School of Economics and Finance

Curtin Business School

The Evaluation of Australian Labour Market Assistance Policy

Alfred Michael Dockery

This thesis is presented as part of the requirements for

the award of the Degree of Doctor of Philosophy

of the Curtin University of Technology

December 2002

Contents

1. INTRODUCTION

This thesis is comprised of a series of published papers relating to the evaluation of active assistance measures for the unemployed in Australia. It offers both applied evaluations of active assistance measures as well as critical assessment of the evaluation approaches that have dominated the literature and policy formation in Australia. "Active" assistance for the unemployed is distinguished from "passive" assistance, such as income support.

The motivation behind the work lies in the fact that a very large amount of public expenditure is directed to active assistance for the unemployed. Over $2 billion dollars was spent on labour market programs at the height of the *Working Nation* package in each of 1995-96 and 1996-97, and $1.5 billion was allocated to "labour market assistance to jobseekers and industry" in the most recent (2001-02) Commonwealth budget. Despite this considerable past and ongoing expenditure, the evaluation effort in Australia has been far short of international best practice. As a consequence, there is no convincing empirical evidence as to how effectively these public resources are being used, or of the relative merits of various options in the design of active interventions for the unemployed.

Ultimately, the goal of the research is to improve supply-side policies designed to address unemployment. As stated, it aims to do this through original empirical evaluations of programs and through critical assessment of existing evaluations and institutional arrangements. The objectives of the research can be stated a follows:

- to improve the understanding of the effectiveness of labour market programs and other assistance measures in Australia through both applied evaluations and critical assessment of existing evaluations undertaken in Australia;
- to contribute to the development of design features of labour market assistance measures and service delivery models in Australia so as to assist public policy in improving the efficiency and equity outcomes of the labour market;
- to contribute to the domestic and international literature on employment service delivery models through an evaluation of Australia's "Job Network" reforms.

The focus of the research is on the micro-economic impacts of assistance measures, or what is often referred to in the literature as "the effect of the treatment on the treated". As such, broader questions of macroeconomic effects are not specifically analysed, although related issues such as substitution effects and deadweight loss associated with programs are canvassed. The initial research involved empirical evaluation of the major labour market programs that were in place under *Working Nation*. This work was made possible by the availability of unit record data from the Australian Bureau of Statistics' Survey of Employment and Unemployment Patterns (SEUP). Attention later shifted to assessment of the Job Network, the framework for the competitive employment services market that replaced the vast majority of Commonwealth Government delivered labour market programs upon its introduction in May of 1998.

The thesis consists of seven central published papers. These papers address a series of research questions relating to the objectives above, including:

- Were the main programs in place under Working Nation — wage subsidy programs, brokered employment programs, job search programs, training programs and the New Enterprise Incentive Scheme — effective in assisting the unemployed into work?
- Does the ongoing "post-program monitoring" methodology used by the government adequately measure the impact and cost-effectiveness of these programs?
- From theoretical considerations and the limited empirical evidence available, what can we say about the likely success of the Job Network model and how might the model be improved?
- From the existing body of evidence, what lessons can be drawn for the design of active assistance measures?
- How can the evaluation approach be improved in order to provide a better guide for policy makers?

Copies of the seven papers make up Section 4 of the thesis. They appear in the chronological order in which they have been published or accepted for publication, which is as follows:

[1]   Stromback, T. and Dockery A. M. (1998), "The Job Compact Mark II?", *Australian Economic Papers*, Vol. 17, No. 2, June, pp. 24-34.

[2]   Stromback, T., Dockery A. M. and Ying, W. (1999), "Labour market programs and labour force status", *Australian Bulletin of Labour*, Vol. 25, No.2, pp. 159-178.

[3]   Dockery, A. M. (1999), "Evaluating the Job Network", *Australian Journal of Labour Economics*, Vol. 3, No. 2, pp. 131-158.

[4]   Stromback, T. and Dockery, A. M. (2000a), *Labour market programs, unemployment and employment hazards: an application using the 1994-1997 Survey of Employment and Unemployment Patterns*, Australian Bureau of Statistics Occasional Paper, ABS Catalogue No. 6293.0.00.002, February.

[5]   Dockery, A. M. and Stromback, T. (2001), "Devolving public employment services: preliminary assessment of the Australian experiment", *International Labour Review*, Vol. 140, No. 4, pp. 429-451.

[6]   Dockery, A. M. and Webster, E. (2002), "Long term unemployment and work deprived individuals: issues and policies", *Australian Journal of Labour Economics*, Vol. 5, No. 2, pp. 175-194.

[7]   Dockery, A. M. (2002), "The New Enterprise Incentive Scheme: an evaluation and a test of the Job Network", *Australian Journal of Labour Economics*, Vol. 5, No. 2, pp. 351-371.

The thesis is set out as follows. The Section 2 contains abstracts of each of the seven published papers. Section 3 contains a literature review, canvassing first the international literature and then the Australian literature. The literature review is

geared to setting the Australian evaluation efforts within the context of international best practice and evidence, and also to highlight the contributions of the papers that make up this thesis. To help in assessing the contribution of the Doctoral research, the corresponding numbers in square parenthesis above are included with each citation of the papers. The Appendix includes letters by co-authors Stromback and Webster indicating the respective contributions of each author for the relevant publications.

In addition to the publications included in this thesis, a number of peripheral but related papers have arisen out of the program of Doctoral research. These are:

Kelly, R., Lewis, P., Dockery, M. and Mulvey, C. (2001), *Findings in the NEIS evaluation*, report to the Department of Employment, Workplace Relations and Small Business, Centre for Labour Market Research, May.

Flatau, P. and Dockery, M. (2001), *How do income support recipients engage with the labour market?*, Policy Research Paper No. 12, Department of Family and Community Services, Canberra.

Stromback, T. and Dockery, M. (2001), *The duration of unemployment benefit spells: a comparison of Indigenous and non-indigenous persons*, Policy Research Paper No. 10, Department of Family and Community Services, Canberra.

Stromback, T. and Dockery, A. M. (2000), "Evaluation of labour market programs: an assessment", paper presented to the 29th Annual Conference of Economists, Gold Coast (3-6 July, Queensland).

Dockery, A. M. (2000), "Regional Unemployment rate differentials and the mobility of the unemployed: an analysis of the FaCS longitudinal dataset", *International Journal of Manpower*, Vol. 21, Iss. 5, pp. 400-421.

2. COMPENDIUM OF ABSTRACTS

[1] Stromback, T. and Dockery A. M. (1998), "The Job Compact Mark II?", *Australian Economic Papers*, Vol. 17, No. 2, pp. 24-34.

Following the 1993 Green Paper of the Committee on Employment Opportunity, the Commonwealth Government committed itself to a substantial increase in assistance for the unemployed. The centrepiece of this strategy was the Job Compact, which was to provide job placements for 160,000 long term unemployed persons in 1995-96 and 1996-97 through a range of labour market programs, and with a "reciprocal obligation" upon jobseekers to accept such offers of temporary employment or forego benefits for a specified period. In the event, actual placements and funding under the Job Compact fell well short of projections for a variety or reasons, and the Compact was scrapped in 1996 by the incoming Coalition Government. In taking stock of the situation, we argue that the central tenets of the Job Compact are alive and well. Despite rhetoric suggesting a "get tough" stand on the unemployed, the Coalition Government has basically streamlined the Job Compact in line with what would have been expected with the benefit of performance evaluation and review.

[2] Stromback, T., Dockery A. M. and Ying, W. (1999), "Labour market programs and labour force status", *Australian Bulletin of Labour*, Vol. 25, No. 2, 159-178.

This article uses data from the first wave of the SEUP to estimate the effect of Australian labour market programs on labour market outcomes in terms of labour force participation and employment. It is the first such microeconomic evaluation of these programs outside of DEETYA's post-program monitoring system and allows for far greater controls for individual characteristics. Participation in wage subsidy and employment training programs are found to have a positive impact on the likelihood of being employed. A fair degree of correspondence is found with DEETYA's evaluations, however the results suggest that employment training programs have a more positive effect on outcomes than indicated by the post program monitoring data.

[3]   Dockery, A. M. (1999), "Evaluating the Job Network", *Australian Journal of Labour Economics*, Vol. 3, No. 2, pp. 131-158.

The introduction of the contracted employment services market, the Job Network, was radical not only in terms of Australian policy, but also in terms of approaches to the delivery of employment services throughout the world. This paper reviews some of the possible benefits and pitfalls of the reforms, and preliminary evidence on the Job Network's performance is presented. Although it is too early to conclude whether the Job Network is outperforming the previous Commonwealth Employment Service, evidence to date at least shows the Job Network is not the disaster that many of its detractors have claimed it to be. Further evaluation must pay attention to the fortunes of the very hard-to-place job seekers, as it appears the funding levels under the Job Network will require either a considerable increase in cost effectiveness in delivering assistance to this group, or a considerable drop in the level of support offered.

[4]   Stromback, T. and Dockery, A. M. (2000a), *Labour market programs, unemployment and employment hazards: an application using the 1994-1997 Survey of Employment and Unemployment Patterns*, Australian Bureau of Statistics Occasional Paper, ABS Catalogue No. 6293.0.00.002, February.

This paper provides estimates of the effectiveness of the major labour market programs that comprised the *Working Nation* set of measures using a duration framework and data from the SEUP. The key indicator is the estimated impact of program participation on the rate at which persons exit episodes of "job search" into either the "working" or "absent from the labour market" states. The data and estimation techniques control for a larger number of individual characteristics than has been possible in other evaluations of these programs, but the possible bias which may arise due to the process of selection into programs is not fully controlled for in this study.

The results indicate that participation in a labour market program is associated with a marked increase in the hazard rate from job search. The magnitude of the estimated effect is very large, suggesting that much of the estimated effect arises from selection bias. Program participation is also estimated to lead to a longer duration of subsequent work spells. Thus it does not appear that programs place people disproportionately into short-term or dead-end jobs. Wage subsidy programs are found to have the most favourable impact. Transition models are then used to replicate DEETYA's post-program monitoring approach. The incorporation of additional variables available in SEUP is found to have little effect on the estimated impacts. Rudimentary controls for the effect of selection into programs are also included in the transition models. These results are inconclusive, but the strong positive effect of participation in wage subsidy programs stands.

[5]   Dockery, A. M. and Stromback, T. (2001), "Devolving public employment services: preliminary assessment of the Australian experiment", *International Labour Review*, Vol. 140, No. 4, pp. 429-451.

Throughout much of the OECD there has been a trend toward devolution in the delivery of active assistance programs for the unemployed. Australia has perhaps gone further than any other by tendering out almost all public employment services under the "Job Network". This paper seeks to contribute to the international evidence on the effectiveness of different models of delivery of public employment services through a review of the Job Network model. There are some salient lessons to be drawn from the Australian experience, particularly with respect to the incentive effects associated with contractual and monitoring arrangements and their implications for evaluation.

[6]   Dockery, A. M. and Webster, E. (2002), "Long term unemployment and work deprived individuals: issues and policies", *Australian Journal of Labour Economics*, Vol. 5, No. 2, pp. 175-194.

Labour market programs are a major policy tool used to assist the long term unemployed and considerable interest in the Job Network and the old *Working Nation* programs has focussed on how well they address the needs of this work deprived group. Despite past efforts, the paper presents evidence that there remains a core group of very long term unemployed who have been work deprived for a large portion of their working lives. While the cross sectional focus of most Australian data collections makes this difficult to document, there is reasonable evidence that this group exists.

So what can we learn from the experience to date? After a review of the recent history of Australian labour market programs and their assessment, the unfortunate answer is 'not much'. In part, this is due to several shortfalls in the evaluation effort, including the lack of rigorous research designs and access to data for independent researchers. There has been a focus on outcomes, with little research to tell us how successes have been achieved. Labour market programs are seen as a 'black box' into which the unemployed enter and come out at the other end either employed or not. A recurring theme from existing evaluations is that net impacts are either negligible or smallish. While the evaluation critiques, quite correctly, focus on estimation and sample selection issues, in the end they are really debating whether the net effects are small or very small. More attention should perhaps be paid to the in-program benefits of participation. Suggestions are offered for a new approach to delivering assistance, and the case put for an enhanced evaluation effort in order to guide future policy.

