CHAPTER ONE

Introduction

1.1 THE PROBLEM

A fundamental problem that has beset Australian firms, whether privately or publicly owned, for decades is the poor information base on which they are required to make important decisions. A corollary to the problem is that Governments are also required to utilise this poor information when they make policy decisions which affect the future of firms. The inevitable and frequent consequence of this is that the decisions and policies are later found to be flawed.

In attempting to come to terms with this fundamental problem, this thesis first conceptualises a model of the firm, which identifies and measures the firm's cost efficiency, service (or product) effectiveness and resource utilisation efficiency. The thesis then proceeds to test cost efficiency concepts, in a case study of New South Wales rail passenger services from 1951/52 to 1991/92. The model is presented as a "corporate" model. This term is deliberately chosen for its comprehensiveness and because it may be seen as a box of tools for measuring key elements of any firm's performance. Like all models, it needs to be viewed as an analog of a firm's activities.

One primary point, however, needs to be reiterated strongly. Like the world of natural science, the business firm is a complex organism. Like some practitioners in the natural sciences, both business people and government functionaries have constantly trapped themselves in over-simplifications. An example is the "tyranny of the bottom line" (Kaye 1993) in the pursuit of short-term profitability with no clear long term strategies. The causes of this current tyranny are to be found in the long, economic boom of the 1950s, leading to the collapse in the 1980s. Poor information fed the inherent gambler's mentality of many businessmen and government functionaries. With good luck, the business decisions yielded substantial benefits; with bad luck the firm could generally come up with a damage control strategy (often involving government diversion of taxpayers' money) which would tide the corporation over until boom times returned. Public scrutiny of the firm's actions could be obscured because the only available documentation of business performance was an annual profit and loss account and

balance sheet. The collapse of the 1980s highlighted the problem of the paucity of information that created a climate of self-delusion among some businessmen who endeavoured to cover their actions, sometimes criminally, by falsifying the public records which they were required to produce.

One aftermath of the collapse of the 1980s boom was the emergence of the philosophy of the "bottom line". Here key aspects of the firm's economic performance are ignored in favour of a sharp focus on net profit in the short term. The fundamental problem of poor information still exists with this philosophy. The long term consequences of this are hard to determine.

Information poverty is particularly severe when it comes to measuring the Australian firm's economic performances especially the efficiency with which the firm utilises the resources at its disposal to produce given levels of output. The Industry Commission (IC) has systematically addressed the problem in its various Inquiries, with the Inquiry into the Rail Industry (IC 1991) exemplifying the problems in attempting to measure economic performance at the industry or firm level. Essentially it comes down to the point that the needed information is only obtainable with considerable difficulty. Measures of *economic* performance are essential to identify future strategic options for the firm as well as the current condition of the firm. All disciplines—accounting, management science and engineering etc.—have methods which measure some aspects of the firm's performance, but they all fail to identify some key elements of the firm's *economic* performance: the emphasis of this thesis and the core issue that it attempts to address.

For example, if we examine the most scientifically advanced area of the accounting discipline, financial theory and capital markets (Copeland and Weston 1988), we still find a focus on the short run. Also, in such models, past data tends to be heavily discounted as a guide to the future. The view holds that past data contains no new information which can be used in the formulation of the final business decision. At best, accounting data is seen as a device for monitoring the performance of a firm's management (Copeland and Weston 1988).

On the other hand, management science has provided a very substantial literature on corporate performance indicators. These are usually expressed in the form of ratios, indices etc. Most commonly, they relate resource input identities to measures of the firm's output. A good example of this approach is performance indicators developed for the rail industry by the Australian Railway Research and Development Organisation (ARRDO). It was this work (ARRDO 1986) which provided the initial impetus for this thesis. At best, such indicators are one-dimensional, linear, descriptions of single points

in the past history of the firm. A major drawback is that it cannot relate more than two variables in a firm's performance in any meaningful way, at any one point in time. The univariate methodology implies that the two variables are related in a linear way, in the sense that:

$$f(a \times I) = a \times f(I)$$

where "a" is any value, "f" is a linear system, and I is the input variable (Weinberg 1975). Systems theory clearly demonstrates that few systems in the physical and social universes operate in a strictly linear fashion. The most limiting factor, however, is that performance indicators are only descriptors of a firm's performance and hardly have a causal content. Experience in their use shows that the analyst is inevitably forced to generate a whole new list of questions arising from interpreting the possible answers given by the original performance indicator, notably—what are the reasons for variations in economic performance between comparable business entities?

Management science, in collaboration with accounting and engineering, has lately created a new variant of performance indicators called "benchmarking" (Leibfried and McNair 1992). It focuses upon measurement of a firm's performance as the key to the development of sound business strategies. However, for sheer "gobbledegook", the following definition of benchmarking is hard to beat (Leibfried and McNair 1992):

> Benchmarking is becoming the basis for this measurement process; it is forward-looking, holistic, participative, quality focused, and above all, stakeholder driven.

Benchmarking suffers from all of the disabilities of performance indicators with some additional ones of its own. First, it focuses on a paragon performer in its own firm, industry etc. and uses the paragon to set the standard for across the board performance in the industry. There is little evidence in the literature on benchmarking to indicate that its methods enable the analyst to understand the reasons why the paragon performer reached its status, or indeed whether apples are being compared with pears. Second, the choice and use of benchmarks (and, indeed, any performance indicators) depends upon the honesty of the analyst. Many North American commentators on benchmarking ethics agree that a good anti-trust lawyer be consulted before publishing benchmark data (Leibfried and McNair 1992).

If existing methods, at best, are only partially useful to business managers, they are even less useful to Governments, which have to make decisions and policies affecting the future of the firms, from the vantage point of the outsider who is the only one removed from the reality of the market. This, in turn, leads to flawed policy decisions. In this critical period of Australia's economic history, there is an urgent need to provide

business managers and government policymakers with better tools for measuring economic performance including guidance on actions to enhance such performance.

1.2 HOW IS THE PROBLEM ADDRESSED?

This thesis proposes a better set of tools for measuring performance. In the first place, to achieve this, a strong conceptual foundation needs to be built. This concept is called the "Quadrae", based on a synthesis of contributions by Hensher (1987) and Fielding (1987). The Quadrae is predicated on the following assumptions:

- Any firm can be analysed as a system composed of sub-systems, processes and sub-processes.
- Any method for measuring a firm's economic performance must incorporate three elements: cost efficiency, service/product effectiveness, and efficiency of resource utilisation.

The Quadrae is a group of three sub-models of the firm built around the key component model containing the Fielding Triangle. The first of these analyses the firm's total factor productivity (TFP) and includes Management and Innovation as quantifiable factors of production. TFP is an index of the amount of aggregate output produced by a unit of aggregate input. The second sub-model analyses pricing efficiency and includes as factors marketing effectiveness of management and service/product innovation (specifically improvement in the quality of the service/product). The third sub-model analyses how well the firm uses the economic resources at its disposal including the second order effects of the firm as it conducts its business. Included as measurable factors are government management initiatives and innovations which impact upon the firm's performance.

For the purpose of the thesis, and to test the methods used in the model, the sub-model is developed in a case study of NSW rail passenger services from 1951/52 to 1991/92. The physical determinants of these railway services are commented upon in Appendix 1 (including Maps of the system), and subsumed throughout the thesis.

The first step in the study is to develop, within the theoretical framework, a robust database, called RAILDATA. Whilst, in most cases, the best available sources are used, much of RAILDATA is not official either because no official data exists or because that which does exist is demonstrably in error. The unit costs which are developed in RAILDATA are derived from audited Annual Accounts of the NSW Railways, tabled in the NSW Parliament. Whilst the unit costs have been validated against other independent sources, they depend on allocative techniques based on rail operations

identities such as locomotive kilometres, train kilometres and the like. The big lesson which emerged from the development of RAILDATA is that good data is absolutely essential to the measurement of the economic performance of firms. The good data will lead to the correct conclusions being drawn up in the diagnosis of past performance. In turn this will assist the prognosis of future performance and the development of appropriate strategies for the firm.

RAILDATA is organised on a systems basis. It views NSW rail passenger services as a system, made up of sub-systems, which in turn may be comprised of many processes and sub-processes. Data has been collected on four distinct market segments serviced by NSW Rail: Suburban, Interurban, Country and Interstate. Four service output indicators are developed for each market segment: passenger journeys, passenger revenue, seat kilometres and passenger kilometres. As explained in the Chapters below, passenger revenue was dropped as a meaningful output variable because a substantial part of it is an economic rent.

The passenger railway system is dissected into four sub-systems: train running, corridor, terminals and stations, and business and corporate management. Data is collated in cost and quantity form for each sub-system for each factor of production, i.e. labour, fuel, materials and capital. Management and technology change are also included. The concept of total quality management is also introduced. To develop a set of testable hypotheses on economic performance, we commenced with the development of partial productivity indicators (PPI), which relate factor inputs (by sub-system) to factor outputs (by segment). The econometric analysis has two objectives. The first objective is diagnostic. Using a "bottom-up" approach, measures of TFP are developed for rail passenger services in NSW, by market segment. The "bottom-up" approach is cross-checked against independently calculated "top-down" estimates only derived from published Annual Accounts. This diagnostic objective may also be termed the "ex post" objective, where a robust database is used to develop strategic insights into the cost efficiency of a business firm. However, in line with most commentators on strategy formulation in the business firm, this activity is considered necessary but not sufficient for deciding on future business strategies. In the concluding Chapter, the study is given an "ex ante" objective as well: the consideration of strategy options aimed at improving cost efficiency in NSW rail passenger services. The details regarding the construction of RAILDATA are provided in Appendix 2.

The other two components of the Quadrae, pricing efficiency and resource utilisation efficiency, are not developed in this thesis. Nevertheless, it is anticipated that this thesis will make a number of important contributions towards an overall understanding of the

concept of efficiency and effectiveness, when considered in the context of the business firm:

1. In an attempt to present a new paradigm for assessing the economic performance of any business firm, this study synthesises, at least in concept, the five major themes in micro-economics: productive efficiency, which is measured by TFP, pricing efficiency, utilisation efficiency (measuring economic gain/loss), management quality as the organising element in the business firm and innovation in its widest sense. Considerable emphasis is placed upon the last two elements.
2. The model eventually derived from this study remains firmly anchored in the traditional theory of the firm.
3. The model heavily borrows from systems theory which enables us to logically disaggregate any firm into its component sub-systems, processes and sub-processes. This enables us to gain greater insight into TFP, which would not be possible if the firm is looked at in the aggregate only.
4. The study demonstrates the usefulness of econometrics in handling the analysis of any firm's performance, to any level of disaggregation required by management.
5. The creation of RAILDATA will act as a useful tool for measuring the efficiency of NSW rail passenger services. The methods used in the development of RAILDATA will also have a wider application. Whilst these data are based on rail operations, Chapters 8 and 9 demonstrate the importance of market knowledge in the proper management of the firm.

