

# The Evaluation of Economic Policy<sup>1</sup>

By

Robin Johnson<sup>2</sup>

## Introduction

There is an increasing concern that economic policy decisions are taken in a vacuum without adequate investigation and evaluation of past decision making. While it would be normal for policy advisors to review problems with a particular policy, it is not customary (nor efficient) to embark on a large research programme to evaluate the real effectiveness of every policy proposal. This may be due to the myriad of policy programmes in place at any one time (thus making choice of projects for evaluation more difficult), but also to a lack of tradition in requiring systematic evaluation before consideration by decision makers. Not unrelated to these issues, is the problem of finding the necessary skills to do the necessary research at an adequate level.

If standards of presentation and adequate background work on economic policy papers is to be improved, then greater attention needs to be paid to evaluation of past policies and instruments. Good presentations should be based on well-researched analysis and facts from an objective point of view and analysis of alternative options for decision makers. In social policy circles, these objectives are summarised in the phrase 'evidence-based policy formation'.

It has to be recognised that decision-making in government is carried out in a political framework and that some issues under discussion derive from previous declarations of intent rather than well-intentioned research. Furthermore, decision-making is in the here and now and there may not be adequate time to prepare analysis in depth. Thus quality may depend on what well-skilled people can do in a restricted time rather than deep-seated research.

In some cases the sheer volume of evaluation work may require specialist services and groups. Such groups need to have the confidence of CEOs and those who write policy advice papers. This assumes that policy advisors do not have the time to do all their own research or even a significant proportion of it. Those with experience will know that such an approach must *anticipate* the policy issues before they occur and some selectivity shown in prioritizing the resulting research projects.

More broadly, government departments need to more seriously engage in information collection activities and/or support national collection agencies such as Statistics NZ. Such agencies may be able to supply the raw materials of good research on policy issues but not always. Within departments, there is a need to develop information systems alongside policy formation and implementation if feedback for evaluation is to be developed properly.

This paper is about the benefits of policy evaluation and the range of research and evaluation techniques needed to back up systematic research on policy options in

---

<sup>1</sup> Paper prepared for Annual Conference of the NZ Association of Economists, Auckland, June 25-27, 2003.

<sup>2</sup> Consulting Economist, Wellington (johnsonr@clear.net.nz)

government. Some of the organisational problems that emerge are discussed. The paper takes the political system as a given and examines how better results can be achieved by prioritizing work programmes in departments and using appropriate techniques for the task in hand.

### **The Westminster Background**

We should never lose sight of the administrative background and origins of our civil service system. Until the early nineteenth century, the affairs of state in the United Kingdom were administered by public officials who owed their positions to political patronage and influence. There was no common system of pay, bribes augmented official salaries, and officeholders, who viewed their positions as property that could be sold, often engaged and paid their own staff. Although the system did not rule out advance by individual ability, it was not a basis for sound administration.

The blueprint for civil service reform was the Northcote-Trevelyan Report of 1854, which advocated the creation of a modern bureaucracy based on a career civil service. Drawing on ideas advanced for the Indian civil service by Thomas Macauley, it was proposed to divide the Government's work into two classes – intellectual (policy and administration) and mechanical (clerical) – and creating a career civil service to carry it out. Staff capable of performing the intellectual work would be recruited from the newly reformed universities; the best talent would be selected through tough competitive examinations supervised by a board of civil service commissioners<sup>3</sup>.

In New Zealand, the Civil Service Reform Act of 1886 provided for entry into the civil service by a competitive examination. The Public Service Act of 1912 made provision for security of tenure, good behaviour, promotion by merit, pensions on retirement, and abolition of entrance by other than competitive examination<sup>4</sup>. Subsequently, the idea of the bureaucracy as a provider of independent and objective policy advice developed slowly and was firmly put in place by some of the well-known civil servants of the day such as Ashwin, McIntosh and Robinson.

