1351.0.55.020

Australian Bureau of Statistics

Research Paper

Explorations of Innovation and Business Performance Using Linked Firm-Level Data

New Issue

Research Paper

Explorations of Innovation and Business Performance Using Linked Firm-Level Data

Marn-Heong Wong

Productivity Commission

Darren Page, Ruel Abello and Kwok Pan Pang

Analytical Services Branch

AUSTRALIAN BUREAU OF STATISTICS

EMBARGO: 11.30 AM (CANBERRA TIME) FRI 7 SEP 2007

ABS Catalogue no. 1351.0.55.020
ISBN 978 0 64248 364 5

Produced by the Australian Bureau of Statistics

INQUIRIES

The ABS welcomes comments on the research presented in this paper.
For further information, please contact Mr Ruel Abello, Analytical Services Branch on Canberra (02) 6252 6307 or email <analytical.services@abs.gov.au>.

EXPLORATIONS OF INNOVATION AND BUSINESS PERFORMANCE USING LINKED FIRM-LEVEL DATA

Marn-Heong Wong
Productivity Commission

Darren Page, Ruel Abello and Kwok Pan Pang
Australian Bureau of Statistics

MAIN FINDINGS

The two studies described in this paper made use of firm-level data from the 2003 Innovation Survey, but differed with respect to the auxiliary datasets used to augment the number of observations and the available information on firms' business performance.

These studies are intended to be exploratory in nature and should be seen as part of the ongoing overall research into the determinants of innovation expenditure and innovation output, and of the impact of these on associated firm performance in Australia. The studies are work-in-progress and aimed at informing future data collection and research direction, rather than to generate definitive results. Issues related to data quality and econometric methodology, such as the limited scope for lag effects due to the short period covered in the data, may affect the robustness of the results. Thus these findings should be treated as indicative only.

Model and data

Bearing in mind the caveats, the main findings from the two studies include:

- The Crépon, Duguet and Mairesse (CDM) framework has been found to be useful in analysing, at the individual business level, the relationship between innovation inputs, outputs and business performance.
- The Innovation Survey data, when augmented with information from other datasets including taxation data, provide a useful platform for microdata analysis of innovation and business performance.
- The factors that influenced innovation input, innovation output and business performance are outlined below.

Innovation input

- Internal information impacted positively on both the decision to invest and the intensity of innovation. Smaller firms also based their decisions on external sources of ideas.
- Lack of access to information from competitors and the presence of intellectual property protection were associated with innovation intensity.
- Collaborations encouraged firms to invest in innovation but did not influence the level of investment in innovation.
- Government regulations and standards were either insignificant or weakly significant in dampening innovation effort.
- The drive to increase export opportunities, revenue and market shares motivated innovation activities.
- Firms with higher market share were more likely to innovate but had lower innovation intensity.
- Potential growth in market demand raised large firms' propensity to invest in innovation.
- Lack of skilled staff did not seem to adversely impact the decision to innovate.
- Financial resources such as government assistance impacted positively on smaller firms' innovation intensity.
- Larger firms were more likely to invest in innovation, but once having decided to invest, evidence of a systematic relationship between the level of investment and firm size was weak.
- Ownership structure and firm age did not seem to make a difference to firm's innovation investment decisions, but inter-industry differences in technological opportunities had an effect.

Innovation output

- Innovation intensity is significantly and positively related to all types of innovation output.
- Factors affecting the degree of success in product innovation and the probability of a positive outcome were broadly similar, but there were differences.
- Internal information, diversity in types of collaboration, strategic protection of IP and firm size were significant in raising the probability of introducing any innovation output type; information spillover from competitors was also important for small firms.

- Some factors did not affect all innovation output types uniformly; external information sources were important for process and organisational innovations.
- Intramural R&D and the employment of workers with technical skills were important to product and process innovations, and to organisational innovation by smaller firms.
- Ownership structure was insignificant; the impact of firm age was mixed; inter-industry differences mattered for product and process innovations but not organisational innovation.

Business performance

- The studies found some preliminary evidence of a significant and positive relationship between innovations and business performance, although the results were far from conclusive.
- The strength of association between the different types of innovation output and business performance depended on the dataset, industry coverage and also the performance measures used.
- The studies found evidence of a positive relationship between innovation output and labour productivity growth, although statistically weak and only with respect to product and process innovations. Organisational innovation output was found to be consistently insignificant.
- For product innovation, this relationship with labour productivity prevailed across firm size groups. For process innovation, the finding was particularly robust for medium-sized firms.
- Labour productivity growth was found to be strongly correlated with capital intensity, while measures of market competition and inter-industry differences showed varying association.
- Non-constant returns to scale and imperfect competition seemed to be valid assumptions for large firms.
- Firms with lower initial productivity were associated with faster labour productivity growth in the IS-EAS study, which pointed to learning and catching-up.
- The association between labour productivity growth, on the one hand, and innovation output and other explanators, on the other hand, was not clear and robust through various investigations of manufacturing and services. Differences between manufacturing and services were, however, indicated.

- Additional results from the IS-EAS study showed a stronger association between innovation outputs and measures of multi-factor productivity, rather than measures of labour productivity. For example, product innovation influenced appeared to have a greater positive effect on MFP growth than on LP growth; organisational innovation showed positive correlation with the level of MFP.
- The association between a high degree of innovation novelty and productivity performance was found to apply to process innovation only.

DISCLAIMER

Any views or opinions expressed in this publication are those of the authors and are not necessarily those of the Australian Bureau of Statistics, the Productivity Commission or the Australian Taxation Office (ATO). Any errors are the authors' own.

The results of these studies are based, in part, on tax data supplied by the ATO to the ABS under the *Income Tax Assessment Act*, which requires that such data are only used for statistical purposes. No individual information is provided back to the ATO for administrative or regulatory purposes. Any discussion of data limitations or weaknesses is in the context of using the data for statistical purposes, and is not related to the ability of the data to support ATO's core operational requirements.

Careful consideration has been given to the privacy, security and confidentiality issues associated with using tax data in this project. Only people authorised under the *Australian Bureau of Statistics Act* have been allowed to see data about any particular firm in conducting these analyses. Results have been confidentialised to ensure that information relating to an individual business is not identifiable.

CONTENTS

LIST OF TABLES

EXPLORATIONS OF INNOVATION AND BUSINESS PERFORMANCE USING LINKED FIRM-LEVEL DATA

Marn-Heong Wong
Productivity Commission

Darren Page, Ruel Abello and Kwok Pan Pang
Australian Bureau of Statistics

ABSTRACT

Innovation is recognised as a key driver of economic growth. However, because the innovation process is complex, its effects on performance are difficult to quantify. This research applied the Crépon, Duguet and Mairesse (CDM) framework at the individual business level to explore the association between innovation and business performance. It used linked firm-level data from surveys conducted by the Australian Bureau of Statistics augmented with administrative data from the Australian Tax Office.

The use of the CDM framework allowed for the analysis of the association between innovation and business performance in three stages: the determinants of innovation investment; the effect of innovation input on innovation outputs; and, the effects of innovative outputs and other firm, industry and market characteristics on business performance. Three different types of innovation output were studied: (i) product, (ii) process and (iii) organisational.

The analysis clearly indicated that the relationship between innovation and performance was complex and variable across different groups of firms. There were differences in the explanations of the decision to invest and intensity of innovation investment. Innovation output was positively influenced by innovation intensity. There was evidence of a positive relationship between innovation and business performance, although the strength of this relationship varied depending on the dataset used, industry coverage and also the performance measures used.

1. INTRODUCTION

1.1 Background

Innovation is perceived as central to economic growth. It involves the commercial production of new or significantly-improved goods and services, or the commercial use of new or significantly-improved methods to produce goods and services. Successful innovations enable businesses to produce higher-value output and to produce more efficiently.

Despite its importance, researchers have had difficulty in measuring the extent to which innovation improves economic performance. Case studies based on individual firms may offer reasonably precise measures, but there are doubts about the extent to which their results are representative of larger populations of firms. Industry- or economy-wide studies, whilst obviously stronger on representativeness, have not always been able to provide precise, robust and unambiguous results.

1.2 Recent developments

A number of related developments in empirical analysis of innovation have occurred over the past two decades. First, better data on innovation have become available. Traditionally, lack of direct measures has meant that innovation has been represented in empirical studies by measures of R&D expenditure or patents, which have shortcomings in this context. Undertaking R&D activity is only one input to the innovation process and, while patents can be considered an output of R&D activity, measures of the numbers of patents granted do not indicate the full extent of innovation in commercial operations.

The OECD initiated a study to establish guidelines, known as the *Oslo Manual* (OECD, 1992), for collecting data that would enable in-depth analysis of business innovation and allow for international comparisons. In the 1990s, innovation data were collected in Canada, and the Community Innovation Surveys (CIS) were run in 13 European countries. Major developments also occurred in Australia. The ABS conducted its first comprehensive innovation survey of the Australian manufacturing industry in 1993, with standards consistent with the *Oslo Manual* and the CIS. Other surveys followed and the ABS now has an ongoing program of innovation surveys conducted every two years, with some key variables collected annually.

A second development was the information and communications technology (ICT) boom of the mid-1990s, which delivered productivity gains, but more so in some countries than others and more so in some industries than others. Clearly, the mere existence of new ICT technologies was not sufficient to deliver productivity gains. It was the ways in which commercial users applied the technologies that made the

difference. Firm-level investigation was needed to explore the role of ICT in enabling new business models, user-developed innovations and the ensuing productivity gains. Firm-level analysis provided insights into the range of other complementary factors that condition the degree to which ICT can deliver productivity gains.

A third development has been in empirical frameworks to support analysis of innovation. The framework of particular interest in this study is the one developed by Crépon, Duguet and Mairesse (1998). Their framework explicitly distinguishes between innovation inputs and innovation outputs, and emphasises that it is innovation outputs that affect business performance. It is applied to data from many firms and can be used to investigate factors that help and constrain innovation at the firm level.

1.3 The role of firm-level analysis

The developments above enhanced the ability and highlighted the need to explore the diversity in the innovation behaviour and performance of firms.

That there is this diversity in behaviour and performance across firms, even within the same industries, is by no means a new idea (see, for example, Schumpeter, 1937). However, for practical reasons, mainstream analysis had been based on the concept of the 'representative' firm. From the mid-1990s, the greater availability of large-scale firm-level (and longitudinal) datasets, together with easier access to the computing power needed to analyse them, enabled researchers to capture and to explain the firm-level diversity in innovation and performance.

The diversity at the firm level is not just a curiosity, but also has considerable policy relevance. The productivity performance of the economy at large depends on the production decisions made by the range of individual firms. If firms are behaving and performing in a mostly uniform manner, policy makers can more easily get a handle on factors that might help or hinder innovation and better productivity performance. Aggregated or 'representative' data – or even anecdotal evidence – might serve quite well. But, if there is diversity in behaviour and performance, the linkages between policy levers and performance outcomes are much less clear.

Investigations based on firm-level data not only reveal the extent of diversity, but also provide a foundation upon which better understanding of the complexity of factors that determine behaviour and performance can be built. They can point to important differences in behaviour and performance between firms in, for example, different industries, size classes, regions or ownership structures. They can point to the complementary and interrelated factors that together determine performance outcomes. (With the many more observations they provide, compared with aggregate time-series data, firm-level data also have the potential to identify a larger and more robust range of influences.) In the case of ICT and its effect on performance, for

example, firm-level analysis pointed to such influences as competitive incentives, the availability of complementary skills and the flexibility that firms needed to restructure work arrangements and organisational structures as important pre-conditions that strengthen the positive effects of ICT on productivity performance (OECD 2004).

1.4 The project

The ABS and the Productivity Commission initiated a collaborative research project in 2006. The project was designed to bring together expertise in data development and productivity analysis. The analytical component focussed on the development and application of the Crépon, Duguet and Mairesse (CDM) framework to Australian data. The data component was focussed on bringing together firm-level data from different survey and administrative sources in order to provide linked information on firm behaviour (innovation) and performance (productivity).

Two complementary studies were undertaken for the project. The studies are intended to be exploratory and to inform future data and analytical work, rather than to generate definitive results. The project had four objectives related to the construction and use of linked datasets:

- to establish a framework suitable for the analysis of innovation using firm-level data;
- to assess the value of linked firm-level data to inform research and policy development;
- to explore factors affecting innovation and performance among Australian firms; and
- to draw any lessons for future data collections and linking exercises.

There have been a number of studies on innovation among Australian firms, but the majority of them concentrated on the innovation process only – for example, DITR (2006), Webster (2003) and Rogers (2000). Where the links between innovation and business performance were analysed – for example, Phillips (1997) and Bosworth and Loundes (2002) – the datasets used either lacked adequate measures of performance or were not as rich in innovation variables compared with the linked datasets constructed for this project.

1.5 The studies

Initial model specification and investigation was carried out with survey data only. Subsequent access to administrative data from the Australian Taxation Office (ATO) enabled investigation of a more representative range of firms. However, additional data-access restrictions meant that the latter project could only be undertaken by a different study team comprised only of ABS staff. As it turned out, one advantage of

undertaking two studies was that it allowed some examination of the robustness of results across the two data sources. This could indicate relative strengths and weaknesses of the two data sources for analytical purposes.

A common starting point was a detailed specification of the general CDM framework. The two studies essentially differed in the linked datasets to which the CDM models were applied. Innovation data were drawn from the ABS 2003 Innovation Survey (IS) in both cases, but the sources of performance data differed. The first study used linked data for common businesses in the IS and the Economic Activity Survey (EAS). This is referred to as the 'IS-EAS' study in this paper. The second study linked the IS data not only with the EAS data but also with ATO's Business Income Tax (BIT) data or Business Activity Statement (BAS) data. This is referred to as the 'IS-EAS-BIT-BAS' study in this paper.

The two studies were the same in broad direction, but differed in some relatively minor respects. Although the IS-EAS-BIT-BAS team drew largely on the specifications and econometric methodologies developed in the IS-EAS study, differences in the details of implementation evolved as the team took independent modelling decisions, largely to account for specific characteristics of their dataset. The other critical differences between the studies were in the type of performance measures that could be constructed and in the coverage and the representativeness of the firms that could be included in the datasets. The IS-EAS dataset is focused on medium and large firms, while the IS-EAS-BIT-BAS dataset is more representative of the full population of firms.

1.6 Outline

Section 2 describes the data sources and the linkages undertaken to produce the datasets for the two studies. The econometric methods are discussed in Section 3. Descriptive statistics are presented in Section 4. The results are presented in Section 5, which starts with a cautionary note on interpreting the results and comparing the findings. Section 6 concludes and identifies areas for future improvements to the research.

2. DATA SOURCES

2.1 Innovation Survey, 2003

Innovation surveys have been conducted periodically by the ABS since 1993–94. This analysis uses data from the 2003 survey (ABS, 2006a).

The reference period for the 2003 Innovation Survey is mainly the three-year calendar period 2001–2003. Some data relate to the calendar year 2003, and financial data relate to the then most recent financial year, 2002–03.

The survey covered the following areas: characteristics of innovating and non-innovating businesses; types of innovations occurring; sales derived from new or changed products; extent of cooperative linkages between businesses and research institutions; source of innovation ideas and funds; drivers and barriers to innovation; and expenditure on innovation.

Defining innovation

The ABS survey defined innovation as the introduction of any new or significantly improved goods or services, the introduction of new operational processes (the methods of producing or delivering goods or services) or the implementation of new organisational /managerial processes (meaning strategies, structures or routines that aim to improve business performance) during 2001–2003. Businesses were considered 'innovators' if they had introduced at least one of these types of innovation during 2001–2003. Firms could report more than one type of innovation.

In this paper these types of innovation will be referred to as 'product', 'process' and 'organisational' innovations.

This definition of innovation means that innovators are not necessarily engaged in the development of goods, services or processes that radical or new to the world. They can be reproducing goods that are already on the market, perhaps using off-the-shelf technology inputs, and making small incremental improvements to their goods and services, or implementing well-understood forms of organisational change. In order to further analyse the 'novelty' of the innovations, questions also asked how 'new' were the innovations introduced by these firms. Responses could be totally new to the world, or new to a country, an industry or otherwise simply new to the business.

2.2 Economic Activity Survey, 2001–02 to 2004–05

The ABS has conducted annual Economic Activity Surveys since 1988–89 (ABS, 2006c). This study uses data from the surveys covering the reference periods 2001–02, 2002–03, 2003–04 and 2004–05. The Economic Activity Survey collects a range of financial information from businesses, such as:

- sales of goods and services;
- expense and inventories items sufficient to calculate intermediate inputs;
- from the above items, nominal value added can be derived;
- depreciation;
- balance sheet items, including non-current assets (not collected in the 2004–05 survey); and
- various detailed other income and expense items.

2.3 Business Income Tax

In Australia, all businesses which are not income tax exempt are required to lodge an income tax return form to the Australian Taxation Office (ATO) on an annual basis to establish their final tax liability. The income tax return form contains information on their income and related information. The changes to the Income Tax Assessment Act made in 2005 enable the ATO to pass the data for all businesses to the ABS to be used for national accounting and statistical research purposes.

Access to the business income tax data enabled the IS-EAS-BIT-BAS study to find a significant proportion of the businesses that were in the Innovation Survey but were not in the EAS. These businesses were mostly small- and medium-sized businesses.

The BIT contains important information that can be used for the CDM model, such as:

- income, including sales of goods and services;
- expenditure items, including depreciation; and
- other variables, such as change in inventories and assets held.

2.4 Business Activity Statement

Since the government introduced the New Tax System in 2000, most Australian businesses are required to submit a Business Activity Statement to the ATO. As in the case of BIT data, legislation allows the ABS to have access to BAS data, which are used in this study to fill some remaining gaps.

Data items that can be sourced from the BAS include:

- total value of sales of goods and services;
- total wages, salaries and other payments;
- value of purchases of capital goods; and
- value of purchases of non-capital goods.

The above data items can be used to derive some measures of business performance such as changes in the total values of sales of goods and services and changes in value added.

2.5 The linked datasets

2.5.1 IS-EAS linked data

Both surveys used common statistical units and sampling frames, based on the ABS Business Register (ABSBR). As such, the linking of data from the various sources was fairly straightforward.

However, ABS policies aimed at reducing the load placed on businesses providing data to the ABS resulted in a relatively low number of small and, to a lesser extent, medium-sized businesses being common to both surveys. As such, the resulting IS-EAS linked dataset summarised in table 2.1 is weighted towards large businesses.

2.5.2 IS-EAS-BIT-BAS linked data

The inclusion of taxation data in the linkage process significantly increases the number of firms in the Innovation Survey for which data on economic activity and business performance are available. It also gives a more even distribution of firms by size grouping (see last column in table 2.1).

2.1 Number of firms in the linked datasets, by employment size

			IS-EAS linked dataset		
Employment	*In-scope population*	*Innovation Survey 2003*	*Sample used in the innovation input & output equations*	*Sample used in the productivity equation*	*IS-EAS-BIT-BAS linked dataset*
0–19	103,416	2,335	73	49	1,329
20–199	34,735	2,188	524	433	1,233
200+	2,646	1,675	1,083	949	1,029
Total	140,797	6,198	1,680	1,431	3,591

3. MODEL SPECIFICATION AND ESTIMATION

This paper applied a modelling framework proposed by Crépon, Duguet and Mairesse (1998), now known as the CDM model, to two sets of Australian data. This model can be used for micro-data at the individual business level and encapsulates the innovative process from firms' decision to invest in innovation activities to the impact of innovations on their business performance. Precisely, the relationships are represented by three sets of equations:

1. the innovation input equation links a firm's innovation investment to its determinants;
2. the innovation output equation relates innovation input to innovation output measures; and
3. the business performance equation examines the effect of innovative output on business performance.

Crépon, Duguet and Mairesse (1998) estimated their model on a French dataset that merged sources of innovation, R&D and accounting data as a simultaneous equation system, on the premise that innovation inputs, innovation outputs and business performance are endogenously determined. Since then, the CDM model (or its variants) has been applied to data in an increasing number of countries including, for example, the Netherlands (Klomp and Leeuwen, 2001), the United Kingdom (Criscuolo and Haskel, 2003), Sweden (Lööf and Heshmati, 2006), Germany (Janz and Peters, 2002) and Chile (Benavente, 2006). While sharing the same structural modelling approach to assessing the links between innovation and firm performance, they differ in their choice of endogenous and explanatory variables and estimation methods.