1.3 OUTLINE OF THE THESIS

This thesis is organised in Chapters. The first Chapter contains an introduction to the thesis. It explores the reasons why this study was undertaken. The objective was to explore the recent history of performance measurement in the public and private sectors. In particular, it considers the inadequacies of existing performance measurement methods and the consequences that these inadequacies may have for strategic decision making by a firm's management. It concludes with a brief assessment of the benefits and costs of using TFP to measure corporate performance. Finally, it spells out the "raison d'etre" for the study: the formulation of an holistic approach to corporate performance measurement which also manages to keep faith with the theory of the firm.

Chapter 2 is a detailed essay on the nature and significance of performance measurement.

Chapter 3 defines the Quadrae Corporate Model. It represents the contextual framework within which the thesis is written and in which TFP of NSW rail passenger services are measured. TFP is presented as one of these key measures of a firm's performance in

economic terms: productive efficiency (measured by TFP), marketing effectiveness (measured by pricing efficiency criteria), and utilisation efficiency (measured by total economic loss/gain). In this Chapter, the three measures are brought together and integrated into a conceptually linked "corporate" model called the Quadrae Model, as it consists of four distinct, but linked elements.

Chapter 4 provides a detailed examination of TFP developed for the thesis. It contains a detailed analysis of the measures and meanings of output together with extensive statistical research findings on four output measures: seat kilometres, passenger journeys, passenger kilometres and passenger revenue. The Chapter concludes with a dissertation on the measures and meanings of input. The statistical analysis of input measures is relegated to Appendix 2: RAILDATA. This is by no means meant to be dismissive of input statistics. It is simply a way to get around the complex methods used to derive the input measures. Also, the veracity of the thesis depends upon the quality of the input statistics. This quality also supports a major theme of this thesis, namely a high quality information base that is a crucial requirement in the measurement of productive efficiency.

Chapter 5 considers, in detail, the process of innovation as a factor of production determining TFP. This is considered to be a major contribution that the thesis makes to TFP.

Chapter 6 considers, in detail, the management of NSW rail passenger services as a factor of production determining TFP. Here again the thesis makes a major contribution to sources of variation in gross measures of TFP.

Chapter 7 presents the empirical results from the case study. It develops partial productivity indicators (PPIs) for the rail passenger system as a whole and for the component sub-systems: train running, terminals and stations, the rail corridor (dissected into variable track maintenance and other corridor activities) and business and corporate administration. Whilst PPIs are essentially descriptive in nature, they are a robust platform for constructing the hypotheses which are tested empirically in Chapter 8.

In Chapter 8, an index number analysis of Gross Total Factor Productivity (GTFP) of NSW rail passenger services, by market segment, is conducted from 1951/52 to 1991/92. In order to produce Net Total Factor Productivity, management changes, technological changes and traffic densities are tested as sources of variation. Management emerges as the pre-eminent source.

Based on the results in Chapter 8, Chapter 9 develops strategic conclusions on rail passenger services which can act as a guide for transport policymakers.

CHAPTER TWO

Factors and Other Studies Motivating This Study

2.1 NEED BY GOVERNMENT FOR PERFORMANCE MEASURES

The issue of productive efficiency has exercised the minds of Australian economists at least as far back as economic commentary has existed in the country. At the first meeting of the Australian Economic Association in Sydney on 21 February 1888, Professor Walter Scott noted in his critique of the English Corn Laws (Scott 1888):

> ... and a truer political economy has taught us that the way of escape was by increasing the efficiency of labour. In whatever way the result of the labourer's work is increased —whether by external circumstances, by improved instruments, or by his own intelligence and perseverance—his share in the wealth produced, other things being equal, will be greater.

In *One Nation* (Keating 1992), the issue is again addressed, as a matter of high priority. By the time of *One Nation*, the issue of Australia's economic performance had been a matter of concern for twenty years. In March 1979, an eminent Study Group on Structural Adjustment reported to the Prime Minister on the long term impact of structural adjustment in the Australian manufacturing sector. Policies on productivity were not explicitly mentioned, although the Study Group's report noted "On average, productivity in Australian factories is not good by international standards" (Crawford et al. 1979).

The Report is unclear on the diversity of input productivities of concern. However, at that time, the focus was very much upon labour productivity, that is specifically, the growth of production per hour worked for the non-farm market sector of the economy. This was found to be in decline in the second half of the 1970s (Norton 1982). No clear reasons could be given for this, although much anecdotal and media reporting was given to the general notion that the "work ethic" was somehow evaporating in Australian workplaces. This lopsided ideological assumption was of considerable concern to the union movement, and perhaps helped to bring it into a strong, proactive co-operation with government and business to eliminate any impediments to labour productivity that did exist. But some analysts continued to pursue the issue of whether, in truth, Australia's productivity decline was due to worsening labour productivity

(Burgess 1989). These commentators suggested that, whilst market deregulation, industrial reform and elimination of restrictive work practices were fine, as far as they went, there were deeper, longer run influences on productivity as well. These were identified as the long term shift of resources from the highly productive resources sector to less productive manufacturing, services and government administration, the decline in the capital intensity of production (also wrought by the shift in resources), and the general slow-down in the world's major economies. After the 1981/82 recession, structural adjustment in the manufacturing sector began in earnest with factory shut-downs, lay-offs, movement of plants off-shore, and reconstruction of plants in new locations. Productivity tended to recede as a policy issue, as the firms were restructured or ceased business altogether.

The issue of public sector efficiency had long exercised the minds of economic policymakers, but political intervention generally kept it off any reform agenda. However, this changed in the mid 1980s. Early in 1987, the Economic Planning and Advisory Committee (EPAC) presented a precursor document on efficiency in the public sector (EPAC 1987). In part, it erected a policy stage on which more authors of technical papers presented their findings later on. In part it advocated an "in house" approach to the pursuit of efficiency by Government Business Enterprises (GBEs). Some of the EPAC ideas were subsequently formalised by the NSW Government.

At the Federal level, the shift in focus onto the public sector led to the release of an avalanche of reports from various Federal agencies on the matter. These varied in quality and emphasis but need mention below because they hold a mirror to Government thinking on the issue of public sector efficiency and structural adjustment. Also they suggest possible future trends in the development of economic performance measurement. Nevertheless, for various reasons, policy strategies and actions have been found to be hesitant. This is in contrast to the Greiner Government in NSW where action was decisive from the time of its election in 1988 and where performance measurement was essentially commenced soon after the reforms were introduced.

In 1988 the Bureau of Transport and Communications Economics (BTCE) reported on a comparison of overseas railway systems (BTCE 1988). This was a precursor document for railway performance measurement. International comparisons, comprising a few simple non-parametric indicators, for different systems were reported. The data was heavily qualified and led to some unwarranted conclusions. However, it did present a good example of the great difficulties of measuring economic performance of railway systems across international boundaries.

In 1988, the Federal Government launched its Reform Package for GBEs involving restructuring, privatisation, etc. of some of the largest firms in the public sector. The pace of these reforms is only now beginning to gather momentum. The election of the Greiner Government in NSW in 1988 was the main impetus to productivity measurement. The Government had been elected on a platform of public sector reform and, in particular, reform of the State's GBEs. That the GBEs were a major cause of the State's poor financial position had been clearly established by a Commission of Audit in 1988 (NSW Commission of Audit 1988). It was perhaps the sense of looming financial crisis which gave urgency to the GBE (or Government Trading Enterprises, GTEs, as they are known in NSW) reform program. Moy (1993) noted:

> In a general sense, public sector reforms are directed toward improving the principal/agent relationship between voters (principals) and their elected representatives (agents) and, in turn, Government (principals) and public sector management (agents).

In such a setting, financial indicators become paramount. Performance targets are financial ones and the concomitant monitoring program concentrates on key financial variables such as revenue/expenditure, earnings before interest and tax, net profit after tax, dividend forecasts and payments, rates of return on equity and total assets and key financial "ratios" such as interest cover, gearing/debt payback, cash flow and current ratio. The ratios are then compared with financial benchmarks in similar organisations or industry averages (Moy 1993).

Finally, the Special Premiers Conference held in October 1990 accepted the idea of national performance monitoring for GBEs. Six industry groups were identified: rail, urban transport, ports, water, sewerage and drainage, electricity, and gas and pipeline transmission. Sixty three GBEs are participating in the publication of financial, non-financial and economic indicators (Moy 1993). Whilst the priority given to performance measurement of GBEs is to be applauded, a number of issues present themselves for consideration. First, there are the concerns expressed by Copeland and Weston (1988) about financial indicators presented by any business firm and the severe limitations on such indicators. These indicators contain only historical information which may be of little guidance to future performance and they provide only a very limited focus of information for those authorities charged with the responsibility of protecting the public interest of investors, creditors and the like, from future business fraud, mismanagement, etc. In the case of government-owned firms in Australia, there is the added problem that the data is presented in a political arena. This tends to place constraints on data released for publication. It must also be noted that benchmarking of government enterprises has similar problems to benchmarking of private firms.

At the Federal level, productivity measurement may have initially been considered the bailiwick of the Australian Bureau of Statistics (ABS). Certainly, the concern by that organisation for statistical rectitude would place its work well ahead of some other attempts at measurement. The ABS has concentrated upon labour productivity, as measured by an annual estimate of gross product at constant prices per person employed for selected industries (Aspen 1990). However, in 1989 and 1990, it accepted the policy limitations of publishing only macro-economic labour productivity estimates and issued an occasional paper titled *Estimates of Multifactor Productivity: Australia* (Aspen 1990). In commenting upon possible future developments, the paper considered that its future efforts would be concentrated upon improving the data underlying output and input measures rather than changing theoretical models. In particular, the ABS proposed to address the issue of sectoral output measures. Thus, the ABS has largely reduced the macro-economic productivity measurement issue to one of data base improvement.

Meanwhile, the Economic Planning Advisory Council (EPAC) had commissioned ad hoc studies on specific sectors such as the resources sector and primary industries in the late 1980s. Conclusions were tentative stating that many factors impinge on productivity (ABARE 1990). The Industry Commission (IC), as a later player, has concentrated its efforts on measuring the economic "performance" of GBEs. Two measures of performance were used:

- real rates of return, which measure the net return to the capital employed, and
- total factor productivity.

The avalanche of reports then gathered momentum. They are briefly outlined and discussed below, in chronological order:

1. Bureau of Industry Economics

Dowrick (1990) for the Bureau of Industry Economics (BIE) investigated Australian labour productivity growth trends and causes. The study used parametric techniques on macro-economic data in an attempt to develop international comparisons of labour productivity. The results are inconclusive.

2. EPAC

Lansdown (1990), for EPAC, studied public sector efficiency from the viewpoint of a Government Business Enterprise. This presentation showed a series of non-parametric performance indicators (including international comparisons) in an attempt to demonstrate that Australia Post had become more efficient since the Federal Government's implementation of the 1988 Reform Package for GBEs.

3. Industry Commission

The Industry Commission (1990) measured total factor productivity in Australia Post, QANTAS and Telecom. Measures of economic performance were also developed for the Australian National Line and the Australian Water Industry. Apart from TFP, real rates of return were also used as a measure of economic performance. The latter were defined as the net return to the capital employed. This study represented the first major attempt to develop methods and databases on each GBE for future analyses of economic performance.