[7]   Dockery, A. M. (2002), "The New Enterprise Incentive Scheme: an evaluation and a test of the Job Network", *Australian Journal of Labour Economics*, Vol. 5, No. 3, pp. 351-371.

The New Enterprise Incentive Scheme provides support for unemployed persons in Australia to start their own business. It is one of the few employment services to be retained in comparable form following the 1998 implementation of the competitive employment services market known as the Job Network. This paper uses Departmental post-program monitoring data to provide an independent evaluation of

NEIS and to add to the body of evidence on the impacts of the transition to a competitive market for employment services. NEIS achieves high outcomes in terms of the proportion of participants that secure employment and cease claiming benefits. However, participation it is not targeted to disadvantaged job seekers and deadweight loss would mean that the *net impact* of participation would be considerably lower. Some noted exceptions aside, gross outcomes, the profile of those entering the scheme and the relative success between target groups have remained very similar under the Job Network. Thus, while the available data cannot support rigorous estimates of net impacts of NEIS under the two regimes, it provides little evidence of the existence of negative effects on equity that may have been feared from the transition to a more profit driven, competitive system of delivery.

3. LITERATURE REVIEW

3.1   Background

Labour market programs (LMPs) are publicly funded measures to assist the
unemployed to find work. The major forms of LMPs are job brokerage, assistance
and training in job search, skills training, job placement and employment subsidies
and direct job creation for unemployed persons. The term "active" labour market
program is often used to make the distinction from passive support for the
unemployed, such as unemployment insurance and unemployment benefits.

Most industrialised countries experienced significant increases in unemployment in
the 1970s and again in the 1980s to which policy makers responded with an
expansion of LMPs. With this expansion came the need for evaluative research both
to guide policy and to meet pressures for greater accountability of government
expenditure. Policy thinking at the time was dominated by the idea of the Phillips
Curve and the belief in a long run non-accelerating inflation rate of unemployment
(NAIRU). High unemployment was considered a temporary phenomenon arising
from a mismatch of skills and LMPs as a way of reducing unemployment without the
inflationary effects associated with aggregate demand instruments - a way of
"cheating the Phillips Curve". This led to a shift in the focus of programs towards
skills training and other supply-side measures and their objectives were primarily
ones of economic efficiency. In evaluating the cost efficiency of supply-side
programs, the cost of direct placements in job creation programs provided a useful
yardstick. Estimating the aggregate effect of LMPs, however, was confounded by
uncertainty regarding the extent of deadweight losses, substitution and displacement
effects. Deadweight loss refers to the proportion of participants in a LMP who gain
employment but who would anyway have found work had they not participated.
Substitution and displacement effects occur when participants in a LMP gain
employment at the expense of other jobseekers (substitution) or existing workers
(displacement) such that there is no net employment gain.

As the more permanent and deep-rooted nature of the rise in unemployment in the
1970s and 1980s became evident, particularly the emergence of long-term

unemployment, LMPs became increasingly concerned with equity outcomes when compared to the earlier manpower programs. Typically, LMPs are now targeted at the most disadvantaged in the labour market, such as the disabled, persons with language difficulties, sole parents, the youngest and oldest categories of workers and those displaced from declining occupational or industry sectors. The long-term unemployed (LTU) provide a convenient "catch all" target group, as those with either observable or unobservable disadvantages will, by definition, be disproportionately represented in this group. Further, where there is state dependency, long-term unemployment in itself constitutes a disadvantage in job search (see OECD 1988, 1991; Disney and Carruth 1992).

Over these years the evaluation of LMPs has evolved in response to changes in labour market theory, policy objectives and methodological challenges, led largely by work from the United States (see Schmid, O'Reilly and Schömann [eds.]1996). At the micro level, rigorous evaluations now typically use longitudinal data to compare the fortunes of program participants with those of a group of non-participants while controlling for individual characteristics and initial conditions, on the assumption that the difference in outcomes between the two groups represents the impact of the program. However, because of the diverse range of potential effects of programs, aggregative studies also play an important role (see Calmfors 1994).

In terms of methodology, two important developments have been, firstly, the recognition of the limitations of using a comparison group of non-participants as a control group, largely because the processes by which individuals select themselves into LMPs can result in a major over-estimation of the program's effects. Led by the work of James Heckman, this has seen the development of sophisticated econometric methods to deal with "selection bias" in data and the advocacy of "experimental" research designs in which persons are randomly allocated into participant and non-participant groups. The second has been the application of the techniques of survival analysis in the evaluation of social and labour market program data.

An important change in labour market theory and policy thinking was the development of the notion of hysteresis. The natural rate hypothesis formulated by Phelps (1967) and Friedman (1968) posits that there is an underlying rate of

unemployment, determined mainly by supply side factors, towards which the economy will gravitate over time. The natural rate of unemployment represents an equilibrium rate, as opposed to an actual rate, that may change over the long-term with changes in institutional arrangements and the matching efficiency of the labour market. The sustained periods of high unemployment in the 1970s and 1980s presented a challenge to the natural rate hypothesis — the economy simply did not revert back to its previous levels of unemployment as predicted following an initial increase (see, for example, Layard, Nickell and Jackman 1994; Cross, Hutchinson and Yeoward 1990, Blanchard and Summers 1987). Attempts to reconcile the hypothesis with the empirical reality led to suggestions of a shifting (in this case rising) natural rate. Clearly, unless the source of such shifts in the natural rate could be explained, this undermines the usefulness of the concept of a "natural" rate.

Hysteresis provided such an explanation. Applied to unemployment, hysteresis is the idea that the *equilibrium* rate of unemployment depends on the history of the *actual* rate (Blanchard and Summers 1987). The idea of hysteresis had long been applied to time series in other fields and to other economics series. According to Cross *et al* (1990: 94), Edmund Phelps was the first to apply the terminology to unemployment, noting in his 1967 article that the current rate will have lingering effects on the natural rate in the future. The notion gained popularity in the 1980s as studies of the time series properties of unemployment provided evidence of a high degree of persistence (Nelson and Plosser 1982, Campbell and Mankiw 1987) and theoretical foundations were formulated. The two main explanations to have emerged are the insider/outsider hypothesis attributed to Lindbeck (see Lindbeck and Snower 1987) and further developed by Blanchard and Summers (1987) and Gregory (1986); and the human capital explanation that focuses on the negative effects of increasing duration in unemployment on jobseekers' skills, motivation and appeal to employers (see Hargreaves Heap 1980, Blanchard and Summers 1988: 317-319). Under either view, it has become largely accepted that the mechanisms of hysteresis can be observed via the durational composition of the unemployed pool, usually measured as the proportion of LTU in total unemployment. As the duration of an individual's unemployment spell increases, their ability to compete for jobs diminishes such that the LTU constitute a less effective supply of labour in terms of moderating wage rises. Under the conditions of hysteresis the problems of substitution and

12

displacement, which can usually only be accounted for through macro-level evaluations of LMPs, may be considered to be of lesser importance, since there are efficiency as well as equity gains from simply "shuffling the queue" of unemployed. The work of Layard, Nickell and Jackman (1994, 1991) was instrumental in highlighting the possible role of LMPs in improving the aggregate trade-off between unemployment and inflation.

Macro-economic or aggregate evaluation

Although not the focus of this research, a note on aggregate evaluation is needed. To estimate the aggregate effect of LMPs requires the evaluator to take account of a wide range of potential impacts. Calmfors (1994), for example, uses a framework to analyse nine different potential effects of labour market programs, including the effect on the efficiency of the matching process of the labour market; impacts on labour demand and supply for varying groups; work incentives and wage effects for the employed ("insiders") through the competitiveness of outsiders and, possibly, through the relative utility of being unemployed; productivity effects if programs are successful in improving the skills of the workforce; tax effects and interactions with other policies.

From this perspective, the evaluation of active measures entails the comparison of two states of the world, S_{0t} and S_{1t} across all periods t within the evaluation horizon. S_0 is the state of the world that prevails if the measure has not been implemented, and S_1 the state of the world which prevails if the measure has been implemented. We see from the above that the two states of the world must be compared across a large range of potential outcome measures and affected populations, and necessarily involves value judgements regarding the relative weights used to compare different outcomes. The evaluation challenge arises from the fact that at least one of these states of the world must be counterfactual and cannot be observed directly. Moreover, if the measure has been implemented, counterfactuals exist for each of the set of states of the world relating to the alternative allocations of those resources.

International reviews of aggregative or macro-economic evaluations can be found in OECD 1988, Calmfors 1994 and Bellmann and Jackman 1996. As mentioned,

Layard, Nickell and Jackman (1994, 1991) provided a seminal analysis. Recent Australian studies include Leeves (2000), Webster and Summers (2000), Webster (1999b), Connolly and Nicol (1997) and reviews can be found in **[6]** Dockery and Webster (2002) and Webster (1997).

3.2   Micro-evaluation – the international experience

This doctoral research has concentrated on a limited number of outcomes for one specific group – namely labour market outcomes for the persons who directly participate in the program or intervention. This more limited aspect of program evaluation is still of utmost importance. Firstly, the participant groups often encompass, with varying degrees of targeting, the most economically and socially disadvantaged persons in the labour market. Thus, even if there are no macro-economic or aggregate efficiency gains, improved outcomes for participants are still of considerable value from the perspective of equity. Indeed, the achievement of improved labour market outcomes for such groups are often clearly stated among the policy objectives for labour market programs. Secondly, it would be generally agreed that the existence of positive impacts for the participant group is a pre-requisite for the existence of wider macro-economic impacts.

Following a standard formulation of the evaluation problem (see, for example, Heckman and Smith 1996) consider an outcome variable Y, say employment status, which is observed for participants and non-participants in a given program. Let P be a dummy variable indicating an individual's participation ($P = 1$) or non-participation ($P = 0$) in the program.

(1) $Y = Y_1$ if $P = 1$

$Y = Y_0$ if $P = 0$

A principal evaluation question is: "What is the effect of the program on those who participated?" [1]. Ideally, we would like to observe Y_1 and Y_0 for the same individual, and the difference between the two would represent the effect of the program. In reality, at any point in time, we can observe only Y_1 or Y_0 for any one individual — they have either participated or they have not. We do not observe the counterfactuals of $Y_0|P=1$, the outcomes that participants would have achieved in the absence of the program, or $Y_1|P=0$, the outcomes that non-participants would have achieved had they been on the program. The usual approach to this problem of missing information is to take Y_0 observed for non-participants as a proxy for what the outcome would have been for the participants in the absence of the program, while controlling for other individual specific characteristics, X, which may affect the outcome, such as human capital variables. If the vector X were to include the treatment received by the participants, we can specify the problem using the expectations operator:

(2) $E(Y_1|X) = g_1(X) = X\beta_1$ if $P=1$

$E(Y_0|X) = g_0(X) = X\beta_0$ if $P=0$

The non-participant group for which Y_0 is observed represents the control or comparison group. Assuming linear relationships and certain other conditions, the effect of the program can be econometrically estimated via methods such as ordinary least squares:

(3) $Y_i = \alpha + \delta P_i + \beta'X_i + \mu_i$

The parameter δ represents the estimated impact of program participation provided that X includes all the other variables that impact upon Y and $\beta'|(P=1) = \beta'|(P=0)$ and μ_i is a well behaved error term.

---

[1] But this is not the only evaluation question. Heckman and Smith (1996: 40-41) provide a discussion of the "parameters of interest" in program evaluation.