The two studies in this paper are a first application of the CDM model on two sets of linked Australian data. As such, they are intended to be exploratory in nature, with the aim to develop a better understanding of the data and appropriate empirical methods and to generate preliminary insights into the relationship between innovation and business performance in Australia. Each equation in the CDM framework was estimated separately to enable a first examination of the conditional behaviour of innovation input, innovation output and business performance using these datasets.

3.1 CDM model

The three sets of equations in the CDM model that examine the links between innovation input, innovation output and business performance are specified as follows.

$$y_{1i} = x_{1i}\beta_1 + \varepsilon_{1i} \tag{1}$$

$$y_{2i} = \alpha_1 y_{1i} + x_{2i}\beta_2 + \varepsilon_{2i} \tag{2}$$

$$y_{3i} = \gamma_1 y_{2gs,i} + \gamma_2 y_{2op,i} + \gamma_3 y_{2om,i} + x_{3i}\beta_3 + \varepsilon_{3i} \tag{3}$$

Innovation input equation

In equation (1), y_{1i} is a measure of 'innovation intensity' or input made by the i-th business in its innovative activity. x_{1i} is a vector of variables that determines the innovation intensity of the i-th business and β_1 is a vector of corresponding coefficients. ε_{1i} is an error term.

Innovation output equation

In equation (2), y_{2i} is a measure of the innovation output produced by i-th business, x_{2i} is a vector of variables that determines the innovation output, β_2 is a vector of corresponding coefficients and ε_{2i} is an error term. Innovation intensity (y_{1i}) is used as an explanatory variable for innovation output. Equation (2) can be estimated using various types of innovation outputs and the two studies in this paper estimated three separate equations using measures of innovation outputs in the areas of 'goods and services' (henceforth referred to as product innovation) ($y_{2gs,i}$), 'operational processes' (process innovation) ($y_{2op,i}$) and 'organisational or managerial processes' (organisational innovation) ($y_{2om,i}$). An alternative equation for product innovation was also estimated, using a different dependent variable (i.e. turnover attributed to product innovation).

Business performance equation

In equation (3), y_{3i} is a measure for the performance of i-th business, x_{3i} is a vector of variables that determines the performance and it includes labour and capital inputs. ε_{3i} is an error term. This equation includes all three types of innovation outputs (i.e. $y_{2gs,i}$, $y_{2op,i}$ and $y_{2om,i}$) on its right-hand-side.

In the CDM model, the links between innovation inputs, outputs and business performance are established through y_{1i} in equation (2) and $y_{2gs,i}$, $y_{2op,i}$ and $y_{2om,i}$ in equation (3).

3.1 List of key dependent variables and selected explanatory variables

Dependent variables	*Examples of explanatory variables*
1. Innovation input equation	
Innovation intensity (Innovation expenditure / Total sales)	Technological opportunities to innovate: – Sources of ideas for innovation – Collaboration arrangements – Regulations and standards Incentives to innovate: – Market power (individual market share; industry concentration) – Potential for sales growth – Appropriability conditions (formal and informal methods of intellectual property protection) Financial resources: – Profitability – Government financial support Other firm characteristics: – Size – Age – Share of foreign ownership Industry dummies (to control for inter-industry differences)
2. Innovation output equations	
a. Product innovation (Share of innovative sales: Turnover attributed to product innovation) b. Product innovation (binary indicator) c. Process innovation (binary indicator) d. Organisational innovation (binary indicator)	Innovation input: – Innovation intensity Technological opportunities to innovate: – as in equation (1) Incentives to innovate: – Appropriability conditions Firms' human capital and absorptive capacity: – Engage in intramural R&D – Share of ICT employees in total employment – Recruit workers with scientific and engineering skills to undertake innovations Other firm characteristics: – as in equation (1) Industry dummies
3. Business performance equations	
a. Labour productivity [1] (Value added per employee) i. annualised growth (2001–02 to 2004–5) ii. level (2003–04) b. Multi-factor productivity [2] (Ratio of gross output to sum of factor inputs) i. annualised growth (2001–02 to 2004–5)	Innovation output: – Share of innovative sales – Process innovation (0/1) – Organisational innovation (0/1) Capital intensity Control variables for competition, skills and other firm characteristics: – Individual market share – Industry concentration – Share of ICT employees – Size – Age – Share of foreign ownership Industry dummies

1 The term 'labour productivity' is used loosely here. See Section 3.2.

2 The term 'multi-factor productivity' is used loosely here. See Section 3.2.

Dependent and explanatory variables

An extensive number of dependent and explanatory variables were specified in the three sets of equations. Table 3.1 sets out the key dependent variables used (and their approximate definitions), and provides examples of the explanatory variables used under broad categories. The complete list of variables and their precise definitions are provided in Appendix A.

3.2 Measuring business performance

Business performance can be measured by a variety of indicators, both at a point in time and over time. Productivity level and growth are popular measures, because of their links to per capita income growth and economic well-being at the aggregate level. Other indicators include changes in employment, duration of survival, and financial variables such as profitability and market valuation. The two studies reported in this paper constructed productivity measures for each firm to examine business performance.

3.2.1 Qualifications

The term 'productivity' is used loosely here, as data limitations mean that the measures could not be constructed strictly according to theoretical requirements. While consistent with the broadest definition of productivity as the ratio of output to factor inputs, there are several qualifications:

- Nominal output values were not deflated to obtain output in volume terms, due to a lack of suitable deflators. This means that some of the changes in output could reflect price changes.
- Labour inputs were measured from available data as the number of persons employed or wages and salaries paid instead of the ideal measure of hours worked adjusted for quality differences.
- Capital input figures that proxied for capital services were crude approximations of productive capital stock. They were derived by applying a rough version of the Perpetual Inventory Method, using non-current asset values as the initial value and depreciation rates reported by businesses.

For a more detailed presentation of these issues, see ABS (2007).

3.2.2 Description of productivity measures

The two studies differed in their coverage of productivity measures. The IS-EAS study explored the relationship between innovation output and a more extensive range of productivity measures in both growth and level, while the IS-EAS-BIT-BAS study focused mainly on labour productivity growth measures.

The IS-EAS study constructed both gross-output and value-added based labour productivity and multi-factor productivity (MFP) measures, in growth rates (annualised between 2001–02 and 2004–05) and in levels (in 2004–05). These productivity measures are not independent of each other and their interrelationships can be shown using the economic theory of production (see OECD (2001) for more). Each measure has it own appeal. Labour productivity is easier to measure, while there is more measurement issue around the construction of a MFP index. However, while labour productivity growth can result from changes in factor intensity ratios as well as MFP growth, a MFP change has the attractive property of isolating how factors other than input mix affect output growth. These factors are often interpreted as changes in technology, managerial ability and/or organisational efficiency. The use of diverse productivity measures in the estimations of the business performance equation was also intended to facilitate approximate comparisons with international studies, which employed different measures. Only the results for selected productivity measures are presented in Section 5 of this paper although references are made to the other results where appropriate.

The IS-EAS-BIT-BAS study constructed two value-added based labour productivity growth indicators (annualised between 2001–02 and 2004–05) – one based on employment and the other based on wages. A measure of annualised growth in value-added was also used. The IS-EAS-BIT-BAS study team has opted to use fewer and relatively simple productivity measures because of data constraints. While the dataset used in this study had a more comprehensive coverage of businesses by size, some data were unavailable from taxation data sources. One of these was information on employment, which meant that labour productivity defined as value-added per employee was derived using employment figures estimated from a model developed within ABS. This introduced greater imprecision in the measure compared with a similarly defined measure derived in the IS-EAS study using directly collected employment figures. As a robustness check, the IS-EAS-BIT-BAS study also carried out estimations with an alternative measure based on available wage data although that measure itself was not ideal and seldom used. The construction of any MFP measure faced the additional issue that non-current asset values and depreciation rates were unavailable from tax data sources to enable the derivation of a rough and ready capital input measure. For this reason, MFP measures were not used.

Details on the construction of productivity measures in the two studies are described in Appendix C.

3.3 Estimation methods and issues

This section outlines the estimation methods used in the two studies based on linked IS-EAS and IS-EAS-BIT-BAS data. More technical details can be found in Appendix D.

3.3.1 Innovation input equation

The equation for innovation input was applied to the full sample, which included businesses that had reported expenditure on innovation – whether the innovation processes had been completed or not [1] – and businesses that did not report such expenditure but reasonably could be assumed to have zero expenditure. Both studies estimated a two-tiered model: the first equation models the probability that the measure of innovation input (innovation intensity) is positive, while the second equation assumes that for positive values, the conditional distribution of innovation intensity on the vector of regressors is lognormal. This two-step approach allows for separate processes to determine a firm's decision to invest in innovation activities and the level of investment. This means that a different set of factors can affect the two elements of choice or that the same factors can have effects in opposite directions. The model was suggested in Wooldridge (2002) and adopted by Criscuolo (2004) in her estimation of an innovation input equation on a full sample of UK manufacturing firms.

Our approach differs from most studies based on the Community Innovation Survey (for example, Crépon et al 1998 and Lööf and Heshmati 2006). In these studies, the CDM model was applied exclusively to the sub-sample of businesses that engaged in innovation and incurred expenditure on innovation, because the data about innovation expenditure and business characteristics were not collected from other firms. In these studies, a generalised tobit model was used to correct for possible sample selection biases.

3.3.2 Innovation output equation

The equations for innovation output were estimated on the full sample. In both studies, the probit model was used to estimate the three equations that had binary indicators of innovation output as the dependent variable. However, the equation

1 There was initial uncertainty over whether the questions on innovation expenditure in the 2003 IS had captured accurate information on innovation expenditure from all respondents. Precisely, the issue was the extent to which businesses that had started but abandoned or not completed innovation activities between 2001 and 2003 did not report any innovation expenditure when it was positive. This would introduce measurement error in the data, or a selection bias if the estimations were carried out on a sample of innovators only. Further examination of the data and verification using the latest 2005 IS data indicated that the assumption of zero innovation expenditure for businesses that did not report innovation expenditure was valid. Since most of the information in the 2003 IS on characteristics such as collaboration and sources of ideas for innovation activities were asked of all respondents, this allowed the estimation of the innovation input equation on the full sample.

using share of innovative sales as the dependent variable ('innovative sales equation') was estimated using different econometric methods in the two studies.

In the study based on IS-EAS data, the innovative sales equation was estimated using the fractional logit regression technique, which has been developed by Papke and Wooldrige (1996) for models with a dependent variable that spans the range from 0 to 1. This method ensures that the estimates will always take values within the range. When the dependent variable has a significant proportion of zeros (and/or ones), which is the case with our full sample, fractional logit regression is more appropriate compared with applying a log-odds transformation to the dependent variable. A logistic functional form is assumed for the conditional mean of share of innovative sales and the parameters are estimated by quasi-maximum likelihood estimation.

The innovative sales equation using the IS-EAS-BIT-BAS data was estimated using a censored regression model (type-1 Tobit model). The model was applied to take into account the fact that there were a significant number of observations with zero values for the dependent variable.

3.3.3 Business performance equation

The two studies in this paper estimated a number of business performance equations. These include estimations on the full sample using a range of business performance measures, as well as on sub-samples by manufacturing and service industries. The IS-EAS-BIT-BAS study ran further regressions by firm size classifications. The ordinary least squares estimation method was used in both studies.

The business performance equations specified in the two studies were derived from a Cobb–Douglas production function. However, instead of a standard derivation that assumes perfect competition in the product market, our specifications accounted for imperfect competition, along the lines of van Leeuwen and Klomp (2006) and Criscuolo (2004). The latter two studies suggested that innovating firms are often operating in markets characterised by horizontal or vertical product differentiation and can be expected to possess market power. There are two implications for firms operating in such markets. First, innovations may impact on firm performance via their effects on demand conditions, rather than serve as a knowledge capital input into production. Thus, product innovation can be thought to proxy for the change in product quality that affects a firm's market share, while operational and organisational process innovations measure the change in knowledge capital input into production. Second, if endogenous firm-specific prices (due to market power) are unobserved and not taken in account in the model derivation, this will lead to biased estimates for the coefficients of the production function. (Details of the derivation of various equations based on different business performance measures – for example, labour productivity and MFP growth and level – are given in Appendix B.)

It should be noted that in our regressions, the values of any financial variables were not deflated because of the lack of price deflators at the required disaggregation level. This is due to the belief that deflation using broad industry-level deflators across different variables might introduce more errors and give rise to more imprecise estimates compared with using nominal values. Nonetheless, this means that the regression results reported in this paper would reflect the effect of industry-level price changes that have not been accounted for. As such, our empirical implementation is not exactly aligned with the theoretical derivation, and this could add to imprecision in the interpretation of the coefficient estimates.

3.3.4 Estimating the CDM equations as a full system

The original CDM paper estimated the three equations as a full system, rather than singly or independently. Subsequent studies using the CDM framework have used different econometric techniques, ranging from running single equation estimations or running only certain parts of the model as a system, to full system estimation. In the time available and with a range of complex data and variable- and model-specification issues to confront in the project, the studies reported in this paper concentrated on single equation estimation. Some preliminary exploration of full-system estimation indicated that it could not be undertaken for the current project in a straightforward way that generated clear and satisfactory results. A thorough investigation of the relationship between innovation and business performance among Australian firms according to a systems implementation of the CDM framework remains an important area for further research. Such an investigation would reveal whether, and the extent to which, estimates from single equation estimation might be biased by the failure to account for simultaneity among the endogenous variables.

4. DESCRIPTIVE STATISTICS

4.1 Businesses by size, industry and innovation status

Table 4.1 presents key statistics for the two linked datasets used in this study. It shows that augmentation of the IS data with ATO data more than doubles the number of observations available for analysis and modeling compared with the case when IS data are augmented with EAS data alone (3,591 and 1,431 firms, respectively).

The linked IS-EAS dataset is skewed towards large businesses (66%) while the IS-EAS-BIT-BAS linked data has a good representation of small businesses (37%).

The linked IS-EAS data has a slightly larger proportion of manufacturing firms (51% compared to 46%).

A larger proportion of firms in the linked IS-EAS data are innovators (63%) compared to the linked IS-EAS-BIT-BAS data (47.5%).

In both datasets, the proportion of innovators rises as the size of the firm increases. In both datasets, about 3 out of 10 small businesses (0–19 employees) are innovators; about 5 out of 10 medium sized businesses (20–99 employees) are innovators; and over 6 out of 10 large businesses (100 and over) are innovators.

Looking at the major industry groupings, in both datasets, manufacturing has a larger proportion of innovating businesses than the service industries.

4.2 Businesses by innovation type and novelty of innovation

Of the firms who reported undertaking an innovation, process innovation has been cited as the most common type undertaken (table 4.1). Around 77 percent and 74 per cent of innovators in the linked IS-EAS dataset and linked IS-EAS-BIT-BAS datasets, respectively, have undertaken this type of innovation.

The second most-cited type of innovation is product innovation. Around 64 per cent of the innovating firms in both the linked IS-EAS dataset and linked IS-EAS-BIT BAS dataset undertook product innovation.

For organisational innovation, the figures for the IS-EAS and IS-EAS-BIT-BAS linked datasets are also roughly similar, at 63 percent and 62 per cent, respectively.

4.1 Distribution of business units in the linked datasets

	Linked IS-EAS dataset	*Linked IS-EAS-BIT-BAS dataset*
No. of business units (observations)	1,431	3,591
Distribution by size (%):		
0–19 employees	3.4%	37.0%
20–99 employees	16.1%	26.3%
100–199 employees	14.1%	8.0%
200 or more employees	66.3%	28.7%
Total	100.0%	100.0%
Distribution by industry group (%)		
Manufacturing	51.4%	45.8%
Services [1]	44.2%	43.7%
Others	4.5%	10.6%
Total	100.0%	100.0%
No. of innovators	901	1,706
% innovators	63.0%	47.5%
% innovators, by size		
0–19 employees	30.6%	30.6%
20–99 employees	51.5%	48.8%
100–199 employees	63.4%	60.1%
200 or more employees	67.3%	64.7%
% innovators, by industry group		
Manufacturing	66.8%	52.8%
Services	59.2%	43.5%
Others	56.3%	41.2%
Of innovators, % innovated in:		
Goods and services	64.5%	64.1%
Operational processes	76.8%	73.9%
Organisational /managerial processes	63.2%	61.5%
Of GS and OP innovators, % innovated by degree of novelty: [2]		
New to business	45.5%	52.4%
New to industry	17.8%	18.5%
New to Australia	23.6%	18.2%
New to the world	13.1%	10.9%

1 The IS-EAS study had a broader coverage of service industries encompassing the following ANZSIC divisions that were in the scope of the Innovation Survey: construction, electricity, gas and water supply; wholesale trade; retail trade; accommodation, cafes and restaurants; transport and storage; communication services; finance and insurance; property and business services; and cultural and recreational services. On the other hand, the IS-EAS-BIT-BAS study's definition of service industries excluded construction and the electricity, gas and water supply divisions. Hence, in the IS-EAS study, 'others' referred to the mining industry only, while in the IS-EAS-BIT-BAS study it was inclusive of mining, construction, electricity, gas and water supply.

2 This refers to the highest degree of innovation novelty of a firm, thus, an innovator in the 'new to the world' category could have introduced innovations with lower degrees of novelty at the same time.

In terms of the novelty of product and operational process innovation, 'new to business' has been the most common novelty degree. Almost half of product and process innovators introduced innovations for which the degree of novelty was no higher than 'new to business'. Close to 1 in 5 of the product and process innovators in either datasets have 'new to industry' as their highest degree of innovation novelty, that is, these innovators could have introduced innovations that were 'new to business' as well, but not innovations with novelty degree higher than 'new to industry'. A larger share of product /process innovators in the linked IS-EAS data introduced innovations where the highest degree of novelty was 'new to Australia' (23.6%) or 'new to the world' (13.1%), compared to the linked IS-EAS-BIT-BAS data (18.2% and 10.9%).

4.3 Business performance indicators

4.3.1 Linked IS-EAS data

Table 4.2 presents summary statistics on various measures of productivity performance and selected firm characteristics by innovators and non-innovators. The spread in quartile values shows that there is considerable heterogeneity in the productivity growth, market power and size of firms, for the groups of innovators and non-innovators alike. For innovators, there is also a varying degree of innovation intensity across quartiles. R&D intensity, however, tended to be low for the majority of the firms, which underlines the weakness in using R&D expenditure as a proxy for business innovations in traditional studies.

Innovators tend to have higher productivity growth rates and greater market power than non-innovators.

Innovators whose highest degree of innovation novelty is 'new to the world' have much higher productivity growth on average (table 4.3).[2] However, productivity growth rates do not seem to rise linearly with the degree of innovation novelty, and especially so for process innovation.

2 This could possibly be a reflection of the pricing power that comes with a 'new to the world' innovation, given that output values rather than volumes were used to calculate the productivity measures.

4.2 Business performance indicators by innovation status, linked IS-EAS data

	Innovators				*Non-innovators*			
	Median	*Q1*	*Q3*	*Mean*	*Median*	*Q1*	*Q3*	*Mean*
Growth rates (%):[1]								
Labour productivity (VA-based)	5.3	–1.7	12.2	4.6	5.0	–1.8	12.4	3.3
Labour productivity (Output-based)	4.3	–1.0	9.5	3.8	3.9	–1.0	9.9	2.7
MFP (VA-based)	5.3	–2.7	12.3	4.2	4.7	–2.8	12.3	3.0
MFP (Output-based)	1.7	–0.6	4.3	1.3	1.9	–0.6	4.7	1.0
Levels								
Innovation intensity (%)	1.3	0.3	4.4	4.1	–	–	–	–
R&D intensity (%)	0.0	0.0	0.8	1.5	–	–	–	–
Initial market share (%) (2001–02)	2.1	0.8	6.7	6.7	1.1	0.4	2.9	3.8
Employment ('000) (2001–02)	331	157	711	784	240	88	468	550

1 Annualised growth rate, 2001–02 to 2004–05.

4.3 Business performance indicators, by degree of product /process innovation's novelty, linked IS-EAS data

	Innovation novelty				
	None	*New to business*	*New to industry*	*New to Australia*	*New to the world*
Growth rates (%): [1]					
Product innovation:					
LP (VA)	3.5	4.7	4.2	4.5	7.6
LP (Gross output)	2.9	3.8	3.5	4.8	4.8
MFP (VA)	3.1	4.1	4.1	4.4	7.0
MFP (Gross output)	1.0	1.0	1.8	1.4	2.0
Operational process innovation:					
LP (VA-based)	3.6	5.5	3.8	1.6	8.7
LP (Output-based)	2.7	4.3	5.0	2.8	5.6
MFP (VA-based)	3.0	5.4	2.7	1.8	8.7
MFP (Output-based)	0.8	1.8	1.7	0.2	3.4

1 Annualised growth rate, 2001–02 to 2004–05.

4.3.2 Linked IS-EAS-BIT-BAS data

Under most performance growth measures, the average innovating business has a higher growth rate than the average non-innovating business (table 4.4). On average, innovating firms' sales have grown annually by 7.6 per cent, compared to 5.5 per cent in non-innovating firms. Value added has grown on average by 9.9 per cent annually among innovating firms, relative to 7.5 per cent among non-innovators.