4. Institute of Applied Economic and Social Research

The Melbourne Institute of Applied Economic and Social Research was commissioned by EPAC to assess four propositions (Dixon and McDonald 1991):

- that debt stabilisation in the mid 1990s would involve a cost in consumption standards in the short/medium term, but would yield benefits in the long term;
- that productivity improvements would make debt stabilisation less painful;
- that if we have more consumption in the short/medium term, there will be less consumption subsequently; and
- that productivity improvements would not directly assist the current account deficit but would improve the growth/balance of payments trade-off.

Orani Model simulations were used to form the basis of the analysis. The hazards of making international comparisons were identified (Dixon and McDonald 1991: 43): where the objective is to gear real wage increases to growth in labour productivity, the lack of an accurate economy wide productivity measure is seen as a serious problem, with neither of the main available measures being considered satisfactory.

5. BTCE

The BTCE (1991) undertook an analysis of total factor productivity with an application to Australian National (Railways). This report attempted to present methodological issues associated with productivity measurement and performance appraisal and an empirical analysis of productivity growth in Australian National Railways in the eight years to 1987/88. As a measure of performance, it concluded that TFP "... will provide a valuable adjunct to financial measures of performance". It noted also that labour productivity and partial productivity indicators had, in the past, been used to prove any case, by "judicious selection". Its analysis of productivity in AN showed that TFP growth had averaged between 5 and 6 per cent between 1979/80 and 1987/88.

6. Industry Commission

The Industry Commission (1992) prepared a set of TFP case studies for a Meeting of the Steering Committee on National Performance Monitoring. In the first study on TFP methods, three sets of performance indicators were considered by the Steering Committee: financial, non-financial and economic. The two major indicators under consideration by the Steering Committee were TFP and the Economic Rate of Return (ERR). The basis for supporting these two indicators as management aids were:

- it helps management to gauge how well the organisation is performing overall,
- it provides information (to GBE management and Government) on how the organisation is performing through time and in relation to its peers,
- it is useful in identifying areas requiring improvement as well as helping to determine appropriate pricing and investment policies.

The paper was particularly supportive of the use of TFP monitoring as a means of assessing progress in microeconomic reform. It provides much information which cannot be obtained from either accounting or non-financial indicators. Extending the performance monitoring framework to include TFP indicators should be considered an important priority (IC 1992). Five case studies of TFP measurements are provided: Australian National, State Rail Authority of NSW, Melbourne Water, Port of Brisbane Authority and Pacific Power.

7. BTCE

In response to a request by the Federal Minister for Land Transport and Shipping Support, the Bureau of Transport and Communications Economics in 1992 investigated the relative efficiencies in the transportation of bulk commodities (BTCE 1992). The Report analyses, in particular, the relative efficiencies of road and rail in the Australian transport of coal, petroleum and minerals, grain, fertiliser and limestone. This request followed the Royal Commission on grain storage, handling and transport which presented its findings in 1988. A model, based on a mixed integer linear program, was developed for each commodity, modelling the lowest resource cost mode for each commodity.

8. Bureau of Industry Economics

In 1992 the Bureau of Industry Economics reported on a study of international performance indicators in rail freight (BIE 1992). In March 1991, the Prime Minister directed the Bureau of Industry Economics to develop international performance indicators for the main infrastructure services in Australia including: rail freight, transport, electricity, telecommunications, waterfront, coastal shipping, water and gas. Univariate partial indicators were developed but the need for a global measure of productivity (i.e. TFP) was recognised.

The Rail Freight Report's method focused on the concept of "World's Best Practice ...the best observed international performance...". Usually these are North American railways. Three broad categories of non-parametric indicators were used: customer oriented indicators (freight charges and service quality), operating efficiency (non-parametric measures of operating efficiency), and "cost structures", which did not include unit costs. An attempt was also made to standardise the comparative data for "non controllable" factors, by adjusting input costs. These factors were identified as: traffic density, traffic mix, terrain and climate. The resulting analysis then aimed to estimate the extent to which Australian systems must reduce "controllable" operating costs to reach achievable world best practice. TFP analysis was not undertaken on the grounds that a number of other organisations were already doing this work.

By 1992, some five years into serious reforms of the public sector, a few tentative conclusions could be drawn concerning the emerging needs for performance measures by government. The NSW and Federal Governments have established new paradigms for other governments to follow. The reforming Governments have pursued financial and fiscal strategies which emphasise deficit reduction, balanced budgets and other "bottom line" financial indicators. This is in line with their broad objective of instilling commercial principles into the managements of GBEs. However, if the views of Copeland and Weston (1988) are to be accepted, then such data has major limitations. At the Federal level, the outcomes from the various studies have been varied. In a few cases like that of the 1990 Telecom Discussion Paper (Lansdown 1990), non-parametric data has been selectively used to make a particular case. At the other end of the scale of excellence, the Industry Commission has vigorously sought to develop an understanding of economic efficiency by emphasising the development of two methodologies, namely the real rates of return and TFP. These measures have now been applied to a number of GBEs and the work is likely to continue. While the issue of the adequacy of data bases has emerged as a major problem, no attention has been given to the potential uses of performance indicators as tools for examining *future* business strategies which, in the end, is the major practical use of any performance indicator analysis.

The third area of development at the Federal level has been the Bureau of Industry Economics' work on international benchmarking. However, there are some serious methodological concerns about this work so far. Nevertheless, potentially, this area of study offers big rewards to those concerned with reforms because such measures provide windows on the "world's best" firms and their activities.

University professionals are also providing useful advice on how economic performance measurement should be developed in Australia. The important contributions are discussed below. Developments in North America are also discussed. For example, there is a benchmarking approach which offers potential benefits (Diewert 1993). Diewert's paper explores a whole range of approaches available to the analyst interested in making international comparisons. Three methods of international comparison are mentioned: index numbers, non-parametric (data envelopment) analysis, and econometrics.

2.2 METHODS OF COMPARISON

2.2.1 Index Numbers

Two types of index number comparisons are identified in the literature, namely bilateral and multilateral.

Bilateral index numbers

Diewert (1993) explains bilateral index numbers as output indexes of production (in quantitative terms) for given vectors of input quantity and input price. These examples are described as the formulas (of input and output) of Paasche's Index, Fisher's Index and Translog.

The explanation is illustrated by a data set aimed at comparing four diesel electric power plants (Nantucket USA, Boulia Australia, Newman Mines Australia, and Aqaba Jordan) for four years from 1987 to 1990. Inputs measured are fuel, lubrication oil, hours of labour, materials and capital. Output is measured in megawatt hours. All prices are in constant 1987 US dollars. Translog, Fisher and Paasche Indexes are then calculated, using the same data set. Diewert's (1993) conclusion is that, in making international comparisons of productivity, both the index number formula and the choice of base unit (or benchmark unit of comparison) are important.

These are important observations, but the validity of the analysis is open to challenge on the issue of the sparseness of the data. First, only four years of data are presented, and no statement is made about relative technical efficiencies of the plants compared or management practices.

Multilateral index numbers

Diewert (1993) then proceeds to a consideration of Multilateral Index Numbers, which are based on averaging bilateral comparisons. Normalised relative productivities for each plant, which are independent of the benchmark unit of production, are obtained by averaging across the productivities obtained for the Bilateral Indexes.

Another interesting approach is the use of the concept of "share of 'world' input and output". Bilateral input and output indexes of one plant are expressed as shares of all plants included in the sample (or "world"). This obviates the criticism that smaller plants are given equal weight to large plants where, presumably, scale economies would influence productivity.

Finally the Geary-Khamis method is described in which plant productivities are expressed in terms of the purchasing power parities of each of the countries in which these plants are located (Diewert 1993). This allows for a "normalisation" of unit costs and output prices across international boundaries. However, Diewert's (1993) general conclusion is that each method yields approximately similar results.

2.2.2 Non-Parametric (Data Envelopment Analysis—DEA)

Diewert (1993) puts forward an approach which uses non-parametric methods for comparing the efficiency of different production units. The case examined is that of a "linearly homogeneous production function with convex, conical or constant returns to scope technology". The relative efficiency of each production unit (known as a Decision Making Unit or DMU) is defined as a non-linear programming model and a scalar measure of efficiency is obtained by solving an equivalent linear programming problem (Hensher and Waters II 1993). Under cost minimisation, the "best behaviour practice" total cost of producing the output/s resides on an efficient, frontier cost function envelope with a value equal to 1. Diewert tested the invariance properties of this method by comparison with the index number methods (Diewert 1993). To do this, the second best and the least efficient plant in the data set is removed from the calculations. The invariance properties of the DEA method are stronger than the index number methods but the former's efficiency estimates are higher. This is probably due to the fact that the least efficient plant has been excluded.

Hensher and Waters (1993) conclude that the choice among TFP Index Number methods and DEA fundamentally rest with data requirements and availability. DEA is parsimonious with data, and this is an advantage where data is hard to acquire or where outputs are not priced. Also, the preferred configuration for a DEA is when cost

information replaces information on physical inputs. Thus, in most cases, only technical efficiency is measured.

These conclusions give some cause for concern with Diewert's results, where only four plants are compared over a period of only four years. Whilst the best and the worst plants are clearly identifiable from a technical efficiency viewpoint, this can be determined intuitively from an analysis of the raw data. From a practical stand-point, it is the upper two quadrants (and the plants in them) that would be of concern to managements interested in establishing benchmarks for their own organisations. Also, the inclusion of only four years of data would be insufficient to measure relative changes in the rankings of plants over time, given that the purpose of any benchmarking is to initiate improvements in efficiency. It usually takes some years to bring about efficiency improvements which would lift a firm's performance towards its benchmark target.

2.2.3 Econometric Methods

These are not covered in Diewert's paper, but a very good overview is presented in Hensher and Waters (1993). These take two forms, one measures a production function and the other, a cost function. The singular advantage of the econometric approach is that (Hensher and Waters II 1993: 13-14):

> ...(it) enables us to embed within a single model the relationship between inputs, outputs and the full range of environmental influences such as institutional and regulatory factors. It enables us to test for the presence of scale and scope and other useful items such as elasticity of substitution between inputs and the extent to which TFP is influenced by the level of output and input prices. The main advantage of the fully parametric approach is that the productivity measure is a derivative of a model which incorporates all the observed and unobserved sources of variation in efficiency across the sample of organisations studied.

The production function measures the technical relationship between outputs and inputs. Traditionally the econometric method has been used to find the line of fit through the *middle* of a data point set. The measure of productivity is the shift between two isoquants from one time period to the next. Production functions identified by the method of stochastic frontier alteration are a richer specification than perhaps this thesis assumes, where the cost function approach is the preferred method. In this latter approach, traditional "middle of the data" regression may be used, although stochastic cost frontier approaches are feasible. An important section of the Hensher and Waters (1993) exposition, for user and practitioner alike, relates to an identification of the practical problems and issues in performance measurement.

First, there are limitations on accuracy. Parametric methods measure error of estimates. Whilst confidence intervals are provided, they can be very wide. This problem can be minimised by rigorous attention to the data and a generous number of observations over time. Non parametric performance measures have no theory of error so that errors are carried into the results.

Second, productivity measures vary with the type of method used even when the data is the same. For example, the use of multilateral index numbers will give a different result to a single index number series. Also, as we saw above when discussing Diewert's comments on index numbers, the choice of formula will also influence results. In practical terms then, great care must be taken when comparing different studies of the same firm.