Variants of this formulation underlie a vast volume of empirical studies of the effectiveness of LMPs, including **[2]** Stromback *et al* (1999) and **[4]** Stromback and Dockery (2000a). In the United States, empirical work has concentrated on the impact of program participation on earnings, particularly with regard to the major training programs instigated under the Manpower Development and Training Act (MDTA) of 1962, the Comprehensive Employment and Training Act (CETA) of 1972 and the Job Training Partnership Act (JTPA) of 1982 (see Park, Power, Riddell and Wong 1996, Devine and Heckman 1996, Warburton 1996, Riddell 1991, Barnow 1987, Ashenfelter and Card 1985 and Ashenfelter 1978 for assessments of the North American literature).

As in Australia, European studies have concentrated less on earnings and more on other labour market-related outcome measures, such as labour force status (OECD 1988, OECD 1991, Disney and Carruth 1992, Bjorklund and Regner 1996). Haskel and Jackman (1988) and Disney and Carruth (1992) discuss evaluations in the United Kingdom. The British Restart program, introduced in 1986 and seen by many as heralding a major paradigm shift in policy away from welfare to "workfare", has been the subject of a number of evaluations that have generally found a positive effect of program participation on the transition rate out of unemployment (Dolton and O'Neill 1996, White and Lakey 1992). Similar findings have been observed for active labour market programs in Germany, the Czech Republic, Hungary and Poland (Beori and Burda 1996, Micklewright and Nagy 1996, Disney and Carruth 1992, Schellhaass 1991). With relatively high levels of expenditure on active labour market measures, evaluations of Scandanavian programs feature prominently in the European literature (Rosholm 1997; Carling, Edin, Harkman and Holmlund 1996; Korpi 1995; Raaum and Wulfsberg 1995; Bjorklund 1991).

There have been considerable advances in the methodologies employed in these evaluations as evaluators recognised and attempted to address shortfalls in specifications of the evaluation problem such as that in (3). These arise largely from the fact that participation will often not be exogenous, but endogenous. In the case of heterogeneous impacts of the treatment, it is likely that those who stand to benefit most from the program are also more likely to participate. If the reasons for this higher expected gain from participation cannot be observed or controlled for by the

evaluator, then a projection of the expected gain observed for participants onto the wider population of the eligible target group will clearly inflate the actual expected benefits from program participation. Further, even with homogenous treatment effects, there may be factors which make individuals both more likely to achieve superior outcomes, Y, and more likely to participate in the program. If this selection process is not controlled for, the endogeneity will result in biased estimates of the coefficient δ.

A sizeable literature has since emerged on the specification of the evaluation problem and the development of estimators for identifying the parameters of interest under various conditions. James Heckman has made a leading contribution to this literature (see as examples Heckman 1976, 1979, 1990; Heckman, Ichimura and Todd 1997, Heckman and Smith 1996; Heckman and Hotz 1989; Heckman and Robb 1985). A useful recent review can also be found in Blundell and Costa Dias (2000). A common statement of the selection problem is to consider participation to be determined by an underlying index variable, P^*, which is dependent upon a set of variables Z and an error term:

(4) $\qquad P^*{}_i = \lambda Z_i + v_i$

The index variable is not observed as such. What is observed is whether or not the individual does participate. To relate the latent variable to the participation outcome we assume that participation occurs when the index value exceeds some threshold, usually standardised to equal zero.

(5) $\qquad P_i = 1$ if $P^*{}_i > 0$

$\qquad\quad P_i = 0$ if $P^*{}_i \leq 0$

To relate the independent variable to the participation outcome, this formulation then lends itself to estimation of the probability that an individual is observed to participate in the program:

(6) $\qquad \Pr(P = 1)_i = h(Z_i)$

17

Selection bias arises in the case of $E(\mu|P=1,X) \neq 0$ or $\text{cov}(\mu_i, \nu_i) \neq 0$.

The difficulty in establishing the counterfactual for participants, and thus the conditions necessary to identify the program effect, led to the advocacy of randomised experiments. In theory, social experiments, in which persons are randomly assigned to either a treatment or control group, solve the evaluation problem by removing any other selection process and establishing the counterfactual. Given true randomisation, we have $E(X|P=1) = E(X|P=0)$ and $E(Y_0|P=1) = E(Y_0|P=0)$; and hence an unbiased estimate of the impact of the program on those who participate is given by $E(Y_1) - E(Y_0)$. The ease of interpretation of this "difference in means" estimator of the program effect in the policy arena is a further great advantage of experimental data.

However, a pure experiment is not always easy to achieve. Issues include the cost and timeliness of trials; how to deal with those allocated to the treatment group but either do not complete or perhaps not even commence treatment, and those who are allocated to the control group who seek alternative treatment. Experiments also face the ethical dilemma of excluding some individuals from treatment. Thus, despite their advantages, evaluations based on experimental data are not abundant in the literature. Considerable attention has also had to be paid to methods of estimating the impact of programs using non-experimental data. Before returning to the relative merits of experimental research designs, the role of longitudinal data in addressing the evaluation problem is discussed.

Longitudinal Data

The preceding formulation of the evaluation problem abstracts from any temporal dimension other than to assume that the outcomes are observed in some post-program period such that program participation status is known. In reality, there are timing issues relating to the measurement of the outcome; the timing and duration of the intervention and the individual's prior duration in various states. The application

18

of longitudinal data both assists in controlling for such factors and in providing additional information that can assist in answering the evaluation problem (See Schömann 1996, Heckman and Singer (eds) 1985, Ashenfelter and Card 1985). Following Chamberlain (1985: 5), the importance of longitudinal data can be considered in the following context. Rewriting (3) explicitly as a cross-sectional regression for period t:

(3`) $E(Y_t|X_t) = \beta'X_t$

we can see that potentially the vector of independent variables could be expanded to include all past and future values of variables in X, and such a regression could also be specified for all past and future Ys for, say, T periods. This gives:

(7) $E(y|x) = \Pi x$

where $y' = (Y_1,...,Y_T)$, $x' = (X_1',...,X_T')$, x and Π have dimensions $T \times k$ and $T \times Tk$, respectively. Estimating the cross-sectional model (3`) as a reduced form of (7) is effectively equivalent to assuming that $\Pi = I_T \otimes \beta'$, which is to say that no past or future values of the elements of X provide additional information on Y in the current period. Clearly this is a very strong assumption, particularly given the trend toward the specification of fuller behavioural models rather than quasi-experimental comparisons in addressing selection issues in evaluations based on non-experimental data. "It is not sufficient to have data from one period before and one period after the programmes, but data from several periods either before or after the programme are needed." (Bjorklund 1991: 85).

Longitudinal data also provides the means to control for unobservables that may affect the outcome variable. Taking the case of $T=2$ in (7) — in which the system consists only of two consecutive cross-sections — differencing those equations yields a model for the change in the outcome variable:

(8) $E(Y_t - Y_{t-1}|X_t - X_{t-1}) = B'(X_t - X_{t-1})$

Fixed unobservables and observables in X will be differenced out of the equation. The seminal work of Orley Ashenfelter provides an excellent exposition of the value of longitudinal data. He was also one of the early advocates of the experimental approach. Using a 1964 sample of participants in the MDTA, Ashenfelter (1978) estimated annual earnings functions as an autoregressive function of prior years' earnings, age and a fixed individual effect. A control group of non-participants was selected from the Continuous Work History Sample to control for the effects of changing economic conditions over time, invoking the implicit assumption that the structure of the earnings functions are the same for participants and non-participants. A dummy variable is included taking on a value of 1 after the year of training and 0 otherwise. By taking the difference between the estimated earnings functions from a post training year and a base year (either the training year or a prior year) and estimating this "difference equation" across all individuals, an estimate of the impact of training is obtained for that year. The equation is estimated for a number of post-training years to give an indication of the "permanency" of the impact of training on earnings.

Two important findings from Ashefelter's work for the discussion at hand are:

(i) the results are sensitive to the base year used. Since, at the time, participants were selected from the unemployed, they typically had low earnings in the year prior to training (1963) and this gave a higher bound to the estimate of the effect of the training on future earnings. If individuals' earnings in 1962 or 1961 were used as the base year, the estimated permanent increase in earnings as a result of training is reduced significantly. As would be expected, earnings were also considerably lower in the year of training, indicating an opportunity cost to participation. Hence, to the extent that such a fall in earnings is temporary and earnings would have recovered anyway, evaluations may overestimate the impact of program participation. The observation that persons experience lower earnings just prior to program participation became known in the literature as "Ashenfelters' dip"[2].

---

[2] This has further implications for selection. If participation in a program requires foregone earnings, persons are likely to participate in a period in which the opportunity cost is lowest, which will be during periods of lower than usual earnings.

(ii) if the autoregressive nature of the earnings function is ignored, that is the coefficients on Y_{t-1}, Y_{t-2}, Y_{t-3} ... in the earnings function are assumed to be zero, the estimated impact of training on earnings is much higher. Yet this is precisely what evaluations assume when comparing only pre and post-training earnings levels.

Heckman and Singer warn that longitudinal data should not be considered as a panacea, and carries with it it's own potential complications such as the exaggeration of measurement error, length-bias sampling and truncation bias (1985: xi). Heckman and Smith also argue:

> "… different and not necessarily more plausible assumptions can be
> invoked in longitudinal analyses than in cross-section analyses. The fact
> that more types of minimal identifying assumptions can be invoked with
> longitudinal data (because longitudinal data can also be used as cross–
> section or a repeated cross-section) does not increase the plausibility of
> those assumptions that uniquely exploit longitudinal data." (1996: 53)

However, nearly all rigorous evaluations of LMPs now utilise longitudinal data and this popularity would suggest general acceptance that longitudinal evaluation designs do offer major advantages in solving the evaluation problem, particularly in the absence of experimental data. There is a large number of alternative specifications for exploiting longitudinal properties of available data within which a variety of classes might be distinguished, including discrete and continuous time models; single or multiple period models; transition models and "difference in differences" matching estimators.

When movement between labour market states is the principal concern, models that perhaps best utilise the properties of longitudinal data are duration models with event history data, or hazard models. These are commonly applied to data on the duration of spells of receipt of benefits and of spells of unemployment. In the latter context, the hazard function that is approximated describes the duration specific exit rate out of unemployment, or from unemployment to one of a number of alternative states in a competing risks framework, conditional on program participation status. A good

overview of the technical foundations of this class of models can be found in Cox and Oakes (1984). A large number of studies have used estimation of the hazard function to assess the impact of specific labour market programs. These include **[4]** Stromback and Dockery (2000a), which found positive effects of participation in the evaluation of Australia's *Working Nation* programs on the exit rate from unemployment, and studies which identified positive impacts of Britain's Restart program on the exit rate from unemployment (Dolton and O'Neill 1996) and, additionally, on the exit rate from benefit claimant status (White and Lakey 1992). Rosholm (1997) found that subsidised employment programs had a positive effect on the unemployment hazard when the placement was in the private sector, but a detrimental effect on most groups when the placement was in the public sector. Gritz (1993) finds a similar result for government-versus-private provided training programs with respect to their impact on participants' subsequent employment and non-employment spells, though many of the training programs included in this study would not have been considered as labour market programs as such.

Korpi (1995) investigates previous findings of a non-negative duration in the unemployment hazard in Sweden, contrary to that observed in most other countries. He shows that this results from estimating the hazard using a monotonic specification which yields an average hazard. A more flexible specification reveals that negative duration dependence does take affect after around 5 months, but that this effect is masked when a monotonic hazard is used by the higher likelihood of transition to a LMP as duration increases for the more disadvantaged. Carling *et al* (1996) also suggest that the generosity of payments to persons on LMPs in Sweden – some offering wages above unemployment benefits — may reduce transitions from unemployment to employment as persons approach the duration required for program eligibility. There is also an extensive associated literature utilising duration models to assess the impact of the parameters in unemployment benefit or unemployment insurance schemes, which has highlighted that exit rates from unemployment tend to increase markedly as persons approach exhaustion of support.