This does not mean that the elected officials have to take the advice given by non-elected officials, but it does represent an important distinction derived from the Northcote-Trevelyan reforms in the UK that the civil service has a tradition of independent advice to government which should not be discarded lightly. In an operational sense, the tradition is maintained by non-elected officials always offering alternative courses of action to Ministers so that their neutrality is preserved. [Many Ministers do not want to understand this principle].

### **Evaluation and Outcomes<sup>5</sup>**

The main problem with policy evaluation is the issue of defining the appropriate set of objectives being sought. If serious studies are to measure the *effectiveness* of policy programmes then the focus has to change from departmental *outputs* to policy

---

<sup>3</sup> Adapted from 'Building Institutions for a Capable Public Sector', World Development Report, World Bank 1997, p.80

<sup>4</sup> NZ Year Book, 1993, p.26.

<sup>5</sup> For a discussion of the role of the Public Finance Act 1989 in New Zealand, see R.Johnson, Improving the policy advice process, *IPS Newsletter* 70, August 2002.

*outcomes*<sup>6</sup>. As the US authorities put it “The Results Act [the relevant legislation] seeks to improve the management of federal programs by shifting the focus of decisionmaking from staffing and activity levels to the results of federal programs”<sup>7</sup>.

The US General Accounting Office (GAO) administers the Government Performance and Results Act of 1993 that *mandates* annual performance plans in the federal system. Congress was seeking to reduce the cost and improve the performance of the federal government, holding agencies more accountable for results and better management. Congress had found that the lack of adequate information on federal agencies’ performance was handicapping congressional policy making, spending decisions and oversight and diminishing federal accountability for program results.

The Results Act seeks to improve the management of federal programs by shifting the focus of decision-making from staffing and activity levels to a results basis. Executive agencies are required to prepare 5-year strategic plans to set general directions and then prepare annual performance plans that establish the connections between the long-term strategy goals and the day-to-day activities of the program managers and staff. Finally, the Act requires that each agency report annually on the extent to which it is meeting its annual performance goals and the actions needed to achieve or modify those goals that have not been met.

In 1999 GAO reviewed evaluation studies in annual performance reports<sup>8</sup>. Agencies used evaluation studies to both improve their measurement of program performance and understanding *and* how might program performance be improved. Not all evaluations were initiated in response to the Results Act; most were self-initiated in response to concerns about a program’s performance or about the availability of outcome data. Commonly agencies all faced challenges in collecting outcome data on an ongoing basis. These challenges included: the time and the expense involved, grantees’ concerns about their reporting burden, and substantial variation in states’ data collection abilities.

In New Zealand, the State Services Commission has reviewed the issues around the evaluation of outcomes in the context of improving the quality of policy advice<sup>9</sup>. Some other recent NZ research has been focussed on refining the meaning given to outcomes<sup>10</sup>. The Strategic Policy Group of the Ministry of Social Policy has been working on the conceptual foundations of cross sectoral social policy and social development<sup>11</sup>.

---

<sup>6</sup> Simply put, efficiency would be measured by some ratio of outputs to inputs, while effectiveness would be measured by some ratio of outcomes to inputs required.

<sup>7</sup> General Accounting Office of the USA (1997), *Guide to Assessing Agency Annual Performance Plans* ([www.gao.gov/special.pubs](http://www.gao.gov/special.pubs)).

<sup>8</sup> GAO (2000), *Evaluations help measure or explain performance* ([www.gao.gov/evaluation](http://www.gao.gov/evaluation)).

<sup>9</sup> SSC (1999a), *Looping the Loop: Evaluating Outcomes and other risky feats*; (1999b), *Essential Ingredients: Improving the quality of policy advice*, ([www.ssc.govt.nz](http://www.ssc.govt.nz)).

<sup>10</sup> Proctor, R., *An Inclusive Economy*; Rea, D., *The Social Development Approach*, papers presented at the Institute of Policy Studies 21 August 2001.

<sup>11</sup> D.Rea, *Evidence-based Policy and Practice in Social Policy*, *IPS Newsletter 70*; August 2002.