Third, the higher the level of aggregation of input and output data, the greater the blurring of the linkages between individual inputs and outputs. This blurring can lead to distortions in the results. Thus, subject to data constraints, disaggregation should be pursued as a worthwhile objective.

Fourth, some studies focus upon measurement of intermediate outputs rather than the final output, because better data is available and/or because it is the intermediate output which is of most interest to a firm, is most understood by a firm's management and can be controlled by them. In the context of this thesis, for example, an intermediate output would be seat kilometres travelled by passenger trains. Measurement of intermediate outputs can, however, give misleading results if they are not linked to final output.

Fifth, productivity measurement focuses on quantitative changes, whereas qualitative changes can often have an important bearing on performance.

Sixth, productivity measurement often excludes the measurement of input and output prices which provide an enhancement to the understanding of the firm's productivity. (Indeed, index number approaches assume price equals marginal cost). Input and output price indices, for example, will show how productivity gains are shared between the firm and its customers.

Seventh, productivity measurement methods presuppose that the firm has commercial goals only. Many government enterprises have community service obligations (CSOs) as well as commercial goals. This raises a host of issues which need to be carefully considered within the context of the particular firm.

The work of Diewert and Hensher and Waters is timely in the Australian context since there has been a burgeoning interest in the productivity measurement of GBEs, of

varying quality, in recent times. These authors remind us of the appeal and pitfalls of productivity measurement.

It is perhaps at this stage in the chronology of public sector re-structuring in Australia that the student of economic performance measurement needs to ask: where is this ferment of activity leading Australian policymakers, and, does it matter in the long run (when GBEs will supposedly be on the same footing as private firms)? To enable us to partially answer such questions, we should undertake a brief overview of overseas developments.

2.2.4 Consideration of Economies of Scale, Density and Scope

An important development in recent years in the application of the translog function was made by Caves et al. (1985) in their consideration of network effects and the measurement of returns to scale and density for US Class 1 railways from 1951 to 1975. The database was robust: 820 observations representing 43 different Class 1 carriers. Three inputs were specified: labour, fuel and an aggregation of capital with materials. Capital was measured as an economic cost. Two outputs were specified: freight ton miles and passenger miles, with route miles included in the cost function as a measure of network size, and each of the 43 railways assigned its own firm effect to represent unobserved network effects. The estimating method used was the generalised translog multiproduct cost function which represents the total cost function.

The results supported earlier studies' conclusions of constant or nearly constant returns to scale when haul and trip lengths are fixed and some increase in returns when haul and trip lengths are increased. More importantly, Caves et al. (1985) found, contrary to most previous studies, returns to density increase substantially when correctly measured ie when panel data is used and controls are established for unobserved network effects. This is a clearly important contribution to the study of railway productivity and is dealt with in Chapter 8, section 8.5.3.

2.3 THE MEASUREMENT OF OVERSEAS RAILWAYS' PERFORMANCE

A view is held by some economists that railways have some unique characteristics which make them different from the general run of industrial firms. This is certainly historically true, and may also be true for the future. Nevertheless, despite their unique qualities, railways may still be amenable to analysis (which allows for the unique qualities), and measurement as to their productive efficiency.

The railways in the industrial economies of the world present themselves as an excellent case study in the utility and pitfalls of productivity measurement and the actions of governments and their bureaucracies resulting from their interpretations of these measurements. The same applies to privately owned railways.

By the end of World War II, most railway systems in the industrial economies of the world were either partially destroyed or worn down by wartime requirements. All were over-resourced in human terms, were poorly managed, subject to generally poor industrial relations and generally incapable of meeting competition from other modes.

The privately operated, mostly North American railways were the first to reform their systems. In being the first they also established the paradigm for railway reform throughout the industrialised world. Essentially, this was to focus on labour productivity largely to the exclusion of other factors of production. In the United States, privately owned railways reduced labour by 61 per cent between 1947 and 1972. But the rate of return on investment in Class I railways still only averaged 2.8 per cent in the period 1962 to 1972. Independent studies (US National Center for Productivity and Quality of Working Life 1973) identified many other factors affecting poor performance including poor railway management, work rules which perpetuated jobs that technology had made obsolete, the Interstate Commerce Commission's (ICC) use of regulatory powers to keep open lightly used branch lines and subsidisation of some shippers of bulk commodities, discriminatory property taxes, passenger train deficits and investment in technologies which were already obsolete in Europe. It is particularly interesting to note that this study identified poor management and poor investment decisions as important factors leading to poor performance. The critical importance of these has been examined, at length, in this thesis. Subsequent de-regulation and intensified competition has forced improvements in management performance and generated more innovation.

In the 1970s and 1980s, European Railways have tended to view the issue of productivity as largely one of labour productivity alone (Whelan 1988):

> Railways (in Europe and North America) tend to see improved productivity as a simple equation of reducing the number of employees.

In truth, the simplistic approach of shedding large numbers of employees can achieve spectacular results for labour productivity in the short run. It will do little for other factor productivity. It may also mask poor management practices and poor investment decisions. Organisation behaviour studies of firms in other mature industries which were "downsized" in the great restructurings of the 1980s also show that morale and efficiency of remaining employees is seriously damaged. In sum, and in the long term, the emerging evidence suggests that a short-term, exclusive focus on labour

productivity may cripple a firm's future ability to compete effectively as new markets appear. The following Chapters of this thesis also show that capital productivity is of greater long-term concern. Also, labour productivity can deteriorate again, after "downsizing" is complete.

In Europe, international comparisons of railway performance had been published on a number of occasions for the British Select Committee on Nationalised Industries in 1959/60 and 1977. However, these comparisons were too aggregated and/or built on too partial a basis to be of much analytical use.

Probably the most comprehensive work on rail productivity measurement was commissioned at different times, by British Rail since 1977 (British Railways Board 1980). This applied particularly in the 1980s, as the organisation seemed to be moving inexorably towards privatisation.

In 1980, British Rail published a comprehensive range of partial productivity indicators, physical indicators and financial indicators (British Railways Board 1980). A methodological innovation was to arrange the financial indicators hierarchically (common practice in analysis of financial statements). This was done in an attempt to isolate the impact of staff costs, staff productivity and average train loading on the residual of total costs and total receipts. The study was repeated in 1982 (British Railways Board 1982). By 1986, some 200 indicators were collated, many on a regional and sectoral basis. They were used mainly for internal management and for consolidation into a productivity and performance index for the railway as a whole, for business sector and expenditure categories. They were also published in the Annual Report to measure Public Service Obligation (PSO) Costs (British Railways Board 1986). The exercise is a tour de force in non-parametric indicator analysis, clearly influenced by traditional Westminster-system Treasury methodology. However, it contains all the limitations of financial and non-parametric indicator analysis.

A study for the World Bank used aggregated non-parametric data to compare the performance of 59 railways of the world, grouped into five regions: Latin America, Africa (Sub-Saharan), Emena (Algeria, Morocco, Tunisia, Egypt, Portugal, Hungary, Poland, Yugoslavia, Turkey and Pakistan), Asia and Developed Countries. In addition to its high level of aggregation, it suffered from all the problems of non-parametric data (Thompson et al. 1990).

Nevertheless, on another level, advances are being made on methodology in North American Universities. Whilst railways are a focus for measurement, there is a broader milieu of investigation, which may be traced back to the 1960s. In 1967, a conference

held in Canada considered the measurement problems of economic activity in the service sectors of Canada and the USA, specifically production, employment and productivity. This was in response to growth in the service sectors of both countries and in Europe. Some important issues were raised including the definition and measurement of "output", the relationship between growth in productivity and growth in industry size and the instability of output per man hour at certain points in a business cycle (Fuchs 1969). This study was followed, in 1972, by a Conference on the Measurement of Economic and Social Performance at Princeton University, USA (Moss 1973). Influential papers on performance measurement were presented by Christensen and Jorgensen and Olson.

Adding to the general milieu of concern about the measurement of economic performance outside primary and secondary industries were the debates in North America regarding acceptable levels of government intervention including regulation in industries generally. Older industries such as railways were usually at the forefront of consideration in these debates. And, any question of regulatory interference very quickly led to the questions of measurement of economic performance. This appears to be one major reason why analysts of railway productivity in North America have been in the forefront of research on economic performance measurement. All the major issues of productivity analysis seemed to occur in the railway industry.

At that time in the UK and Europe the issues of economic performance had not yet emerged. In the great nationalisation of industries following World War II, many of the largest business firms had passed into the public sector and were, therefore, subject only to whatever accounting controls were applied by government treasuries. And, as we saw with British Rail, even now, accounting measures of performance are the preferred tools.

In 1980, Caves and Christensen examined the relative efficiency of public and private firms in a competitive environment using the two Canadian railways as a case study. TFP indexes were used for the period 1956 to 1975. The principal conclusion was that because of the competitive environment in which the two railway systems operated, Canadian National (publicly owned) was not inherently less efficient than Canadian Pacific, the privately owned railway (Caves and Christensen 1980).

In 1987, a major study was published on the growth and economic performance of the Canadian transcontinental railways 1956 to 1981 (Freeman et al. 1987). This represented, at the time, the most comprehensive use of TFP indexes to evaluate railway productivity—and indeed the service sector as a whole. It seems to have established the paradigm for future studies certainly in the North American setting. There is also some

evidence that the US Interstate Commerce Commission has engaged consultants to develop parametric measures for US railroad productivity utilising the methods developed by Caves and Christensen (Reebie Associates 1988).

In 1991, Oum and Yu undertook an international comparison of the economic efficiency of passenger railway systems. 17 systems were studied, for the period from 1978 to 1988. Data Envelopment Analysis (DEA) was used to measure the productivity of labour, rolling stock and energy. A Tobit regression model was established to control for the effects of the differences in operating environments such as networks and market size (Oum and Yu 1991).

However, it is in the area of market regulation that a new application for productivity analysis has been developing. It is known as "price level", "price cap", or "incentive" regulation. It was developed in the UK for British Gas, Telecom and the privatised British Airports Authority, as the "RPI-X" Model. Here, a single product firm is allowed to increase its price in line with inflation (as measured by the retail price index) less an amount "X" which is representative of an expected productivity gain. This approach has also been adopted by the US Federal Power Commission, and is used by the US Interstate Commerce Commission (Waters and Tretheway 1991).

In this brief survey of economic performance measurement at the Government level, all analysis has been ex post, that is historical data is used to derive conclusions about the economic performance of a GBE, industry etc. However, as was noted by Copeland and Weston, historical data has limitations in financial decision making about the future. This applies particularly in the rapidly changing economic climate of today. Similar assertions are made by other authors (e.g. Beesley and Kettle 1986). Beesley and Kettle developed their views within the context of an examination of Victorian Railways costs up to 1980. Also, the assertions were made within the context of a need to develop strategies leading to pricing efficiency by the Victorian railways.

The first assertion is that there is no "one time cost" of providing a rail service, in the future. It all depends upon the context within which a strategic business decision is being made about the future: about a particular rail service over a specified period of time in the future. Three points are made to support this idea:

- costs are measurable only with respect to some specific output change in the future,
- the decision to change future output must also involve a simultaneous decision regarding what revenue this output is expected to generate, and
- the final choice of future options will involve consideration of more than one estimate of cost (Beesley and Kettle 1986).