As in the linked IS-EAS data, there is a wide spread in quartile growth values, in both innovators and non-innovators alike. Innovation intensity also varies considerably across quartiles.

4.4 Business performance indicators, innovators and non-innovators, linked IS-EAS-BIT-BAS data

	Innovators				*Non-innovators*			
	Median	*Q1*	*Q3*	*Mean*	*Median*	*Q1*	*Q3*	*Mean*
Growth rates (%): [1]								
Sales	6.1	–0.2	13.2	7.6	4.8	–2.3	12.4	5.5
Value added	7.1	–1.7	18.0	9.9	5.5	–5.1	16.4	7.5
Labour productivity (a)	4.3	–3.9	13.0	6.4	3.7	–3.9	12.4	6.1
Labour productivity (b)	0.7	–5.3	7.9	3.2	1.0	–5.5	8.0	3.4
Levels:								
Innovation intensity (c)	1.7	0.5	5.0	5.0	0.0	0.0	0.0	0.3

1 Annualised growth rate, 2001–02 to 2004–05.

(a) using value-added divided by employment estimates (as proxy for labour inputs). Model based estimates for employment were derived for businesses where taxation data (which does not directly collect employment information) was used as the source for performance information.

(b) using value-added divided by wages (as proxy for labour inputs).

(c) Innovation expenditure /total sales (%) in 2002–03.

5. RESULTS: MODELLING INNOVATION INPUT, OUTPUT AND BUSINESS PERFORMANCE

5.1 Cautionary notes

As discussed in previous sections, there are a number of similarities and differences between the dataset compositions and the empirical implementation in the two studies.

The findings of the two studies are not fully comparable. A major difference is that the IS-EAS data comprises predominantly large firms with employment above 200 (66 percent), while the IS-EAS-BIT-BAS data had a more even distribution of small, medium and large firms. The analysis below shows that a considerable amount of variation in the results between the two studies can be attributed to firm size differences, although it has not been possible to identify precisely all the reasons behind the differences in results.

There are issues related to data quality and econometric methodology that might affect the robustness of these results. Consequently, the findings contained in this paper should be used for indicative purposes only. Some of these issues are as follows:

- **Lags not captured**. At the time of this analysis, data used as performance indicators were only able to cover the period from 2001–02 to 2004–05 – i.e. up to only a short period after 2003, the primary reference period for innovation data in this study. This may not be enough time to allow for lags in businesses recouping innovation costs and then going on to harness any benefits from that innovation. Future research into the possible effect of lags is required.
- **Nominal prices, instead of volume terms**. All financial data are available in nominal prices only. At this stage, sufficient work has not been undertaken in Australia or internationally to be able to accurately reflect individual firm level performance in volume terms. As such, some of the changes in firm performance over time which are measured in this analysis may be due to the impact of price changes. See ABS 2007 (Chapter 3) for a full discussion of this issue.
- **Crude proxies for labour and capital inputs**. The linked data contains measures relating to the workforce of each business such as wages and salaries paid and the number of persons employed. Either of these measures is only a crude proxy for labour inputs for productivity measurement. Similarly, only a crude proxy for the input of capital services is enabled from the source data available to this analysis.

- **Crude approximations of market share**. There is no geographic or product detail available in the source data which would enable an accurate measurement of market share. For those businesses which operate across the whole of and only within Australia, where there is high product specialisation within an industry and where there is low import penetration, the calculation of market share will be quite accurate. Where one or more of these conditions do not apply, the calculation of market share will be less accurate.
- **Simultaneity bias**. Each of the equations in the CDM framework was estimated separately to enable a first examination of the conditional behaviour of the dependent variables. This single equation estimation (rather than simultaneous equation estimation) is not taking into into account any possible cross-correlations between equations. This could result in biased estimates if there is considerable simultaneity among the endogenous variables. Any future refinement to these studies will look closely into the use of simultaneous equation estimation of the CDM model.

Given the above, the findings from these studies need to be treated as indicative only and should not be used in isolation or as indicating definitive causal relationships. Nevertheless, the study provides valuable insights in this field in the Australian context.

5.2 Results from the IS-EAS and IS-EAS-BIT-BAS studies

The focus of this section will be on examining the relative significance and direction in correlation between the explanatory variables and the dependent variable. The level of significance of each variable is identified by an asterisk, where three, two and one asterisk(s) denote significance at the 1, 5 and 10 percent level respectively. For example, significance at the 1 percent level means that the probability of rejecting the null hypothesis that a parameter is zero when it is true is 1 percent of the time. The size of the coefficient estimates will not be presented nor interpreted in this paper. A comparison of coefficient estimates between the two studies is difficult as there are some differences in equation specifications and variable definitions. It is noted that the precision of estimates might be also affected by data measurement and methodology issues, as set out in the section above.[3]

In discussing the results from the three sets of equations, the firm performance equations are more extensively analysed, as a primary aim of linking Innovation Survey data to economic performance data is to enable the exploration of how innovations affect firm performance. Also, in discussing the factors that impact on firm's performance, this paper has chosen to first analyse the results for the innovation

3 The tables with the actual coefficient estimates are available on request. The reader is cautioned though of comparing and interpreting the magnitudes between the two studies, for the reasons explained in Section 5.1.

output variables, and then to go on to interpret the results on other variables, in order to highlight how other factors could be more significant in affecting performance in our samples.

The box below gives a short summary of the results. The detailed description of the results follow in the next sections.

BOX 1. OVERVIEW OF RESULTS

The modelling provided strong support for the structured approach of the CDM framework. There were meaningful distinctions between the explanations of the decision to invest in innovation, the level of investment, the innovation outputs produced and the effects on business performance.

Generally speaking, the modelling was 'well behaved' in the sense that it provided results that accorded reasonably with prior expectations of the effects that various measured influences would have on innovation behaviour and business performance. There was a reasonable degree of overall explanation of variation for this type of analysis.

While the analysis has not been definitive, it has clearly indicated that the relationship between innovation and performance is complex and variable across different groups of firms.

- There were major differences in innovation behaviour and performance effects according to the size of firms. For example, larger firms are more likely to invest in innovation, but not necessarily more intensively. Medium-sized firms were more likely to see improved productivity performance from process innovation.
- The innovation input and output equations were better determined than the performance equations. Estimation results from the performance equations were sensitive to the dataset, industry coverage and the performance measure used.
- In terms of general findings specific to the three links specified in the model:

Innovation input

There were differences in the explanations of the decision to invest in innovation and the decision on how much to invest. Factors representing the opportunities to innovate, incentives to innovate, financial resources and other firm characteristics could differ in their significance or directions of influence in the two decision stages.

Innovation output

Innovation outputs—in general and, specifically, outputs of product, process and organisational innovations—were positively associated with innovation intensity.

There were other firm-specific and technological environment factors associated with different innovation outcomes. Many of them affected the different types of innovation outputs similarly, although some differences were also observed.

Business performance

There were positive associations between innovation output and performance but they were relatively weak statistically.

The modelling indicated that product and process innovation had a clearer effect on performance than managerial and organisational innovation. (However, there were also indications of sensitivity of this finding to the productivity measure used.)

There were also indications of different innovation behaviour and performance effects in manufacturing and services, although specific distinctions did not come through clearly.

5.2.1 Innovation input equations

The two studies applied the same two-tiered model to estimate the innovation input relationship. The first step modelled a firm's propensity to invest in innovation activities, and conditional on investment, the second step estimated the level of its innovation investment (measured as innovation intensity). The two studies differed in their inclusion and definition of a few regressors. Table 5.1 shows the estimation results on the full sample of both datasets.

The first thing to note is that there is a difference in the significance of factors affecting a firm's decision to invest (equation 1) and the amount of its innovation expenditure (equation 2).

Internal information impacted positively on both the decision to invest and innovation intensity; smaller firms also based their decisions on external sources of ideas.

Among the information sources, internal information (from within a firm or enterprise group) seemed to be important in a firm's decision on whether to invest in innovation activities in both studies. The IS-EAS-BIT-BAS study found in addition that commercial information and free information sources were also significant. Further examination of sub-sample regression results by firm size categories showed that these external sources of information were significant for the small (0–19 employees) and medium (20–199 employees) firms but not among large firms (200 or more employees). The greater reliance on external sources of ideas is not surprising for businesses with a smaller internal knowledge pool.

After a firm had decided to invest, the IS-EAS study showed that innovation intensity was positively related to internal information and the IS-EAS-BIT-BAS study showed

that the coefficient was insignificant in the full sample but significant in the large firm sub-sample. Vertical information was significant and positive in the IS-EAS-BIT-BAS study but not in the IS-EAS study.

Lack of access to information from competitors and the presence of intellectual property protection were associated with innovation intensity.

Both studies found that innovation intensity was strongly negatively related to information from competitors. This result is probably a reflection of the phenomenon where firms that compete strongly prefer to protect knowledge and spend more heavily on innovation as a competitive strategy. Or the result may also suggest that the lesser the knowledge spillover from competitors, the more a firm needs to spend on innovation investment. This is also related to appropriability conditions, where firms that are confident that competitors cannot easily copy from them would tend to invest more in innovation. The generally positive coefficient estimates of intellectual property variables support this point. The IS-EAS study observed that diversity in strategic (or informal) methods of intellectual property protection, such as secrecy and complexity of design had a strongly significant and positive effect on both innovation investment propensity and intensity compared with diversity and/or intensity in formal methods of protection such as patents and trademarks. The IS-EAS-BIT-BAS study identified both strategic and formal methods of protection, measured as yes/no binary indicators, as having a positive impact.

Collaborations encouraged firms to invest in innovation but did not influence the level of investment.

Estimates on the collaboration variables showed that collaborations with a diverse number of organisations were linked to a higher propensity to innovate. The IS-EAS study further found that diversity in the types of collaboration, which was not included in the regression on IS-EAS-BIT-BAS data, also mattered.

However, conditional on investment, innovation intensity did not appear to be highly correlated with collaboration arrangements. The IS-EAS study found that only the coefficient on collaboration with the wider enterprise group was significant. This effect was negative, which suggested that a transfer of knowledge specifically through this channel was linked to a decrease in innovation expenditure for a firm in the IS-EAS sample, presumably because it could draw on resources from other parts of the group. This coefficient was insignificant in the IS-EAS-BIT-BAS study. The only significant collaboration variable in the intensity equation in the IS-EAS-BIT-BAS study was collaboration with suppliers and clients, which was positive. Sub-sample regression results indicated that this factor was likely to be significant among small and medium firms.

5.1 Results from the estimation of innovation input equations [1,2]

	Propensity to innovate		*Level of investment*	
	Dependent variable: Invest in innovation (yes/no)		*Dependent variable: Innovation intensity*	
Explanatory variables	*IS-EAS*	*IS-EAS-BIT-BAS*	*IS-EAS*	*IS-EAS-BIT-BAS*
Internal information	+++	+++	++	.
Vertical information	.	.	.	++
Information from competitors	.	.	−−−	−−−
Commercial information	.	+++	.	.
Information from universities	.	.	.	.
Information from government	.	.	.	.
Free information	.	+++	.	.
Collaboration with wider enterprise group	++	+++	−	.
Collaboration with suppliers and clients	++	+++	.	++
Collaboration with competitors	.	.	.	.
Collaboration with commercial sector researchers	+	+++	.	.
Collaboration with universities, gov't and non-profit institutes	.	.	.	.
Collaboration with other types of organisations	.	.	.	.
Collaboration diversity	+++		.	
Collaboration intensity	.		.	
Barrier: government regulation or standards	−	.	.	.
Formal protection of IP (diversity)	.		.	
Formal protection of IP (intensity)	.		−−	
Formal protection of IP (0/1 dummy)		+++		++
Strategic protection of IP (diversity)	+++		+++	
Strategic protection of IP (0/1 dummy)		+++		+++
Drivers: improve productivity or reduce costs	.		.	
Drivers: improve productivity		+++		.
Drivers: reduce costs				−−−
Drivers: increase revenue	+++	+++	.	.
Drivers: increase export opportunities	++	+++	++	+++
Other market-related drivers (incl. increase market share)	+++		.	
Drivers: increase market share		+		+
Other market-related drivers (excl. increase market share)		+++		.
Initial market share (log-level)	++	.	−−−	
Initial market share (level)		+++		−−−
Herfindahl index	.	.	.	.
Industry sales growth	++	.	.	−−
Barrier: lack of skilled staff	++	+	.	.
Barrier: cost or availability of finance	.	.	++	+++
Government financial support	.	.	.	+++
Profit as a share of total sales	.	+++	.	.
Firm size (employment)	++		.	
20 to less than 200 employees		++		−
200 or more employees		+++		.
Foreign ownership (>0% and <50%)	.	.	.	.
Foreign ownership (50% or more)	.	.	.	+
Age (1 to <4 years)	.	.	.	.
Age (4 to <9 years)	.	.	.	+
Age (9 years or more)	.	.	.	.
Non-profit institute dummy	.		.	
2-digit industry dummies[3]	0.020		0.000	
1-digit industry dummies[3]		0.000		0.000
Number of observations	*1,679*	*3,591*	*1,105*	*1,904*

+, ++, and +++ indicate positively significant at 10%, 5% and 1% respectively; −, −− and −−− indicate negatively significant at 10%, 5% and 1% respectively.

1 Where the cell is blank, the variable was not used in the modelling in that particular study.

2 "." indicates that a variable has returned an insignificant coefficient.

3 p-value for joint significance of industry dummies.

Government regulations and standards were either insignificant or weakly significant in dampening innovation effort.

In the IS-EAS study, barriers to innovation in the form of government regulation and standards was shown to be weakly significant (at the 10 percent level) in dampening a firm's propensity to innovate, but had an insignificant effect on the innovation expenditure of investors. This variable was insignificant in both propensity and level equations in the IS-EAS-BIT-BAS study.

The drive to increase export opportunities, revenue and market shares motivated innovation activities.

Cost-push and demand-pull drivers identified by businesses as being the reasons for innovation showed varying strengths of association with the propensity to invest and innovation intensity.

Both studies found that the driver to increase export opportunities was important in both decision stages. The coefficient estimates for this variable support the hypothesis that firms which are driven by increased export opportunities and thus facing greater market competition, are more likely to engage in innovation activities.

Firms in the IS-EAS sample did not seem to be strongly driven by the motive to improve productivity or reduce costs, while the intent to increase revenue and other market-related drivers (including increasing market share) were important in a firm's propensity to innovate. The findings in the IS-EAS-BIT-BAS study on these variables were similar. In addition, the latter study yielded a significant estimate on the driver to improve productivity in the propensity equation. Sub-sample results suggested that this reflected a strongly significant estimate among small firms. The IS-EAS-BIT-BAS study also found that the driver to reduce costs was significantly and negatively correlated with innovation intensity. Sub-sample results showed that this relationship was significant among small and medium firms.

Firms with higher market share were more likely to innovate but had lower innovation intensity.

In terms of quantitative indicators of market demand and competition, both studies found that firms with higher market share at the beginning of the period were more likely to invest, but having decided to do so, they invested less. The result for the propensity equation is consistent with the Schumpeterian (1942) view that ex ante market power favours innovation, while the second result implies that firms with greater market power spend relatively less on innovation activities, possibly because they can earn greater rents from similar amounts spent on innovations.

The Herfindahl index measuring industry concentration had insignificant coefficient estimates in both decision stages.

Potential growth in market demand raised large firms' propensity to invest in innovation.

Potential growth in market demand (as represented by 4-digit specific industry sales growth) was significant in raising the probability of innovation investment in the full sample of IS-EAS data and in the large firm sub-sample of the IS-EAS-BIT-BAS data. This variable was insignificant in influencing innovation intensity in the IS-EAS study, but significantly and negatively correlated with innovation intensity in the IS-EAS-BIT-BAS study. Sub-sample results indicated that this link was particularly strong among medium firms.

Lack of skilled staff did not seem to adversely impact the decision to innovate.

The coefficient on the variable 'lack of skilled staff as a factor hampering innovation', which is intended to proxy for human capital and a firm's absorptive capacity, was significant and positive in the propensity equation of both studies. The positive sign seemed counter-intuitive if interpreted as a causal relationship. However, as an associated phenomenon it is plausible, as demand for skilled labour is likely to be high if innovation is occurring in addition to normal business activity.

Government financial support had a positive effect on the innovation intensity of small firms but did not appear to influence the innovation intensity of large firms.

Government financial support appeared to be an insignificant factor in influencing innovation intensity using IS-EAS data, but was significant using IS-EAS-BIT-BAS data. Sub-sample results showed that the significant relationship was likely to predominate among small firms. It could be expected that small businesses might face greater financial constraints and benefit from government subsidies in innovation activities. Profitability did not seem to impact significantly on firms' innovation investment behaviour in the IS-EAS sample, but was found to be significantly and positively correlated with the probability to invest in innovation in the IS-EAS-BIT-BAS sample. By sub-samples, this coefficient was significant for small and large firm categories but not medium firms. Estimates on the cost or availability of finance as a barrier to innovation investment variable were significant in the innovation intensity equation in both studies. However, the positive sign means that firms spend more on innovation when it faces financial constraints, which may at first glance appear as counter-intuitive, but may make more sense in terms of Schumpeter's business cycle where it is argued that, all things being equal, innovation emerges from difficult times.

Larger firms were more likely to invest in innovation but innovation intensity was unrelated to firm size.

Results for firm size variables (log-level employment in IS-EAS and size dummies in IS-EAS-BIT-BAS) showed strongly that larger firms were more likely to invest in

innovation, but conditional on investment, evidence of a systematic relationship between innovation intensity and size was weak. This finding is similar to that in several CDM-related studies, and is in line with stylized fact three of Cohen and Klepper (1996).

Ownership structure and firm age did not seem to make a difference to firm's innovation investment decisions, but inter-industry differences had an effect.

Only the IS-EAS-BIT-BAS study found a weakly significant and positive association between foreign ownership share of 50 percent or more and firms aged between four and nine years and innovation intensity. Non-profit institutes, which were identified by a dummy variable in the IS-EAS study, did not show significantly different behaviour from the other firms.

Industry dummies at the 2-digit level (IS-EAS) and 1-digit level (IS-EAS-BIT-BAS), which proxied for inter-industry differences in technological opportunities, were jointly significant in both equations. This suggests that firms' innovation investment decisions in both stages are influenced by the industry they are in.

5.2.2 Innovation output equations

Four innovation output equations were estimated in each study: two for product innovation, with different measures of product innovation outcomes (share of innovative sales and a binary indicator), and one equation each for process and organisational /managerial innovation. The difference between the two product innovation output equations is that the innovative sales variable measures the degree of innovation success, while a binary yes/no variable is a cruder measure. The latter was estimated so that the results for product innovation could be more easily compared with those obtained from the process and organisational innovation output equations, where the only output measures available are binary responses. The results from estimations on the full sample are presented in table 5.2.

The main difference between the two studies with regard to innovation output equations is the application of different econometric methods in their respective estimation of the innovative sales equation. However, for the three equations with binary dependent variables, the two studies applied a common estimation method, but differed in the use of some explanatory variables, which is further described in the analysis below.

