at the stage where they could immediately benefit from assistance.

This leads us to a discussion of the time-varying component of the treatment effect. Is there any evidence that the impact increases or decreases in time? Unfortunately, the slope coefficient is not as stable. In one of our preferred specifications (5) it is negative and insignificant (suggesting no effect, or that it may decline over time); in another (7) it is positive and significant (suggesting that some of the benefits take a while to appear). Because of this variability (and results presented in the next section), we conclude that we are not able to estimate the time varying component of the treatment effect at this stage.

The impact on sales due to receiving other types of grant assistance from NZTE is also positive and significant (about 5%) and for firms accessing MkDS services (also 5%). The impact on sales due to assistance from other government agencies (primarily FRST) also has a significant positive impact on sales (7%). This shows the importance of having information on all types of government assistance in order to correctly estimate the treatment effect for any particular scheme, to reduce the risk of benefits due to one scheme being falsely attributed to another.

In summary, our results show that firms receiving GSF grants and associated services between 2002 and 2005 have on average around 9%[36] higher sales than similar unassisted firms, after controlling for the effects of previous sales, employment, capital, exporting behaviour and previous government assistance. The impact is strongest for firms treated before 2004 and is insignificant (and possibly negative) for firms treated in 2005.

These estimates probably err on the generous side – as some selection bias may remain. This is because the firms that are classified as high growth potential firms and provided with GSR assistance by NZTE are also the firms most likely to respond to that assistance. If we have not controlled for all the characteristics that determine selection into treatment then we will have bias. Of greatest concern would be those characteristics that do not change over time (or change slowly). However, the approach we use is able to reduce this source of selection bias relatively well, by first-differencing to remove fixed effects as well as by instrumenting the endogenous variables.

Intensity model

The first set of results assumed a relatively simple model for measuring the impact of treatment. More realistically, the impact of assistance might depend on its intensity. Table 12 shows the results for the intensity model, which is based on the cumulative grant amount received by the firm prior to that year. We also use a cumulative dollar model for the other government assistance control variables.

[36] We reran the regressions with a mean α_0 coefficient to verify this estimate

The results using this model are less conclusive. The treatment coefficient is positive for both preferred specifications, however it is only significant for one (specification (7)). We find no evidence of more complex dynamics. The impact is statistically insignificant for longer time lags (i.e., *L2_grant_gsf* coefficient is insignificant). The linear time-varying component in this model was insignificant so we do not show those results. We also did not include MkDS services in the intensity models because there was no accurate data to reflect the intensity of services provided by NZTE staff.

Our preferred specifications give treatment coefficients *L_grant_gsf* between 0.0005 and 0.0011. In order to assess the treatment effect in terms of dollar impact on sales, we need to make some assumptions about the typical size of a change in treatment, in this case, the average size of a GSF grant. To see this consider the following regression equation: $\ln Y = \alpha D$ where α is the regression coefficient. Differentiating using the chain rule gives $(1/Y_0)\Delta Y = \alpha \Delta D$. In the case of the dummy variable model $\Delta D = 1$, so that α gives the percentage increase in *Y* due to treatment *D* ($= 100 \times \Delta Y / Y_0$). In the intensity model, the change in treatment from one year to next, ΔD, is the average value of a GSF grant paid between 2002 and 2005 ($51,300). The treatment coefficient for our preferred model is 0.00082[37]/1000 (recall that the cumulative dollars were divided by 1000). Thus the percentage increase in sales (the treatment effect on sales) $= 100 \times 8.2 \times 10^{-5} \times \$51{,}300 = 4\%$. This impact estimate is about half that estimated using the dummy variable model.

Interestingly, other types of business assistance to firms from NZTE (*Lgrant_other*) do not appear to have a significant impact on sales when the intensity of assistance is taken into account. This may be because the likelihood of receiving other NZTE assistance is positively related to the size of their GSF grant in dollar terms (our variable measures the intensity of treatment *within* a programme, and this is correlated with intensity of all NZTE treatment across all schemes). However, the impact due to assistance from other agencies does have a significant impact, although the long-run impact is smaller. The assistance from other agencies is primarily targeted at technology development and we would therefore expect a longer term impact. Since this is the not the primary focus of the evaluation, we only include terms for the impact two years post treatment but acknowledge that this is not the most realistic model for assistance of that type.