Financial economists would suggest also the application of risk factors to the future output and future revenue calculations. The second assertion is that there is a time dimension to future cost, and its avoidability, if future output is changed (assuming, always, that cost minimisation remains an underlying goal). Most studies assume that future costs can be changed instantaneously, which does not happen in the real world. Future costs need to be determined in accordance with a future timetable for their avoidance or retention. Also, if future costs are being considered, then future input prices and factor productivities also need to be considered. The inclusion of the time dimension has enormous practical importance in an econometric analysis where the analyst is searching for causal relationship. For example, the past switch from steam to diesel traction provides useful insights for future assumptions on technical innovation.

The third assertion by Beesley and Kettle concerns the measurement of costs (and, by implication, factor prices and productivities). These requirements are necessary in the measurement of costs:

1. their relevance, i.e. each strategy for changing future output must have an array of costs attributed to that output change,
2. their correct valuation, i.e. it is not simply the correct values applied to unit prices of factors employed in achieving the proposed output, but it may also include the opportunity forgone in hiring new resources to meet the change in output,
3. their structure be known, i.e. how the costs of the factor of production behave with changes of output in the different sectors of the business. Here again the concept of risk may be included. There are two types of risk associated with future costs: first, that output will have a greater variance around the mean than originally anticipated and, second, that output will not be great enough to bring about a total reduction in costs (or by implication, an increase in factor productivity).

The current literature on financial economics, and Beesley and Kettle's study on railway economics clearly demonstrate that ex post measurement of economic performance is of limited use to economic policymakers. However, if their results are of sufficient detail and disaggregation, these results may be fed into decisions about future strategies relating to output, revenues and costs. This is the real future challenge for economic performance measurement. Strategic modelling is now gaining acceptance in the private sector overseas, as a tool used in the process of strategy making. Only a modelling technique which maintains the rigor of econometrics, driven by an economic theory of productivity, has the capacity to provide the insights needed for effective corporate strategies for government enterprises.

2.4 NEED FOR PERFORMANCE MEASURES IN THE PRIVATE SECTOR

The impetus for the measurement of economic performance generally, and the promotion of parametric methods in particular, has generally come from Governments. This has certainly been the case in Australia. The main reason for this appears to be the central role played by accounting systems (including cost accounting) in the generation of management information. Accounting has its own conventions and discipline for dealing with forecasts, budgets etc. With the advent of corporate, business and strategic planning, management has turned to its accounting systems to provide the information for its corporate and strategic models. The more sophisticated risk analysis and other techniques still tend to focus on cash flows and other accounting identities. Where econometric analysis is used in the private sector, it is usually a skill which is bought for a specific assignment to respond to a government request or to make a submission in a public inquiry. Matters such as total factor productivity measurement, by non-accounting methods, are considered far too esoteric for more than occasional consideration by corporate management in the private sector.

The situation overseas has been a little different, but not much. In North America, industries dealing with regulators, such as the Interstate Commerce Commission, have needed to come to grips with productivity measurement issues. The question remains, for the future, as to the extent and direction in which economic performance measures should be developed by firms in the private sector. There are trends in events which both oppose and support the development of economic performance measurement.

First, there is the general movement towards deregulation of economic activity. Where such deregulation has been comprehensive and quick, there has been an element of chaos in the restructured markets. Events have generally moved too swiftly for even the most sophisticated information systems. The Australian paradigm for this was the deregulation of banking in 1983 (Wallace 1993):

> If anything, the 1980s credit explosion demonstrated the extent to which previous regulatory arrangements had relieved Australian banks, business borrowers and individual Australian consumers from the responsibility of making mature, considered, borrowing and lending decisions.

In the future, however, it is more likely that deregulation will occur at a more politically constrained pace and be monitored more closely by the government. This is a function of politics in Australia. Of necessity this means that the Government will need to have at its disposal much more sophisticated tools of analysis than hitherto.

Second, the fundamental economic problem which is emerging in the 1990s in Australia and other developed economies, is the ability of all sectors of the economy to achieve acceptable levels of efficiency so that the economy as a whole is able to generate sufficient export markets and improvement in other sectors to bring about a sustainable economic growth. This is likely to shift the focus of consideration away from labour productivity, in isolation. Instead, concern is likely to focus on the performance of other factors of production, capital in particular. But apart from cost efficiency, there is also a need to monitor pricing efficiency and the efficiency of resource allocation. If the private sector cannot match government policymakers in strength of case or cogency of argument then it is likely to have government solutions imposed upon it.

Third, the restructuring of the government sector itself will create an important need for monitoring all aspects of efficiency. Some GBEs will be privatised and others will be opened up to competition. Managements of all firms will need to carefully monitor progress. As explained earlier, accounting systems are not adequate for monitoring economic efficiency changes to the level of sophistication which will be necessary.

Fourth, there are indications that the primary, secondary and tertiary sectors of the economy may see a more pro-active role by Government towards the promotion and the development of some industries. Again, it requires the managements in these industries to have a reasonably sophisticated understanding of economic efficiency and its measurement.

2.5 CONCLUSIONS ON THE NEED FOR PERFORMANCE MEASURES

It is becoming very clear that, in Australia in the 1990s, the overwhelming issue is one of economic survival, i.e. the achievement of acceptable levels of efficiency which will lead to export growth and sustainable economic development. Industry markets and Governments need to develop a more sophisticated understanding of productive efficiency.

There are basically two valid and useful ways of measuring TFP. The first is the non-parametric, index number approach wherein productivity is measured using data on inputs and outputs derived from the historical records of the firm. Provided the rules of aggregating inputs and outputs are observed, valid measures of TFP are achievable. The second measure of TFP is parametric or statistical requiring the estimation of a production or cost function which recognises a number of variables which influence performance.

But all performance measures tend to fall short of managerial needs on two counts. First, they are always ex post. However, if they are well constructed they will be potent diagnostic tools for future corrective action. Second, they are often presented "in vacuo". This can be overcome by analysing them in a corporate strategic modelling context.

CHAPTER THREE

The Quadrae Corporate Model: A Context for Measuring TFP

3.1 INTRODUCTION TO THE MODEL

In Chapter 2, a strong case was made for estimating TFP in order to obtain a meaningful measure of the efficiency of the firm in any industry. This Chapter expands on this theme by postulating that TFP is one of three key measures of a firm's performance in economic terms: productive efficiency (measured by TFP), marketing effectiveness (measured by pricing efficiency criteria), and resource utilisation efficiency (estimated by a measure of total economic gain/loss).

The three measures are brought together, conceptually, as a "corporate" model (i.e. a model of the business firm) called the Quadrae Model, because it consists of four distinct, but linked elements. Because the aim of this thesis is to develop the notion of TFP analysis in a case study of NSW rail passenger services, the other notions of marketing effectiveness and resource utilisation efficiency are not developed beyond this Chapter. Nevertheless, they have a solid basis in micro-economics and several suggestions are made as to what further useful research may be done in those areas. Thus the Model is contextual only, at this stage.

The term Quadrae refers to a pictorial representation of the model of the firm which comprises four interlinked triangles even though these have no geometric significance. The term is derived from the medieval Latin "quadra" which forms the base of many words in the English language and describes concepts containing four parts (*The Concise Oxford Dictionary of Word Origins*, Hoad 1986).

Another purpose of Chapter 3 is to set down a philosophical basis for the Quadrae Corporate Model and to place it within the context of economic and management theory in general. A statement also needs to be made about the choice of the case study: rail passenger services run by the NSW Government (i.e. a public utility) as a heavily regulated monopoly, but manifestly interfered with by all kinds of pressure groups and lobbies. On the one hand, this makes the NSW Railways a much more complex study

than the incorporated industrial firm, financial corporation or trading house, run for private profit. As noted in Chapter 1, railways do have some unique characteristics. On the other hand, its activities over a long period of time have been a matter of the public record. This record has been tinkered with at times, and history revised to obscure the truth. But the quality and wealth of gathered data enables the anomalies to be identified and adjusted.

The Quadrae Model, when fully developed, will contain some unique features:

1. It will be firmly anchored in the traditional theory of the firm.
2. It will synthesise five major themes in microeconomics to produce a paradigm for assessing the economic performance of the firm:
   - cost efficiency as measured by TFP,
   - pricing efficiency where a suitable measure is still to be determined,
   - utilisation efficiency which may be estimated by a measure of total economic gain/loss,
   - management quality examined as the organising element in the firm, and
   - innovation which is seen as the principal determinant of the firm's growth.

The Model borrows from two other disciplines for its view of things. First, systems theory, which enables a modeller to view any firm as a system, which may then be disaggregated into its component sub-systems, processes and sub-processes. Second, that the Model becomes an integral tool for the application of Total Quality Management (TQM). It will be contended in this thesis that TQM and related disciplines are the only management methods that systematically focus attention upon the manager in her/his innovating function to improve productivity in the first instance and then to expand this to improve pricing efficiency and utilisation efficiency. Also, TQM and related disciplines focus attention upon measurement of management actions and their consequences. If this thesis imparts anything, it is that TFP analysis requires scrupulous attention to data. Nevertheless, a great deal of additional research would be required to take the Quadrae beyond its current contextual stage of development.

3.2 THE FIRM AS A SYSTEM

Gordon Fielding rightly perceived the transit organisation (and, by implication, any firm), in the context of systems theory (Fielding 1987). As Weinberg (1975) puts it:

> Some have said that the general systems movement was born out of failures (i.e. inability to keep pace with the second order effects of their first-order victories) of science, but it would be more accurate to say that the general system is needed because science has been such a success...they have revealed a complexity with which they are not prepared to deal.

Or, as Paul Davies (1992) has put it in relation to the history of science:

> The study of computation has enabled us to recognise that the world is ordered both in the sense of being algorithmically compressible, and in the sense of having depth. The order of the cosmos is more than mere regimented regularity, it is also organised complexity, and it is from the latter that the universe derives its openness and permits the existence of human beings with free will. For three hundred years, science has been dominated by the former: the search for simple patterns in nature. In recent years, with the advent of fast electric computers, the truly fundamental nature of complexity has been appreciated. So we see that the laws of physics have a twofold job. They must provide the simple patterns that underlie all physical phenomena, and they must also be of the form that enables depth—organised complexity—to emerge. That the laws of our universe possess this crucial dual property is a fact of literally cosmic significance.

The firm, and the markets in which it operates, may be looked at as analogs of the universe even though they are creations of the mind of man. Nevertheless, the theory of the firm enables us to view the activities of the business enterprise through the broad spectrums of TFP, price theory etc. Systems theory and the practices of TQM will allow the essential complexities of business behaviour to emerge, be examined and measured.

Fielding identifies a number of good reasons for using systems thinking (Fielding 1987). First, it characterises the corporation as a unified system of interrelated parts. And, if we are to accept the statements of scientists such as Weinberg and Davies (Weinberg 1975; Davies 1992), none of these parts will interact linearly with each other (although Davies (1992) asserts that some physical systems are so infinitesimally non-linear that a close approximation may be given by a linear explanation).

Second, systems analysis enables the organisation to be described in its historical and socio-political context, making the influence of government and the community on its operations readily identifiable. Third, it enables a distinction to be made between service production and service consumption.