In particular the information needs of ongoing monitoring of policy initiatives has to be identified by departments. Ryan<sup>12</sup> distinguishes between immediate and intermediate outcome monitoring. He also talks about “ultimate policy impacts”. The latter are societal-level outcomes and tend to be rather unspecific, unattributable and usually several years after the event. “What staff and providers need for ongoing management is information regarding the immediate and intermediate outcomes of service delivery; what changes client status or their conditions of existence are created as a more-or-less direct result of services delivery? Improvement – or change in the desired directions – will tell them they are doing something right and to keep going”.

The problem then becomes, as Ryan recognises, one of devising appropriate indicators and setting up a suitable monitoring framework. He recognises, further, that the time involved and the intellectual difficulty in generating such indicators makes for slow progress. There is thus a fundamental need to address these issues at the policy development stage and the need for enhanced training and recruitment to achieve the follow-up.

### **Purposes of Policy Evaluation**

Evaluation of past and present policy programmes is not only about outputs and outcomes. Chelinsky<sup>13</sup> makes the point that the purpose may also differ between end-users. She identifies three general perspectives:

1. evaluation for accountability (e.g. the measurement of results or efficiency);
2. evaluation for development (e.g. the provision of evaluative help to strengthen institutions);
3. evaluation for knowledge (e.g. the acquisition of more profound understanding in some specific area or field).

In evaluation for results the evaluator is faced with answering the question whether a particular intervention *caused* a particular result, or put another way, whether a change observed is attributable to the intervention. This kind of cause-and-effect question usually calls for methods that allow findings or estimates to be linked to interventions as closely and conclusively as possible. On the other hand, purposes such as strengthening institutions, improving agency performance, or helping managers think through their planning, evaluation and reporting tasks call for evaluation methods that will improve *capacity* for better performance. The third category, knowledge seeking evaluations, involve gaining greater understanding of the issues confronting public policy. The effort to gain such explanatory insights requires strong designs and methods, usually involving both quantitative and qualitative approaches, and advanced levels of both substantive and methodological expertise<sup>14</sup>.

---

<sup>12</sup> Ryan, B., *Death by Evaluation ? Reflections on Monitoring and Evaluation in Australia and New Zealand*, BIIA Conference on Public Sector Performance, Wellington, October 17 2001.

<sup>13</sup> Chelinsky, E. (1997), The coming transformation in evaluation, in *Evaluation for the 21<sup>st</sup> century*, eds Chelinsky and Shadish, Sage Publications.

<sup>14</sup> Op cit, p.10.

## The Need for Incentives

In 1999 the SSC view was that there has been a neglect of the evaluation side of the policy process in the past<sup>15</sup>. There was not a formal requirement in the cabinet manuals to carry out such evaluation. Treasury requirements were apparently not explicit enough. The problem was recognised but not made mandatory.

The SSC took the view that a good deal of evaluation already takes place. But most is focussed on evaluation for the purpose of better delivering and implementing of programmes. Less emphasis had been placed on evaluating the impact of interventions on broader outcomes or on how departmental activities contributed to Government's stated policy priorities<sup>16</sup>. Evaluation was typically not built into the policy document at the outset, thereby making future review problematic.

Reasons why evaluation of outcomes was not a strong feature in the NZ context were summarised as:<sup>17</sup>

- low perceived demand ("ministers not interested", information has an upside and a downside);
- fluctuating policy settings (rapid changes in policy settings at elections and other times);
- funding (reflecting lack of priority in the past);
- methodological issues (lack of clear cut answers in social research);
- dearth of skills and capabilities (lack of competent evaluators); and
- self interested stakeholders (ministers and public servants' unwillingness to be scrutinised, each for their own reasons)<sup>18</sup>.

The Public Finance Act has clearly raised the profile of monitoring and evaluation in government processes though decision makers have not moved to the ultimate sanction of compulsory reporting<sup>19</sup>. The inherent logic of the outcome/output distinction means that departments must be able to justify the outputs they put forward for ministerial agreement and budgetary approval. One Treasury official defines the relationship in these terms "Assessments need to be drawn about the relationship of inputs to outputs (technical efficiency or value for money), outputs to outcomes (allocative efficiency or effectiveness) and on changes in departmental capability (physical or intangible investments).....An intention of the reforms was to create greater incentives for departments to assess and reveal performance in each of these dimensions. The spur would come from Ministers acting as a discerning customer.