This section aims to examine the possible impact of innovation intensity, which measures firms' discretionary expenditure on innovation activities, and other factors on the probability or success of various types of innovation output. Other studies have mainly analysed the possible influence of innovation (or R&D) intensity and other factors on product innovation (as measured by the share of innovative sales),

but seldom for process innovation and not for organisational innovation due to unavailable data. The estimation of equations by types of innovation output enabled the two studies in this paper to analyse the innovation input-innovation output relationship from the interesting perspective of comparing similarities and differences in the influence exerted by the same set of factors on three types of innovation output.

Innovation intensity is significantly and positively related to all types of innovation output.

A primary observation from comparing across equations in both studies is that innovation intensity was significantly and positively related to innovation output. The results here support international findings of a significant and positive return on innovation investment for product innovation, and also highlight that the probability of positive outcomes in process and organisational innovations increases with the intensity of spending on innovation activities.

Besides innovation intensity, a firm's technological environment and firm-specific characteristics are thought to have a direct impact on innovation outcomes. There is considerable similarity in the significance and direction of association between any particular factor and the different types of innovation output, although some differences are also observed.

Factors affecting the degree of success in product innovation and the probability of a positive outcome were broadly similar, but there were differences.

Firstly, with regard to the two product innovation equations, factors that were significant in raising the probability of product innovation as well as the expected share of innovative sales in both studies were: internal information, vertical information (from suppliers and clients), strategic protection of intellectual property, intramural R&D and recruitment of workers with scientific, engineering and IT skills. In addition, inter-industry differences were significant in affecting both the share of innovative sales and the probability of a positive innovation outcome. Several factors appeared to have a significant impact on either a positive outcome or share of innovative sales only. This was particularly so in the IS-EAS study where collaboration with universities, government agencies and non-profit research institutes, diversity in the types of collaboration, formal methods of intellectual property protection and firm size were significantly and positively linked to the introduction of product innovation, while older firms and non-profit institutes seemed to be associated with a lower share of innovative sales. The product innovation equations in the IS-EAS-BIT-BAS study had a few more estimates that were significant but they were mainly negative, which implied that variables such as information from universities and other information sources were negatively associated with product innovation success. As mentioned

above, the econometric method applied in the IS-EAS-BIT-BAS study to the innovative sales equation might not be most suited to a fractional dependent variable and coefficient estimates should be interpreted bearing this in mind.

Internal information, diversity in types of collaboration, strategic protection of IP and firm size were significant in raising the probability of introducing any innovation output type; information spillover from competitors was also important for small firms.

Turning to the comparison of results by types of innovation output measured as binary responses, it was observed that some factors besides innovation intensity were significant in raising the probability of positive innovation outcomes across all types of innovation output in both studies. One of these was internal information. The IS-EAS-BIT-BAS study found in addition that the information from competitors variable was significantly and positively associated with all types of innovation output. Further investigation of the results from sub-sample regressions suggested that this relationship was dominant among small firms only. It would seem reasonable that small businesses with less internal resources would find the information spillovers from their competitors particularly useful.

The IS-EAS study found that diversity in the types of collaboration – which was not included as an explanatory variable in the IS-EAS-BIT-BAS study – was also significant and positive. On the other hand, coefficient estimates on the collaboration intensity variable that measured the frequency of collaborations were insignificant across all three equations. This finding on the relative effects of collaboration diversity versus intensity supports the observation in a study by DITR (2006) that the diversity of collaboration appears to be much more important than the intensity of collaboration to achieving positive innovation outcomes (in DITR's case it's the degree of novelty of the innovation outcomes).

Strategic protection of intellectual property – whether it is in place (IS-EAS-BIT-BAS study) and the diversity of methods (IS-EAS study) – was shown to be a significant factor across all types of innovation output. This contrasted with formal protection of intellectual property, which seemed to have a significant impact on the introduction of product innovation only.

Firm size was significantly and positively related to the introduction of all types of innovation output.

Some factors did not affect all innovation output types uniformly; external information sources were important for process and organisational innovations.

The influence of some other factors was not uniform across all three types of innovation output. It seemed that the successful introduction of process and

organisational innovations were reliant on a wider mix of external knowledge flows in addition to internal information sources compared with product innovation, particularly in the IS-EAS study. These included information from commercial sources (consultants and commercial R&D enterprises) and free sources (professional conferences, websites and journals).

Intramural R&D and the employment of technical skills were important to product and process innovations, and to organisational innovation by smaller firms.

Intramural R&D, which proxied for firms' capacity to absorb scientific knowledge, and the recruitment of workers with scientific, engineering and IT skills had a significantly positive correlation with product and process innovation in both studies, while their links with organisational innovation were positive but insignificant in the IS-EAS study and significant in the IS-EAS-BIT-BAS study. Sub-sample results showed that the significant link with organisational innovation seemed to be dominant among small and medium firms (for intramural R&D) and small firms (for recruitment of workers with scientific skills). It is likely that in smaller firms, the same team of people would be involved in both technological and non-technological innovation activities, while large businesses might have a separate team with more generic management skills rather than scientific and engineering skills working on organisational innovation activities.

Ownership structure was insignificant; the impact of firm age was mixed; inter-industry differences mattered for product and process innovations but not organisational innovation.

The share of foreign ownership was found to be an insignificant factor in all types of innovation output in both studies. The evidence on firm age was mixed. Firms between the age of 1 and 9 years seemed to be associated with a lower probability of introducing process and organisational innovations in the IS-EAS study, but estimates on firm age dummies in the IE-EAS-BIT-BAS study were all insignificant.

Estimate on the non-profit institute dummy included only in the IS-EAS study indicated that non-profit institutes did not differ significantly from commercial firms on their probability of introducing any type of innovation.

Industry dummies were jointly significant in the product and process innovation output equations, especially in the former, but insignificant in the organisational innovation output equation in both studies. This difference is not surprising as organisational innovations would tend to be more generic in nature and less industry-specific.

5.2(a) Results from estimation of innovation output equations, IS-EAS[1,2]

	Dependent variables			
Explanatory variables	*Innovative sales (%)*	*Product innovation (yes/no)*	*Process innovation (yes/no)*	*Organisational innovation (yes/no)*
Innovation intensity	+++	+++	+++	+++
Internal information	++	+++	+++	+++
Vertical information	++	+++	.	++
Information from competitors	.	.	.	.
Commercial info	.	.	++	+++
Information from universities	.	.	.	.
Information from government	.	.	.	.
Free information			+++	+++
Other information source				
Collaboration with wider enterprise group	.	.	++	.
Collaboration with suppliers and clients	.	+++	+++	.
Collaboration with competitors	–	.	.	.
Collaboration with commercial sector researchers	.	.	.	.
Collaboration with universities, govt and non-profit institutes	.	.	.	.
Collaboration with other types of organisations	.	.	.	.
Collaboration diversity	.	+	++	+++
Collaboration intensity	.	.	.	.
Barrier: govt regulation or standards (to product innovation)	.	.	.	.
Barrier: govt regulation or standards (to process and org innov)	.	.	.	.
Barrier -govt regulation (to product OR process and org innov)				
Formal protection of IP (diversity)	.	++	.	.
Formal protection of IP (intensity)	.	.	.	.
Formal protection of IP (0/1 dummy)				
Strategic protection of IP (diversity)	+++	+++	+++	+++
Strategic protection of IP (0/1 dummy)				
Other firm characteristics	.	.	.	.
Barrier: lack of skilled staff (for product innovation)	.	.	.	.
Barrier: lack of skilled staff (for process and organisational innov)	.	.	++	.
Barrier: lack of skilled staff (for product, process or org innov)	.	.	.	.
Intramural R&D	+++	+++	+++	.
Recruit workers with scientific, engineering or IT skills (for product innovation)	+++	+++	.	.
Recruit workers with scientific, engineering or IT skills (for process and organisational innovation)		.	+++	.
Recruit workers with engineering, scientific or IT skills (for product, process or organisational innovation)				
Share of ICT employees in total employment	.	.	.	.
Firm size (log-level employment)	.	+++	+++	+++
20 to 199 employees				
200 or more employees				
Foreign ownership (>0% and <50%)	.	.	.	.
Foreign ownership (50% or more)	.	.	.	.
Age (1 to <4 years)	––	.	–	–
Age (4 to <9 years)	––	.	.	––
Age (9 years or more)	–––	.	.	.
Non-profit institute dummy	–	.	.	.
2-digit industry dummies[3]	0.000	0.000	0.009	0.489
1-digit industry dummies[3]				
Number of observations	1,673	1,673	1,673	1,673
Pseudo R-squared	0.188	0.337	0.298	0.239

+, ++, and +++ indicate positively significant at 10%, 5% and 1% respectively; –, –– and ––– indicate negatively significant at 10%, 5% and 1% respectively.

1 Where the cell is blank, the variable was not used in the modelling in that particular study.

2 "." indicates that a variable has returned an insignificant coefficient.

3 p-value for joint significance of industry dummies.

5.2(b) Results from estimation of innovation output equations, IS-EAS-BIT-BAS[1,2]

	Dependent variables			
Explanatory variables	*Innovative sales (%)*	*Product innovation (yes/no)*	*Process innovation (yes/no)*	*Organisational innovation (yes/no)*
Innovation intensity	+++	+++	+++	+++
Internal information	+++	+++	+++	+++
Vertical information	+++	+++	++	.
Information from competitors	.	++	++	+++
Commercial information	.	.	+++	+++
Information from universities	−−−	−−	.	.
Information from government	.	.	−	.
Free information	+++	+++	+++	+++
Other information source	−−	.	.	.
Collaboration with wider enterprise group	.	.	++	++
Collaboration with suppliers and clients	+++	+++	+++	+++
Collaboration with competitors	.	.	.	.
Collaboration with commercial sector researchers	.	.	.	+
Collaboration with universities, govt and non-profit institutes	.	.	.	.
Collaboration with other types of organisations	.	.	.	.
Collaboration diversity				
Collaboration intensity				
Barrier: govt regulation or standards (to product innovation)				
Barrier: govt regulation or standards (to process and org innov)				
Barrier -govt regulation (to product OR process and org innov)	.	.	.	.
Formal protection of IP (diversity)				
Formal protection of IP (intensity)				
Formal protection of IP (0/1 dummy)	+++	+++	.	.
Strategic protection of IP (diversity)				
Strategic protection of IP (0/1 dummy)	+++	+++	+++	+++
Other firm characteristics	.	.	.	.
Barrier: lack of skilled staff (for product innovation)				
Barrier: lack of skilled staff (for process and org innovation)				
Barrier: lack of skilled staff (for product, process or org innov)	+	.	+++	+++
Intramural R&D	+++	+++	+++	+++
Recruit workers with scientific, engineering or IT skills (for product innovation)				
Recruit workers with scientific, engineering or IT skills (for process and organisational innovation)				
Recruit workers with engineering, scientific or IT skills (for product, process or organisational innovation)	+++	+++	+++	++
Share of ICT employees in total employment				
Firm size (log-level employment)				
20 to 199 employees	.	.	++	+++
200 or more employees		+++	+++	+++
Foreign ownership (>0% and <50%)	.	.	.	.
Foreign ownership (50% or more)	.	.	.	.
Age (1 to <4 years)	.	.	.	.
Age (4 to <9 years)	.	.	.	.
Age (9 years or more)	.	.	.	.
Non-profit institute dummy				
2-digit industry dummies[3]				
1-digit industry dummies[3]	0.000	0.000	0.007	0.114
Number of observations	3,591	3,591	3,591	3,591
Pseudo R-squared		0.480	0.440	0.360

+, ++, and +++ indicate positively significant at 10%, 5% and 1% respectively; −, −− and −−− indicate negatively significant at 10%, 5% and 1% respectively.

1 Where the cell is blank, the variable was not used in the modelling in that particular study.

2 "." indicates that a variable has returned an insignificant coefficient.

3 p-value for joint significance of industry dummies.

5.2.3 Business performance equations

The two studies estimated the relationship between innovation output variables and business performance using various performance measures. These measures differ across the two studies, with the exception of a labour productivity growth rate annualised between 2001–02 and 2004–05, where labour productivity is measured as nominal value added per employee (henceforth referred to as LP growth). The results of estimations on LP growth are the most comparable between the two studies, even though they are still not exactly so because slightly different sets of explanatory variables were used. Both studies also estimated productivity equations using manufacturing and services sub-samples.[4] The IS-EAS study estimated both growth and level equations, using gross-output and value-added based MFP and LP measures, and also ran regressions that distinguished innovation output by their degree of innovation novelty. In addition to using LP measured as value added per employee as the dependent variable, the IS-EAS-BIT-BAS study also ran growth regressions using value-added and an alternative measure of labour productivity (value-added per dollar of wages) as the dependent variable.

The analysis in this section is based on results presented in three tables. Table 5.3 reports the results of estimations on LP growth on the full sample in both studies and by firm size categories in the IS-EAS-BIT-BAS study, so as to identify differences in results between the two studies that may arise from firm size variations. Table 5.4 presents the results of estimations on LP growth for the manufacturing and services sub-samples. Table 5.5 reports results from selected equations of interest estimated on IS-EAS data as they are similar to the specifications used in several international studies. These are regressions using growth and level measures of MFP and a level measure of LP as the performance indicators, and innovation output distinguished by their degree of innovation novelty as the explanatory variables.

5.2.3.1 Results from labour productivity growth regressions and differences by firm size

Labour productivity growth and innovation output showed evidence of a positive relationship, although statistically weak and only with respect to product innovation (across firm size) and process innovations (particularly for medium-sized firms); organisational innovation was insignificant.

4 The IS-EAS study had a broader coverage of the service industries that excluded the mining division only, while the IS-EAS-BIT-BAS study excluded this and the construction and the electricity, gas and water supply divisions.

5.3 Results from estimation of labour productivity growth equations (full sample and by size classification)[1,2]

	Dependent variable: Annualised labour productivity growth (2001–02 to 2004–05)[3]				
	Linked IS-EAS data	*Linked IS-EAS-BIT-BAS data*			
Explanatory variables	*Full sample*	*Full sample*	*Small firms*	*Medium firms*	*Large firms*
Share of innovative sales	+	++	.	.	.
Process innovation	.	+	.	++	.
Organisational innovation	.	.	.	.	.
Log change in capital intensity	++	++	+++	+++	+++
Log change in employment	−−−	.	.	.	.
Relative LP	−−−				
Share of ICT employees	.				
Initial market share (log)	+				
Initial market share	.	−−	−−−	.	−−−
Herfindahl index	.	.	+	.	
Log change in industry demand	+++	.	.	.	+++
Firm size: (log-level employment)	.				
Firm size dummy: (20–199 employees)		+++			
Firm size dummy: (> 200 employees)		+++			
Foreign ownership (>0% and <50%)	.	+	.	.	.
Foreign ownership (50% or more)	.	+	+++	.	.
Age (1 to <4 years)	.	.	.	.	.
Age (4 to <9 years)	.	.	.	.	.
Age (9 years or more)	.	.	.	.	.
Non-profit institute dummy	−−−				
2-digit industry dummies[4]	0.000				
1-digit industry dummies[4]		0.000	0.123	0.000	0.000
Observations	1,426	3,591	1,329	1,233	1,029
Adjusted R-squared	0.413	0.181	0.135	0.125	0.428

+, ++, and +++ indicate positively significant at 10%, 5% and 1% respectively; −, −− and −−− indicate negatively significant at 10%, 5% and 1% respectively.

1 Where the cell is blank, the variable was not used in the modelling in that particular study.

2 "." indicates that a variable has returned an insignificant coefficient.

3 Annualised log change between 2001–02 and 2004–05

4 p-value for joint significance of industry dummies.

Table 5.3 shows that both studies found some evidence of a statistically positive relationship between innovation output and LP growth, but the association was relatively weak (between 5 and 10 percent level of significance) and applied only to product and process innovations. The binary variable of organisational innovation was consistently insignificant.

The coefficient estimate on product innovation, as measured by the share of innovative sales, was significant and positive in regressions on the full sample in both

studies. Sub-sample estimations by firm-size classification using IS-EAS-BIT-BAS data yielded insignificant estimates across the board. This indicated that the positive relationship between product innovation and LP growth prevailed across businesses of different size, and was not concentrated in any particular firm size. Share of innovative sales was preferred as a proxy for product innovation over the alternative of a binary indicator because it measured the degree of success. The IS-EAS study found that the use of a binary variable consistently yielded insignificant estimates across specifications. This suggested that a more refined measure might be more effective in capturing the association between innovation output and productivity growth.

Process innovation, as measured by a binary indicator, was weakly significant and positive in the full sample estimation using IS-EAS-BIT-BAS data, but not on the IS-EAS data. Sub-sample estimates showed that there was a significantly positive correlation between process innovation and LP growth only among medium firms but not for small or large firms. The insignificant estimate in the large firm sub-sample corroborated the finding from IS-EAS data, and suggested that the association between process innovation and LP growth did not appear to be strong among large firms.

Of the other variables, capital intensity was strongly correlated with labour productivity growth while measures of market competition and inter-industry differences showed varying association.

Turning to the other explanatory variables, growth in capital intensity was strongly significant and positively related to LP growth in both studies. This is not surprising given that increased capital intensity is an important component of LP growth.

Results on the variables measuring the state of market competition, namely, firms' market share and industry concentration, are mixed. The coefficient estimate on the market share variable was weakly significant and positive in the IS-EAS study but strongly significant and negative in the full sample, as well as in the small-firm and large-firm sub-sample estimations, in the IS-EAS-BIT-BAS study. This difference between the two studies could arise because of the different measures used for the market share variable. The IS-EAS study defined a log-share variable, while the IS-EAS-BIT-BAS study defined a level-share variable. Log transformation of a variable would make the estimate less sensitive to the influence of outliers, and the difference in results might reflect this. The Herfindahl index was largely insignificant. Only the estimate in the small-firm sub-sample was weakly positive.

Industry dummies were jointly significant in the IS-EAS study and in the full and sub-samples of the IS-EAS-BIT-BAS study except for the small firm sub-sample. This suggested that the average productivity growth rates of small firms were less likely to exhibit inter-industry differences compared with large and medium firms.

Non-constant returns to scale and imperfect competition seemed to be valid assumptions for large firms.

The LP equations in both studies included variables that controlled for possible non-constant returns to scale and imperfect competition (refer to Appendix B for the derivation of these proxies). Results from the IS-EAS study appeared to support the inclusion of these controls, while results from the IS-EAS-BIT-BAS study implied that these controls might be more valid in a predominantly large-firm sample. Estimation on IS-EAS data showed that the growth in employment variable was strongly significant, which implied the existence of non-constant returns to scale and/or imperfect competition. The coefficient estimate on industry demand growth, which controlled for firm-specific price differences, was also strongly significant.

Regressions were run in the IS-EAS study assuming constant returns to scale and perfect competition, that is, excluding these variables, as a robustness check. It was found that such specifications did not affect the results of the innovation output variables and other control variables in terms of their sign and significance. However, the explanatory power of the regressions, as indicated by the value of the adjusted R-squared, was lowered, which supported estimations that accounted for non-constant returns to scale and imperfect competition using IS-EAS data. The IS-EAS-BIT-BAS study found that the employment growth variable was insignificant in both full and sub-sample estimations, and that industry demand growth was strongly significant only in the large-firm sub-sample.

The significance of initial firm size and foreign ownership differed across datasets.

Initial firm size as measured by log-level employment did not have a significant correlation with LP growth using IS-EAS data. However, size classification dummies included in the full sample estimation on IS-EAS-BIT-BAS data showed that medium and large firms had higher average LP growth compared with small firms.

Coefficient estimates on the share of foreign ownership structure dummy were insignificant in the IS-EAS study, but weakly significant and positive in the IS-EAS-BIT-BAS study. Sub-sample regressions indicated that small firms that were majority foreign-owned were particularly associated with higher LP growth, as the coefficient estimates on foreign ownership dummies were insignificant in the medium and large firm categories. All firm age dummies were insignificant in both studies.

Firms with lower initial productivity was associated with faster labour productivity growth in the IS-EAS study, which pointed to learning and catching-up.