An introductory statement also needs to be made about the choice of the case study: rail passenger services run by the NSW Government.

Figure 3.1 describes the case study firm, NSW rail passenger services, as a system. For the most part of the study period 1951/52 to 1991/92, rail passenger services have been provided by the monolithic NSW Railways: in reality a "federation" of functional groups organised into branches providing specific skills for the running of railway trains. This basic structure had existed since the 1860s and had been borrowed from the British Railways which in turn had borrowed from the British Army. The NSW rail passenger services system had first to be abstracted out of the monolith. The details of this abstraction are contained in Appendix 2.

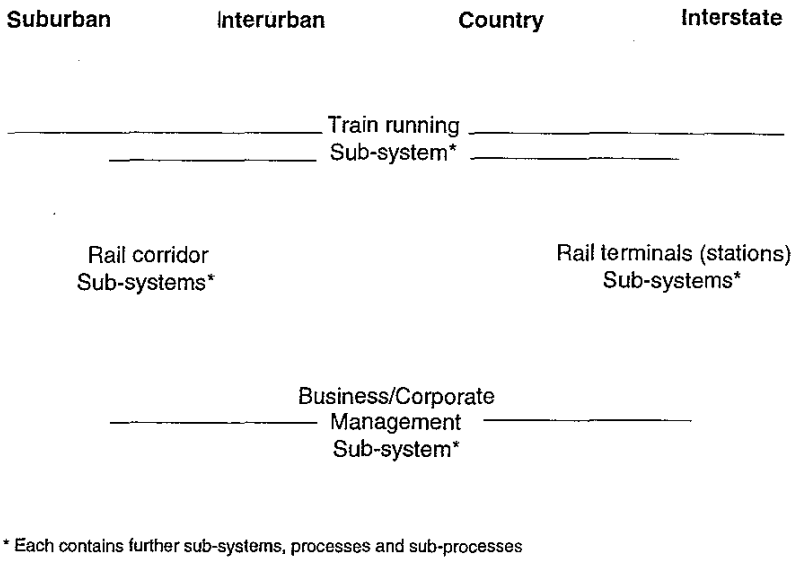

Figure 3.1 The NSW rail passenger services system

Once this abstraction is done, we are able to describe passenger services as shown in Figure 3.1, namely four major sub-systems comprising train running, rail corridor, rail terminals (and stations) and business/corporate management. These sub-systems come together to provide a service to four distinct segments of the market, namely suburban, interurban, country and interstate.

The train running sub-system involves functions associated with the crewing and operation of trains between their origin and destination and with the maintenance of the rollingstock used in these operations. The rail terminal (and stations) sub-system involves functions associated with the staffing and operation of terminals, stations, yards and signalling installations but not their maintenance. The rail corridor sub-system involves functions associated with the staffing, operations and maintenance of the permanent way (i.e. railway lines, etc.), the electrical reticulation system and all buildings and other infrastructure. The business/corporate management sub-system involves functions associated with the management of rail passenger services.

Within each of these sub-systems, there are hundreds and, perhaps, thousands of processes and sub-processes which comprise the sub-systems. It is possible, with some effort, to "map" these, to cost them, and to determine what cost efficiencies may be achieved given inputs and outputs.

Within each of the sub-systems, groups of activities are carried out which have a functional linkage. Engineering activities, for example, will be found in all four sub-systems, but will tend to dominate the rail corridor sub-system; railway operations

functions dominate train running and terminals (stations); and commercial functions dominate the management sub-systems. Each functional group has its own culture and ways of creating information.

A railway system is a complex interlocking of engineering, train running and commercial activities. On a day-to-day basis, a vast range of activities are undertaken. Every train or carriage moved in the system generates a record as does every ticket sold and every freight consignment shipped. Since their beginnings in the 19th Century these activities have created recorded information. Until the 1970s, the records were manually compiled. Manual compilation has several drawbacks. Data entry onto a form or rough pad is often undertaken under stress such as the sale of tickets in morning peak times. Data entry may also be undertaken at night or in wet weather. The data may also be subject to a number of collations and transfers from one document to another. Each time this is done manually, the chance of error is increased. Management sometimes imposes arbitrary computations and re-definitions on data, without documenting the fact. When the manager moves to another job, the reasons for that person's decisions are lost and the data remains biased or in error. As the data becomes progressively more summarised and more readily accessible to public scrutiny, so greater pains are taken to sanitise the data. In the past, this has led to censorship or presentation in an ambiguous form.

The art of analysing railway data is to understand its compilation process and to assess its relative freedom from tampering over the years. When tampering has occurred, the analyst must patiently unravel the data and re-constitute it in the form required. The data then becomes the analyst's best estimate and stands until more information emerges to either verify or refute the estimate. Generally speaking, however, errors and biases have not been so widespread as to invalidate broad comparative analysis over long time periods.

The railway system can also be viewed as an integrated set of engineering functions which enable vehicles to operate safely to deliver people and goods from an origin to a destination. The integration and interlocking of the functions to produce safe transport always has, and always will, make a railway a complex engineering system. Apart from its complexity, other factors have influenced the way in which information on the engineering activities of the railway has been collated and reported.

First, the railway is publicly owned, so the engineers must report annually that they have spent money wisely in maintenance and new construction. Until the major reorganisation of the 1980s, each branch of engineering activity (civil, for the permanent way and other infrastructure; mechanical, for the rollingstock, maintenance

depots and workshops; electrical, for electrical infrastructure and supply of electricity; and signals, for communication) has maintained a relative exclusivity about its activities. This has meant that one engineering branch was readily distinguishable from every other branch. Also, railway accounts, like most public authority accounts until the 1980s, were established on a cash basis.

Consequently, expenditure was recorded as a set of relatively straightforward cash disbursements classified under a generally logical Chart of Accounts. However, this accounting system did have major drawbacks including, in particular, that it was not amenable to cost accounting and other management accounting practices. Nevertheless, the data at source is in sufficient detail to provide a robust framework for developing total factor productivity indices. Additionally, between 1960 and the 1980s, all railway systems adopted a relatively standard Chart of Accounts, making comparative analyses easier.

Second, because railways were a major engineering enterprise in the public sector, annual reporting contained general information about the latest assets acquired. Because the Australian political culture has always needed reassurance that government is technologically progressive, railway engineers have been encouraged to demonstrate their capacity to build, use or maintain technologically advanced railway equipment and machinery. Traditionally, however, railway engineers have been "risk averse" to both bad news and good news about the assets for which they are responsible. The consequences of this are a tendency to down play both bad and good news, in case the good news is only temporary, and the asset in question is really a "lemon". Also, the performance of assets in operational conditions can be at considerable variance with nominal performance data supplied by the manufacturer and others. The consequence of these factors is that annual reports contain the blandest of quantitative engineering data. This requires the analyst to have considerable background information to interpret what is written in the annual reports and other archival documents. *Third*, a great deal of technical data is collated by engineers for distribution amongst their peers. Very little of this is filtered out to non-engineers.

The railway system can also be viewed as an integrated set of train operating functions. The actual operation of trains was traditionally the responsibility of a Branch usually known as the Traffic Branch. In particular, the officers of this Branch have been concerned to have the trains run safely to predetermined timetables. And, when things have gone wrong, to respond effectively to the restoration of the orderly running of trains. In a system as complex and inter-related as a railway, things always go wrong. Consequently, train running is a minute to minute proposition. Over the years, a

massive information gathering process has been developed to record train running. It has had two inherent deficiencies. First, it has relied on manual clerical effort, usually after the event, at the end of a shift. Second, it is geared to the assignment of responsibility for "out of course" events, that is, the assignment of blame for things that go wrong. Both factors have had the effect of distorting the veracity of operational data. Some operational data, such as train kilometres, however, tends to be robust statistically if it has been collected over a long period of time in as raw and untampered a form as possible. It is especially robust where it can be verified from engineering data. The use of sophisticated computer systems, to date, has not solved the two fatal flaws of operational data collection naivety, clerical procedures to prepare the data for entry and the assignment of blame. The computer has made it difficult to trace data errors and rectify them. It is only when operational data can be collated in "real time", without human intervention, will that data be usable by analysts without concern for its accuracy.

Finally the railway system can also be viewed as a set of business activities. Since their inception in the nineteenth century and until the 1980s, railways have operated some of Australia's largest businesses on a Colonial Treasury accounting system. This system is a cash-flow system accounting for every cent earned or disbursed. Very effective auditing systems have been introduced to ensure that very little defalcation can occur undetected.

Whilst accounting conventions used in the past have distorted some elements of the financial position and whilst key data has sometimes been deliberately omitted in a particular year, a "golden thread" of consistent records of financial transactions runs through the Annual Accounts of railways, year after year. The art in making sense of the financial data is to identify the "golden thread" in the notes and qualifications to accounts.

There are many limitations to the traditional accounting systems as a means of providing full information on the business health of railway systems. These were highlighted in the Report of the NSW Commission of Audit (1988). The wealth of financial data contained in the Annual Accounts of Earnings and Working Expenses, supported by archival material, is an adequate starting point for securing the global financial statistics needed for the construction of input and output indices. The analyst is confronted by a whole range of database issues arising from accounting conventions adopted in the past and present. A challenge of the current study is to unravel some of the inconsistencies in definition both across systems and over time.

Rail systems must carefully measure three aspects of their business decisions: the effect of their decisions upon productivity (i.e. cost efficiency), the effect of their decisions upon service effectiveness (measured by service quality and pricing efficiency), and the effect of their decisions upon the efficient utilisation of economic resources such that the economic benefits provided by services exceed the economic costs imposed on the community. The present study emphasises the first dimension of the decision process. In the case of service effectiveness, the emerging discipline of total quality management (TQM) has the potential for introducing measurement precision. There is a large body of literature on pricing efficiency but this needs to be integrated with cost efficiency and utilisation efficiency which is beyond the scope of the present study.

The concept of the "firm as a system" has two major practical advantages. First, it facilitates the application of costing down to the minutest sub-process and therefore permits a "bottom-up" approach to TFP measurement. Second, it facilitates the use of monitoring of TQM measures designed to improve productivity.

3.3 THE NEXUSES BETWEEN THE QUADRAE MODEL, TFP AND TQM

This thesis follows Alfred Marshall's injunction that the study of economics should be "largely directed by practical needs" (Marshall 1920). Therefore, it seeks actively to provide meaningful nexuses between: the more useful business management practices (such as TQM, and there are others) aimed at enhancing productive efficiency, an insightful model of the firm and a practical and useful key measure of the firm's economic efficiency.

Some analysts and practitioners of TQM tend to put the management technique within the context of an essentially Eurocentric cultural setting. Thus, for example, where TQM is applied to an American firm, American values may be stressed to employees required to implement TQM. Likewise, Australian practitioners seem to stress the notion of teamwork drawing on sporting analogies and Australia's international sporting prowess. The danger of this approach is that many of the subtleties of KAIZEN, and its application in other Asian settings, apart from Japan, will be missed. KAIZEN's fundamental philosophical roots are in a universal, pre-industrial craft workshop where continuously improving production techniques created the foundations for all industrial revolutions which were to follow. We are just now gaining an understanding of the fact that Japan and other Asian countries possessed these pre-industrial craft skills in abundance much before the Western world.