---

<sup>15</sup> For a penetrating discussion of evaluation in the Aid Division of the Ministry of Foreign Affairs and Trade, see *Toward Excellence in Aid Delivery*, Report of the Ministerial Review Team. See also footnote 16.

<sup>16</sup> SSC (1999a), op cit.

<sup>17</sup> For another view see: S Van Evera, 'Why States Believe Foolish Ideas: Non-Self-Evaluation by States and Societies' ([www.mit.edu](http://www.mit.edu)). Organisations are poor self-evaluators. Myths, false propaganda, and anachronistic beliefs persist in the absence of strong evaluative institutions to test ideas against logic and evidence. Organisations turn against their own evaluative units as they threaten jobs and the status of incumbents. Organisations attack their own thinking apparatus if that apparatus does its job!

<sup>18</sup> SSC (1999a), pp.6-7.

<sup>19</sup> The Australians tried such a system in the early 1990s. See Di Francesco (1998), The Measure of Policy: Evaluating the Evaluation Strategy as an Instrument for Budgetary Control, *Australian Journal of Public Administration* 57, 33-48.

Ineffective outputs would be cut out, and if prices were too high other suppliers would be sought, or if this was not possible changes might be sought to management"<sup>20</sup>.

"To date this goal of departments proving their performance has only been partly achieved. One reason has been the time needed to make and embed changes of this magnitude: lifting the levels of technical and management skills, introducing new systems of funding, reporting, and performance measurement. Changing departmental cultures takes years. Another reason is that greater onus could have been placed on departmental managers to prove their efficiency and effectiveness. Placing the onus of proof onto spending proponents would have increased the incentives on departmental Ministers to seek information from their departments"<sup>21</sup>.

The SSC view of these matters was that existing incentives in the system tended to cause managers to avoid open scrutiny of the relative merits of existing programmes; that the budget focus on new initiatives avoided zero-based evaluation of past programmes; and that departments tended to protect their vote even when another department's programme was found to be more effective. In addition the short-terminism in the system is not conducive to outcome evaluation. Evaluations – especially in the social policy area - typically require a longer time-frame or some ongoing commitment to monitoring progress over time<sup>22</sup>. Sometimes policy formation cannot wait for the necessary research to be carried out. These are powerful disincentives and they will require considerable encouragement and direction from CEOs and Ministers if they are to be reversed.

### **Evaluation: Statutory and Indicative Requirements**

In this section, I want to discuss statutory and informal requirements for better economic evaluation. Out of this sort of discussion, it is possible to get some idea of the skills and capability required. I start with Treasury requirements for policy papers; then regulatory impact statements, and finally more informal areas where policy evaluation is more likely to be required.

#### *Treasury Requirements*

The Secretary to the Treasury has recently noted (*Dominion Post* April 19), that in the past there was insufficient focus on getting results (outcomes); problems in contracting for service delivery; weak links between strategy and spending; reduced departmental ability to provide services (capability), and uneven agency performance. An advisory group has recently expressed concerns about cooperation between public service agencies: too little attention to overall results; legalistic contracting; and insufficient training and leadership. The resulting "The Review of the Centre" found there was a need for: better-integrated, citizen-focussed service delivery; improvements to people management and public sector culture; and a need to deal with fragmentation between agencies.

Treasury as the budgetary control agency is clearly interested in the outcomes of departmental proposals and their costs. Substantial cost-benefit analysis has not been

---

<sup>20</sup> Bushnell P. (1998), p.2, Does Evaluation of Policies Matter?, Foreign and Commonwealth Economic Advisors, London.