The productivity equation in the IS-EAS study included a few explanatory variables that were not in the IS-EAS-BIT-BAS study. These were: relative LP, which measured a firm's initial productivity level relative to the frontier; share of ICT employment, which

proxied for a firm's human capital; and a dummy variable that controlled for a possible difference in the average productivity growth rate of non-profit institutes. Coefficient estimate on the relative LP variable was negative and significant. This suggested that firms with lower initial productivity would catch-up and grow faster. The share of ICT employment was insignificant, while there was evidence that non-profit institutes were associated with lower LP growth.

5.2.3.2 Results from labour productivity growth regressions by manufacturing and services sub-samples

The first thing to note is that there is considerable divergence in the results obtained from the two datasets (table 5.4). Thus, it is not possible to generalise findings for the manufacturing or service industries based on both sets of results. The differences are not surprising. Although the IS-EAS-BIT-BAS study did not carry out manufacturing and service industry estimations by firm size categories, some of the differences with the IS-EAS study could be due to the dominant effects of small and medium firms, as can be seen from the earlier analysis. The two studies also adopted a different coverage of services industries in their services sample. The two manufacturing samples covered the same industries, but there were still variations in the set of regressors used.

The association between labour productivity growth and innovation output did not seem significant for manufacturing; for labour productivity growth versus other explanators, most of the significant relationships were unique to each study.

Results for manufacturing from both studies showed that all types of innovation output did not have a significant association with LP growth.

Of the variables that were significant, most of them were unique to each study. In the IS-EAS study, these were relative LP, non-profit institute dummy, and industry dummies at the 2-digit level. The IS-EAS-BIT-BAS study had significant firm-size dummies. Of the variables that were common to both studies, only the industry demand growth variable that controlled for changes in relative output prices among firms was significant in both. Even in that case, the level of significance differed. In the IS-EAS study, industry demand growth was strongly significant. This, coupled with a strongly significant employment growth variable, seemed to imply that the manufacturing sub-sample of the IS-EAS data supported the assumptions of non-constant returns to scale and imperfect competition more strongly than IS-EAS-BIT-BAS data. The IS-EAS manufacturing sub-sample also showed a significantly positive relationship between initial market share and LP growth.

5.4 Results from estimation of labour productivity growth equations (manufacturing and services sub-samples)[1,2]

	Dependent variable: Annualised labour productivity growth (2001–02 to 2004–05)[3]			
	Linked IS-EAS data		*Linked IS-EAS-BIT-BAS data*	
Explanatory variables	*Manufacturing*	*Services*	*Manufacturing*	*Services*
Share of innovative sales	.	.	.	+++
Process innovation	.	.	.	+
Organisational innovation	.	.	.	.
Log change in capital intensity	.	+++	+++	+++
Log change in employment	−−−	−−−	.	.
Relative LP	−−−	−−−		
Share of ICT employees	.	.	.	.
Initial market share (log)	++	.		
Initial market share			.	−−−
Herfindahl index	.	.	.	.
Log change in industry demand	+++	.	+++	.
Firm size: (log-level employment)	.	.		
Firm size dummy: (20–199 employees)			++	++
Firm size dummy: (> 200 employees)			+	+++
Foreign ownership (>0% and <50%)	.	.	.	+
Foreign ownership (50% or more)	.	.	.	+++
Age (1 to <4 years)	.	.	.	+
Age (4 to <9 years)	.	.	.	.
Age (9 years or more)	.	.	.	.
Non-profit institute dummy	−	−−−		
2-digit industry dummies[4]	0.001	0.000		
1-digit industry dummies[4]				0.011
Observations	721	641	1,643	1,569
Adjusted R-squared	0.171	0.553	0.050	0.301

+, ++, and +++ indicate positively significant at 10%, 5% and 1% respectively; −, −− and −−− indicate negatively significant at 10%, 5% and 1% respectively.

1 Where the cell is blank, the variable was not used in the modelling in that particular study.

2 "." indicates that a variable has returned an insignificant coefficient.

3 Annualised log change between 2001–02 and 2004–05

4 p-value for joint significance of industry dummies.

The coefficient estimate on growth in capital intensity was insignificant in the IS-EAS study but strongly significant in the IS-EAS-BIT-BAS study. The former result was rather puzzling as growth in capital intensity was unlikely to be an insignificant contributor to LP growth. It should be noted that the two studies constructed different proxies for capital stock. However, it is unlikely that the insignificant estimate on capital intensity growth yielded by the IS-EAS manufacturing sample was due to capital measurement issue, as this variable was strongly significant in the full sample and services sub-sample estimations. It is also noted that the explanatory

power, as measured by the adjusted R-squared value, was particularly poor in the IS-EAS-BIT-BAS manufacturing sample estimation. Overall, these results were probably indicative of some data anomalies that particular affected results in a smaller manufacturing sub-sample compared with the full sample.

There was evidence of a significant and positive relationship between labour productivity growth and innovation output in services in the IS-EAS-BIT-BAS study; capital intensity, foreign ownership and firm age were among the significant variables.

Services sub-sample estimations showed that innovation output variables all had insignificant correlations with LP growth in the IS-EAS study, but product innovation and process innovation coefficients were significant and positive in the IS-EAS-BIT-BAS study.

Of the other variables that were common to both studies, the coefficient estimate on growth in capital intensity was strongly significant and positive in both.

Results from the IS-EAS-BIT-BAS study also suggested that businesses with a higher share of foreign ownership and firm age between one and four years were linked to higher LP growth. These were not evident in the IS-EAS study. The IS-EAS-BIT-BAS study also found that the initial market share variable was significant and negative, while the estimate on IS-EAS data was insignificant. The industry demand growth variable was insignificant in both studies, while the growth in employment variable was significant only in the IS-EAS study. Variables that were particular to each study were mostly significant. In the IS-EAS services sample, these were relative LP, non-profit institute dummy and 2-digit industry dummies. In the IS-EAS-BIT-BAS sample, these were firm size dummies and 1-digit industry dummies.

Differences between manufacturing and the service indutries were indicated; industry demand growth was significant for manufacturing, but not services.

Comparing the results between the manufacturing and services sub-samples within each study and across both studies, it can be seen that there are differences between the two industry groups in terms of the factors that were significantly linked to LP growth, although this was more obvious in the IS-EAS-BIT-BAS study. An interesting observation that applied to both studies is that the industry demand growth variable, which controlled for firms' ability to charge different prices, was significant in the manufacturing sub-sample but not in the services sub-sample. This could indicate that service industries were more competitive, but could also reflect that the variable had captured differences in the manufacturing and service industries that arose from other unobserved factors.

The innovation output variables were all insignificant in the IS-EAS industry sub-samples but product innovation was significantly positive in the full sample, while product and process innovations were significantly positive in the IS-EAS-BIT-BAS services sub-sample as well as the full sample. This could imply that the significant relationship between product innovation and productivity growth in the full sample regression on IS-EAS data was not clearly delineated along the lines of manufacturing and services businesses, while the association might have been stronger for services firms in the IS-EAS-BIT-BAS dataset. In fact, it is noted that overall the results of the IS-EAS-BIT-BAS services sub-sample estimation were quite similar to the full sample results.

5.2.3.3 Additional results from IS-EAS study

MFP growth equation

The first column in table 5.5 gives the results of regression on gross-output based MFP. The results will be compared with those obtained from the LP growth equation (in table 5.3), in order to identify possible differences in the linkage between the explanatory variables and the different productivity measures, bearing in mind that MFP growth measures productivity changes due to factors other than factor (capital and intermediate input) intensities.

Product innovation, as represented by the share of innovative sales, appeared to have a greater positive effect on MFP growth than LP growth.

Of the innovation output variables, only the share of innovative sales (product innovation) contributed positively to MFP growth, which was same as the finding on LP growth. Further examination showed that the partial effect of successful product innovation on MFP growth appeared to be greater than on LP growth. This is not surprising if innovation output variables are considered to be capturing the impact of 'disembodied' knowledge growth (that is, not embodied in physical and human capital).

A few factors such as skills and industry concentration seemed to have stronger links with MFP growth rather than LP growth.

The share of ICT employment and Herfindahl index had a weakly significant and positive association with MFP growth but were insignificant in the LP growth estimation. Initial market share showed a strongly significant and negative correlation with MFP growth. This was unlike its weakly significant and positive link with LP growth.

5.5 Additional results from estimation of productivity growth and level equations, IS-EAS data[1,2]

	Growth[3]	*Level*[4]		*Innovation novelty*
Explanatory variables	*MFP[-op]*	*MFP[-op]*	*LP*	*LP growth*
Share of innovative sales	+	.	.	
Product innovation (new to business/ industry/Aus)				.
Product innovation (new to world)				.
Process innovation (new to business/ industry/Aus)				.
Process innovation (new to world)				+
Organisational innovation	.	++	.	.
Growth in capital stock	–			.
Growth in capital intensity				+++
ln(capital level)		–––		.
Capital intensity level			+++	.
Growth in employment				–––
Relative MFP (or LP)	–––			–––
Share of ICT employees	+	+	+++	.
Initial ln(market share)	–––	––	+++	+
Herfindahl index	+	.	+++	.
Industry demand growth	.			+++
ln(Industry demand level)		–	–––	.
Firm size: ln(employment)	+++	.	–––	.
Foreign ownership (>0% and <50%)	.	.	.	.
Foreign ownership (50% or more)	.	+	.	.
Age (1 to <4 years)	.	–	–	.
Age (4 to <9 years)	.	–	–	.
Age (9 years or more)	.	–	–	.
Non-profit institute dummy	–––	.	.	–––
2-digit industry dummies[5]	0.000	0.000	0.000	0.000
Observations	1,426	1,413	1,416	1,426
Adjusted R-squared	0.248	0.540	0.738	0.414

+, ++, and +++ indicate positively significant at 10%, 5% and 1% respectively; –, –– and ––– indicate negatively significant at 10%, 5% and 1% respectively.

1 Where the cell is blank, the variable was not used in the model.

2 "." indicates that a variable has returned an insignificant coefficient.

3 Annualised log change between 2001–02 and 2004–05.

4 Productivity (log) level in 2003–04.

5 p-value for joint significance of industry dummies.

Initial firm size was strongly significant and positively correlated with MFP growth, although it appeared to be insignificantly linked to LP growth. It should be noted that a firm size (log-level employment) variable is a control for scale in the derivation of a productivity level equation. In a productivity growth equation, a firm size control is not derived from production theory. However, the IS-EAS study found that the inclusion of a firm size control improved the explanatory power of regressions across specifications in terms of the value of adjusted R-squared. It is possible that the initial

firm size variable captures the effects of unobserved factors such as managerial abilities on MFP growth.

The coefficient on industry demand growth was insignificant in the MFP growth equation, unlike in the LP growth equation. However, this was probably an anomaly as this variable was significant in most other specifications using different productivity measures and variations in the set of regressors.

The other results were similar to those obtained in the estimation on LP growth.

MFP and LP level equations

Organisational innovation showed significant and positive correlation with MFP level.

The third and fourth columns in table 5.5 show that all innovation output variables were insignificant in the LP level estimation, but organisational innovation showed a significant and positive correlation with MFP level in 2003–04. This result applied regardless of whether the LP or MFP measure was gross-output or value-added based. It again indicated that innovation output variables might have a stronger effect on MFP performance rather than labour productivity, which is plausible, since labour productivity is expected to reflect a significantly positive association with capital intensity and it does. The significant relationship between organisational innovation and productivity level may indicate that the bulk of organisational innovations are incremental improvements that produce relatively quick results. Alternatively, it could simply mean that firms that implement organisational innovations are those with higher MFP level. Unlike the productivity growth equations, the innovative sales variable was found to be insignificant across specifications.

Share of ICT employment and productivity levels are positively correlated.

There was firmer evidence of a significant and positive relationship between the share of ICT employment and productivity level compared with productivity growth. The coefficient estimate on the market share variable was significant but, as found in the MFP and LP growth regressions on IS-EAS data, it was negatively correlated with MFP level and positively correlated with LP level.

The coefficient on industry concentration as measured by the Herfindahl index was significant and positive in the LP equation but insignificant in the MFP equation.

Assumptions of non-constant returns to scale and imperfect competition were supported by level estimation results

As with the productivity growth estimation results in the IS-EAS study, the strongly significant association between the level of capital stock in the MFP equation or level

of employment in the LP equation and level of industry demand seemed to substantiate the assumptions of non-constant returns to scale and imperfect competition for the IS-EAS sample.

Older firms were associated with lower productivity levels

Unlike the productivity growth equations, there was evidence of significant and negative correlations between firm age dummy variables with the level of MFP and LP. This meant that older firms were associated with lower productivity level relative to firms that were less than 1 year old. It might imply that entrants equipped with new technology were on average more productive than incumbents instead of the premise that a firm's productivity rose as it learnt from experience.

The coefficient estimate on majority foreign-owned firms was significant and positive in the MFP equation, but not in the LP equation.

Non-profit institutes did not seem to be significantly different from corporations in terms of productivity level. However, the joint significance of industry dummies pointed to the existence of inter-industry differences in productivity level.

Innovation novelty

The fourth column in table 5.5 presented coefficient estimates for the LP growth equation that differed from the main specification used in table 5.3 only in its measures of product and process innovation. The share of innovative sales variable is replaced by two dummy variables that measure whether the highest degree of innovation novelty of product innovations is 'new to the world' ('novel product innovations') or 'others' (which means new to business, industry or Australia). The binary process innovation variable is replaced by two dummy variables that measure whether the highest degree of innovation novelty of process innovations is 'new to the world' ('novel process innovations') or 'others'.

The association between a high degree of innovation novelty and productivity performance was found to apply to process innovation only

The estimates suggested that among the innovation output variables, there was a significantly positive relationship only between novel process innovation and LP growth. In estimations using other dependent variables such as MFP growth, and MFP and LP levels, it was also found that any significant association between a high degree of innovation novelty and productivity performance applied to process innovation only. It is interesting to note that product innovation, when measured using a binary indicator, did not show a significant association with productivity growth, regardless of the degree of innovation novelty.

The magnitudes, significance and signage of the other variables were little changed from the results in the main specification.

5.2.3.4 Brief summary and comparison with other studies

To summarise, both studies found preliminary evidence of a significant and positive relationship between innovations and business performance, although the results were far from conclusive. The strength of association between the different types of innovation output and business performance depended on the dataset, industry coverage and also the performance measures used. The overall finding is in line with the broad thrust of results in international studies that applied a CDM-type model, although they are not directly comparable. Furthermore, the results in other studies are also not uniform and are sensitive to model specifications, variable definitions and econometric methods.

To give a sense of the diversity, for example, the original CDM (1998) paper found a significantly positive correlation between product innovation (share of innovative sales) and the level of labour productivity among French manufacturing firms in 1990. Lööf and Heshmati (2006) tested the sensitivity of results to changes in specification and variables on Swedish data in the mid-1990s. They showed that among different performance measures, sales was a crude proxy for value added. They also observed homogeneity in the estimated positive relationship between share of innovative sales and both labour productivity level and growth in the manufacturing and service industries. In Criscuolo (2004), which is closest to the model specification applied in the studies here, particularly the IS-EAS study, it was found that novel product innovations, as well as managerial and organisational changes, were significantly correlated with higher MFP growth among UK manufacturing firms between 1998 and 2000. Estimates on process innovation were insignificant. Van Leeuwen and Klomp (2006) found that the impact of the share of innovative sales on firm performance depended on the performance measure used, for Dutch firms between 1994 and 1996. The estimates were insignificant using value-added per employee, but significant and positive using revenue per employee growth.

It should be noted that the point estimates on the innovation output variables in this paper appeared small compared with the estimates in other studies (again not directly comparable). This could indicate a downward bias in the innovation coefficients from using simple least squares estimation and not taking into account possible simultaneity between innovation output and input. For instance, Crépon et al. (1998) and van Leeuwen and Klomp (2006), who tested various econometric methods, found that full system estimators yielded the most robust estimates.

There are a few other points of interest to note when briefly comparing the results in this paper to those in international studies. A feature of the productivity equations

specified here is that they take into account possible non-constant returns to scale and imperfect competition. These assumptions were made and substantiated in Criscuolo (2004) and van Leeuwen and Klomp (2006). This paper found that these assumptions seemed to be more valid for the predominantly large firm sample in the IS-EAS study. The IS-EAS study included a variable that measured a firm's initial MFP/LP relative to the frontier to proxy for a firm's scope for learning and catching up and found, as in Criscuolo (2004), that the estimates were significant and negative. Several international studies (for example, Crépon et al. 1998 and Lööf and Heshmati 2006) found that there was a significant and positive correlation between the share of graduates (engineers and administrators respectively) in total employment and productivity performance. The Australian datasets did not have such measures of workers' skills or quality, and the share of ICT employees was used as a proxy in the IS-EAS study while no proxy for human capital was used in the IS-EAS-BIT-BAS study because that variable was not available in BIT-BAS data. The significant estimates on the skills variable found in other studies suggested that estimations using Australian data would benefit from the inclusion of an adequate measure of workers' skills in future.

6. CONCLUSIONS

This paper has reported on the conduct and findings of two studies of innovation and its effects on the performance of Australian firms.

It was not an objective of this paper (or the studies themselves) to generate definitive quantitative measures of how various factors affect innovation and firm performance. Rather, as outlined in the Introduction, there were four objectives related to the construction and use of linked datasets.

Each of these objectives is now considered, drawing on the material presented in previous chapters.

6.1 Establish a framework suitable for analysis of innovation using firm-level data

An objective of this project was to specify econometric models that would enable rigorous analysis of firm-level innovation and its effects on performance. There have been previous Australian studies of innovation and business performance (see Section 1), but the linking of firm-level data from a range of sources, as part of this project, has opened up the possibility of more elaborate model specification. Since the CDM framework has become an important focus of international analyses of innovation and productivity, it presented itself as the prime analytical vehicle to use in this project. Nevertheless, considerable work was needed to specify suitable models (taking account of available data) and to establish that the CDM framework could be successfully implemented on data from individual Australian firms.

The modelling undertaken for this project has established that the CDM framework is well suited to the quantitative analysis of firm-level innovation and performance on Australian data. The CDM framework presents a structure – decision on innovation and intensity, production of innovation outputs, and effects of innovation outputs on performance – that is empirically supported and provides useful insight into distinct dimensions of the innovation process. For example, modelling within the CDM framework has reinforced the importance of distinguishing between innovation inputs and innovation outputs (and the limitations in using expenditure on R&D as a measure of innovation).

The models specified in this project provide a sound platform for further elaboration. While they have provided useful indicative results in their own right, there are acknowledged limitations in the current analysis that would warrant further investigation in the future. The main areas are: simultaneous system estimation, the possibility of interactions between explanatory variables and specification of lag structures (especially for innovation outputs to affect performance) and feedback effects.

6.2 Assess the likely potential for unit record data, linked from different Australian surveys, to provide policy relevant information

The study provided further clear evidence of the importance of taking a firm-level view and of the value of linked innovation and performance micro-data and the analysis that they support. As discussed in the Introduction, evidence of diversity in behaviour and performance across firms – even within the same industry – implies that the ability to capture and investigate that diversity is extremely valuable, if not indispensable, in developing a thorough understanding of the ways and extent to which policy and other influences affect innovation and performance.

The data illustrate that there is much diversity in the innovation behaviour and performance of firms. Innovation activities vary by firm size and, in the data used, larger firms are more likely to innovate than smaller firms. Innovation also tended to be more in the operational processes than in new products or services. The spread in quartile values of the business performance indicators used in this study also show that there is considerable heterogeneity in the productivity growth and market power of firms, for the groups of innovators and non-innovators alike.

Modelling within the CDM framework has shown that the relationship between innovation and the performance of Australian firms is not simple. Some of the indicated complexities and nuances in the relationship would be of considerable policy interest. For example, there was evidence that government assistance has more effect on the innovation behaviour of small firms than of large firms. The effect of innovation on performance depends on the type of innovation, size of firm and industry of operation. Depending on the characteristics of firms, any government measures that support innovation may be more or less effective.

To some extent, the performance effects associated with innovation did not come through in the results as strongly as might be expected. However, further analysis of simultaneity in relationships and the existence of lagged performance effects is perhaps needed before clearer magnitudes become evident.

6.3 Explore factors affecting innovation and performance among Australian firms

For reasons that have been spelt out, the results of the two studies should be treated as indicative, rather than definitive. There is more model and data development to do to improve the model results. With that caveat, the data and analysis point to the main findings already summarised at the beginning of this paper and in Section 5.

