Appendix A

Survey and modelling methodology

A.1 The mail survey

The major source of data for the study was a survey sent to approximately 5000 firms in five industries, namely:

- Clothing and footwear;
- Engineering;
- Information technology and telecommunications;
- Scientific and medical equipment; and
- Processed foods and beverages.

The sample was structured to be broadly representative of Australian industry – with a mix of new and traditional industries, large and small firms, varying importance of technology, exports and so on.

The other main consideration in constructing the sample was relevance to the Department of Industry, Science and Technology. We consulted with the various industry divisions throughout the Department to identify those areas where it was felt that the research project could most usefully make a contribution to understanding business operations in particular industries.

The sample was drawn by the Australian Bureau of Statistics (ABS) and included all management units employing between 5 and 500 persons in 18 ANZIC codes[1]. The ABS was also responsible for the mail out, collection and data entry, as well as providing advice on the structure of the survey form.

The structure of the sample and the response rate by industry is shown in Table A.1. The overall response rate of just over 30 per cent is considered by both the ABS and the BIE as being satisfactory for a non-compulsory survey[2].

[1] The ANZIC codes making up the sample were 2241, 2242, 2249, 2250, (Clothing and footwear); 2769, 2851, 2864, 2869 (Engineering); 2841, 2842, 2849 (IT&T); 2543, 2832, 2839 (Sci/med); 2111, 2130, 2172, 2183 (Food).

[2] The response rate would have been higher but for the fact that less than one in four Clothing and footwear firms returned the form.

Table A.1 Sample structure and response rate

Industry	*Sample size*[a]	*Returns*	*Response rate (%)*
Clothing and footwear	1250	291	23.3
Engineering	1694	578	34.1
IT&T	343	131	38.2
Scientific and medical equipment	330	122	37.0
Processed foods and beverages	477	156	32.7
Total	*4094*	*1286*[b]	*31.4*

Notes: (a) Sample size is equal to the number of forms dispatched less number of firm "deaths" identified by returns.
(b) Industries could not be assigned for 8 respondents.
Source: BIE survey.

A draft of the survey was pilot tested on 10 firms to identify any areas of confusion or weakness in the survey design. This resulted in some small modifications. The final survey form is reproduced at the end of this appendix.

A.2 Response bias check

Even with a relatively good response rate, the possibility of bias in the survey results is very real. Around 60 per cent of those firms who were sent a survey form did not complete and return it. Many of these firms may well have no cooperative arrangements at all and therefore did not see the survey as relevant to them. Those that did return the survey may have been those that had arrangements or were particularly interested in them. If this were the case, the results of the survey would be biased, overestimating the proportion of firms who have cooperative arrangements.

To check for this type of bias, we contacted 100 non-respondents from the original sample and asked them a subset of the questions from the main survey. The break up of the 100 was roughly the same as the structure of the original sample.

The main finding from the exercise was that only 22 per cent of non-respondents indicated they had one or more cooperative arrangements. This conflicted with the findings of the mail survey, which suggested that 40 per cent of firms were involved in cooperation, indicating that bias was a problem[3].

It is possible to correct for the response bias however. For example, we can estimate the proportion of firms involved in cooperation across Australia using the data presented above. We know that 41 per cent of respondents cooperate (from the survey returns) and that 22 per cent of non-respondents also cooperate. Across the whole sample, therefore, we estimate that 28 per cent of firms are involved in business cooperation[4].

[3] The main problem is in examining issues around cooperating versus non-cooperating firms (Chapter 4). When comparisons are made between cooperating firms (as in Chapter 5 onwards), this form of bias is not of concern.

[4] $\frac{0.41 \times 1286 + 0.22 \times 2808}{4094}$ = 28 per cent

It would be possible to adjust all the results in this report (for example cooperation by industry, state, product type) provided the degree of bias is known. We chose not to adjust throughout the report because the scale of the non-response survey needed to adjust all the factors considered in the report is prohibitively large. Rather than impose an across-the-board adjustment of 32 per cent[5], regardless of whether it actually describes the bias for a particular group of firms, we consider it better to use the unadjusted data.

A.3 Testing the interpretation of 'cooperative business arrangements'

As noted in Chapter 2, there is a plethora of terms used to describe business cooperation. What someone might call a joint venture another might call a consortium and yet another might call a strategic alliance. Some may not view these as cooperative at all if they have a contractual basis, even though cooperation is very much present and vital for success.

Moreover, the distinction between cooperative and arm's length relationships becomes very blurred at the latter end of the spectrum. A simple arm's length transaction might be the purchase of an 'off the shelf' product from a supplier. If, for example, the two exchange information about future requirements to assist one or both in planning production runs on an informal basis, the relationship starts to take on cooperative aspects.