<sup>21</sup> op cit, p.2

<sup>22</sup> SSC (1999a), p.11.

required for this purpose in the past although in the 1970s Treasury employed zero-based budgeting procedures which necessitated sharper operational categorisation and clear delineation between existing programmes and new proposals. Treasury expects tight reasoning from departments and clear demarcation of options available to reach a given goal. There are clear directives for departments to define what outcome is being sought, whether all the options have been considered (including the status quo), and what assumptions lie behind the declared costs and benefits. The directives ask whether there are clear criteria for analysing the options, and whether the criteria include effectiveness and efficiency considerations. While Treasury is in favour of increased evaluation of efficiency and effectiveness, it is wary of such requirements becoming a compliance activity with no bite. The aim would be appropriate, but not at the cost of excessive resourcing, and effort should be directed to the highest priorities<sup>23</sup>.

There are new initiatives to promote result-based management. An example is “Managing for Outcomes” which is jointly led by the central agencies and TPK and asks all public service departments to define the results they are to achieve, and state the measures they will use to check their success, how their own services will contribute, the capability they need, and how they will manage risks.

Departments are also being asked to report in a readable form how they are managing for outcomes in a “Statement of Intent” to be published with the Budget. SOIs are to focus on medium and long-term planning. They will cover 3 or more years; should meaningfully describe the relationship between an agency’s services and the results the government wants; how this will be evaluated; the relationship with related agencies; and the capability available for delivery. Treasury states that ‘the move to SOIs will take time’ ([www.treasury.govt.nz/briefing2002/chap4.asp](http://www.treasury.govt.nz/briefing2002/chap4.asp)). In sum, Treasury is now engaged with other agencies in developing evaluation – getting evidence for whether policies work, researching how to improve innovation, and improving the governance arrangements for Crown entities.

### *Regulatory Impact Statements*

As of July 1 1998, all policy proposals submitted to cabinet which called for government bills or statutory regulations had to be accompanied by a Regulatory Impact Statement (RIS)(A Ministry of Commerce suggestion). Such Statements were required to consistently examine potential impacts arising from government action and provide an assurance that new or amended regulatory proposals had been subject to proper analysis and scrutiny as to their necessity, efficiency, and net impact on community welfare<sup>24</sup>. The requirements include “a statement of feasible options (regulatory and/or non-regulatory) that may constitute viable means for achieving the desired objective(s)” and “a statement of the net benefit of the proposal including the total regulatory costs and benefits of the proposal and other feasible options”. While these requirements have generally been met over the years since they were introduced, the convention was soon adopted that such Statements needed only outline the problems involved and not provide a full-scale analysis of necessity, efficiency and

<sup>23</sup> Personal communication, David Galt, Treasury.

<sup>24</sup> Ministry of Commerce (1998), *A Guide to preparing Regulatory Impact Statements*.

impact<sup>25</sup>. There is always a danger that such requirements can easily be ticked off at submission time.

### *Merger Requirements*

In the competition field, the Commerce Commission is required by statute to examine policy issues involving acquisitions and mergers. The methodology utilised is very instructive in terms of high level economic analysis of the national interest. In terms of the Commerce Act 1986 the Commission is required to look at any acquisition or merger in terms of increased market dominance (s 67(3)(a)), and failing that test in terms of whether the public benefits that might follow would justify the acquisition or merger (s 67(3)(b)). "The authorisation procedure requires the Commission to identify and weigh the detriments likely to flow from the acquiring of a dominant position in the relevant markets, and to balance those against the identified and weighed public benefits likely to flow from the proposed merger"<sup>26</sup>. This is equivalent to the national point of view or the national interest.

The preciseness of this approach is illustrated by a strict concern for comparing bananas with bananas. Reports and analysis are based on a thorough discussion of what is known as the *status quo scenario*, as compared with any changes brought about by regulatory means or policy changes. In the case of the Dairy Determination, two base scenarios (counterfactuals) are developed and utilised.