6.4 Draw any lessons for future data collections and linking exercises

The studies highlighted some issues for the ABS to consider in its design of surveys, some of which have already (and independently) been incorporated into subsequent collections. These included further tightening the sequencing of innovators and non-innovators within the collection vehicle and improvements to the collection of innovation expenditure data, both issues that have been addressed for the 2005 Innovation Survey. Although not always easily accommodated, where possible gradated responses are useful in sensitising the data for analytical purposes. Any future Australian studies on innovation and performance will greatly benefit from the integration of a range of business surveys into the newly Integrated Business Characteristics Strategy which will add to the range of potential explanatory variables available.

The use of different datasets also provides an opportunity to assess differences in analysis and results between the relatively simple performance measures enabled by BIT-BAS data and the more-precise performance measures that can be constructed from EAS data. The analysis and results of the two studies are not exactly comparable, because data differences lead to some variations in model and variable specification, and there are important differences in samples. Nevertheless, it appears that there is a trade-off in using different sources for data on firm performance. The results indicate better explanation of performance can be obtained when the EAS-based performance data are used. On the other hand, the BIT-BAS data provide information on a more comprehensive range of firms.

ACKNOWLEDGEMENTS

The authors would like to thank both the Australian Bureau of Statistics (ABS) and the Productivity Commission (PC) for allowing Dr Wong to be seconded to the ABS, as part of a collaborative arrangement, to develop the framework for the analysis and to undertake the modelling work on the IS-EAS dataset. The authors would like to acknowledge the contribution of Dean Parham (PC) who helped manage the two complementary studies. Important contributions from the following ABS staff are also acknowledged: Steve Crabb, Marion McEwin, Shiji Zhao, John Blanchette, Glyn Prichard and Don Brunker. The authors are also grateful for the input on econometric methodology from Trevor Breusch (Australian National University), Chiara Criscuolo (London School of Economics) and Ben Dolman (PC).

REFERENCES

Australian Bureau of Statistics (2002) *Australian National Accounts: Non-Profit Institutions Satellite Account, 1999–2000*, cat. no. 5256.0, ABS, Canberra.

—— (2006a) *Innovation in Australian Businesses*, cat. no. 8158.0, ABS, Canberra.

—— (2006b) *Patterns of Innovation in Australian Businesses, 2003*, cat. no. 8163.0, ABS, Canberra.

—— (2006c) *Australian Industry, 2004–05*, cat. no. 8155.0, ABS, Canberra.

—— (2007) *Discussion Paper: The First Iteration of the Business Longitudinal Database*, cat. no. 8164.0, ABS, Canberra.

Benavente, J. (2006) "The Role of Research and Innovation in Promoting Productivity in Chile", *Economics of Innovation and New Technology*, 15(4–5), pp. 301–315.

Bosworth, D. and Loundes, J. (2002) *The Dynamic Performance of Australian Enterprises*, Melbourne Institute Working Paper No. 3/02, Melbourne Institute of Applied Economic and Social Research, University of Melbourne and Manchester School of Management, UMIST.

Cohen, W. and Klepper, S. (1996) "A Reprise of Size and R&D", *Economic Journal*, 106, pp. 925–951.

Crépon, B., Duguet, E. and Mairesse, J. (1998) "Research, Innovation and Productivity: An Econometric Analysis at the Firm Level", *Economics of Innovation and New Technology*, 7, pp. 115–158.

Criscuolo, C. (2004) "Innovation, Knowledge Flows and Productivity Growth: Evidence from the UK Community Innovation Survey", in *Explaining Firms' Heterogeneity in Productivity and Wages, Ownership, Innovation and Size*, PhD Thesis, University of London, unpublished.

Criscuolo, C. and Haskel, J. (2003) *Innovations and Productivity Growth in the UK: Evidence from CIS2 and CIS3*, Discussion Paper, Centre for Research into Business Activity, Office for National Statistics, UK.

DITR (Department of Industry, Tourism and Resources) (2006) *Collaboration and Other Factors Influencing Innovation Novelty in Australian Businesses – An Econometric Analysis*, DITR, Canberra.

Hall, B. and Mairesse, J. (1995) "Exploring the Relationship between R&D and Productivity in French Manufacturing Firms", *Journal of Econometrics*, 65(1), pp. 263–293.

Janz, N. and Peters, B. (2002) *Innovation and Innovation Success in the German Manufacturing Sector: Econometric Evidence at Firm Level*, Paper presented at EARIE 2002.

Klette, T. (1996) "R&D, Scope Economies and Plant Performance", *The Rand Journal of Economics*, 27(3), pp. 502–522.

Klomp, L. and van Leeuwen, G. (2001) "Linking Innovation and Firm Performance: A New Approach', *International Journal of the Economics of Business*, 8(3), pp. 343–364.

Lööf, H. and Heshmati, A. (2006) "On the Relationship between Innovation and Performance: A Sensitivity Analysis", *Economics of Innovation and New Technology*, 15(4–5), pp. 317–344.

Martin, R. (2002) *Building the Capital Stock*, Centre for Research into Business Activity (CeRiBA), mimeo.

Organisation of Economic Cooperation and Development (1992) *Proposed Guidelines for Collecting and Interpreting Technological Innovation Data, (Oslo Manual)*, OECD, Paris.

—— OECD (2001) *Measurement of Aggregate and Industry-Level Productivity Growth: OECD Manual*, OECD, Paris.

—— (2004) *The Economic Impact of ICT: Measurement, Evidence and Implications*, OECD, Paris.

Papke, L. and Wooldridge, J.M. (1996) "Econometric Methods for Fractional Response Variables with an Application to 401(K) Plan Participation Rates", *Journal of Applied Econometrics*, 11(6), pp. 619–32.

Phillips, R. (1997) *Innovation and Firms Performance in Australian Manufacturing*, Industry Commission Staff Research Paper, AGPS, Canberra, September.

Rogers, M. (2000) *Understanding Innovative Firms: An Empirical Analysis of the GAPS*, Melbourne Institute Working Paper No. 08/00, Melbourne Institute of Applied Economic and Social Research, University of Melbourne.

Schumpeter, J.A. (1937) *The Theory of Economic Development*, Harvard University Press, Cambridge, MA.

Schumpeter, J.A. (1942) *Capitalism, Socialism and Democracy*, Harper & Row, New York.

Van Leeuwen, G. and Klomp, L. (2006) "On the Contribution of Innovation to Multi-factor Productivity Growth", *Economics of Innovation and New Technology*, 15(4–5), pp. 367–390.

Webster, E. (2003) *Firms Decisions to Innovate and Innovation Routines*, Melbourne Institute Working Paper No. 5/03, Melbourne Institute of Applied Economic and Social Research, University of Melbourne.

Wooldridge, J.M. (2002) *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge.

APPENDIXES

A. VARIABLE DEFINITIONS

A.1 Innovation investment equation

Variables	*Definition / relevant question in the Innovation Survey questionnaire*	*Range of values*
Dependent variable		
Innovation intensity *	Share of innovation expenditure in total sales (innovation expenditure/sales)	0–1
Explanatory variables		
Information sources	Based on responses to Q15: *Which of the following are the key sources of ideas or information which help this business (to innovate)?*	
Internal information	– Within this business – Other parts of a wider enterprise group to which this business belongs	0/1 dummy
Vertical information	– Clients or customers – Suppliers of equipment, materials, components or software	0/1 dummy
Information from competitors	– Competitors and other businesses from the same industry	0/1 dummy
Commercial information	– Consultants (including paid professional advice of all kinds) – Commercial laboratories /research and development enterprises	0/1 dummy
Information from universities	– Universities of other higher education institutes – Private non-profit research institutions	0/1 dummy
Information from government	– Government agencies	0/1 dummy
Free information	– Professional conferences, meetings, fairs and exhibitions – Websites, journals	0/1 dummy
Other information *	– Other sources of information	
Collaboration	Based on responses to Q14: *Please indicate the types of organisation with which this business collaborated*	
Collaboration with wider enterprise group	– Other parts of wider enterprise group to which this business belongs	0/1 dummy
Collaboration with suppliers and clients	– Suppliers of equipment, materials, components or software – Clients or customers	0/1 dummy
Collaboration with competitors	– Competitors and other businesses from the same industry	0/1 dummy
Collaboration with commercial sector researchers	– Consultants – Commercial laboratories /research and development enterprises	0/1 dummy

Collaboration with universities, government and non-profit research institutes	– Universities or other higher education institutions – Government agencies – Private non-profit research institutes	0/1 dummy
Collaboration with other types of organisations	– Other types of organisations	0/1 dummy
Collaboration diversity **	Based on Q13: *Number of types of collaborations that business was actively engaged in*	(0–6)
Collaboration intensity **	Sum of number of collaborations across all types adjusted by collaboration diversity index	
Industry dummies	1-digit * or 2-digit ** industry dummies	
Regulation	Based on Q21: *Which of the following factors hampered this business (in innovation)? (a) Cost related barriers (iv) Government regulation or standards*	0/1 if ticked (a)(iv)
Drivers of innovation	Based on responses to Q20: *What are the key reasons that drive this business to innovate*	
Drivers: improve productivity or reduce costs **	Profit-related drivers: – Improve productivity – Reduce costs	0/1 dummy
Drivers: improve productivity *	Profit-related drivers: Improve productivity	0/1 dummy
Drivers: reduce costs *	Profit-related drivers: Reduce costs	0/1 dummy
Drivers: increase revenue	Profit-related drivers: – Increase revenue	0/1 dummy
Drivers: increase export opportunities	Market-related drivers: – Increase export opportunities	0/1 dummy
Other market-related drivers **	Market-related drivers: – Be at the cutting edge of industry – Increase responsiveness to customer needs – Increase market share – Establish a new market – Exploit new ways to manage this business's supply chain – High degree of price competition in this business's product markets	0/1 dummy
Drivers: increase market share *	Market-related drivers: Increase market share	0/1 dummy
Other market-related drivers *	Market-related drivers: – Be at the cutting edge of industry – Increase responsiveness to customer needs – Establish a new market – Exploit new ways to manage this business's supply chain – High degree of price competition in this business's product markets	0/1 dummy
(Initial) market share	Firm sales / Industry sales, by 4 digit ANZSIC industry in 2001-02*; Log(Firm sales / Industry sales) by 4 digit ANZSIC industry in 2001–02**	
Herfindahl index	Industry concentration index by 4-digit industry	0–1

Industry sales growth	Annualised growth in log nominal sales by 4 digit industry (2001–02 to 2004–05)	
Formal protection ** (diversity index)	Intellectual property diversity index (number of types of methods used to protect IP)	1–4
Formal Protection ** (intensity index)	Intellectual property intensity index (total number of methods used)	1–n
Strategic protection **	IP diversity index (total number of type of informal method used)	1–4
Strategic IP protection indicator *	Firm is using at least one informal method to protect intellectual property	0/1 dummy
Formal IP protection indicator *	Firm is using at least one formal method to protect intellectual property	0/1 dummy
Lack of skilled workers	Based on Q21: *Which of the following factors hampered this business (in innovation) (c) Lack of skilled staff*	0/1 dummy
Availability of finance	Based on Q21: *Which of the following factors hampered this business (in innovation)? (a) cost related barriers (iii) cost or availability of finance*	0/1 dummy
Government financial support	Based on Q34: *Please estimate the expenditure (total innovation expenditure excluding research and development) by the source of funding (funds from government)* *– Commonwealth* *– State or local*	0/1 dummy
Profit share	Profit / Total sales	
Foreign ownership (between 0% and 50%)	Foreign ownership is between 0% and 50%	0/1 dummy
Foreign ownership (more than 50%)	Foreign ownership is more than 50%	0/1 dummy
Firm size **	Log (employment)	
Employment * (20 to 199 employees)	Number of employees is between 20 and 199	0/1 dummy
Employment * (more than 200 employees)	Number of employees is greater than 199	0/1 dummy
Age (between 1 and 4 years)	The firm is between 1 year and 4 years old	0/1 dummy
Age (between 4 and 9 years)	The firm is greater than 4 years old, but less than 9 years old	0/1 dummy
Age (>=9 years)	The firm is 9 years old or older	0/1 dummy
Non-profit institute dummy **	Non-profit institute indicator	0/1 dummy

Note:

* used in modelling the linked IS-EAS-BIT-BAS data;

** used in modelling the linked IS-EAS data only;

If no asterisks, variable was used in both studies.

A.2 Innovation output equation(s)

Variables	*Definition*	*Range of values*
Dependent variable		
(1a) Product innovation	Percentage sales due to new and improved goods and services. Q8(a): *Please estimate how the turnover of this business in calendar year 2003 was attributed new goods or services introduced during the past three calendar years*	Share: 0–1
(1b) Product innovation	Based on Q5(a): *Did business introduce new goods and services in the past three calendar years?*	0/1 dummy
(2) Process innovation	Based on Q9(a): *Has this business implemented any new or significantly improved operational processes during the past three calendar years?*	0/1 dummy
(3) Organisational innovation	Based on Q12(a): *Has this business implemented any new or significantly improved organisational/managerial processes during the past three calendar years?*	0/1 dummy
Explanatory variables		
Information sources	Based on responses to Q15: *Which of the following are the key sources of ideas or information which help this business (to innovate)?*	
Internal information	– Within this business – Other parts of a wider enterprise group to which this business belongs	0/1 dummy
Vertical information	– Clients or customers – Suppliers of equipment, materials, components or software	0/1 dummy
Information from competitors	– Competitors and other businesses from the same industry	0/1 dummy
Commercial information	– Consultants (including paid professional advice of all kinds) – Commercial laboratories/research and development enterprises	0/1 dummy
Information from universities	– Universities of other higher education institutes – Private non-profit research institutions	0/1 dummy
Information from government	– Government agencies	0/1 dummy
Free information	– Professional conferences, meetings, fairs and exhibitions – Websites, journals	0/1 dummy
Other Information *	Other sources of ideas or information	
Collaboration	Based on responses to Q14: *Please indicate the types of organisation with which this business collaborated*	

Collaboration with wider enterprise group	– Other parts of wider enterprise group to which this business belongs	0/1 dummy
Collaboration with suppliers and clients	– Suppliers of equipment, materials, components or software – Clients or customers	0/1 dummy
Collaboration with competitors	– Competitors and other businesses from the same industry	0/1 dummy
Collaboration with commercial sector researchers	– Consultants – Commercial laboratories/research and development enterprises	0/1 dummy
Collaboration with universities, government and non-profit research institutes	– Universities or other higher education institutions – Government agencies – Private non-profit research institutes	0/1 dummy
Collaboration with other types of organisations	– Other types of organisations	0/1 dummy
Collaboration diversity **	Based on Q13: *Number of types of collaborations that business was actively engaged in*	(0–6)
Collaboration intensity **	Sum of number of collaborations across all types adjusted by collaboration diversity index	
Industry dummies **	1-digit* or 2-digit** industry dummies	
Regulation ** (goods/service)	Based on Q21: *Which of the following factors hampered this business (in innovation)? (a) Cost related barriers (iv) Government regulation or standards*	0/1 dummy
Regulation ** (process)	Based on Q21: *Which of the following factors hampered this business (in innovation)? (a) Cost related barriers (iv) Government regulation or standards*	0/1 dummy
Regulation * (process/goods/services)	Based on Q21: *Which of the following factors hampered this business (in innovation)? (a) Cost related barriers (iv) Government regulation or standards*	0/1 dummy
Formal protection ** (diversity index)	Intellectual property diversity index (number of types of methods used to protect IP)	1–4
Formal protection ** (intensity index)	Intellectual property intensity index (total number of methods used)	1–n
Strategic protection **	IP diversity index (total number of type of informal method used)	1–4
Strategic IP protection indicator *	Firm is using at least one informal method to protect intellectual property	0/1 dummy
Formal IP protection indicator *	Firm is using at least one formal method to protect intellectual property	0/1 dummy
Skilled workers ** (goods/services)	Based on Q21: *Which of the following factors hampered this business (in innovation) (c) Lack of skilled staff*	0/1 dummy
Skilled workers ** (process)	Based on Q21: *Which of the following factors hampered this business (in innovation) (c) Lack of skilled staff*	0/1 dummy

Skilled workers ** (goods/services/process)	Based on Q21: *Which of the following factors hampered this business (in innovation) (c) Lack of skilled staff*	0/1 dummy
Firm size **	Log(Employment)	
Employment * (20 to 199 employees)	The number of employees is between 20 and 199	0/1 dummy
Employment * (more than 200 employees)	The number of employees is greater than 199	0/1 dummy
Foreign ownership (between 0% and 50%)	Foreign ownership is between 0% and 50%	0/1 dummy
Foreign ownership (more than 50%)	Foreign ownership is more than 50%	0/1 dummy
Age (between 1 and 4 years)	The firm is between 1 year and 4 years old	0/1 dummy
Age (between 4 and 9 years)	The firm is older than 4 years, but less than 9 years old.	0/1 dummy
Age (>=9 years)	The firm is 9 years old or older	0/1 dummy
Non-profit institute dummy **	Non-profit institute indicator	0/1 dummy
ICT employment share **	% of ICT employees in total employment (2002–03)	
Intramural R&D	Based on Q4: *Did this business engage in any of the following activities during the calendar years 2003? (b) Development related activities (iii) Research and experimental development performed by this business.*	0/1 dummy
Engineering and scientific skills (goods/services) **	Based on Q23: *Which of the following skills and capabilities does this business look for if engaging people (for innovation activities): Engineering, Scientific, Information Technology.*	0/1 dummy
Engineering and scientific skills (process) **	Based on Q23: *Which of the following skills and capabilities does this business look for if engaging people (for innovation activities): Engineering, Scientific, Information Technology.*	0/1 dummy
Engineering and scientific skills (goods/service/process) **	Based on Q23: *Which of the following skills and capabilities does this business look for if engaging people (for innovation activities): Engineering, Scientific, Information Technology.*	0/1 dummy

Note:

* used in modelling the linked IS-EAS-BIT-BAS data;

** used in modelling the linked IS-EAS data only;

If no asterisks, variable was used in both studies.

A.3 Productivity equation

Variables	*Definition*	*Range of values*
Dependent variable		
Labour productivity (value-added-based or gross output-based) **	Annualised growth in labour productivity between 2001–02 and 2004–05	
	Labour productivity level in 2003–04	
Total factor productivity (TFP) (value added-based or gross output-based) **	Annualised growth in TFP between 2001–02 and 2004–05	
	TFP level in 2003–04	
Value added growth *	Annualised growth of value added between 2001–02 and 2004–05	
Labour productivity (Employment) **	Annualised log growth in labour productivity between 2001–02 and 2004–05, where Labour productivity = Value added / Employment	
Labour productivity (Wage) **	Annualised log growth in labour productivity between 2001–02 and 2004–05, where Labour productivity = Value added / Wage	
Explanatory variables		
Innovation output		
Product innovation	Innovative sales / binary	
Process innovation	Firm has introduced process innovation	0/1 dummy
Organisational innovation	Firm has introduced organisational innovation	0/1 dummy
Firm employment size		
Firm size **	Number of employees	
Employment (20 to 199 employees) *	Number of employees is between 20 and 199	0/1 dummy
Employment (More than 200 employee) *	Number of employees is greater than 200	0/1 dummy
Ownership variables		
Foreign ownership (between 0% and 50%)	Foreign ownership is between 0% and 50%	0/1 dummy
Foreign ownership (more than 50%)	Foreign ownership is more than 50%	0/1 dummy
Firm age variables		
Age (between 1 and 4 years)	The firm is between 1 year and 4 years old	0/1 dummy
Age (between 4 and 9 years)	The firm is older than 4 years, but less than 9 years old	0/1 dummy
Age (>=9 years)	The firm is 9 years old or older	0/1 dummy
Non-profit institute dummy **	Non-profit institute dummy	0/1 dummy
2-digit industry dummies **	2-digit industry dummies	

Other explanatory variables

Capital growth **	Log change in capital stock (used only when the dependent variable is TFP growth)	
Growth in capital stock / wage *	Log growth in Capital stock per employee (used only when the dependent variable is Labour productivity (Wage))	
Growth in capital intensity	Log change in capital stock per employee (used only when the dependent variable is labour productivity growth)	
Log capital level **	Log capital level (used only when the dependent variable is TFP level)	
Log capital intensity level	Log capital intensity level (used only when the dependent variable is labour productivity level)	
Growth of employment **	Annualised growth in employment between 2001–02 and 2004–05 (used only when the dependent variable is Labour productivity value added-based))	
Log change in employment *	Annualised growth in employment between 2001-02 and 2004-05 (used only when the dependent variable is labour productivity growth))	
Log growth in wage *	Log growth in wage (used only when the dependent variable is Labour productivity (Wage))	
Relative TFP **	Initial TFP level relative to median firm in 4-digit industry (used only when the dependent variable is TFP growth)	
Share of ICT employee **	% of ICT employees in total Employment (2002–03)	
Initial market share *	Firm sales / Industry sales by 4 digit ANZSIC industry in 2001–02	
Log initial market share **	Log initial market share	
Herfindahl index	Industry concentration index by 4-digit industry	0–1
Log change in industry demand	Annualised log change in sales between 2001–02 and 2004–05	
Log Industry demand level **	Log Industry demand level	

Note:

* used in modelling the linked IS-EAS-BIT-BAS data;

** used in modelling the linked IS-EAS data only;

If no asterisks, variable was used in both studies.