KAIZEN and TQM seem to offer a useful philosophical and analytical basis for understanding productive efficiency. Perhaps one of the best interpretations of the historical context in which the Asian craft ethic has developed over the centuries is the Toynbee edition of the history and culture of China and Japan (Toynbee 1973).

3.4 THE FIELDING MODEL

Figure 3.2 shows the Fielding Model of a public transit corporation, as it might be seen for the purposes of measuring its performance. It is one analog of how a transport corporation might be seen to be performing. It is, therefore, an abstraction and not a representation of full reality. It is depicted as a triangle.

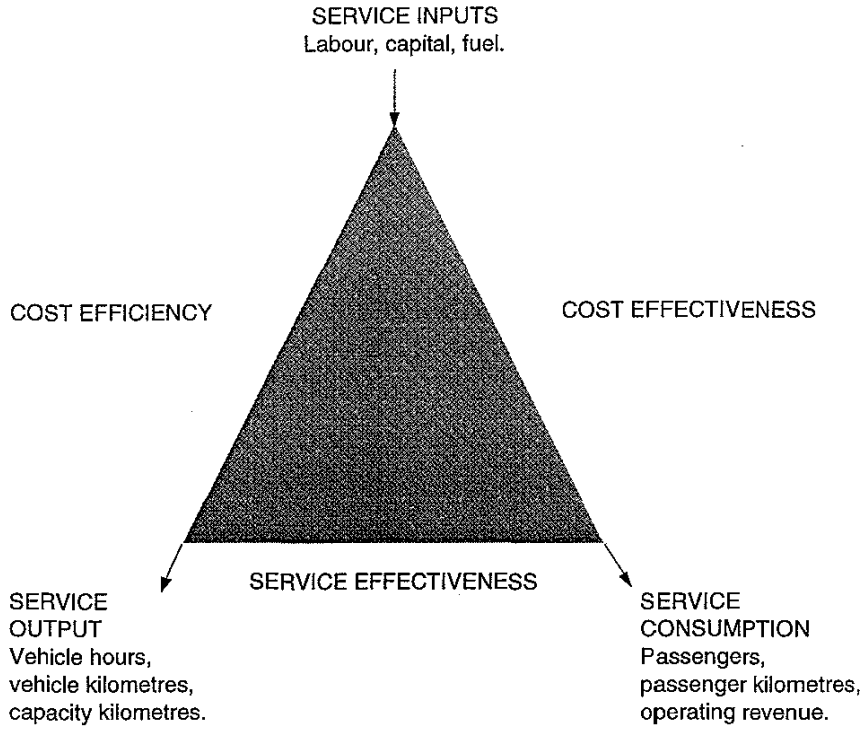

Figure 3.2 The Fielding model of transit performance

The points of the triangle show three types of statistics which Fielding considers should be used to calculate transit performance, namely service input, service output, and service consumption. Each of these points is linked by a side of the triangle, called cost efficiency, cost effectiveness and service effectiveness. Each side then depicts the relationship between the connected points of the triangle. For example, cost efficiency

depicts the relationship between input and output, service effectiveness measures the consumption of service outputs, and cost effectiveness measures the relationship between the cost of producing the service and consumption (Fielding 1987).

A most important distinction is made between "efficiency" and "effectiveness", both terms being used in their economic context. "Efficiency" shows how well factors of production are used to produce outputs. "Effectiveness" measures the consumption of transit output as well as the impact of transit on societal goals such as reducing traffic congestion. Overall indicators integrate efficiency and effectiveness measures as when costs of service inputs are related to consumption. Cost per passenger and the ratio of revenue to the cost of producing the service are overall measures. Cost efficiency, service effectiveness and cost effectiveness are the terms used to describe the three dimensions of transit performance (Fielding 1987).

Fielding identifies three types of data, which are readily available in the US, under Section 15 of the Urban Mass Transportation Act. The three are:

1. *Service input data.* These measure the quantity of resources used to produce a transit service and may either be expressed in monetary or non-monetary units (e.g. employee hours, fuel cost).
2. *Service output data.* These measure the quantity of service produced by a transit firm expressed in non-monetary units (e.g. vehicle kilometres, capacity kilometres).
3. *Service consumption data.* These measure the quantity of service actually used by customers and may either be expressed in monetary or non-monetary units (e.g. fare-box revenue, passenger kilometres).

From these three types of performance measures, forty eight measures were calculated from Section 15 data (Fielding 1987). It is not the purpose of this Chapter to cover Fielding's research which is well documented in his book. However, one dimension of his model needs to be looked at in some depth, namely "service effectiveness". This links service outputs and service consumption. Service consumption is measured as passenger trips. Either vehicle hours or vehicle kilometres may then be used as the denominator in ratio measures of service utilisation. Statistics of revenue generated are also used to measure service consumption.

Three further groups of measures are included by Fielding under service effectiveness, which are as follows:

1. *Operating safety.* Fielding regards it as a difficult decision whether safety is an efficiency or an effectiveness indicator. He decides upon the latter noting that operating safety is related to the way in which service is provided, but it also influences consumption.

2. *Public assistance.* These measures provide a way to assess equity in the distribution of funds as well as the efficiency of assistance in terms of vehicle hours produced or effectiveness in terms of passengers carried.
3. *Social effectiveness.* These measures indicate how well the transit service is utilised by the intended beneficiaries.

The third element of the Fielding Triangle is labelled "cost effectiveness" which is defined as measuring the relationship between the cost of producing a service and the consumption of that service. This creates a logical inconsistency for anyone trying to generalise the Fielding approach to incorporate measures of the broader economic efficiency of a firm's activities. The inconsistency is created by the fact that one point of the "cost effectiveness" side of the Triangle measures service consumption (passenger numbers, passenger kilometres etc.) whilst the other measures service inputs (labour, capital, fuel). However, when generalised to the Quadrae, the cost efficiency side of the Triangle measures output in two ways—as a "supply side" output (seat kilometres) *and* as a "demand side" output (passenger kilometres). Both are significant. Consequently both are shown as outputs in Figure 3.3 depicting the Quadrae.

A further inconsistency is created by the concept of "cost effectiveness" and its definition. Fielding (1987) opts for a very narrow definition where:

> If a single overall cost-effectiveness measure is required, then operating deficit per passenger or subsidy per passenger is preferable. These measures approximate what economists consider to be minimisation of cost economic efficiency. A more positive measure is the additional trips per dollar of net cost.

Anyone familiar with government policymaking in transport or indeed in any major economic activity will readily understand that the measure of a deficit or a subsidy of a transport service is much too narrow a definition of efficient (or inefficient) utilisation of resources. For example, in this day and age of relatively integrated policy making, impacts on employment, regional economic growth, GDP and balance of payments, have to be factored in. Recognition of elements such as this creates a further logical problem for generalisation of the Fielding Triangle, namely while measuring a broader concept of "cost effectiveness" it is useful to weigh service consumption against service input costs. *It is, therefore, critical to measure service output against service input costs.* This is revealed in later Chapters of this thesis where a gross level of overservicing of off-peak rail passenger service is discussed.

The conclusion suggested is that too strict an adherence to the Fielding Triangle, elegant as it is, leads to logical blind alleys, when an attempt is made to generalise the model and its application to public transport services. Nevertheless, the Triangle is a useful conceptual building block.

3.5 THE QUADRAE MODEL

Figure 3.3 shows the Quadrae Model, in its general case of any private or public corporation producing goods and/or services. As in the case of the Fielding Model, it is an analog of how the corporation might be seen to be performing. It is, therefore, another abstraction and not a representation of full reality. It is depicted as a quadrae of four triangles, with the Fielding Triangle as its foundation triangle (Figure 3.3). The three new triangles are numbered HF1, HF2 and HF3.

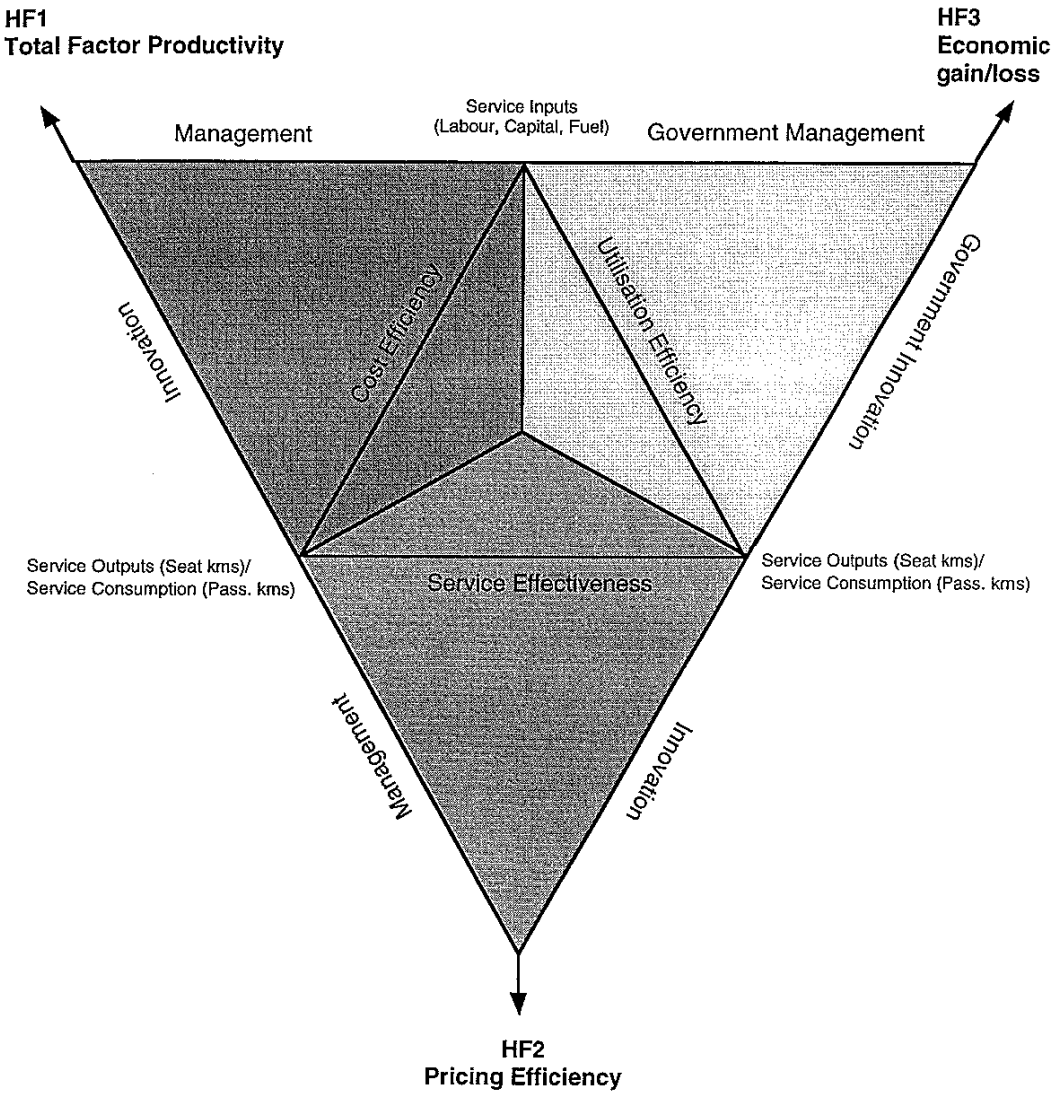

Figure 3.3 The Quadrae

3.5.1 Triangle HF1

Triangle HF1 measures the cost efficiency of the corporation. As with the Fielding Triangle, it links three dimensions of corporate performance, namely cost efficiency, management quality, and innovation. Cost efficiency depicts the relationship between input and output. We may also use the traditional measures of service output or, in the case of a manufacturer, goods output. Again, these should be expressed in non-monetary terms, wherever possible.