- (1) The *status quo* counterfactual is not regarded as a standstill in time. Rather, the Commission asks what events would have taken place *in the absence of the merger* and then puts them in the *status quo* scenario. Thus the Commission assumes a continuation of the current industry structure (pp.90-91) and of the single desk seller for exports, and that efforts at further structural change within that structure and regulatory framework will continue.
- (2) The *deregulation counterfactual*. This envisages the deregulation of the Dairy Board export monopoly, and the emergence of two powerful competing cooperative companies, possibly with overseas links, heavily involved in marketing in competition. In such a process, this counterfactual would include the unbundling of payout to suppliers, improvement to the internal pricing systems, and joint venture developments.

The Commission then proceeded to an examination of the *detriments* of the proposed merger and the possible benefits against these *two* counterfactuals. They looked at allocative and productive efficiency changes in the NZ domestic market as a whole<sup>27</sup>.

The same approach is evident in the Qantas-Air New Zealand merger proposals. In this case, the definition of the counterfactuals has proved very tricky. On the basis of

---

<sup>25</sup> For a full analysis, see B.Wilkinson, The problem of inadequate regulatory impact statements, *IPS Newsletter* 70, August 2002.

<sup>26</sup> Commerce Commission (1999), *Draft Determination....in the matter of an application for authorisation of a business acquisition involving New Zealand Dairy Board...and others*, p.84.

<sup>27</sup> In this particular case the Commission found that the detriments exceeded the benefits of the dairy company merger compared with both counterfactuals and recommended against government approval. This methodology is a good example of sound economic reasoning based on available industry information or that supplied to the Commission. It is an *ex ante* rather than an *ex post* analysis and is a sound basis for examining future government policies.



confidential information provided by the airlines, the Network Economics Consulting Group constructed a counterfactual based on the possible competitive situation without an alliance. Qantas would increase its capacity on both NZ domestic and trans-Tasman routes and Air NZ would have to match these increases to maintain market share. Over a 5-year period Air NZ would be forced to withdraw its international services, eventually shrinking to a domestic airline. Critics have thus raised the issue whether the strategy of the airlines is to paint a counterfactual so dire for AirNZ that the alliance becomes more acceptable than it would otherwise be<sup>28</sup>. These issues involve serious examination of the possible trends in the air passenger market over a considerable period as well savings derived from cost efficiencies. Thus choice of the counterfactual is fundamental to any analysis of a merger proposal and involves serious consideration of possible future scenarios – perhaps more than one?

#### *Indicative Requirements: The Structure, Conduct, Performance Approach (SCP)*

Many policy issues are susceptible to an institutional approach where structural and conduct issues are separated out from performance issues. The Institute of Economic Research has been at the forefront in applying SCP to industry and policy issues. This approach asks three different questions about economic behaviour which complement each other. Structure refers to the institutional environment in which a firm finds itself. This may set a number of constraints on any actions for reform or improvement. Conduct refers to how the the management conducts its business including governance issues. Performance refers to efficiency in production including returns on shareholders funds. Following this schema in some detail illuminates industry behaviour and can be used to make comparative observations on policy change from a national point of view.

Government has commissioned NZIER and other agencies in a number of policy reviews of this kind, the dominant feature of which was an analytical approach based on the SCP<sup>29</sup>. These reviews include studies of the marketing arrangements for wheat, meat products, dairy products, freight transport and others. Regulatory issues are discussed in terms of their impact on industry from a national point of view. These studies are very thorough and are instructed by the national point of view throughout. One criticism is that there are not enough of them, nor of the people who can carry them out, to make a significant contribution to the formation of government policy on a day-to-day basis.

#### *Investment Requirements*

Indicative information on the return on public investment aids decision making. Cost-benefit analysis (CBA) is also based on an appropriate national interest counterfactual and then forecasts of costs and returns (suitably defined for the issue in question) are subject to a discounted cash flow analysis. In the past, much government time was taken with discussions of the appropriate interest rate, when perhaps more focus should have been placed on the counterfactual(s). CBA focusses on the national interest part of government investment – though it can be used in financial analysis

---

<sup>28</sup> 'Air cartel's critics counter the counterfactual' *The Independent*, 5 March 2003.