B. DERIVATION OF THE PRODUCTIVITY (GROWTH AND LEVEL) EQUATIONS

The following derivation simplifies and combines features in both Klomp and van Leeuwen (2001) and Criscuolo and Haskel (2003).

We start by assuming an augmented Cobb–Douglas function where knowledge capital (K) enters as a separate input to the production of gross output (Q), in addition to the inputs of physical capital (C), labour (L), material inputs (M) for each firm i:

$$Q_i = A C_i^{\alpha} L_i^{\beta} M_i^{\lambda} K_i^{\gamma} \tag{1}$$

The function can be expressed alternatively as:

$$q_i = \alpha c_i + \beta l_i + \lambda m_i + \gamma k_i \tag{2}$$

where lower case letters denote the logarithms of variables. Subsequent equations in this section will be presented in log-linearised form unless otherwise specified.

We further assume that the demand curve (in log-linearised form) facing the firm in a monopolistic competition setting is:

$$q_i = q_I + d_i + \frac{\mu}{1-\mu}(p_i - p_I) \tag{3}$$

where q_i, q_I, p_i and p_I denote, respectively, the demand and own price (index) of firm i and total sales and average output price of market I. d_i is a 'demand-shifter' which represents all other effects on demand except price effects. This includes the effect of innovation on demand through improved product quality. μ is the mark-up of firm specific prices over marginal cost, and $\frac{\mu}{1-\mu}$ represents the price elasticity of demand.

Since data on the quantity of physical output, q_i, are seldom collected, q_i is typically obtained as deflated total sales or revenues, (r_i):

$$r_i = q_i + p_i - p_I \tag{4}$$

When prices differ across firms, firm-specific price deflators should be use to obtain real output. If industry price deflators are used in their absence, this may lead to biased estimates of the coefficients of the production function that have to be corrected for.

Substituting (3) into (4), we obtain:

$$r_i = \frac{1}{\mu} q_i + \frac{\mu - 1}{\mu} q_I + \frac{\mu - 1}{\mu} d_i \tag{5}$$

Substituting (2) into (5) gives:

$$r_i = \frac{1}{\mu}(\alpha c_i + \beta l_i + \lambda m_i + \gamma k_i) + \frac{\mu - 1}{\mu} q_I + \frac{\mu - 1}{\mu} d_i \tag{6}$$

Scaling both sides by l_i, and rearranging terms, (6) can be reformulated as follows:

$$r_i - l_i = \frac{\alpha}{\mu}(c_i - l_i) + \frac{\lambda}{\mu}(m_i - l_i) + \left(\frac{\alpha + \beta + \lambda}{\mu} - 1\right) l_i + \frac{\gamma}{\mu} k_i + \frac{\mu - 1}{\mu} q_I + \frac{\mu - 1}{\mu} d_i \tag{7}$$

which relates labour productivity – measured as gross output per employee, $(r_i - l_i)$ – to capital intensity $(c_i - l_i)$, intermediate input intensity $(m_i - l_i)$, knowledge capital (k_i), industry sales (q_I), and the demand shifter (d_I). The labour input term (l_i) on the right hand side allows for imperfect competition and non-constant returns to scale.

Time differencing (7) gives the formulation for the labour productivity growth equation:

$$\begin{aligned} \Delta r_i - \Delta l_i = {} & \frac{\alpha}{\mu}(\Delta c_i - \Delta l_i) + \frac{\lambda}{\mu}(\Delta m_i - \Delta l_i) + \left(\frac{\alpha + \beta + \lambda}{\mu} - 1\right) \Delta l_i + \\ & \frac{\gamma}{\mu} \Delta k_i + \frac{\mu - 1}{\mu} \Delta q_I + \frac{\mu - 1}{\mu} \Delta d_i \end{aligned} \tag{8}$$

This is the primary equation used in van Leeuwen and Klomp to estimate the relationship between innovation and productivity (growth).

Criscuolo and Haskel reformulate (6) to define a total factor productivity (TFP) measure on the left hand Side. ln(*TFP*) is calculated as:

$$r_i - S_l l_i - S_m m_i - (1 - S_l - S_m) k_i$$

where S_l and S_m denote the shares of the costs of labour and intermediate inputs in total revenue. This calculation assumes constant returns to scale and makes use of the firm's first order condition for profit maximization (as shown in Klette, 1996), which implies that:

$$\beta = S_l \mu \text{ and } \lambda = S_m \mu$$

That is, the output elasticities of the variables inputs are equal to the product of the factors' revenue shares and the mark-up.

A TFP level equation can be expressed as follows:

$$\ln TFP_i = \left(\frac{\alpha + \beta + \lambda}{\mu} - 1 \right) c_i + \frac{\gamma}{\mu} k_i + \frac{\mu - 1}{\mu} q_I + \frac{\mu - 1}{\mu} d_i \tag{9}$$

Time differencing (9) gives a *TFP* growth equation:

$$\Delta \ln TFP_i = \left(\frac{\alpha + \beta + \lambda}{\mu} - 1 \right) \Delta c_i + \frac{\gamma}{\mu} \Delta k_i + \frac{\mu - 1}{\mu} \Delta q_I + \frac{\mu - 1}{\mu} \Delta d_i \tag{10}$$

Criscuolo and Haskel estimated a TFP growth equation. In both TFP level and growth equations, the inclusion of a capital input term on the right hand side allows for imperfect competition and non-constant returns to scale.

If the revenue variable, r, is redefined as value-added (output towards final demand), then the above specifications can apply to value-added based labour productivity and TFP equations, except that terms involving intermediate inputs will be removed.

C. DATA CLEANING AND DERIVATION OF NEW DATA ITEMS

C.1 Data cleaning and treatment in the linked datasets

The initial IS-EAS linked dataset consisted of 1874 business units, while the IS-EAS-BIT-BAS linked dataset comprised 5,061 units. However, data anomalies necessitated the removal of some of these observations.

The following section elaborates on the data cleaning process for the two datasets.

C.1.1 IS-EAS linked data

Since each equation was estimated separately, the data cleaning process was implemented in two stages: firstly with respect to variables used in the innovation input and output equations, which used primarily data from the Innovation Survey, and secondly, with respect to additional variables in the productivity equation. Thus, the first two equations were estimated on a larger sample than the productivity equation.

Innovation variables

The initial dataset had 13 observations with innovation intensity greater than 1, that is, these units had innovation expenditure that is greater than total sales. It is possible that in the short term, as in a particular year, innovation expenditure, which is inclusive of R&D expenditure, may exceed total sales. Furthermore, even if some of these innovation intensity values are wrong, the omission of these observations may bias the estimation downwards to the extent that they are businesses with 'true' innovation intensity at the high end of the spectrum. Upon closer examination, there was no evidence that observations with innovation intensity greater than 1 were systematically associated with characteristics which might support the assumption that they are firms with consistently high innovation intensity, such as firm size or profitability. Also, some of these observations had implausibly large values in the hundreds. Moreover, the inclusion of observations with innovation intensity below 5 in our preliminary estimations showed that these observations exerted an undue influence on the results. Given that there is no strong ground for including them, these observations were dropped from our final estimation sample. Eight observations with missing or negative innovation intensity (due to zero or negative sales values) were also removed.

About 6 percent of businesses indicated that they had introduced innovations, but reported zero innovation expenditure. While it may be the case that some businesses may be innovating for free (e.g. costless imitation of other businesses), there are also two reasons that could lead to measurement errors for these observations. Firstly,

there may be measurement error due to businesses not keeping records of non-R&D innovation expenditures, which may be more pronounced due to the 2003 Innovation Survey being the first iteration of this survey instrument. Secondly, innovation expenditure is requested for the financial year 2002–03 while the reference period for implementation of innovations relates to the 2001–2003 calendar year period.

While it is expected that the innovation input would precede the innovation output, due to data limitations it is assumed that the 2002–03 expenditure is indicative of average annual innovation expenditure over this period. However, some investment may be lumpy, for example the acquisition of machinery for the implementation of product innovation, and may have occurred outside of financial year 2002–03. In that case, a zero value in 2002–03 is an inaccurate representation of innovation expenditure over the 2001–2003 calendar year period. Due to the high likelihood of measurement error for these two reasons, these observations were also removed from the linked dataset.

Production and performance variables

Variables in the EAS dataset or constructed from EAS data with the following characteristics were cleaned out, partly based on some of the criteria used in Hall and Mairesse (1995):

- Negative or zero gross output, value-added, employment or capital stock.
- Average annual growth rate in value added, gross output or material inputs of more than 300 percent or less than –90 percent.
- Average annual growth in labour and in capital of more than 200 percent or less than –50 percent.
- Extreme values in market share and profit share variables, as identified from data plots.

It was verified that the boundary values set by Hall and Mairesse for the removal of observations with extreme variability in growth rates were appropriate for this dataset, as the observations removed were outlier values that were either in the 1st percentile or above the 99th percentile. These cuts further remove about 300 observations from the sample, with some observations being dropped for several reasons simultaneously.

In total, under 25 percent of observations from the initial dataset were removed in the 'cleaned' final sample used in the productivity equation.

Treatment of non-profit institutions

The EAS dataset contains a flag that identifies whether a business unit is a non-profit institution (NPI). NPIs are legal or social entities created to produce goods and services but their status does not permit them be a source of income, profit or other financial gain for the units that establish, control or finance them (ABS, 2002). Around 3 percent of observations in the IS-EAS dataset were identified as NPIs. If there is ground to believe that these NPIs are driven mainly by factors that are at odds with the assumed profit maximizing motive of corporations, then NPIs should be removed from the sample. However, there is no firm evidence that NPIs engaged in market production do not aim to maximize surpluses. Furthermore, there is some uncertainty over the accurate identification of NPIs, that is, some units might be corporations and not NPIs. Given these considerations, observations identified as NPIs are retained in the sample. However, a dummy variable is included in the estimations to take into account possible differences in the behaviour between NPIs and corporations.

Reconciling inconsistencies in ANZSIC codes

Both IS and EAS datasets contain information on the industry (ANZSIC) classification of the business units. The ANZSIC code in the IS records a business' industry classification at the end of 2003. There are some 5 percent of observations where this code is different from the industry classification given by the EAS data for financial year 2002–03. The discrepancy in these cases has been reconciled by adopting the following criteria:

- If the ANZSIC (industry classification) for an observation is consistent either across all four years of EAS data (from 2001–02 to 2004–05), or for three years between 2002–03 and 2004–5, but different from the ANZSIC in IS, the ANZSIC code in EAS is used. The basis for doing so is that the consistency likely reflects that the industry classification is more accurately recorded in EAS.
- Otherwise, the ANZSIC codes given in the IS are used.

C.1.2 IS-EAS-BIT-BAS linked data

Before commencing the modelling process, extreme observations according to the following criteria were removed from the initial linked dataset. Firstly, 1013 observations with zero, negative or missing levels in sales, value added, labour productivity, capital stock or employment for 2001–02 or 2004–05 were removed. Secondly, 34 observations with annualised growth rates in these variables of less than 90% contraction or more than 300% expansion were removed. Finally, 16 units showing innovation intensity greater than 1 were removed.

In addition, for reasons discussed in the description of the data cleaning process for the linked IS-EAS dataset in the previous section, observations representing businesses that reported the introduction of innovation but zero innovation expenditure were removed from the IS-EAS-BIT-BAS data.

The table below summarises the data cleaning process.

C.1 Summary of additional data cleaning, IS-EAS-BIT-BAS data

		Data cleaning			
Data source	*Observations before data cleaning*	*Zero levels & unrealistic growth**	*Innovation intensity >1*	*Innovators with zero innovation expenditure*	*Observations after data cleaning*
EAS	1,873	188	5	164	1,516
BIT	727	61	6	61	599
BAS	2,461	804	5	176	1,476
All sources	5,061	1,053	16	401	3,591

* Levels & annualised growth in Sales, Value added, Labour productivity, Capital stock and Employment.

C.2 Construction /derivation of new data items

C.2.1 IS-EAS linked data

Gross output, value-added and intermediate inputs

Gross output, value-added and intermediate input expenses are derived as defined in the glossary in *Australian Industry* (ABS cat. no. 8155.0).

Capital stock

Theoretically, the appropriate measure of capital input for production is capital services, that is, the flow of productive services from the cumulative stock of past investments. Since flows of the quantity of capital services are not usually directly observable, it is typically assumed that services flows are in proportion to a stock of productive capital, which is capital adjusted for efficiency declines and retirement.

There are two broad approaches to the measurement of capital stock: direct measurement, which involves obtaining estimates directly from capital owners, and the perpetual inventory method (PIM), which involves the compilation of a rolling inventory of capital stocks. In Australia, aggregate and industry-level capital stock measures are constructed using the PIM. At the firm level, a 'standard' capital stock measure has yet to be developed.

It is not uncommon to find firm-level productivity studies using the total book value of non-current assets as a quick and dirty proxy for productive capital stock. However, this option is ruled out for the linked IS-EAS dataset, as data on assets and liabilities are no longer processed or collected in EAS from 2003–04. It is possible though, to derive a capital series using EAS data and applying a rough version of the PIM as follows:

$$K_t = K_{t-1} - D_{t-1} + I_t$$

where K is the capital stock, D is the value of depreciation, and I is net capital expenditure.

The value of non-current assets in 2001–02 is used as the initial capital stock value to start the recursion in PIM. Instead of applying a constant depreciation rate to the series, the value of depreciation in each year is measured as the depreciation expenses reported by businesses in the EAS. Net capital expenditure is derived as acquisitions less disposals.

Given that capital expenditure data are available by asset types in the EAS, an alternative is available to construct capital stock for each industry by asset type, and then aggregate the series using an appropriate weight to obtain total capital stock, following the methodology developed in Martin (2002). Since the value of non-current assets in EAS is not broken down by asset types, initial capital stock is derived by using some firm-specific weight (such as employment or investment shares) to apportion sectoral capital stock aggregates. The reason for trying this alternative is that an overall capital stock variable derived in this way may be more precise than the first method outlined above. However, preliminary investigations show that the capital stock figures derived using the second method give mainly negative capital growth rates by industry between 2001–02 and 2004–05, when calculations using industry capital stock figures show mainly positive growth. Since the growth rates given by capital stock figures derived from the first method are broadly consistent with industry trends, those figures are used in our analysis.

Labour productivity and total factor productivity measures

Labour productivity (value-added based) and labour productivity (gross-output based) are derived as value-added per employee and gross output per employee respectively.

TFP (value-added based) relates value-added to a combination of labour and capital inputs, while TFP (gross-output based) takes into account the contribution of intermediate inputs in producing output as well.

The appeal of using a TFP measure over a labour productivity measure is that labour productivity growth can result from increases in factor intensity ratios as well as TFP

growth. A TFP change is attributed to factors other than a change in combined inputs, and is often interpreted to reflect the impact of changes in technology, managerial ability and/or organisational efficiency on output growth. Thus, TFP is a 'purer' indicator of how factors other than input mix affect productivity growth.

Annualised growth rates used in the estimations are derived as the annualised difference of the log of labour productivity or TFP between 2001–02 and 2004–05.

C.2.2 IS-EAS-BIT-BAS linked data

Value-added

For businesses where financial data could be sourced from EAS data supplemented by BIT, industry value-added was derived as defined in the glossary in *Australian Industry* (ABS cat. no. 8155.0).

Where BAS has been used as the source of financial data for a business, value-added has been derived by subtracting non-capital purchases from sales. 'Non-capital' purchases is an item similar in concept to 'Other operating expenses' which commonly features in ABS economic survey outputs. This is not strictly the definition for value-added, as it should also include changes in inventories in the calculation. Previous ABS analysis has highlighted that, for most small businesses, inventories do not change significantly over time. Given this, while the derived item is a proxy for value-added, it is considered fit for the purpose of this analysis.

Labour productivity

Two proxies for labour productivity have been derived for this analysis. Firstly, a value-added to wages ratio has been produced by dividing value-added by wages. This is a one proxy for labour productivity, recognising that wages are not the ideal measure of labour inputs but addressing the unavailability of employment information in the taxation data sources.

Secondly, a value-added to employment ratio has been derived by using model-based estimates for employment where this information is not available (BAS and BIT data). Derivation of labour productivity growth requires employment estimates for the start year (2001–02) and final year (2004–05) of our period of analysis.

Capital stock

Similarly to the derivation of capital stock in the IS-EAS study, a rough version of the perpetual inventory method (PIM) is used to construct firm-level capital stock for this analysis. Data limitations inherent in using administrative data, though, in particular BAS data, prevent the same application of PIM as used in the IS-EAS study. The BAS only contains very limited data related to capital inputs for a business, with capital

purchases a measure of investments contributing to a cumulative capital stock. Specifically, compared to the data items available in the IS-EAS dataset, the following three EAS data items are unavailable: non-current assets, disposal of assets and depreciation. The methodologies used to overcome the unavailability of these three data items is outlined below.

Firstly, BAS does not include information on non-current assets which is used in the IS-EAS study as the measure of initial capital stock value in 2001–02 to start the derivation using PIM. Initial value of k_t at the individual business level as at June 2002 is derived as the investment share (averaged over three years) of aggregate productive capital (chain volume measure) at the ANZSIC division 1-digit industry level. The measure of aggregate capital is real productive capital stock obtained from unpublished National Accounts data at the 1-digit industry level.

Secondly, BAS does not include information on disposals, required for deriving net investments. This would result in an upward bias in growth rate in capital stock. Offsetting this bias, the initial value of k_t as at June 2002 is derived as the investment share (where the particular business's investment also does not account for disposals) of productive capital stock at the ANZSIC industry level. For businesses where financial data is sourced from the EAS-BIT dataset, there is high correlation between growth rates in capital stock using net (of disposals) investment vs gross investment.

Finally, depreciation is derived by applying a constant depreciation rate at the ANZSIC (1-digit) division level equal to the average of the ratio of consumption of fixed capital to net capital stock over the four years 2001–02 to 2004–05. Consumption of fixed capital is valued at current prices, and measures the difference in economic value of an asset (due to wear and tear and foreseen obsolescence) between the beginning the period and the end of the period. Net capital stock is also measured at current prices at the end of each financial year.