HF1 also adds two further dimensions to the basic Fielding Triangle, namely management and innovation. Management effectiveness depicts the relationship between service inputs and TFP. Innovative efficiency depicts the relationship between service outputs and TFP.

TFP measures the productivity of all heterogeneous inputs to the production process (Hensher et al. 1992). Its strength lies in its ability to recognise that many heterogeneous inputs are used to produce a number of heterogeneous outputs. Of particular importance is the following (Hensher et al. 1995):

> Two broad empirical approaches have been used in measuring TFP: the non-parametric index number approach and the parametric econometric cost model approach. Although the latter is more general, it is more complex. All empirical studies undertaken by government inquiries in Australia have opted for the simpler Index number approach. Inputs and outputs are aggregated and compared using an index weighting procedure based on an underlying economic theory of index numbers. This is not an arbitrary procedure. It is based on an acceptable economic interpretation of cost efficiency as capturing the way in which multiple inputs are combined to produce multiple outputs at minimum cost. We adopt this approach in this report with ongoing research extending into the parametric approach.

Importantly, for the expanded Quadrae Model used in the Australian Railways productivity study, managerial and technological changes have been identified as having a "statistically significant impact on the variations in gross TFP" (Hensher et al. 1995). Inclusion of "management" and "innovation" as the dimensions which link cost efficiency with TFP, in the general case, requires further comment.

Management

In the Railways TFP Study, it was clear that the appointment of some leaders (be they "Chief Executives", "Commissioners" or "Chief Commissioners") had a major impact on corporate performance, whilst others had a minimal impact. This study confirms that management practices are both identifiable and measurable. But, in the general case of the Quadrae Model, it is useful to explore the many facets of management. One of the most important insights into management is provided by Penrose (1959: 31-32):

> Entrepreneurial services are contrasted with managerial services which relate to the execution of entrepreneurial ideas and proposals and to the supervision of existing operations. The same individuals may, and more often than not probably do, provide both types of service to the firm. The management of a firm includes individuals supplying entrepreneurial services as well as those supplying managerial services, but the "competence of management" refers to the way in which the managerial function is carried out while the "enterprise of management" refers to the entrepreneurial function.

Penrose acknowledges, however, that the latter is a "slippery concept, not easy to work into formal economic analysis, because it is so closely associated with the temperament or personal qualities of individuals".

A number of entrepreneurial qualities are then considered, including:

- "entrepreneurial versatility" involving imagination and vision on the part of the firm's leadership. This is not to be confused with the "vision statements" so popular in current management writings and which tend to reflect the confusions afflicting so many business leaders.
- "fund-raising ingenuity" involves the ability to attract capital and ensure the continued growth of the firm.
- "entrepreneurial ambition" involves management leadership being possessed of the capacity to make the firm grow in line with its long term vision of the future.
- "entrepreneurial judgement" implies the ability to make correct business decisions.

Another notion explored by Penrose is that of the "management team" where management of the firm has experience of working together as a unit. The accelerated growth of the firm can lead to the overloading of the management team resulting in poor quality of management.

These notions typify the general perception of Western capitalism of management as an elite, "officer class" of corporate personnel who are charged with the responsibility of being entrepreneurs/business administrators. But it is antithetic to the perception held by the KAIZEN Movement. This is illustrated in Figure 3.4. Here, "innovation" ("entrepreneurial activities", in the Penrose definition) continues to be undertaken by the relatively small group, "Top Management". Some entrepreneurship (innovation) is also delegated to middle management. The remaining entrepreneurship (improvement and innovation) is exercised by all levels of the corporation in KAIZEN: the continuous improvement and maintenance of standards. As Imai Masaaki (1986) puts it:

> The higher up the manager is, the more he is concerned with improvement. At the bottom level, an unskilled worker working at a machine may spend all his time following instructions. However, as he becomes more proficient at his work, he begins to think about improvement in the way his work is done, either through individual suggestions or through group suggestions.

Once the improvement has been implemented and established as a standard procedure then the maintenance function of the corporation takes over (Penrose's "managerial services"). In the maintenance function, the task of management is to establish the policies, rules, directives and procedures to ensure that the corporation does not backslide into its old ways of doing things.

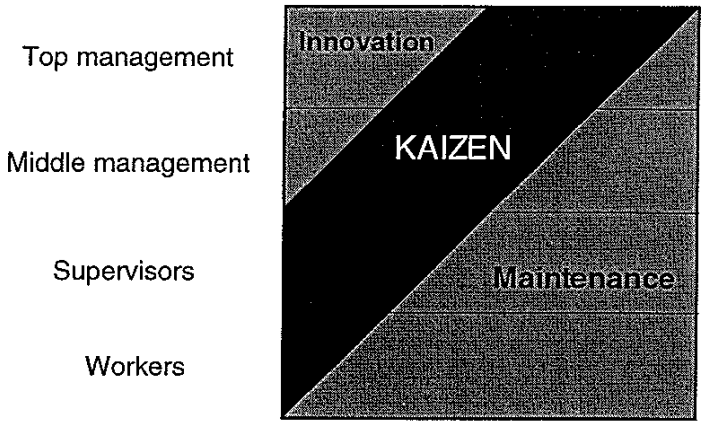

Figure 3.4 KAIZEN in the management process of any corporation

Over the long haul, innovation and improvement tend to disappear from the corporation and fellow stagnated corporations follow each other like lemmings in succeeding waves into the sea to drown. Over the long haul only a process such as KAIZEN can provide the firm with constant invigoration.

The introduction of the KAIZEN method into the management process seems to make for a much more systematic measurement of management quality. The statistical quality control techniques used to measure changes in the operating or other processes of the corporation can be applied at all levels of management. This eliminates the element of chance and differentiates the good manager from the bad manager blessed with occasional luck.

Innovation

In the Railway TFP Study, it was clear that some technological change had the potential for a major impact on productivity improvement. Some had positive impacts whilst many had negative impacts. The important thing in performance measurement is to identify the positive from the negative.

The term innovation best describes the linkage between TFP and outputs. It describes the whole process of change of products or services wherein every facet of the corporation may be altered. The KAIZEN Movement makes a particular point of this change, thus (Imai Masaaki 1986):

> Western management worships at the altar of innovation. This innovation is seen as major changes in the wake of technological breakthrough, or the introduction of the latest management concepts or production techniques. Innovation is dramatic, a real attention-getter...innovation is generally a one-shot phenomenon.

Again, the gradualist, marginalist and measured, approach of KAIZEN is antithetic to the Western capitalist notion of innovation. Nevertheless, the term "innovation" is preserved as a descriptor of the linkage between TFP and outputs. Whilst the bulk of this thesis is devoted to coming to grips with the elements of Triangle HF1 in Figure 3.3, it is important to outline the concepts behind Triangles HF2 and HF3. These can only be briefly outlined at this stage since it is TFP measurement which is of prime concern.

3.5.2 Triangle HF2

HF2 measures the marketing effectiveness of the corporation. As in the case of the Fielding Triangle, marketing effectiveness depicts the relationship between output and the consumption of that output in the market place. Again, such consumption may be expressed in either monetary or non-monetary terms.

In the case of HF2, the apex of the triangle represents pricing efficiency. This can be handled, in index number format, in the same way as TFP i.e. as an output price index. However, it may be more useful to explain the possibilities of setting up a parametric econometric price model. But, this is for the future.

As in HF1, two similar dimensions link pricing efficiency with output and consumption. First, the management dimension links output with pricing efficiency. Here, all the skills and quality of management come into play. As entrepreneur, the manager must ensure that the pricing of goods/services brought to market are such as to ensure sufficient penetration of markets to enable a sufficient return on investment in the long run. Thus entrepreneurial versatility, fund-raising ingenuity, ambition, and judgement are all brought into play in the decisions leading to the pricing of output. This activity complements the equally important issue of input pricing by management, covered in HF1. The maintenance function ensures that marketing plans are adhered to, sales targets achieved, etc. Again, the introduction of KAIZEN methodology into the management process would make possible the more systematic measurement of

management quality to effectively carry out corporation's activities. Second, the innovation dimension links consumption and pricing efficiency. In the KAIZEN terminology we would refer to it as improvement in quality in the goods/services offered at the price which customers are willing to pay.

3.5.3 Triangle HF3

As was noted above, in the comments on the Fielding Triangle, the exigencies of relatively integrated economic policymaking now require that major activities such as provision of transport services be measured for the efficiency with which they utilise economic resources, to determine whether the provision of service results in an economic gain or an economic loss. Also, whilst it is important to measure service consumption, *it is imperative to measure service output.* This is particularly so in complex markets such as public transport where overservicing is prevalent and caused by many factors.

The apex of the triangle is labelled "Economic gain/loss". This is deliberately broad so as to avoid terms such as "consumer surplus" or "social benefit/cost ratio". This general label allows for the broadest definition and avoids argument at this point in the development of the Quadrae, of what should and should not be included in weighing the benefits of service provision against its costs.

As in the case of the other two triangles, two further dimensions link service input and service output. First there is the management dimension which has been labelled "government management" and covers the whole gamut of intervention including regulation, taxation, subsidisation, etc. Whilst all government management initiatives would, in theory at least, be aimed at maximising the utilisation of resources, it is a most imperfect instrument for the undertaking of this task. Thus, at the very least, it would be useful for the corporation to know if it was winning or losing as the result of the impingement of government management upon its activities.

Second, there is "innovation" which is government sponsored innovation aimed at expanding the efficiency of corporations. This innovation may take the form of major improvements in infrastructure or it may be such far-sighted ventures as the implementation of Total Quality Management principles in the NSW Department of Education aimed at lifting the quality of education of the State's future workforce and consumers. As we have seen, some previous attempts at innovation in Victoria ("Victoria Inc") and Western Australia ("WA Inc") led to the establishment of exclusive "clubs" of selected politicians and businessmen that resulted in corruption on a grand

scale and, in the case of Victoria, the serious undermining of the State's economy. Even Japan, the creator of "Japan Incorporated" and the precursor of enlightened partnerships between government and business corporations, has not escaped major corruption scandals. The development of Triangle HF3 is an immense area of study which is beyond the scope of this thesis.

3.6 CONCLUSION

The purpose of this Chapter is to provide a contextual model within which productive efficiency may be measured. TFP measures, in vacuo, have limited corporate strategic usefulness and are part of an overall framework of performance. A number of areas of further significant research are also identified by the setting up of a contextual model including: the nexus between TFP and TQM and a deeper understanding of its philosophical roots in KAIZEN; a pricing efficiency measurement system; and a measure of resource utilisation efficiency.