<sup>29</sup> Bollard, A., Gale, S., Harper, D., and Savage, J. (1991), *An Introduction to Industrial Organisation*, Ministry of Commerce, Wellington; Nixon, C. (1993), *The Impact of Wheat Deregulation on the Arable Industry*, Ministry of Agriculture, Wellington; Pickford, J.D. and Bollard, A. (1998), *The Structure and Dynamics of New Zealand Industries*, Dunmore Press.

just as well – including social as well as economic outcomes. CBA was useful in government administrative terms in the past because it compressed a large quantity of numbers down to a few criteria like present value and internal rates of return. This enabled quicker decisions to be made when there was a choice of options and/or investments

It was widely used by the Ministry of Works and the Ministry of Agriculture in the past for soil conservation investment and is still used for roading investment.

### *Social Returns to Investment*

Where time permits or the particular analysis has already been undertaken, econometric modelling may be useful in illuminating possible effects of policy changes. Econometric models are useful in circumstances where time series or cross-section data are available and match the entity which is changed by a given policy. Statistical analysis is time-consuming and its execution must anticipate to a large degree the kind of policy problems which might bear such detailed examination.

As an example, the supplementary minimum prices scheme for farmers in the 1970s is a policy stance which is well suited to econometric analysis. The model concerned was developed at Lincoln University and relates farm exports and production to changes in investment and product prices for sheep, beef and dairy farms<sup>30</sup>. From the base year of the model in 1981, investment and outputs were simulated for 1982 to 1985 using product prices *with and without* deficiency payments (the difference between market prices for products and payments actually received). For 1986 actual prices applied to both simulations and continued for the four following years to allow interactions between farm sub-sectors to work themselves out<sup>31</sup>.

If deficiency payments had not been paid, sheep number growth would have been less, and dairy and beef cattle numbers would have been higher. In terms of increased export volumes (the aim of the policy) beef volumes would be higher, sheep volumes lower, and dairy volumes higher in the short term without subsidies. The total level of exports would have been two per cent lower. In 1982 dollars the long term difference in exports was \$184m per year or 2% less than that which eventuated. Deficiency payments cost up to \$1018m in 1982 dollars over 4 years so the benefits would have to be spread over 6 years or more to show a positive national return (undiscounted). This result has been confirmed by other econometric models<sup>32</sup> though the authors did not feel disposed to draw the same conclusions.

This example demonstrates how an econometric model can be sector-wide, provide for interactions between variables, is subject to statistical tests, and is relatively free of observer bias. On the other hand, national policy initiatives may not always lend themselves to clear uni-sectoral and national interest goals.

---

<sup>30</sup> Laing, M.J., and Zwart, A.C. (1983), *The Pastoral Livestock Sector and the Supplementary Minimum Price Policy*, Discussion paper No 70, AERU, Lincoln.

<sup>31</sup> Johnson, R.W.M. (1986), *Livestock and Feed Policy in New Zealand, 1975 to the present*, CAPS, Massey University.

<sup>32</sup> Sandrey, R., and Reynolds, R. (1990), *Farming without Subsidies*, Government Print, p.166.

*Defining Social Outcomes: Poverty measurement and socio-economic disparity*

In the social policy, the measurement of outcomes is particularly difficult. It is one thing to set up programmes of social assistance based on any number of criteria, but quite another to know that the welfare of the individuals concerned has been improved. There is an increasing awareness of these issues and I want to discuss poverty and social disparity measurement of outcomes as opposed to departmental outputs. This is not the final word on a large subject but is indicative of outcomes measurement and its impact on policy.