[37] We have averaged the coefficients from our preferred specifications 5 and 7.

Impact on value-added for firms receiving combined GSF/services

The results for the impact of receiving GSF assistance on value-added (sales less purchases) are less positive (see Table 13). Similar to the impact on sales, the treatment effect is positive for firms treated in earlier years and negative for firms treated in 2005[38]. This trend is seen in all our preferred specifications. However, none of the coefficients are significant. If we pool these dummies and estimate the average impact for firms treated between 2002 and 2005, it is 4%, but this estimate is also statistically insignificant. The time-varying coefficient is insignificant for all specifications. Why might sales increase, but value added not? To understand this, it is useful to consider what value added is: it is sales revenue minus intermediate consumption (in our case, purchases). Sales revenue may have increased, but the cost of the intermediate goods and services required to produce these extra sales is greater. Thus, one potential explanation for this lack of significant impact on value-added is that expanding firms may be constrained by suppliers and so increasing demand for inputs (e.g. purchases) are met with rising prices.

Looking at the other variables, we see evidence of persistence in shocks to value-added. It is unclear what the relationship is between value-added and the number of employees or level of capital, since the coefficients switch signs for different specifications. Exporting activity is linked to reduced value-added in the same year. Surprisingly the impact due to all other types of government assistance is mainly significant and quite high, ranging from about 8% for other NZTE grants, 6% for MkDS services and 11% for assistance from other government agencies. However, all specifications suffer from serial correlation of the residuals which reduces our confidence in these results. This autocorrelation may be due to additional dynamics in the determination of valueate that we have not modelled.

Intensity model

Turning now to the intensity model, we find that one of our preferred specifications, column (7), does show a significant effect of treatment on value-added (Table 14). This is similar to the intensity model results for sales. The point estimates and standard errors across the preferred specifications are similar. The average impact on value-added due to receiving GSF assistance $= 100 \times 8 \times 10^{-5} \times \$53,100 = 4\%$. This result is the same as that estimated using the dummy variable model. However, the statistical significance of

[38] Recall, that *Lgsf2006* refers to firms treated prior to 2006, i.e., in 2005.

the results is low and is not significant at all in columns (5) and (6). There does not appear to be any change in impact over time.

The high values of impact for other types of government assistance for this model have disappeared or been dampened down. In contrast to results from the dummy treatment model, the impact due to other types of NZTE grants is not significant and the impact due to assistance from other agencies is only significant for one of our preferred model specifications. One thing is clear and that is that past shocks to value-added influence the current year's values. All preferred specifications still suffer from serial correlation in the residuals.

Impact on productivity for firms receiving combined GSF/services

Interestingly, when we move to labour productivity (or value-added per employee) we find stronger evidence for a significant effect due to assistance. The dummy treatment results (Table 15) show the same pattern of more positive impacts for firms treated in earlier years and negative impact for firms treated in the last year. The average impact for all treated firms is surprisingly high at 9%, although the estimates are only significant for one of our preferred specifications[39].

The results from the intensity model are quite consistent across specifications and all are statistically significant (Table 16). The average impact on productivity is 6% based on intensity model results, or about half that estimated by the dummy treatment model. Similarly, assistance from other government agencies, other types of NZTE grants and MkDS services all appear to have a positive impact on productivity.

There is less evidence of persistence in labour productivity after removing fixed effects; in particular, the sign of the lagged dependent variable is not constant across specifications. Productivity diminishes with increasing employees (suggesting either declining returns to scale and/or constraints on physical capital) and exporting activity and lower levels of capital.

Impact due to CMS services alone

All our previous results have shown the impact of a combination of assistance from NZTE – the receipt of a GSF grant and associated advisory services. In these cases we are unable to differentiate between the impact due to the grant and that due to services provided by NZTE. However, there are also firms that have received CMS services but

[39] We prefer Specification 6 over Specification 7 because the latter shows evidence of serial correlation in the residuals.

have never received a GSF grant. These estimates of the impact of CMS services are shown in Table 17 to Table 19. We do not have accurate data on the intensity of CMS services, so we only show results using a dummy treatment variable.