It is not clear from the survey results where firms have generally 'drawn the line' between arm's length and cooperative activities. In addition, the survey form contains some specific examples of the types of cooperative activities with which we believed firms might commonly be involved. The risk of providing examples, however, was that respondents may have interpreted them as the only forms of cooperation in which we were interested.

To resolve some of these issues, we conducted a follow-up telephone survey of 100 respondents[6] who indicated in the mail survey that they had no cooperative arrangements. These firms were asked about their relationships with their customers and their suppliers to identify areas where cooperation may take place. In addition they were asked if they had been involved in a range of activities with cooperative aspects[7].

Interestingly, when the definition of cooperation was explicitly broadened to cover those areas that may not have been considered in responses to the mail survey, 50 of these firms indicated that they in fact currently had one or more cooperative arrangements. This result suggests that the findings of the mail survey understate the proportion of firms involved in cooperation in Australia.

It is important to note, however, that the additional arrangements identified in the telephone follow-up survey were typically, though not solely, towards the arm's length end of the cooperative continuum (see Chapter 2). Assuming that these 100 firms are representative, around half of all those who indicated in the mail survey they were not involved in cooperation, are in fact involved in more 'marginal' forms of cooperation.

---

5 Going from 41 per cent to 28 per cent is a fall of 32 per cent.

6 The structure of the sample was 59 small firms, 27 medium firms and 14 large firms.

7 Including: joint ventures or networks; joint production; equipment sharing; joint implementation of quality, best practice, training or benchmarking programs; research and development partnerships; and joint bidding for consortia.

We can now estimate the proportion of firms who are involved in both 'core' and 'marginal' cooperation[8]. Accordingly if a broad definition is adopted, up to two-thirds of firms are currently involved in cooperation. It would seem then that around one-third have no cooperative arrangements at all.

The proportion of firms involved in substantial or 'core' cooperative arrangements is probably somewhere around one-third. That is, the 28 per cent of firms identified in Section A.2 and the additional core forms of cooperation arising from the telephone follow up survey. In sum:

- one-third of firms are not involved in cooperation
- around one-third of firms have 'core' cooperative arrangements
- up to two-thirds of firms have 'core' and/or 'marginal' cooperative arrangements.

A.4 Significance tests

Throughout the report we compare various groups to see if there are any differences between them. For example we examine the propensity to establish cooperative arrangements by firm size in Chapter 4 and find that large firms are the most likely to cooperate, with 63 per cent having at least one arrangement, compared to the average figure of 41 per cent.

But, given that these are survey results, the question arises: to what extent does this finding apply to large firms more generally? After all, this finding concerning large firms is based on only 76 firms. How confident can we be about a result based on these 76 responses? If another survey was taken of a different group of firms would it too find that large firms are more likely than average to cooperate?

Testing the statistical significance of results provides a basis for answering these types of questions. It provides a measure of how sure we can be that large firms are, in fact, more likely to cooperate than average.

Significance testing is particularly useful when important differences are identified but these are based on fairly small samples. In these cases the chances that the result simply reflects sampling variation is relatively high. The testing procedure used takes account of the size of the sample.

The test used is based on the hypothesis that two observed proportions are, in reality, the same. A test statistic is calculated based on the difference between the two proportions and their standard errors[9]. The larger the test statistic, the more confident we can be that this hypothesis is false – that the observed difference between large firms and the average is a 'real' difference.[10]

---

8 $\frac{0.28x4094 + 0.5(0.72x4094)}{4094}$ = 64 per cent

9 Specifically, the formula is : $Z = \frac{p_a - p_b}{\sqrt{\frac{p_a(100-p_a)}{n_a} + \frac{p_b(100-p_b)}{n_b}}}$

where p_i is the observed proportion for group *i* and n_i is the number of firms in group *i*.

10 The value for the test statistic (Z) is compared to the relevant figure from tables of values for the distribution of the standard normal curve to give the level of significance.

In our example, the test statistic is calculated to be 3.83. This is a very high result, and we can be very confident (99.8 per cent) that the observed difference for large firms is not simply the result of the sample. There is only a 0.02 per cent probability that our confidence is misplaced and the two proportions are actually the same – that large firms are as likely as average to cooperate.

Of course, given that the difference in this example was so large (63 per cent compared to 41 per cent) this is perhaps not a surprising result. What about the case where differences are not so large? For example, 36 per cent of micro firms were found to cooperate compared to the average of 41 per cent. In this case a test of significant differences is more useful. The test statistic is calculated to be 1.84, which gives us a confidence limit of 93.4 per cent. We can be quite confident that micro firms are less likely to cooperate (there is only a 6.6 per cent chance that micro firms are as likely to cooperate).