Experimental versus non-experimental research designs

There is no question that the availability of experimental data assists the evaluator in identifying the impact of a program on its participants. Practical constraints, however, ensure that in most cases evaluators will continue to have to deal with non-experimental data. Outside of the pure randomised experiment, Blundell and Costa Dias categorise the approaches using non-experimental data into four types: natural experiments, matching, selection models and structural simulation modeling (2000: 428-30). With the exception of the latter, there is considerable overlap in the way these approaches attempt to identify the counterfactual and hence solve the evaluation problem. Natural experiments make use of aspects of policy implementation to identify a naturally occurring control group. For example, if a program is implemented incrementally by region, or applies only to a certain age group, then a comparison of differences in outcomes before and after implementation for those areas or groups affected with that for unaffected "near neighbours" provides the "difference-in-differences estimator" of the program effect.

A comprehensive assessment of matching methods is provided by Heckman, Ichimura and Todd (1997). One strong advantage of matching is that in many cases no particular parametric or functional model needs to be assumed. Again, however, matching can only be done on the basis of observables. Even if the treatment and comparison groups can be matched precisely on the values of the observable characteristics, X, it needs to be further assumed that the means and distributions Y_1 and Y_0 are independent of the participation decision once conditioned on X. One approach to ensuring this assumption holds is to select the matched controls from a sample of individuals who applied to enter the program but did not ultimately participate in it (see Heckman *et al* 1997). This must still leave doubts as to whether these "no shows" are not inherently different to those who do commence.

The selection model was initially developed by Heckman (1976, 1979), and involves a 2-step procedure of initially estimating a participation equation such as (4) and deriving a correction term that is then included as an additional variable in the outcome equation. To identify the program effect, the approach requires an exclusion restriction in which there is at least one independent variable explaining participation

that does not enter into the outcome equation. The approach can be considered as a special case of the Instrumental Variable approach to controlling for endogeneity (see Bowden and Turkington 1984: 39-48), and is also referred to as "selection on unobservables" as it does not require the assumption that no unobservables enter the participation decision (Heckman and Smith 1996, Heckman and Robb 1985).

The results of evaluations using non-experimental methods have proven to be sensitive to the construction of control groups and the specification of econometric models, raising doubt as to whether results from these methods are sufficiently reliable to do away with the need for social experimental trails to guide policy (Ham and LaLonde 1996: 177-178). To test the performance of various evaluation techniques and the specification tests for choosing between models, LaLonde (1986) compares the results of an experimental evaluation of the earnings effect of a temporary employment program in the US with the results that "would have been produced by an econometrician" using a range of commonly employed comparison groups and estimation techniques. He finds the non-experimental results to be highly sensitive to the choice of control group and the econometric specification of the models. While two-step procedures generally gave estimates of the training effect closer to that for the experimental results, LaLonde argues that the standard specification tests for choosing between models would still have left the researcher choosing from a very wide interval of estimates around the experimental result. It should perhaps be noted that a shortfall in the data used for LaLonde's non-experimental estimates was the availability of only one year of pre-training earnings, which precluded testing a number of estimators that have been advocated in the literature.

Fraker and Maynard (1987) exploit the same experimental study to undertake a similar exercise, though with less sophisticated econometric testing, to show that the estimated impact of the training program is sensitive of the choice of comparison group, and that in a number of cases estimates would have led to qualitatively wrong conclusions about the impact. They likewise conclude that evaluations based upon comparisons groups should be avoided in preference for true experimental controls. However, Fraker and Maynard undertake the evaluations for two separate groups that participated in the program - youth and female recipients of AFDC. While the

results do indeed vary markedly for youth, the estimated effects for the AFDC groups derived from using different models and comparison groups were more consistent with those obtained in the experimental evaluation. Although the authors do not offer this explanation, it seems likely that the unreliability of the comparison group estimates for youth arises because unobservables may have a much larger effect for youth. As individuals age and their time in the labour force increases, what were once "unobservables", such as latent ability or motivation, will manifest themselves in "observables" such as earnings and labour market history.

The debate continues today as to the need for experimental versus non-experimental evaluation designs. Many believe that the diversity of estimates derived from non-experimental methods, and the inability to select between them, makes such evaluation of little or no use in informing policy (see Riddell 1991: 63-64). Other researchers, led by Heckman, are more optimistic. Heckman and Hotz (1989) reanalyse LaLonde's data using the most recent approaches to find that a simple set of tests successfully eliminates all but the non-experimental methods that reproduce the inferences obtained by experimental methods. Heckman and Smith (1996) note there have been further developments in methodology to control for selection bias and that randomised experiments themselves are not free of bias: "In the existing literature, the assumptions required to justify experiments are often ignored or downplayed, while those required to justify nonexperimental methods are often overstated." (Heckman and Smith 1996: 83).

Types of bias which arise in experimental studies include randomisation bias, in which the individuals (or units of observation) alter their behaviour once they know that they are the subject of a randomisation experiment, or alter their behaviour to ensure they avoid being part of the experiment; and substitution bias, in which the individuals randomly excluded from treatment seek substitute treatment from other providers. In their review of experimental evaluations of European labour market policy, Bjorklund and Regner conclude:

> "... problems were encountered in some of the actual experiments, which,
> if they have not rendered them total failures, have at least made the
> experiments less useful than expected. However, a number of

25

experiments have produced valuable new insights. Therefore we conclude that there is a role for classical experiments in future evaluation research …" (1996: 111-112).

Given the state of the debate, it would seem appropriate for policy makers to employ a range of techniques in the evaluation of at least the major labour market programs. With cost and timeliness being crucial dimensions for policy makers, some combination of ongoing evaluations using non-experimental techniques and occasionally reinforced by randomised experiments may prove optimal for both obtaining rigorous estimates of program impacts and for "evaluating the evaluations". Policy makers should also remain alert to opportunities to exploit natural experiments that arise in the implementation of policies and low-cost opportunities to impose randomisation, such as when program places are supply constrained and it is already necessary to ration applicants.

3.3   Labour market assistance and micro evaluation in Australia

Review of Australian labour market programs[3]

A high incidence of unemployment and of long-term unemployment also became entrenched features of the Australian labour market from the early 1980s, with estimates of the natural rate of unemployment in the 1990s as high as 9 per cent (Crosby and Olekalns 1998, Debelle and Vickery 1998). As shown in Figure 1, the Commonwealth Government has responded to each major labour market downturn with increased spending on LMPs. Over the years, there has been a myriad of different programs, far too numerous to document here even ignoring State programs. The Bureau of Labour Market Research (BLMR) notes that public sector job creations schemes have been operated in Australia since at least the 1840s and typically coincided with major economic downturns (1984c: 20). Sloan and Wooden (1987), Stretton and Chapman (1990), Webster (1997), **[1]** Stromback and Dockery

---

[3] A number of programs aimed at encouraging skills formation, such as the National Apprenticeship Assistance Scheme, the Commonwealth Rebate for Apprenticeship Full-time Training (CRAFT) and the Australian Traineeship System have not been included here as they were not targeted specifically at the unemployed.

(1998) and **[6]** Dockery and Webster (2002) discuss more recent developments in Australian LMPs. Stretton and Chapman identify three distinct phases in the delivery of programs between 1970 and 1990: 1973 to 1975, 1976 to 1985 and 1985 to 1990 as summarised below. The main features of the1990s were the *Working Nation* package, the Work for the Dole scheme and, perhaps most significantly, the replacement of publicly delivered programs with the competitive employment services market, the Job Network.

Figure 1: The unemployment rate and real Commonwealth Government spending on labour market programs; Australia 1973-74 to 1999-00.
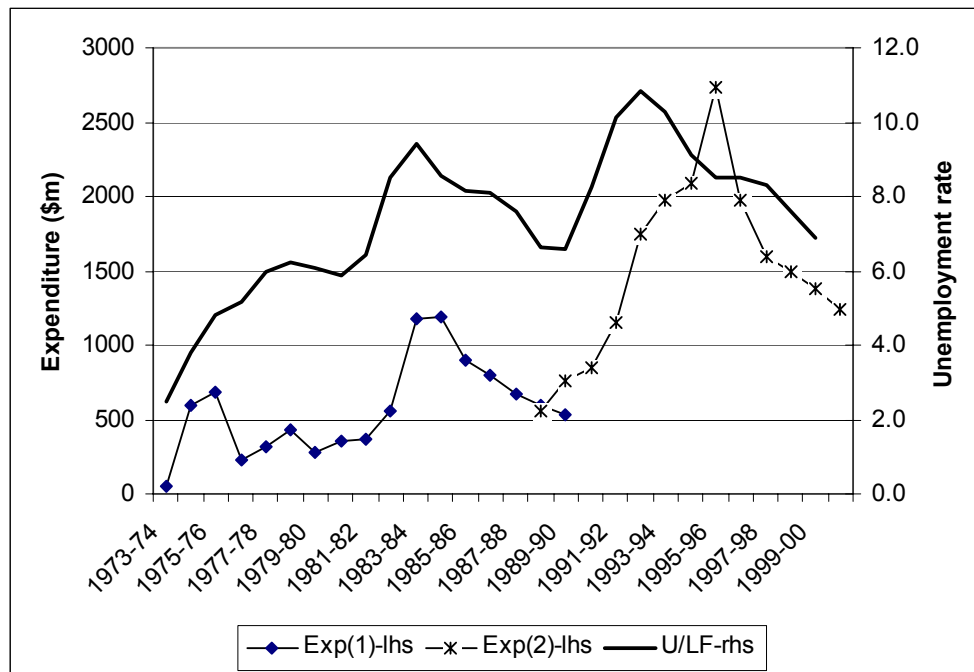
Notes: Expenditure in 1989-90 dollars. Series Exp(1), from Stretton and Chapman (1990), is a sum of expenditure on individual programs. Series Exp(2) is the Cwlth budget expenditure on "labour market assistance to jobseekers and industry". From 1988-89 to 1995-96, expenditure on employment services has also been added in, as these two items appear to have been amalgamated thereafter and to allow greater comparability with expenditure following the introduction of the Job Network.

Stretton and Chapman saw the years of 1973 to 1975 as characterised by a readiness to intervene but an "overzealous" and "simplistic" approach. The main programs at this time were the newly introduced Regional Employment Development Scheme (REDS), a direct job creation program, and the National Employment and Training System (NEAT). Introduced in 1974, NEAT provided funding for unemployed persons to participate in courses at formal education or training institutions and to

subsidise on-the-job training by employers. Between 1976 and 1985 the economy failed to regain full employment before the unemployment began to rise again, this time to double digits in the middle of 1983. A lack of conviction in the effectiveness of programs saw frequent changes in the level of expenditure and mode of delivery plus closer targeting. Following general dissatisfaction with the experience with REDS, no major public sector job creation program was introduced by the Commonwealth in the seven years following the demise of REDS in 1975 (BLMR 1984c). NEAT underwent a number of changes and came to comprise mainly of a range of training schemes aimed at particular target groups before the term was abandoned in 1981 (BLMR 1984a:1). One of those programs introduced under the NEAT umbrella, in 1976, was the Special Youth Employment and Training Program (SYETP), a wage subsidy program that accounted for almost half of all program expenditure by 1978-79 (BLMR 1984a, Vella and MacKay 1986, Stretton and Chapman 1990: Table 2. Hoy 1983 provides a thorough description of the SYETP).

Expenditure on wage subsidies and training programs remained significant throughout this period, however, the emphasis turned again to direct job creation programs as the unemployment rate climbed towards 10 percent in 1982-83. In 1982, $200 million of an anticipated savings of $300 million from a suspension of wage increases for Commonwealth Employees was redirected to the Wage Pause Program, a one-year public sector job creation program targeted to the unemployed who were "especially disadvantaged". The commitment to job creation programs was then extended with the establishment of the Community Employment Program (CEP) in 1983 (BLMR 1984c, Sloan and Wooden 1987: 157-158).