The estimates for impact on sales due to these services are quite variable between preferred specifications and in some cases, unrealistically high. For example, the treatment effect ranges from insignificant and negative to over 35% for firms receiving services prior to 2003. For firms starting treatment a year later in 2003, the effect is more stable and about 14%. This is followed by a jump in impact to over 55% for firms starting treatment in 2004. While the pattern is roughly similar to that found when considering the impact of combined grant and CMS, we do not consider these estimates to be robust. One of our concerns is the lack of high quality information about exactly when a firm began to receive intensive client management. We approximated this date from other information in the historical records. Fortunately, NZTE began collecting more detailed information about recipients from 2007 onwards, so evaluations in the future will be able to draw on a higher quality dataset.

The results for value-added and productivity are similarly variable. All but a few estimates of the treatment effect are insignificant. The exception is the impact for firms that began receiving CMS in 2004; however some of the estimates are unrealistically high. For example, one specification estimates over 70% impact on productivity, although the impact is lower (25%) for the model with no serial correlation in the residuals.

Summary of panel results

In summary, once we control for firm size, capital, exporting behaviour, previous government assistance from other sources and pre-assistance levels of these outcomes, we find a statistically significant positive effect of GSR assistance on sales (and to a lesser extent value-added and productivity). The results show that firms receiving GSF grant and NZTE advisory services have on average about 4-9% higher level of sales, value-added and productivity than unassisted firms the year following transitions in receiving assistance. The estimates are not particularly sensitive to the model assumed for receiving GSR assistance – results for the dummy treatment model and an intensity model are similar although when intensity is taken into account the impact is smaller. The effect of GSR assistance on value-added is the least conclusive – we find some evidence of impact using the intensity model but results from the preferred dummy treatment model are not significant.

An interesting finding is that the impact appears to be higher for firms receiving assistance in the earlier years (prior to 2005). This may be linked to changes in selection processes over time whereby initial selection for assistance favoured firms with established relationships with NZTE and predecessor organisations. It may be that these firms were at the stage where they could more readily benefit from assistance and we have not picked this up with our control variables or by removing the fixed effects. In other words, there is still some selection bias in our estimates because treated firms are meant to be indistinguishable from untreated firms conditional on the controlling characteristics. Another explanation is that the set of firms that can benefit was small and the 'lowest hanging fruit' were picked in the early years.

We have not been able to separately determine the impact of receiving a GSF grant from CMS advisory services. Results for the impact relating to CMS were too variable and insignificant to reach firm conclusions. This could be due to the smaller numbers in this group and the poor quality of information relating to receipt of services.

We do find evidence of significant effect for other types of government assistance using dummy treatment models, primarily those from other agencies (FRST and TPK) and to a lesser extent MkDS services from NZTE.

We have seen that it was necessary to use instruments to reduce the bias caused by troublesome regressors related to employment, exporting and capital where we suspected a two way relationship between the regressor and the outcome variable. It is also clear that there is a high degree of persistence in all our outcome variables; firms with high levels of sales or values added in preceding years were likely to be followed by high levels of sales in the next year. For this reason, we think the first order dynamic model is required and that standard approach with no dynamics, e.g. OLS or fixed effects, will give seriously biased estimates for the treatment effect. Our best results (those with the least evidence of serial correlation in the error terms) were obtained using the difference GMM models.

A4.2 Matching estimators

The results presented in the previous section suggest that there may be some relation between the receiving GSR assistance and certain firm characteristics; and we have attempted to overcome this endogeneity through the use of instrumental variables. Matching estimators deal with this by estimating the probability of treatment conditional on these characteristics. We turn now to the results of using the propensity score matching methodology.

Propensity score estimators

The results of estimating the propensity scores for receiving treatment (i.e. receiving a GSF grant and associated services) are presented in Table 20. We estimated the probit model results by separate industry groups: manufacturing and services. We also estimated the likelihood separately for all grants (Columns 1 and 3) and new grants only (Columns 2 and 4).

As can be seen from the table, the models include measures of firm sales and previous government assistance as well as year dummies. The likelihood of receiving a GSF grant increases with higher sales in the year of receipt and if a firm had received any other form of government assistance since 2000 (including assistance from NZTE, FRST and TPK). Note that the pseudo-R^2 [40] levels are quite high for such models in all cases, indicating a good level of fit.