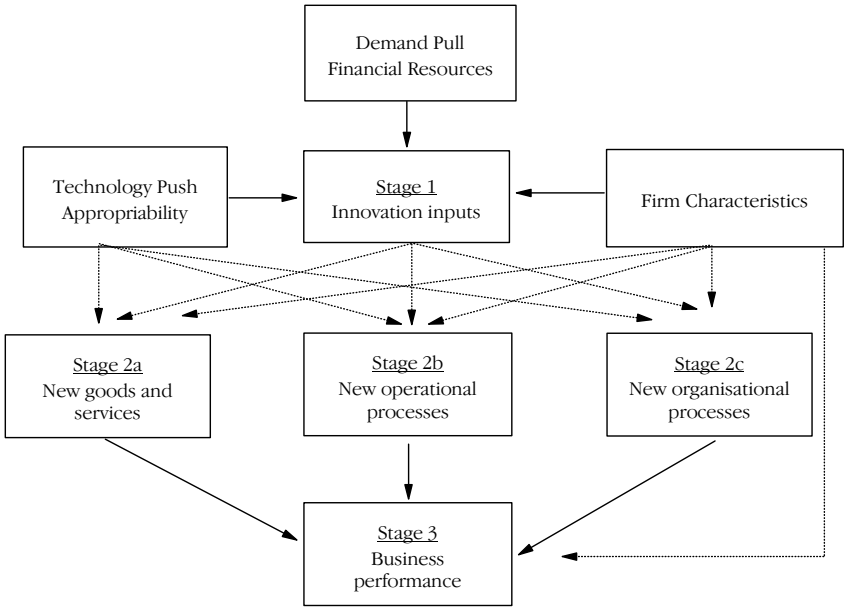
Employment

The derivation of (labour) productivity growth requires employment estimates for the start year (2001–02) and final year (2004–05) of our period of analysis. BAS and BIT data sources do not include employment information, with employment for these sources derived as model-based estimates based on wages and salaries reported. The model used was developed within the ABS and is now used in deriving industry employment estimates derived from the EAS-BIT dataset published in Australian Industry 2004–05 (cat no 8155.0). Parameter estimates for this model were derived for each of the two years using EAS data.

D. TECHNICAL DETAILS OF THE APPLICATION OF THE CDM MODEL IN THE IS-EAS AND IS-EAS-BIT-BAS STUDIES

The Crépon, Duguet and Mairesse (CDM) framework is used to examine the links between innovation input, innovation output and business performance. In particular, it is applied to explain how investing in innovation affects productivity, by decomposing this relationship into an effect from innovation input to innovation output and one from innovation output to productivity (or business performance).

D.1 The CDM model



There are three stages involved in the modelling process, each corresponding to an econometric equation (or set of equations) that explain the above (see figure 1). The first stage models the inputs to innovation to see what factors influence the decision to invest in innovation and the level of this investment. These factors include the firm's characteristics, technology push factors, appropriability, demand pull factors and financial resources. The second stage models innovation output by type (i.e. innovation in goods and services, operational processes and organisational /managerial processes). Separate equations are estimated for each type to determine the factors that influence the particular innovation output. The innovation intensity in stage 1 is included as one of the possible explanators. The third stage relates productivity or business performance to innovation outputs and other factors.

The following sections describe the estimation of the three equations as adopted in these studies (i.e. as applied in the linked IS-EAS data and as applied in the linked IS-EAS-BIT-BAS data).

D.1 Innovation input equation

The first equation explores the determinants of innovation input. The original CDM paper made use of R&D intensity as the measure of innovation input. But because of the low contribution of R&D expenditures to total innovation expenditures in Australia and the importance of service industries in Australia, this study makes use of a more comprehensive measure of innovation input. Innovation input is represented by the ratio of innovation expenditure to total sales in the reference year 2002–03 (covering the twelve months from 1 July 2002 to 30 June 2003). Innovation expenditure includes internal and external R&D expenditures plus expenditures on all non-R&D activities related to the implementation of product innovation, operational process innovation and organisational /managerial process innovation.

In the first equation, the relationship between innovation investment and its determinants is examined, using:

$$y_{1i} = x_{1i} + \varepsilon_{1i}$$

where y_1 is innovation intensity – that is, the share of innovation related expenditure in total sales – and x_1 is the vector of explanatory variables. i is the firm subscript.

The vector of explanatory variables (i.e. influences on innovation intensity) can be broadly categorised as follows:

Opportunities to innovate. These refer specifically to opportunities in the technological environment for innovation. They are proxied by variables on the sources of ideas for innovation, collaboration arrangements, barrier to innovation in the form of government regulations and standards and industry dummies to represent inter-industry differences in technological opportunities.

Incentives to innovate. Demand conditions that affect firms' incentives to innovate are represented by measures of the initial states of competitiveness (initial market shares and industry concentration), potential for sales growth (annualised industry sales growth) and appropriability conditions. Cost-push and demand-pull impetus for innovation as identified by firms are also included under this category.

Financial resources and other firm characteristics. Financial resources affect the cost of innovation, and these are represented by variables on whether cost or availability of finance hampers innovation efforts, government financial support, and profit as a share of total sales. Firms' human capital (more specifically, problems with it) is proxied by the lack of skilled staff as a barrier to innovation. Firm characteristics of size, age and ownership structure proxy for unobservables that might affect innovation activity.

In a number of studies, notably those based on the Community Innovation Surveys in Europe such as in Crépon et al. (1998) and Loof and Heshmati (2006), the main estimations are carried out using a sample of firms engaged in innovation (or R&D) activities only. One reason is the sample design, which collected data on innovation expenditures and characteristics for innovators only. If failure to engage in innovation activities is associated with lower (potential) innovation intensity, estimates based on a sample of innovators will be biased. Thus, these studies implement a generalized tobit model to correct for possible sample selection bias.

In the 2003 Australian Innovation Survey, most of the information on characteristics such as collaboration and sources of ideas for innovation activities were asked of all respondents, that is, both 'innovators' (defined as businesses that have introduced or implemented innovations between 2001 and 2003) and 'non-innovators'. However, there is some uncertainty over whether the questions on innovation expenditure had captured accurate information on innovation expenditure from all respondents. The questions had asked for expenditure 'related to the introduction /implementation' of different types of innovations, instead of the more general 'to introduce or to develop' used elsewhere in the survey. It is possible that some 'non-innovators' that had started and abandoned, or had ongoing but not completed innovation activities did not respond to these questions even though they had positive innovation expenditure. This will introduce measurement error in the data, or a selection bias if the estimations were carried out on a sample of innovators only. An examination of the data shows that the proportion of non-innovators that had entered positive innovation expenditure figures constitutes about 5 percent of all observations. The issue is whether the remaining non-innovators, comprising 34 percent of total observations, can be assumed to have no innovation activities and thus zero innovation expenditure. Subsequent investigation seems to support this assumption.

The 2005 Innovation Survey specifically collects information on businesses that have not yet introduced innovations but nevertheless are engaged in innovation activities (abandoned or innovations yet to be completed). It is found that the net rate of these activities by similar sized businesses is around 5 percent, similar to the share of non-innovators reporting positive innovation expenditure in the 2003 IS. Since there is ground to believe that the data on innovation expenditure is accurately captured for all respondents in the 2003 IS, this study will carry out estimations using the full sample of innovators and non-innovators. In this regard, the study is similar to that of Criscuolo (2004), who estimated the equations using a full sample of UK manufacturing firms.

Since our sample comprises both positive innovation expenditure and a substantial portion of zero expenditure, an appropriate estimation method to use is a standard censored Tobit model for corner solution responses. One limitation of the standard

Tobit model is that a single mechanism determines whether a firm decides to invest in innovation activities and the level of investment. It is possible that a different set of factors affect the two elements of choice or that the same factors have effects in opposite directions. A two-step model that specifies a 'propensity' and a 'level' equation allows for separate processes to determine these two choices. The final results presented in this paper are based on the application of variants of this type of model on the two datasets.

The IS-EAS and IS-EAS-BIT-BAS studies both specify a 'propensity' and a 'level' equation, and assume that the two equations are not interrelated through some omitted variables. Both studies assume that once a firm decides to invest, there is no corner solution and the level of innovation expenditure (or innovation intensity) is always positive. This means that non-investors do not impose restrictions on the parameters of the level equation.

Each study uses the same regressors in both the 'propensity to invest' and 'level of investment' equations, as there are no *a priori* reasons for specifying different sets of regressors. The only exception is the government financial support variable, which is included only in the second equation, since it is always zero for firms that did not invest and thus does not contribute to the explanation of a firm's decision to invest. However, there is some difference in the set of explanatory variables used between the two studies. The linked IS-EAS study included count measures of diversity and intensity that capture the extent of businesses' collaboration activities and intellectual property protection, as well as a dummy variable to control for possible difference in the behaviour of non-profit institutes. A log-level employment variable is used to control for firm size, instead of dummies by firm size range in the IS-EAS-BIT-BAS study.

Below are details of the specifications used in the two studies.

The two studies estimate a two-tiered model as specified in Wooldridge (2002). The same model was applied in Criscuolo (2004), and can be written as follows:

$$g^* = x_{1a}\beta_1 + \varepsilon_{1a} \tag{1}$$

$$h^* = x_{1b}\beta_1 + \varepsilon_{1b} \tag{2}$$

and $\ln(y_1) = h^*$ if $g^* > 0$.

The error terms in equations (1) and (2) are assumed to be normally distributed and independent between equations:

$$\begin{pmatrix} \varepsilon_{1a} \\ \varepsilon_{1b} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & \sigma^2 \end{pmatrix}\right]$$

g^* is a latent innovation decision variable, such as the expected present value of firm profit accruing to innovation investment. The observable counterpart $g = 1$ when $g^* > 0$, that is, if a firm invests in innovation, otherwise Zero. h^* represents a latent, or true level of innovation intensity. y is the observed innovation intensity, and conditional on a firm's decision to invest, the distribution of innovation intensity on the vector of regressors '$\boldsymbol{x}$', is lognormal.

The parameters are estimated by maximum likelihood technique, where the log-likelihood function of innovation intensity for observation i is

$$\log L_{1i}(y_1 \mid X_{1a}, X_{1b}; \beta_{1a}, \beta_{1b}) =$$
$$\sum_0 \log\left[1 - \Phi(X_{1a}\beta_{1a})\right] + \sum_1 \log\left[\Phi(X_{1a}\beta_{1a})\frac{1}{y\sigma}\phi\left(\frac{\log(y) - X_{1b}\beta_{1b}}{\sigma}\right)\right]$$

where Φ and ϕ are the standard normal cumulative distribution function and the corresponding density function, respectively.

D.2 Innovation output equations

In the second step, the transformation process from innovation input to innovation output is explored given:

$$y_{2i} = \alpha_1 y_{1i} + x_{2i}\beta_2 + \varepsilon_{2i}$$

where y_2 is an indicator of innovation output.

Using the basic equation above, we estimated four equations, each referring to a measure of innovation output. The first two are for product innovation, while the last two are for process and organisational /managerial innovation outputs, respectively.

The four output measures are:

1. a binary indicator (i.e. yes/no) of whether the business has introduced any new or significantly improved goods or services in the calendar year period 2001–2003;
2. the share of the business's turnover in calendar year 2003 that was attributed to new goods or services (product innovation) introduced during the calendar year period 2001–2003;
3. a binary indicator (i.e. yes/no) of whether the business has introduced any new or significantly improved operational processes in the calendar year period 2001–2003.

4. a binary indicator (i.e. yes/no) of whether the business has introduced any new or significantly improved organisational /managerial processes in the calendar year period 2001–2003.

The difference between the two product innovation output equations (1 and 2 above) is that the innovative sales variable (2) measures the degree of innovation success, while a binary yes/no variable (1) is a cruder measure. The latter is estimated so that the results for product innovation can be more easily compared with those obtained from the process and organisational innovation output equations, where the only output measures available are binary responses.

Innovation intensity, y_1, enters as an explanatory variable in each of the equations above.

The set of other explanatory variables, x_2, in each innovation output equation are broadly the same groups as in the innovation investment equation. The exceptions are proxies for the incentives to invest in innovation and financial variables, most of which are excluded from the innovation output equations, on the assumption that these factors do not directly affect innovation outcomes, but enter indirectly through innovation intensity. Only the three intellectual property variables from the 'incentives to invest' set of variables are included in the innovation output equation. This is based on the assumption that besides measuring the rent that firms can earn from introducing innovations, these are also measures of the technological environment and unobserved incentive on workers to achieve positive innovation outcomes (if they know that their intellectual property will be well protected).

Three variables are added to the innovation output equations to proxy for firms' human capital and absorptive capacity. These are: a dummy variable on whether a firm engages in intramural R&D, the share of ICT employees in total employment, and whether firms look for workers with scientific and engineering skills to develop innovations. These variables are excluded from the innovation investment equations as they are assumed to be affected by a firm's decision to innovate rather than affecting its decision to innovate.

As with the innovation input equation, estimations are carried out on the full sample so that inferences can be drawn for both innovators and non-innovators. In both studies, all the innovation output equations with binary dependent variables are estimated using the probit model. However, they differ in the method used to estimate the innovative sales equation.

D.2.1 Estimating the innovative sales equation

Linked IS-EAS data

The innovative sales equation is estimated using fractional logit regression, which was developed by Papke and Wooldrige (1996) for models with a fractional dependent variable which spans between 0 and 1. This method ensures that the estimates will take values within the unit interval. It is also more appropriate compared with applying log-odds transformation to the dependent variable when the dependent variable has a significant proportion of zeros (and/or ones), which is the case with our full sample. A logistic functional form is assumed for the conditional mean of share of innovative sales (y_2) and the parameters are estimated by quasi-maximum likelihood estimation, where the quasi-log likelihood function for a given observation i is:

$$\log L_{2i}(y_2 \mid x_2;\beta_2) \;=\; y_{2i}\log\left[\frac{\exp(x_2\beta_2)}{1+\exp(x_2\beta_2)}\right] + (1-y_{2i})\log\left[1-\frac{\exp(x_2\beta_2)}{1+\exp(x_2\beta_2)}\right]$$

The marginal effect is:

$$\frac{\partial E(y_2 \mid x_2;\beta_2)}{\partial x_{2i}} \;=\; \beta_i\left[\frac{\exp(x_2\beta_2)}{\left[1+\exp(x_2\beta_2)\right]^2}\right]$$

Linked IS-EAS-BIT-BAS data

The innovative sales equation using the linked IS-EAS-BIT-BAS data is estimated using a censored regression model (type-1 Tobit). The censoring is done to take into account the fact that there may be significant number of observations with zero values for the dependent variable.

The Tobit (or type-1 Tobit) model is represented by the two equations below.

$$y_2^* \;=\; y\alpha + x_2\beta_2 + \varepsilon_2$$

$$y_2 = \begin{cases} y\alpha + x_2\beta_2 + \varepsilon_2 & \text{if } y_2^* > 0 \\ 0 & \text{otherwise} \end{cases}$$

where y is the innovation intensity, x_2 is a vector containing: source of information variables, type of collaboration variables, appropriability conditions variables, barrier variables, ownership structure variables, firm size and firm age variables.

The first equation is a model for a latent (or unobserved) dependent variable. It is assumed that the disturbance term in this first equation is normally distributed with a mean of zero. The second equation accounts for the left-censoring at zero that exists for the observed dependent variable, the share of the business' turnover attributable to new goods or services.

D.3 Business performance (productivity growth and level) equations

The third set of equations examines the link between productivity and innovation output, using an augmented Cobb–Douglas production function:

$$y_{3i} = \gamma_1 y_{21i} + \gamma_2 y_{22i} + \gamma_3 y_{23i} + x_{3i}\beta_3 + \varepsilon_{3i}$$

where y_3 is a measure of productivity growth and Level. y_{21}, y_{22} and y_{23} are the three types of innovation output (product, process and organizational) where the product output is represented by percentage innovative sales, and both the process and organisational outputs are represented by binary Variables. x_3 is a vector of other explanatory variables.

Other studies have used a number of productivity measures, including TFP (gross-output based) growth (Criscuolo, 2004), labour productivity (gross-output based) growth (van Leeuwen and Klomp), and labour productivity (value-added based) level (Crépon et al., 1998; Loof and Heshmati, 2006). Although our results may not be exactly comparable with these studies because of differences in our sample and units of measurement of some of the variables, we will run a number of estimations using different measures of productivity growth and level to cross check the influence of innovation output variables on firms' productivity performance. These productivity measures are not independent of each other, and their interrelationships can be shown using the economic theory of production. For example, one of the driving forces behind labour productivity growth is the rate of TFP change. The rate of change in value-added based TFP equals the rate of change of gross-output based TFP multiplied by the inverse of the nominal share of value added in gross output. Since the ratio is smaller than one, the growth in value-added based TFP will be higher than the growth in gross-output TFP for the same business or industry (OECD, 2001).

D.3.1 Differences between the methodologies in the two studies regarding the productivity equation

The most significant difference between the two studies was in the selection of performance variables.

(1) Linked IS-EAS data

The study using the linked IS-EAS data constructed eight business performance indicators, namely:

- growth (annualised log change) in total factor productivity or TFP (both value-added and gross-output based) from 2001–02 to 2004–05;

- growth (annualised log change) in labour productivity or LP (value-added and gross-output based) from 2001–02 to 2004–05;
- log-level of TFP (value-added and gross-output based) in 2003–04; and
- log-level of LP (value-added and gross-output based) in 2003–04.

Note that the standard derivation of a productivity relationship from an augmented Cobb–Douglas function assumes perfect competition in the product market. However, van Leeuwen and Klomp (2001) and Criscuolo and Haskel (2003) highlighted that innovating firms effectively are operating in markets characterised by horizontal or vertical product differentiation, and can be expected to possess market power. There are two implications: firstly, innovations may impact on firm performance via their effects on demand conditions, rather than serve as a knowledge capital input into production. Secondly, if endogenous firm-specific prices are unobserved and not taken into account, this will lead to biased estimates for the coefficients of the production function. These two studies accounted for imperfect competition in their derivation of the productivity equation, and we adopt their specifications in our productivity growth equation estimations. We further adapt these specifications to derive specifications for labour productivity and TFP level equations. The detailed derivation is provided in Appendix B.

For example, the TFP growth equation is expressed as:

$$\Delta \ln TFP_{3i} = \delta_0 + \delta_1 \Delta c_{3i} + \delta_2 y_{21i} + \delta_3 y_{22i} + \delta_4 y_{23i} + \delta_5 \Delta q_I + \varepsilon_{3i}$$

where Δc is growth in capital input, y_{21}, y_{22} and y_{23} are the three types of innovation output, and Δq_I is growth in industry demand. Controls for human capital and market competitiveness are also added to the regression. A variable on each firm's TFP level relative to the median firm in the 4-digit industry, 'relative TFP', is also included. This variable captures the scope for learning, as firms that are further away from the frontier have more scope for learning and so catch up and grow faster.

The productivity growth equations are estimated using Ordinary Least Squares method. They are run on the full sample of innovators and non-innovators, (and a sample of innovators only), as well as sub-samples by manufacturing and services sectors to explore possible differences in the behaviour of businesses in manufacturing and services industries. Regressions using indicators that measure the degrees of novelty of product and process innovation were also run to test the hypothesis that innovations with higher degree of novelty are more strongly and positively correlated with productivity performance.

It should be noted that in our regressions, the value of any financial variables have not been deflated because of the lack of price deflators at a suitably disaggregated level. We believe that deflation using broad industry level deflators across different variables might introduce more errors and give rise to more imprecise estimates compared with using nominal values. Nonetheless, this means that our regression results would reflect the effect of industry level prices changes that have not been accounted for. As such, our empirical implementation is not exactly aligned with the theoretical derivation, and this could add to imprecision in the interpretation of the coefficient estimates. Acknowledging that measurement errors in the variables used in our estimation may affect the quality of the results, we also ran regressions assuming constant returns to scale and perfect competition as a robustness check.

(2) Linked IS-EAS-BIT-BAS data

The study using the linked IS-EAS-BIT-BAS data made use of three business performance indicators, namely:

- log annualised growth in value added from 2001–02 to 2004–05;
- log annualised growth in labour productivity, defined as value added/employment, from 2001–02 to 2004–05;
- log annualised growth in labour productivity, defined as value added /wages, from 2001–02 to 2004–05.

FOR MORE INFORMATION . . .

INTERNET **www.abs.gov.au** the ABS website is the best place for data from our publications and information about the ABS.

LIBRARY A range of ABS publications are available from public and tertiary libraries Australia wide. Contact your nearest library to determine whether it has the ABS statistics you require, or visit our website for a list of libraries.

INFORMATION AND REFERRAL SERVICE

Our consultants can help you access the full range of information published by the ABS that is available free of charge from our website, or purchase a hard copy publication. Information tailored to your needs can also be requested as a 'user pays' service. Specialists are on hand to help you with analytical or methodological advice.

PHONE 1300 135 070

EMAIL client.services@abs.gov.au

FAX 1300 135 211

POST Client Services, ABS, GPO Box 796, Sydney NSW 2001

FREE ACCESS TO STATISTICS

All statistics on the ABS website can be downloaded free of charge.

WEB ADDRESS www.abs.gov.au

2000001567401
ISBN 9780642483645

RRP $11.00