Programs had become fragmented as the government targeted assistance to very specific groups and their number increased from six in 1975 (Stretton and Chapman 1990: 25) to 25 Commonwealth funded programs in 1984-85 encompassing 36 separate components, all in addition to 58 State funded schemes (Sloan and Wooden 1987: 147). After 1985 the Kirby Report[4] provided some much needed direction for LMPs, with recommendations for rationalisation of programs, closer integration with welfare policies and a focus on the equity role in the context of an overall social justice strategy (Stretton and Chapman 1990: 26-30). A new wage subsidy program,

Jobstart, replaced SYETP and a number of other programs in December of 1985 (Vella and MacKay 1986: 213). The labour market improved steadily over these years, and job creation programs again fell by the wayside toward the end of the decade.

In 1990 Stretton and Chapman noted current developments, embodied in the introduction of the Active Employment Strategy, as being a shift from passive to active support of the unemployed, closer integration of welfare and labour market policy, earlier identification and intervention for the disadvantaged and the principle of "reciprocal obligation" between the unemployed and the State.   Looking forward to the 1990s, they saw policy challenges ahead as the labour market had again begun to deteriorate, suggesting an increase in the level of wage subsidies and the return of direct job creation programs as appropriate responses to a looming recession. The downturn that followed was to be more severe even than Stretton and Chapman's "pessimistic scenario". Following establishment of the Committee on Employment Opportunity in January of 1993, the government adopted the main recommendations of the ensuing Green Paper (*Restoring Full Employment*) and White Paper (*Working Nation*). These reinforced the notion of reciprocal obligation through the "Job Compact". The government was to guarantee the LTU with employment or training places of sufficient quality to significantly enhance their employment prospects while, on their part, jobseekers were obliged to accept such offers or forego social security benefits.

Under *Working Nation*, as the package of measures became known, outlays on LMPs were to increase from $1.4 billion in 1993-94, to $2.4 billion by 1996-97, with the LTU the main target. This was to come about through an increase in the number of places and in the level of the wage subsidy for the LTU under JobStart; an expansion of existing brokered employment and training placements and the creation of a new direct job creation program called New Work Opportunities. Importantly, *Working Nation* included an evaluation strategy consisting of a longitudinal panel established in 1994, which is discussed further below. In the event, the *Working Nation* strategy proved overly ambitious, with both placement numbers and funding falling well below their targets (**[1]** Stromback and Dockery 1998).

---

[4] Committee of Enquiry into Labour Market Programs (1985).

With economic recovery underway and the election of the Coalition Government in 1996, the bulk of the *Working Nation* measures were dismissed as being expensive and ineffective. The wage subsidy program JobStart was retained as the major LMP on the basis of departmental evaluations that suggested it was the most effective form of assistance. Though New Work Opportunities was scrapped, a new direct job creation program, Work for the Dole, re-emphasised the reciprocal obligation philosophy, now rebadged as "mutual obligation" and with the government's obligation much reduced. The rhetoric of the time clearly suggested a "get tough" stance on the unemployed, one that seemed to carry some favour with the electorate, and a distinction has been maintained between such mutual obligation measures and "true" LMPs. The Work for the Dole scheme was introduced toward the end of 1997. It initially applied only 18 to 24 year old jobseekers who had been on income support for six months or more. Persons allocated to a Work for the Dole project - typically a community or environmental project outside of the market sector - risked sacrificing social security benefits for failing to attend. "Mutual obligation requirements", of which Work for the Dole is the main program, were later extended to older job seekers and the scheme was expected to expand to 50,000 places in 2001.

Undoubtedly the most fundamental shift in the approach to delivering LMPs came in 1997 when the government announced it would "cash out" the funding for publicly delivered programs and called for tenders to provide one or more of three main employment services: job brokerage services, job search training and intensive assistance (see Productivity Commission 2002; **[5]** Dockery and Stromback 2001; OECD 2001; Department of Employment, Workplace relations and Small Business (DEWRSB) 2001a, 2001b, 2000; **[3]** Dockery 1999). The Commonwealth Employment Service was abolished and the new competitive market for employment services, the Job Network, came into effect in May of 1998. While Work for the Dole and a handful of smaller programs such as the New Enterprise Incentive Scheme have been retained, the Job Network is now the major mechanism for delivering assistance to the unemployed, with 320,000 commencements in intensive assistance in 2000.

Evaluations of Australian programs

Overviews of the findings from Australian LMP evaluations can be found in **[6]** Dockery and Webster (2002) and Webster (1997). From its formation in 1980, the BLMR took on increasing responsibility for the evaluation of Commonwealth programs during its life. The BLMR conducted specific evaluations of NEAT (BLMR 1984a), the SYETP (Vella and McKay 1986, Moy 1983, Stretton 1982) and the Wage Pause Program (BLMR 1984c, Curtain 1984), as well as more general reviews of findings and methodological issues (see McKay and Hope 1986, Lewis and Ryan 1985, BLMR 1984b). **[4]** Stromback and Dockery (2000a) and DEETYA (1997: Appendix A) summarise results of post-program monitoring studies for the *Working Nation* programs in the mid-nineties.

The BLMR evaluation of NEAT was based on two mail surveys of persons who had participated or withdrawn from NEAT training. Constructs used to measure labour market outcomes included the time taken to find work following the training, earnings and employment status at the time of the survey. Data on the total time in employment and unemployment in the twelve months prior to the survey and in the twelve months prior to the NEAT training were also collected. No control group was surveyed, rather a comparison between outcomes for those who completed their NEAT placement, those who withdrew during the placement and those who withdrew before training commenced was provided as a possible indicator of the impact of the treatment. This permitted a difference-in-differences estimator for the time in employment and unemployment. On average, completers exhibited slightly better employment outcomes than non-completers as well as larger reductions in time spent in unemployment from the pre-training period. Measured outcomes were superior for formal trainees compared to in-house trainees. As acknowledged in the report the observed differences may well have been due to differences in the attributes of completers compared to non-completers (BLMR 1984a: 39). While using non-completers is a convenient way of controlling for selection *into* programs, it faces the obvious limitation that non-completers differ from completers in non-trivial ways that also influence outcomes. Given also the potential for response bias — the response rate for both surveys was around 53% but markedly lower for non-

completers — the evaluation offers little empirical evidence as to whether or not NEAT had a positive impact upon the participants.

Again in the absence of a control group, Stretton (1982) presents a comparative analysis of five youth employment and training programs, of which three fell within SYETP. Logit models of the probability of being currently in employment and of having had any full-time job since the assistance ceased are estimated based on a post-assistance survey. Controls are included for age, gender, level of schooling, duration of unemployment prior to entering the program and State. The results indicate that youths who participated in the one course based program, the Education Program for Unemployed Youth, achieved significantly worse labour market outcomes than those who had participated in any of the four work-based training programs.

Hoy (1983: 17-20) discusses three attempts at estimating the net employment gains due to SYETP based on employers' responses to survey questions asking whether the vacancy was due solely to the subsidy, to replace workers who left or due to expansion, later combined with questions on what action they would have taken had no subsidy been available. The estimated net gains range from 19 percent to 33 percent of private sector placements. Conversely, between 67 and 81 percent of placements represent deadweight loss, substitution or displacement effects. In both the NEAT evaluation and Stretton's (1982) evaluation of the employment-based training programs for youth, the beneficial effect was limited to participants being kept on by their "host" or placement employer. Obviously retention rates vary with economic conditions, but a review of surveys of SYETP participants shows that at least one-third were retained in employment for six months or longer after assistance ceased (Hoy 1983: 25-26).

Vella and MacKay (1986) model the number of commencements in SYETP to find that demand factors are preferred to supply factors in determining commencements, including a measure of the net wage paid under the scheme (after taking account of the level of the subsidy) relative to junior weekly award rate. The estimates imply that a 10 percent fall in the relative wage would lead to a 6.6 percent increase in monthly SYETP commencements in the immediate term and a longer-term increase

of 21 percent. The results need to be treated with caution on a number of accounts. The demand models use a large number of independent variables, including monthly dummies, in explaining a relatively small number of observations on the independent variable. Not surprisingly the R-squared varies from 0.86 to 0.97 for the three models reported, and the adjusted R-squared statistic is not reported. In the model achieving an R-squared of 0.97, 17 independent variables are used for 33 observations! Further, the authors note that the subsidy level was at times increased in response to a rise in youth unemployment. Thus the results pertaining to the subsidy level may themselves be capturing supply effects with the expected results – a rise in supply and hence the subsidy leads to a rise in commencements. In any case, the finding cannot be used to imply that an increase in commencements represents a net increase in employment for the target group (and indeed, Vella and Mackay do not make this claim) as we know nothing of the outcomes those youth would otherwise have achieved.

Projects approved under the 1983 Wage Pause Program created over 20,000 jobs with an average duration of 24.7 weeks and at an average cost of $11,380 per job. Placements were not as tightly targeted to the long-term unemployed as was expected, with only 71 percent of placements going to persons unemployed for eight months or longer (BLMR 1984c). A number of components of the evaluation of the WPP appear never to have been published, including estimates of the economic impact of the program and of its impact on the participants as foreshadowed in the interim report (BLMR 1984c: appendix 1). This is unfortunate as the methodology was to include a survey of participants and a quasi-control group comprising of persons referred to a WPP project but not placed. The evaluation plan is discussed in more detail in Curtain (1984), including the survey response rates for both participants and the control group.

Anticipating that wage subsidy programs would become a permanent feature of active labour market policy, Lewis and Ryan (1985) noted the shortfalls inherent in previous efforts to assess the employment generating effects of wage subsidies. The bulk of these had relied on surveys of employers that sought in various ways to identify whether workers put on were truly marginal, and did not simply replace existing workers or workers that would have been hired anyway. Focussing on the

impact on aggregate employment rather than on the individual, Lewis and Ryan recommended greater attention be paid to estimating the elasticities of demand and substitution for different categories of labour. It was McKay and Hope who stressed the BLMR's commitment to "...the use of comparison groups because, despite the possible methodological limitations, this is the only basis on which the net effectiveness of program participation can be assessed." (1986: 13). They provide what appears to be the first departmental acknowledgement of the problems of selection bias, the advantage of experimental research designs and the potential methods of retrieving the necessary conditions to obtain unbiased estimates of program impacts in their absence.

In doing so, they pre-empted the reluctance in Australia to use random assignment in the evaluation of LMPs: "It is often argued ... it is not morally or politically defensible to experiment with program delivery by allowing access for some individuals and not others" (1986:13). In fact, it seems the Kirby Report of 1985 and the deliberations of the BLMR at this time were to set the tone for the evaluations for the remainder of the century and beyond. Matched comparisons groups using departmental post-program monitoring (PPM) surveys and/or administrative data were to become the principal source of information on the effectiveness of the various programs throughout the 1990s and the approach has now been carried into the evaluation strategy for the Job Network. All micro-evaluations since have focussed on employment status as the outcome variable, reflecting only one part of the Kirby Report recommendation that the primary objective of any LMP should be to improve the long-term employment and earnings prospects of the target group (McKay and Hope 1986: 9).

Throughout the 1990s the government conducted PPM surveys and reported participants' labour market status three months after completion of a program. These estimates represented "gross outcomes" as they made no allowance for the proportion of participants who would have found work anyway (deadweight loss). The wage subsidy program JobStart consistently displayed the highest gross outcomes, with typically 50 to 60 percent of participants in employment three months after assistance ceased. Gross employment outcomes for the host of other job search, training and employment placement programs fell mostly between 25 to 50 percent.

The higher gross outcome for the wage subsidy program held across sub-groups of jobseekers according to the major characteristics that may reflect labour market disadvantage, namely age, gender, duration of unemployment and other special target groups. More recently PPM data has been used to assess the effectiveness of the New Enterprise Incentive Scheme (Kelly *et al* 2001 and **[7]** Dockery 2002). As with previous government evaluations, these studies find NEIS to achieve high gross employment outcomes, but that it is a relatively expensive form of assistance. From around 1990, the Department also began to use matched control groups drawn from the register of unemployed persons to obtain estimates of Y_0, and thus estimates of the "net impact" of assistance. Combined with data on placement costs, the "cost per net impact" was calculated as a measure of the cost effectiveness of programs. Matching was on the basis of gender, age and duration of unemployment. Twenty to 30 percent of the control group for wage subsidy programs were found to be in unsubsidised employment after three months, giving a net impact of about 30 percent. The prima facie interpretation - that 30 percent of the participants in wage subsidy programs find work and would not have done so had they not participated – is an impact of considerable magnitude. However, for the range of other programs, the net impact was found to be of the order of 5 to12 percent and would have to be considered quite marginal.