Bob Stephens has recently summarised the role of poverty measurement in policy development and formulation<sup>33</sup>. He notes 'there is considerable academic, technical, and policy-related debate over the appropriate conceptualisation of poverty, the generosity of the poverty measure, how that measure should be adjusted for household size and composition, and how to adjust the measure through time. The different concepts, measures and equivalence scales all identify different people and family types as having a higher incidence and severity of poverty, with different groups identified again depending on whether a cross-section or longitudinal analysis is undertaken. If one objective of government policy is the alleviation and amelioration of poverty, then these conflicting results make policy formation in this area a nigh-on impossible task. However, a partial resolution is to recognise that the different concepts of poverty can be used for different policy objectives. Income-based poverty measures are more appropriate for monitoring the impacts of economic and social aid and determining the level of cash assistance, while outcome or living standard measures provide insights into the role of asset accumulation and whether social assistance should be provided in cash or in-kind.. Static measures are more useful for poverty alleviation by assisting in determining the appropriate level of assistance and who should receive it, while dynamic measures provide insights into ultimate causes and thus long-term solutions to poverty'.

Stephens provides an overall summary of the poverty measurement work he has undertaken with Waldegrave and Frater in this paper. Most important, to my mind, is the discussion of living standards in policy development. This research looks at negative outcomes or what activities families are forced to restrict purchase or ownership of due to lack of income<sup>34</sup>. Surveys of elderly, working population and Maori ask questions about restrictions on living standards which can be aggregated up into a master score. Such restrictions are classified by age and size of family. The results are presented as the proportion of the sample population indicating enforced lack of an item, eg. % without adequate heating, % without childcare services, % missing doctors visits, % showing dampness in the home etc. Their importance lies in the establishment of social benchmarks which can continue to be measured through time (assuming adequate sampling). These measures are of course the social 'outcomes' which social policy should be aiming to improve. It is well known that departments have been happy to pursue departmental objectives ('outputs' in the jargon'), such as meeting Ministers' deadlines, rather than persuading Ministers to agree to policies with long term effects on outcomes.

---

<sup>33</sup> R Stephens, The Role of Poverty Measurement in Policy Development and Formulation, paper for presentation at the Policy Network Conference, Wellington, 30-31 January 2003.

<sup>34</sup> Ministry of Social Development, 2002.

Measures of socio-economic disparity are concerned with the influence of cultural factors which explain differences in poverty and access to services. It is of some importance to establish the relative importance of different factors because social perception of these differences often varies from the underlying reasons. Simon Chapple has done work on the commitment to addressing Maori socio-economic disparity<sup>35</sup>.

Chapple examines known statistical collections such as Census data, the Household Labour Force Survey, population projections, and the Income Supplement of the Household Labour Force Survey, to explore the role of ethnicity in socio-economic disparity. He challenges the conventional wisdom on Maori disparity in public policy circles that takes as axiomatic the key importance of macro ethno-cultural differences between Maori and non-Maori. This view de-emphasises non-ethnic cultural differences and down-plays cultural similarities. Such an approach leads to “Maori for Maori” solutions to perceived problems when appropriate analysis indicates that disparity is not an ethnic problem but a cultural problem based on low literacy, poor education and living in geographical concentrations that have socio-economic problems. For reasons of social justice and efficiency, effective policy to close the gaps needs to focus on those most disadvantaged or at risk.

These conclusions shifted the then policy stance on closing the gaps and moved the disparity programs to a broader base. If such a high level of analysis had been available at an earlier stage of policy development, the policy goals could have been adjusted accordingly (perhaps).

### **Summary and Conclusions**

1. Greater encouragement of evaluation work is required from government, both at the Ministerial level and the CEO level. Treasury has a key role to play in its public management role. Departmental cooperation is a key issue. There are great differences between departments.
2. There is a need to organise greater research effort into policy outcomes and find better systems to incorporate research results into policy formation and advice. The capability of departments is an issue. Disincentives to objective and thorough reporting should be removed.
3. Higher levels of training are required for some or all of the policy analyst occupation class. This could be in-house or through the universities. University cooperation is needed to identify suitable courses and class materials. A suitable manual should be developed which showed exactly what standards of analysis can be expected from better and higher training compared with the present.
4. The precepts of the 1912 Act and the examples of Ashwin and others should be preserved.

---

<sup>35</sup> Chapple, S. (2000), Maori Socio-economic Disparity, *Political Science* 52(2), 101-15.