This may seem a sparse set of characteristics to explain whether a firm receives GSR assistance. We tested several other variables to see whether they could improve the models. The other variables included firm size, age, exporting behaviour, group structure, growth in size and sales, and government assistance by agency and type. Some of these other variables were significant for predicting treatment. However, they failed the balancing test in one or more blocks and so could not included. (Recall, the balancing test compares the means of firm characteristics in the control and treated populations. If the test fails, then the two groups are different and so can't be used to calculate the ATT). In fact, balancing the strata turned out to be a very challenging task. We mainly used an ad hoc method to explore various combinations of variables and only included those that were significant in the probit model estimation. Only our success came when we used parsimonious specifications such as those shown here. We estimated the propensity score separately for manufacturing and services, effectively allowing the coefficients to sector. We do not consider these specifications to be ideal and future work will explore further sectoring (e.g. with firm size) to attempt to find specifications that include more dynamics and satisfy the balancing hypothesis.

The distribution of assisted and unassisted firms across the blocks after testing the balancing hypothesis has been met is set out in Table 21 and Table 22 for manufacturing and services firms, respectively. The tables show the advantage of starting from such a large number of potential controls. In many cases, treated individuals/firms in blocks at the upper end of the propensity score distribution are more

[40] Note that unlike in a linear regression, the pseudo-R^2 of a probit cannot be interpreted strictly as the proportion of the variance in the dependent variable explained by the model.

numerous within their block than potential controls. In such cases, if one wishes to use nearest neighbour matching, a tough decision needs to be made about whether to match with or without replacement. If one chooses to match with replacement, single control firms may be matched to many treated firms. Without replacement, treated firms will have to be matched to firms outside their balanced stratum and thus are likely to be rather dissimilar, and hence inappropriate, controls or dropped from the evaluation altogether.

We tried the same approach with firms that only received CMS services and had not received a GSF grant. However, we were unable to find any specification that passed the balancing test. The following discussion refers to firms receiving the combination of GSF grants and CMS advisory services.

Average treatment effects

The estimates of impact based on the propensity score matching are shown in Table 23 to Table 26. Each table sets out the calculated average treatment effect on the treated (ATT) along with the number of treated firms and the number of controls used to calculate this. Each table also presents the bootstrapped standard errors and corresponding *t*-values. Our bootstrapped standard errors are based on twenty replications.

Each set of tables shows results for two approaches: matching on levels (all treatments) and matching on differences (new treatments only). Results are split by industry group.

Sales

We begin with the impact on sales for firms in the manufacturing industry (see Table 23). The stratified and nearest neighbour estimators enable us to identify 290 instances of receiving GSF grants and 225 instances of new grants over the time period. We match to 280 controls with the nearest neighbour methods and over 80,000 firms using the stratified method. The stratified method provides us with two statistically significant estimates at around 15%; one using the levels approach (all treatments) and the other using first differences (new treatments only). It is encouraging that these ATT estimates are similar to each other. The ATTs for the two different nearest neighbour methods are also similar to the stratified matching estimates, although they are only significant when estimating the impact using differences.

These results suggest that firms have on average 15% higher sales one year after receiving assistance. It is important to note that this is not because these firms already had high turnover as the pre-treatment level of firm sales is included in the set of control characteristics. This result says that given comparable levels of sales, broad industry group and similar prior government assistance, firms that receive GSR assistance will have higher sales the following year.

Results for the services industry are similar. There are many more controls (over 300,000) but few treated firms at around 135 for services group. Clearly, manufacturing firms are more likely to have received GSF grants than those in the services or other industries. The impact of receiving GSF on sales is similar for firms in the service industry (11-18%).

Value-added

Because our data requirements are higher than for sales (i.e. we lose data for firms with zero or missing purchases because the log is undefined) the number of treated firms we can consider is 265 for manufacturing and 115 for services. Focusing on the manufacturing firms, it appears that the ATT is sensitive to which approach is used; whilst an ATT value of 17% is obtained using all treatments, the value is less using the first difference approach (9%) and only marginally significant. The impact on value-added for services is statistically insignificant using both approaches.

Productivity

The impact on productivity (Table 27 and Table 28) was statistically insignificant for all industry groups and both types of matching methods (levels or first differences). This appears to be inconsistent to the panel model results which showed impacts similar to those found for sales and value-added. It is unclear why this is the case