Taking the estimate of the net impact on employment together with unit placement costs, the cost per net impact for programs was estimated to range from just over $4,000 for wage subsidy and job search training programs to around $10,000 for training programs, $50,000 for the JobSkills training and employment placement program and almost $150,000 for the New Work Opportunities direct job creation program (DEETYA 1997: 15). These estimates contributed to the scrapping of the Job Compact's strategy of guaranteed job placement for the long term unemployed as being expensive and ineffective.

A major innovation of *Working Nation* was the embodiment of an evaluation strategy within the set of assistance measures of which a longitudinal survey of jobseekers, the Survey of Employment and Unemployment Patterns, was a key element. SEUP ran from September 1994 to September 1997. The sample population included a sub-sample of known LMP participants, a sub-sample of jobseekers and a

representative population reference group. Stromback and Dockery (**[4]** 2000a and 2000b) draw on this data to estimate the impact of participation in four main types of LMPs (training, brokered employment placements, job search assistance and wage subsidies) in a duration framework and to address some of the evaluation issues associated with the Departmental matched-control group approach (see also **[6]** Dockery and Webster (2002)). They find an unrealistically large and positive estimated effect of program participation on the exit rate out of unemployment, leading them to suspect a strong selection bias effect. The relative order of programs in terms of the magnitude of their estimated impact was consistent with Departmental net impact studies with the exception that the brokered employment placements targeted to the most disadvantaged jobseekers were now estimated to have a greater impact than training programs or job search assistance. They also find that participation in LMPs extended the expected duration of subsequent working spells following a period of unemployment, contrary to concerns that programs tended to "recycle" people into and out of short term or dead end jobs.

Using transition analysis to replicate the matched-control groups approach, Stromback and Dockery (**[4]** 2000a, 2000b) find broadly consistent results with departmental evaluations over the same period. The scope for non-response bias is far smaller in the SEUP data, and the availability of a much wider range of controls for individual characteristics had little impact on the estimates of program effectiveness. However, rudimentary efforts to control for selection bias cast doubt over the positive impact for each of the types of program except the wage subsidy program. This is perhaps surprising as **[3]** Dockery (1999: 145) and **[4]** Stromback and Dockery (2000a: 42) point out that the wage subsidy program was, *a priori*, expected to exhibit strong signs of selection bias. As JobStart operated, jobseekers used their eligibility for a wage subsidy as a selling point to potential employers when trying to find work. However, a jobseeker was not considered to have commenced a placement until an employer agreed to take them on for the subsidy period. Thus, from the pool of jobseekers eligible for the subsidy, a strong selection bias must have existed in which those who actually commenced the program had attributes positively associated with the likelihood of gaining employment. **[6]** Dockery and Webster (2002) suggest that for comparability between programs, wage subsidy outcomes should be measured as a proportion of all persons *eligible* for the

subsidy, not as a proportion of those taken on by employers. That would more accurately reflect the effect of the program treatment, which technically is the eligibility for a subsidy. Following the downturns of the early 1980s and early 1990s, take up rates for wage subsidy programs fell well below target level levels, while brokered job placements typically increase during downturns (Stretton and Chapman 1990: 23, [1] Stromback and Dockery 1998: 27). Clearly the effect of this cyclical pattern could produce a favourable bias in measures of the relative effectiveness of wage subsidy programs if they are taken only on the basis of outcomes for those who did secure placements.

The matched-control group approach has carried over to the Job Network, with surveys being used to measure jobseekers' labour market status at set times after conclusion of assistance, and for a comparison group matched from persons registered as unemployed. The evaluation strategy for the new employment services market acknowledged that "a planned control group" would provide better estimates, but this was rejected in favour of the matched comparisons groups on the basis that it "would raise serious legal and ethical difficulties." (DEETYA 1998: 12). There is also now an added focus on outcomes used to determine providers' eligibility for an outcome payment, and these derive mainly from changes in the jobseeker's drawings on social security benefits as determined from administrative data. The evaluation strategy also proposed the use of a longitudinal cohort survey of persons registered as unemployed, combining administrative data and client surveys from late 1998 to mid 2001. This seems not to have been implemented.

The Job Network is undoubtedly a very bold policy reform, even in an international context, and rigorous evaluation of its performance would be of great interest to policy-makers throughout the world. The government has released several reports assessing the success of the transition to the Job Network (DEWRSB 2001a, 2001b, 2000). These have been quite wide ranging, encompassing process evaluation, issues of equity and access and using a number of performance indicators, such as job-seeker and employer satisfaction surveys. However, the "net impact" estimator of the differences in employment and off-benefit outcomes between participants and the matched control group, combined with their resulting implications for the "cost per positive outcome", have been the major focus in measuring the effectiveness of the

Job Network and in comparing it to the previous regime of publicly delivered LMPs. By and large independent researchers have not had access to micro-data that would permit an evaluation other than in cases where it has been provided by individual Job Network providers.[5] Instead, they have had to offer what insights they can through assessment of the Job Network model and critical reviews of Departmental evaluations (**[5]** Dockery and Stromback 2001; Webster 1999a; **[3]** Dockery 1999; Harding 1998).

There is strong evidence that cost efficiencies have been realised from the implementation of the competitive market in employment services and the linking of payments to outcomes. In the case of job search assistance for the shorter-term unemployed, it also appears that outcomes for the unemployed have been maintained. However, it is questionable whether that same can be said of Intensive Assistance targeted to the most disadvantaged jobseekers and whether the evaluation methodology is, in any case, sufficient to make such an assessment (see **[6]** Dockery and Webster 2002; **[5]** Dockery and Stromback 2001 and **[3]** Dockery 1999). In particular, it is pointed out that the profit driven market creates added incentives for providers to invest in identifying and attracting jobseekers with characteristics that are positively associated with the likelihood of achieving a paid (positive) outcome, and that the separation of the referring agency from providers adds further potential selection processes that evaluators need to be aware off (**[5]** Dockery and Stromback 2001). In the most comprehensive review of the existing evidence to date, the Productivity Commission has endorsed the Job Network's underlying "purchaser-provider" model for delivering employment services, largely on the grounds that "… net employment outcomes are small, similar to previous programs. However, the total costs are much less …" (2002: xx).

One study that has been able to utilise departmental administrative data in an assessment of the Job Network is **[7]** Dockery (2002). NEIS was one of the few LMPs retained in comparable form with the introduction of the Job Network. Dockery uses the fact that the available PPM data spans placements both prior to and since the introduction of the Job Network to find similar gross outcomes under each regime, and some cost efficiencies again appear to have been realised under the Job

---

[5] "Providers" is the term used for the agencies contracted to deliver employment services.

Network. But, importantly, little evidence is found that equity outcomes have deteriorated based on the placement ratios and relative outcomes for disadvantaged groups. Given the minor changes in the parameters of the scheme, Dockery describes this as a "weak test" of the impact of the Job Network. Nevertheless, the central tenets of the reform agenda of price competition through the tender process and payments linked to outcomes did come into affect.

A number of more specific limitations of the matched control group approach used in the Departmental evaluations of the *Working Nation* programs and now the Job Network have also been highlighted (see also OECD 2001). As published, the various evaluations have ignored issues of non-response bias to the PPM surveys. The Department also adopted the curious convention of removing persons in further assistance from both the numerator and denominator when calculating the proportion with positive outcomes. Surely the alignment of outcome measures with policy objectives would dictate that those in further assistance should be included as a negative outcome ([4] Stromback and Dockery 2000a: 32). Of particular concern is the measurement of outcomes at a set time after the end of assistance rather than from commencement, and the matching of participants to controls at the time of completion of assistance. In making comparisons across programs of different durations, outcomes at a given duration from commencement or referral would be more meaningful (see Stromback and Dockery 2000b, [6] Dockery and Webster 2002). This practice will also have introduced an upward bias in estimated net impacts for the Job Network in the initial Departmental evaluations, since at that stage a disproportionate number of those completing assistance will be early completers, for whom the main reason for ceasing assistance — and hence being included in the evaluation's participant group — will be that they have gained employment ([6] Dockery and Webster 2002). The Productivity Commission reiterates many of these concerns in a comprehensive review of Departmental evaluations of the Job Network (Productivity Commission 2002: Chapter 5 and Appendix E).

After reviewing the state of knowledge on what does and does not work in offering assistance for the long-term unemployed, Dockery and Webster make a case for greater access to administrative performance data for individual researchers and for

the use of randomised experiments that have been repeatedly rejected in Australia on the grounds of equity considerations. Curtain (1984) and Stromback and Dockery (2000b) have previously stressed the need for independent evaluations. In reference to Table 1, the lack of evidence from experimental evaluation designs increases the chances of both forms of errors being made by policy makers — denying all of the target group access to a program that would have offered them assistance, when in fact following an evaluation period all could have had access to this program; and subjecting all of the target group to a program that is ineffective, or that may even be detrimental to their labour market prospects.

Table 1: Policy maker's decision – potential outcomes

	Introduce program	Don't introduce program
Program is effective	√	X
Program is ineffective	X	√

Depicting the policy choice in this light may help to overcome some of the reluctance by policy-makers and the public, on the basis of ethical concerns, to accept random trials. Randomised trials could also be conducted with no reduction in access by bringing forward assistance by random assignment and using the variation in the timing of the intervention to assess the impact of the program, although this would come at the cost of some loss of applicability of the estimates. Further, under the Job Network, a significant proportion of persons do not select a preferred provider and are assigned to a provider by the Department. Randomisation of this process would enable identification of the relative impact of providers and treatments (**[6]** Dockery and Webster 2002).

Conclusions

From the preceding overview, it can be seen that policy developments in relation to LMPs have been largely driven by the government's need to respond to fluctuations

in aggregate unemployment and, more recently, by ideology. The extent to which they have been informed by evaluation has been limited by a range of factors — technical, practical and political. On a number of occasions the government has drawn on the findings of evaluations in announcing the decision to end programs, notably direct job creation programs. However, it must be said that at the time the comprehensive set of *Working Nation* measures were introduced in 1993-94, policy-makers had little real Australian-based evidence on what forms of assistance to the unemployed did or did not improve the labour market outcomes of the participants. The decision to scrap the Job Compact measures targeted to the most disadvantaged was also made before any findings were available from SEUP, which was to be the centrepiece of the *Working Nation* evaluation strategy.

While the United States has led the charge through the evaluations of the range of training programs provided for the unemployed over the years, the significant institutional differences between the US and Australian labour markets means the results have limited applicability for domestic policy. The evaluation of LMPs in Australia has lacked the rigour in design and the econometric sophistication that has been recognised as necessary by the international literature to provide unbiased estimates of the impact of programs. The recent work by Stromback *et al* ustilising the SEUP data (Stromback and Dockery **[4]** 2000a, 2000b; **[2]** Stromback *et al* 1999) appears to mark the first major Australian LMP evaluation to use longitudinal data or to introduce controls for selection bias beyond matching on observables or using non-completers as a control group, and indeed the first time appropriate data has been available to independent researchers for such a purpose.[6] It should be noted, too, that Australia is not alone in this regard. According the OECD's John Martin most European countries have failed to adequately evaluate their LMPs, and he actually singles Australia out as one of the countries "… beginning to see the light as regards undertaking rigorous evaluations of their labour market programs." (1998: 285). However, this review makes it clear that Australia can be counted among those countries that must, in Martin's view, strive for more systematic and rigorous

---

[6] Although Miller and Volker have extensively analysed other characteristics of the youth labour market using the Australian Longitudinal Survey, including participation in training and the effectiveness of the Commonwealth Employment Service as a job search option (Miller and Volker1996, Miller 1990).

evaluations of labour market programs if we are to expand the international knowledge of what works and why for the unemployed (1998: 293).