Finally, let us look at the case for medium firms. Chapter 4 shows that 43 per cent of medium firms have arrangements, which is slightly above average. But a test of significant differences tells us that we can only be 44 per cent (Z= 0.58) confident that medium sized firms are more likely to cooperate than average. We would therefore conclude that the difference is not statistically significant – that medium-sized firms are as likely as average to cooperate.

A.5 The probit model

A.5.1 Overview

The probit model is a non-linear binary choice model used in situations where a dependent variable has only two possible outcomes. For example, a firm is either cooperating or it is not, or it gains a particular type of benefit from its cooperation or it does not. It allows the determination of the probability of a firm with a given set of attributes (both in terms of its own characteristics and the characteristics of its arrangements) having one outcome, rather than the alternative. Hence the model can also show the strength of relationship between the explanatory variables and the dependent variable.

The probit model is not the only possible choice of model for these situations. The logit model is another. In practice, however, the logit and probit models tend to yield extremely similar results, the only real difference being that the logit model estimates all tend to be larger than the probit estimates (Davidson and Mackinnon 1993). The probit model was chosen because the smaller, more conservative, estimates produced would be less likely to include unimportant variables as being influential.

The model results are more robust than the partial analysis findings as they take into account the interaction between different variables (firm characteristics), whereas a partial analysis looks at only two variables in isolation from all others. Thus a partial analysis of individual survey findings will tend to give the same results as the model but will also additionally give some misleading results. Partial results provide a useful basic test for hypotheses but, where possible, they should be compared with the results from a more robust model.

The data collected concerned:

- whether a firm was in a cooperative business arrangement or not
- which industry the firm was in
- the state the firm was located in
- metropolitan or regional location

- level of employment*
- level of turnover*
- level of exports*
- age of firms*
- competitive advantages of firms
- constraints on firm performance
- ownership of firms
- level of processing of goods
- technology used
- competition faced
- benefits from cooperation
- performance/competitiveness changes due to cooperation.

The four marked '*' are continuous variables. The others are all represented by dummy variables where '1' indicates the presence of the characteristic and '0' means it is absent.

Initially the models were run incorporating all the variables listed above, as well as some which were generated from the primary data (turnover growth, export propensity, log of employment and log of turnover). This was reduced down to contain only significant explanatory variables. Simultaneously, another probit model was built up from *a priori* knowledge. The point at which these two models converged became the final model form.

It is necessary to exclude one of the dummy variables from each exclusive set to avoid singularity. The impact of any single variable can still be inferred.

Many different models were constructed using the principles described above. The dependent variables modelled were:

Chapter 4

- arrangement(s)/no arrangement

Chapter 5

- formal arrangement
- informal arrangement
- arrangement with a customer
- arrangement with a supplier
- arrangement with a non-customer/supplier firm
- arrangement involving one other firm
- arrangement involving two or more other firms
- arrangement involving an overseas-based firm

Chapter 8

- five or more major/critical benefits from all arrangements
- three or more major/critical supplier benefits from all arrangements

Chapter 9

- increase in exports from key arrangement
- increase in exports turnover from key arrangement
- increase in exports productivity from key arrangement
- increase in exports technological competitiveness from key arrangement

A.5.2 Goodness of fit and hypothesis testing

Methods of hypothesis testing for single parameters, joint hypotheses and hypotheses about the goodness of fit of the model as a whole can be used with probit models. These are parallel to the procedures which can be applied to OLS models.

In regression models, it is common to test the hypothesis that all slopes are zero by using an *F* test. For binary response models, the same hypothesis can easily be tested by using a likelihood ratio test. The method produces a statistic that follows, approximately, a chi-square distribution when the null hypothesis is true.

There is no close parallel with the measure of R^2 in OLS. The R^2 test measures the proportion of the variance that is 'explained' by the exogenous variables. There are some surrogate measures, but the nature of the estimation problem renders all such measures problematic. The McFadden R^2 is the simplest surrogate, being:

$$1 - (\text{unrestricted value/restricted value})$$

This is a plausible measure of goodness of fit because it lies between 1 and 0. However the number of surrogate R^2 measures and the lack of a dominant one makes it difficult to use any with great certainty, although it does serve the purpose of allowing comparison between models.

Finally, the percentage of right predictions can be used instead of the surrogate R^2 values to estimate the goodness of fit of the model. This gives the ratio of the number of firms a model predicts to have an attribute to the proportion of firms with an attribute.

A.6 Survey form