3.4   References

Ashenfelter, O. (1978), "Estimating the effect of training programs on earnings", *Review of Economics and Statistics*, 60, 47-57.

Ashenfelter, O. and Card, D. (1985), "Using the longitudinal structure of earnings to estimate the effect of training programs", *Review of Economics and Statistics*, 67, 648-60.

Barnow, B. (1987), "The impact of CETA programs on earnings: A review of the literature", *Journal of Human Resources*, XXI, 4, 606-639.

Bassi, L. (1984), "Estimating the effects of training programs with non-random selection", *Review of Economics and Statistics*, 66, 36-43.

— (1983), "The effect of CETA on the post-program earnings of participants", *Journal of Human Resources*, 18, 539-56.

Bellmann, L. and Jackman, R. (1996), "Aggregate Impact Analysis" in G. Schmid, J. O'Reilly and K. Schömann (eds) *International handbook of labour market policy and evaluation*, Edward Elgar: UK, 143-62.

Beori, T. and Burda, M. C. (1996), "Active labor market policies, job matching and the Czech miracle", *European Economic Review*, 40, 805-817.

Bjorklund, A. (1991), "Evaluation of labour market policy in Sweden" in *Evaluating labour market and social programmes: The state of a complex art*, OECD: Paris, 73-88.

Bjorklund, A. and Regner, H. (1996), "Experimental evaluation of European labour market policy", in G. Schmid, J. O'Reilly and K. Schömann (eds) *International handbook of labour market policy and evaluation*, Edward Elgar: UK, 89-114.

Blanchard, O. and Summers, L. (1988) "Hysteresis and the European unemployment problem", in Cross, R. (ed.) *Unemployment, Hysteresis and the natural rate hypothesis*, Basil Blackwell: Oxford, 306–64.

— (1987), "Hysteresis in unemployment", *European Economic Review*, 31, 288-95.

Bowden, R. J. and Turkington, D. A. (1984), *Instrumental Variables*, Econometric Society Monographs in Quantitative Economics, Cambridge University Press: Cambridge.

Blundell, R. and Costa Dias, M (2000), Evaluation methods for non-experimental data", *Fiscal Studies*, 21, 4, 427-468.

Bureau of Labour Market Research (BLMR) (1984a), *The National Employment and Training System (NEAT): an evaluation*, Monograph Series No. 3, BLMR, Australian Government Publishing Service: Canberra.

— (1984b), *Submission to Committee of Inquiry into Labour Market Programs*, Monograph Series No. 5, BLMR, Australian Government Publishing Service: Canberra.

— (1984c), *Public sector job creation: interim report on the Wage Pause Program*, Interim Report Series No. 1, BLMR, Australian Government Publishing Service: Canberra.

Calmfors, L. (1994), "Active labour market policy and unemployment - a framework for the analysis of crucial design features", *OECD Economic Studies*, 22, Spring, 7-47.

Campbell, J. and Mankiw, N. (1987), "Are output fluctuations transitory?", *Quarterly Journal of Economics*, 102, 857-80.

Carling, K., Edin, P., Harkman, A. and Holmlund, B. (1996), "Unemployment duration, unemployment benefits and labor market programs in Sweden", *Journal of Public Economics*, 59, 3, 313-334.

Chamberlain, G. (1985) "Heterogeneity, omitted variable bias and duration dependence" in Heckman, J. J. and Singer, B. (eds), *Longitudinal analysis of labor market data*, Cambridge University Press: Cambridge, 3-38.

Committee of Enquiry into Labour Market Programs (1985), *Report of the Committee of Enquiry into Labour Market Programs* (The Kirby Report), AGPS: Canberra.

Committee on Employment Opportunities (1993), *Restoring full employment* (Green Paper), AGPS: Canberra.

Commonwealth of Australia (1994), *Working Nation: policies and programs* (White Paper), AGPS: Canberra.

Connolly, G. and Nicol, D. (1997), *The effect of expenditure on labour market assistance and urban and regional development in the incidence of long term unemployment*, Economic Conditions and Forecasting Section, Department of Employment, Education and Training, Canberra.

Cox, D. R. and Oakes, D. (1984), *Analysis of survival data*, Chapman and Hall: Great Britain.

Crosby, M. and Olekalns, N. (1998), "Inflation, unemployment and the NAIRU in Australia", *Australian Economic Review*, 31, 2 117-29.

Cross, R., Hutchinson, H. and Yeoward, S. (1990), "The natural rate, hysteresis and the duration composition of unemployment in the US", *Quarterly Journal of Business and Economics*, 29, 2, 89-116.

Curtain, R. (1984), "The evaluation of the wage pause programs: approaches and methodology", BLMR Conference Paper No. 45, Bureau of Labour Market Research: Canberra, June.

Debelle, G. and Vickery, J. (1998), "Is the Phillips curve a curve? Some evidence and implications for Australia", *The Economic Record*, 74, 227, 384-98.

Department of Employment, Education and Training (DEET) (1995), *Working Nation Evaluation strategy*, EMB Report 1/95, Evaluation and Monitoring Branch, Economic and Policy Analysis Division, Canberra.

Department of Employment, Education, Training and Youth Affairs (DEETYA) (1998), *Evaluation strategy for the employment services market*, EMB Report 2/98, Evaluation and Monitoring Branch, Analysis and Evaluation Division, Canberra.

— (1997), *The net impact of labour market programs*, EMB Report 2/97, Evaluation and Monitoring Branch, Analysis and Evaluation Division, Canberra.

Department of Employment, Workplace Relations and Small Business (DEWRSB) (2001a), *Job Network: a net impact study*, Evaluation and Program Performance Branch Report 1/2001, April.

— (2001b), *Job Network Evaluation Stage Two: progress report*, Evaluation and Program Performance Branch Report 2/2001, February.

— (2000), *Job Network Evaluation Stage One: Implementation and market development*, Evaluation and Program Performance Branch Report 1/2000, February.

Devine, T. J. and Heckman, J. J. (1996), "The economics of eligibility rules for a social program: a study of the Job Training Partnership Act (JTPA) – a summary report", *Canadian Journal of Economics*, 29, Special Issue 1, S99-S104.

Disney, R. and Carruth, A. (1992), *Helping the unemployed: Active labour market policies in Britain and Germany*, Anglo German Foundation, London

Dockery, A. M. (2002), "The New Enterprise Incentive Scheme: an evaluation and a test of the Job Network", *Australian Journal of Labour Economics*, 5, 3, 351-371.

—— (1999), "Evaluating the Job Network", *Australian Journal of Labour Economics*, 3, 2, 131-158.

Dockery, A. M. and Stromback, T. (2001), "Devolving public employment services: preliminary assessment of the Australian experiment", *International Labour Review*, 140, 4, 429-451.

Dockery, A. M. and Webster, E. (2002), "Long term unemployment and work deprived individuals: issues and policies", *Australian Journal of Labour Economics*, 5, 2, 175-194.

Dolton, P. and O'Neill, D. (1996), "Unemployment duration and the Restart effect: some experimental evidence", *The Economic Journal*, 106, 435, 387-400.

Fraker, T. and Maynard, R. (1987), "The adequacy of comparison group designs for evaluations of employment-related programs", *Journal of Human Resources*, 22, 2, 194-227.

Friedman, M. (1968), "The role of monetary policy", *American Economic Review*, 58, 1-17.

Gregory, R. (1986), "Wages policy and unemployment in Australia", *Economica*, 53, s53-s74.

Gritz, R. M. (1993), "The impact of training on the frequency and duration of employment", *Journal of Econometrics*, 57, 21-51.

Groenewold, N. and Taylor, L. (1992), "Insider power as a source of hysteresis in unemployment: tests with Australian data", *The Economic Record*, 68, 200, 57-64.

Ham, J. C. and LaLonde, R. J. (1996), "The effect of sample selection and initial conditions in duration models: evidence from experimental data on training", *Econometrica*, 64, 1, 175-205.

Harding, D. (1998), "What incentives does Job Network create?", in *Mercer-Melbourne Institute Quarterly Bulletin of Economic Trends*, 4.98, 40-46.

Hargreaves Heap, S. (1980), "Choosing the wrong natural rate, accelerating inflation or decelerating employment and growth", *Economic Journal*, 90, 611-20.

Haskel, J. and Jackman, R. (1988), "Long-term unemployment and the effects of the Community Programme', *Oxford Bulletin of Statistics*, 50, 379-408.

Heckman, J. J. (1990), "Varieties of selection bias", *American Economic Review*, 80, 2, 313-318.

— (1979), "Sample selection as a specification error", *Econometrica*, 47, 1, 153-161.

— (1976), "The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models", *The Annals of Economic and Social Measurement*, 5, 475-492.

Heckman, J. J. and Hotz, V. J. (1989), "Choosing among alternative nonexperimental methods for estimating the impact of social programs: the case of manpower training", *Journal of the American Statistical Association*, 84, 862-80.

Heckman, J. J., Ichimura, H. and Todd, P. E. (1997), "Matching as an econometric evaluation estimator: evidence from evaluating a job training programme", *Review of Economic Studies*, 64, 605-654.

Heckman, J. J. and Robb, R. (1985) "Alternative methods of evaluating the impact of interventions: an overview", *Journal of Econometrics*, 30, 239-267.

Heckman, J. J. and Singer, B. (eds) (1985), *Longitudinal analysis of labor market data*, Cambridge University Press: Cambridge.

Heckman, J. J. and Smith, J. (1996), "Experimental and nonexperimental evaluation", in G. Schmid, J. O'Reilly and K. Schömann (eds) *International handbook of labour market policy and evaluation*, Edward Elgar: UK, 37-88.

Hoy, M. (1983), "Review of five years operation of the Special Youth Employment Training Program", BLMR Conference Paper No. 18, Bureau of Labour Market Research: Canberra, September.

Korpi, T, (1995), "Effects of manpower policies on duration dependence in re-employment rates: the example of Sweden", *Economica*, 62, 353-71.

LaLonde, R. (1986), "Evaluating the econometric evaluation of training programs with experimental data", *American Economic Review*, 76, 4, 604-20.

Layard, R., Nickell, S. and Jackman, R. (1994), *The Unemployment Crisis*, Oxford University Press, New York.

— (1991), *Unemployment: macroeconomic performance and the labour market*, Oxford University Press: Oxford.

Leeves, G. (2000), "Duration-specific unemployment outflow rates and labour market programs", *Australian Economic Review*, 33, 3, 221-34.

Lewis, P. E. T. and Ryan, C. A. (1985) "Wage subsidies, their employment effects and how to evaluate them", BLMR Working Paper No. 51, Bureau of Labour Market Research: Canberra, April.

Lindbeck, A. and Snower, D. (1987), "Union activity, unemployment persistence and wage-employment ratchets", *European Economic Review*, 31, 157-67.

Martin, J. P. (1998), "What works among active labour market policies: evidence from OECD countries' experiences", in Debelle, G. and Borland, J. (eds) *Unemployment and the Australian labour market*, Alken Press: Australia, 276-302.

McKay, R. and Hope, C. (1986), "Advances in the evaluation of labour force programs: issues and methodological approaches", BLMR Conference Paper No. 61, Bureau of Labour Market Research: Canberra, July.

Micklewright, J. and Nagy, G. (1996), "Labour market policy and the unemployed in Hungary", *European Economic Review*, 40, 819-928.

Miller, P. (1990), "Training in the youth labour market", *Labour Economics and Productivity*, 2, 1, 1-25.

Miller, P. and Volker, P. (1987), "The youth labour market in Australia", *The Economic Record*, 63, 182, 203-19.

Nelson, C,. and Plosser, C. (1982), "Trends and random walks in macroeconomic time series: some evidence and implications", *Journal of Monetary Economics*, 10, 139-62.

Organisation for Economic Cooperation and Development (OECD) (2001), *Innovations in labour market policies: The Australian way*, OECD: Paris.

— (1991), *Evaluating labour market and social programmes: the state of a complex art*, OECD: Paris.

— (1988), *Measures to assist the long-term unemployed: Recent experience in some OECD countries*, OECD: Paris.

Park, N., Power, R., Riddell, W. C. and Wong, G. (1996), "An assessment of the impact of government-sponsored training", *Canadian Journal of Economics*, 29, Special Issue 1, S93-S98.

Phelps, E. (1967), "Phillips curves, expectations of inflation and optimal unemployment over time", *Economica*, 34, 254-81.

Productivity Commission (2002), *Independent Review of the Job Network*, Report No. 21, AusInfo, Canberra.

Raaum, O. and Wulfsberg, R. (1995), "Unemployment, labour market programmes and Wages in Norway", Department of Economic Memorandum No. 24, University of Oslo, August.

Riddell, C. (1991), "The evaluation of manpower and training programmes: the North American experience", in *Evaluating labour market and social programmes: the state of a complex art*, OECD: Paris, 43-72.

Rosholm, M. (1997), "Evaluating subsidized employment programmes in the private and public sector", CLS Working Paper, Centre for Labour Market and Social Resarch, Aarhus, Denmark.

Schellhaass, H. (1991), "Evaluation strategies and methods with regard to labour market programmes: a German perspective" in *Evaluating labour market and social programmes: The state of a complex art*, OECD: Paris, 89-106.

Schmid, G., O'Reilly, J. and Schömann, K. (eds)(1996), *International handbook of labour market policy and evaluation*, Edward Elgar: UK.

Schömann, K. (1996), "Longitudinal designs in evaluation studies", in G. Schmid, J. O'Reilly and K. Schömann (eds) *International handbook of labour market policy and evaluation*, Edward Elgar: UK, 115-142.

Sloan, J. and Wooden, M. (1987), "Labour Market Programs", in Freebairn, J., Porter, M. and Walsh, C. (eds) *Spending and taxing: Australian reform options*, Allen and Unwin: Sydney, 146-165.

Stretton, A. (1982), "The short term impact on participants of selected youth employment and training programs", BLMR Working Paper No. 15, Bureau of Labour Market Research: Canberra, November.

Stretton, A. and Chapman, B. J. (1990), "An analysis of Australian labour market programs", CEPR Discussion paper No. 247, Centre for Economic Policy Research, Australian National University: Canberra.

Stromback, T. and Dockery, A. M. (2000a), *Labour market programs, unemployment and employment hazards: an application using the 1994-1997 Survey of Employment and Unemployment Patterns*, Australian Bureau of Statistics Occasional Paper, ABS Catalogue No. 6293.0.00.002, February.

— (2000b), "Evaluation of labour market programs: an assessment", 29th Annual Conference of Economists, Gold Coast (3-6 July, Queensland).

— (1998), "The Job Compact Mark II?", *Australian Economic Papers*, 17, 2, 24-34.

Stromback, T., Dockery A. M. and Ying, W. (1999), "Labour market programs and labour force status", *Australian Bulletin of Labour*, 25, 2, 159-178.

Vella, F. and McKay, K. (1986), "The determinants of take-up of SYETP wage subsidies for youth", *Journal of Industrial Relations*, 28, 210-24.

Warburton, W. P. (1996), "What went wrong with the CETA evaluations", *Canadian Journal of Economics*, 29, Special Issue 1, S105-S108.

Webster, E. (1999a), *Job Network: What can it offer*, Melbourne Institute of Applied Economic and Social Research, University of Melbourne, August.

— (1999b), "Labour market programs and the Australian Beveridge curve: 1978 to 1997", *Economic Record*, 75, 231, 405-16.

— (1997), Labour market programs: A review of the literature, Working Paper No 23/97, Melbourne Institute Working Paper Series, University of Melbourne, October.

Webster, E. and Summers, P. (2000), "The effect of labour market programs on wage inflation", *Journal of Industrial Relations*, 42, 3, 383-97.

White, M. and Lakey, J. (1992), *The Restart effect: does active labour market policy reduce unemployment?*, Policy Studies Institute: London.

Note: For copyright reasons Section 4, which contains the following articles, has not been reproduced.

(Co-ordinator, ADT Project Curtin University of Technology, 01/08/03)

[1] Stromback, T. and Dockery A. M. (1998), "The Job Compact Mark II?", *Australian Economic Papers*, Vol. 17, No. 2, pp. 24-34.

[2] Stromback, T., Dockery A. M. and Ying, W. (1999), "Labour market programs and labour force status", *Australian Bulletin of Labour*, Vol. 25, No. 2, 159-178.

[3] Dockery, A. M. (1999), "Evaluating the Job Network", *Australian Journal of Labour Economics*, Vol. 3, No. 2, pp. 131-158.

[4] Stromback, T. and Dockery, A. M. (2000a), *Labour market programs, unemployment and employment hazards: an application using the 1994-1997 Survey of Employment and Unemployment Patterns*, Australian Bureau of Statistics Occasional Paper, ABS Catalogue No. 6293.0.00.002, February.

[5] Dockery, A. M. and Stromback, T. (2001), "Devolving public employment services: preliminary assessment of the Australian experiment", *International Labour Review*, Vol. 140, No. 4, pp. 429-451.

[6] Dockery, A. M. and Webster, E. (2002), "Long term unemployment and work deprived individuals: issues and policies", *Australian Journal of Labour Economics*, Vol. 5, No. 2, pp. 175-194.

[7] Dockery, A. M. (2002), "The New Enterprise Incentive Scheme: an evaluation and a test of the Job Network", *Australian Journal of Labour Economics*, Vol. 5, No. 3, pp. 351-371.

4. PUBLISHED PAPERS

[1] Stromback, T. and Dockery A. M. (1998), "The Job Compact Mark II?", *Australian Economic Papers*, Vol. 17, No. 2, pp. 24-34.

[2] Stromback, T., Dockery A. M. and Ying, W. (1999), "Labour market programs and labour force status", *Australian Bulletin of Labour*, Vol. 25, No. 2, 159-178.

[3] Dockery, A. M. (1999), "Evaluating the Job Network", *Australian Journal of Labour Economics*, Vol. 3, No. 2, pp. 131-158.

[4] Stromback, T. and Dockery, A. M. (2000a), *Labour market programs, unemployment and employment hazards: an application using the 1994-1997 Survey of Employment and Unemployment Patterns*, Australian Bureau of Statistics Occasional Paper, ABS Catalogue No. 6293.0.00.002, February.

[5] Dockery, A. M. and Stromback, T. (2001), "Devolving public employment services: preliminary assessment of the Australian experiment", *International Labour Review*, Vol. 140, No. 4, pp. 429-451.

[6] Dockery, A. M. and Webster, E. (2002), "Long term unemployment and work deprived individuals: issues and policies", *Australian Journal of Labour Economics*, Vol. 5, No. 2, pp. 175-194.

[7] Dockery, A. M. (2002), "The New Enterprise Incentive Scheme: an evaluation and a test of the Job Network", *Australian Journal of Labour Economics*, Vol. 5, No. 3, pp. 351-371.

[1]   Stromback, T. and Dockery A. M. (1998), "The Job Compact Mark II?", *Australian Economic Papers*, Vol. 17, No. 2, pp. 24-34.

[2]   Stromback, T., Dockery A. M. and Ying, W. (1999), "Labour market programs and labour force status", *Australian Bulletin of Labour*, Vol. 25, No. 2, pp. 159-178.

PREFACE

Both this paper and paper **[4]** of this thesis, Stromback and Dockery (2000a), are based upon analyses of data from the Survey of Employment and Unemployment Patterns (SEUP). The current paper was written when the data from the only initial survey and the first wave of the longitudinal surveys became available. It examines transitions between the state of unemployment at the time of the initial survey to other labour market states one year later, including the effect of participation in labour market programs at some point during the intervening year using the bivariate logit model. The scope for more sophisticated analysis was limited as the authors did not have access to the primary data, but were provided with the services of an officer of the Australian Bureau of Statistics to run programs on the data for a short period.

In **[4]** Stromback and Dockery (2000a) access to unit record file for the full three waves of SEUP permitted significant extensions to the analysis. In particular, continuous time models are estimated taking account of the timing of the intervention (ie. participation in a labour market program) and basic controls for selection bias are introduced.

ERRATUM

On page 168, the sentence "To save space, the full results are not included, but Table 8 summarises the parameters pertaining to the LMP effect." should read "… but Table 7 summarises …".

[3]   Dockery, A. M. (1999), "Evaluating the Job Network", *Australian Journal of Labour Economics*, Vol. 3, No. 2, pp. 131-158.

[4]   Stromback, T. and Dockery, A. M. (2000a), *Labour market programs, unemployment and employment hazards: an application using the 1994-1997 Survey of Employment and Unemployment Patterns*, Australian Bureau of Statistics Occasional Paper, ABS Catalogue No. 6293.0.00.002, February.

ERRATUM

In Tables E1 and E2 (page 55), the figures in the final columns are percentages, not numbers of persons as indicated by the column headings.

[5]   Dockery, A. M. and Stromback, T. (2001), "Devolving public employment services: preliminary assessment of the Australian experiment", *International Labour Review*, Vol. 140, No. 4, pp. 429-451.

[6] Dockery, A. M. and Webster, E. (2002), "Long term unemployment and work deprived individuals: issues and policies", *Australian Journal of Labour Economics*, Vol. 5, No. 2, pp. 175-194.

This paper arises from a joint presentation by the authors to the conference *Creating jobs: the role of government* hosted by Centre for Economic Policy Research and the Melbourne Institute in September of 2001. The *Australian Journal of Labour Economics* selected the paper for publication in a special issue containing the proceedings from the conference. As both authors held positions on certain issues that we hoped to put across to the audience in a limited time, it must be acknowledged that the paper is not as balanced and well supported by theory and empirical evidence as may have been the case had it been prepared from the outset for an academic journal.

In particular, it is acknowledged that:

- There are arguments and empirical evidence to counter some of the claims made in the paper. Certainly much of the past evidence on the effectiveness of direct job creation programs would not be supportive of the proposal for a universal employment program for the very long-term unemployed. I refer readers to the review of the Australian programs contained in the introduction to this Thesis, the review by Martin (1998) and to other papers in the Special Issue of the *Australian Journal of Labour Economics* (vol. 5 no. 2), notably the contribution by Gary Burtless.

- The statement on page 186 that "The general consensus of the findings to date suggests the net impacts for the LTU and VLTU are small …" should be interpreted as the authors' own view of the consensus that can be drawn from the range of studies reviewed, and not to imply that this is the consensus view among economists. I do believe, however, that the statement can be well supported, and draw attention to the majority of estimates of the impacts for the training, job search assistance and brokered employment programs comprising *Working Nation* and the useful review of international evidence by Martin (1998). The suggestion that the "…greatest value from labour market programs for the most work deprived, arises not from the permanent or after program effects, but from being in work while on the program" (p. 186) is given as a plausible corollary if

one was to accept that the post-program impacts are negligible for this group. Again, it is not intended to imply that there is a consensus on this point in the literature.

- The aim of the cluster analysis in this paper is not to add further empirical evidence on the link between individual characteristics and long-term unemployment, an area which has been extensively analysed (see Le and Miller 2001 for a recent contribution). At issue is our observation that labour market programs have largely operated as a "black box", with little attention paid to the actual mechanisms by which different types of assistance improve jobseekers' prospects and to tying these to the identified barriers to employment faced by the target group. This has been a central theme of recent OECD research and has been noted before in Australia: "The fundamental problem with the implementation of LMPs is their failure to address the underlying causes of unemployment …" (Sloan and Wooden 1987: 146). The purpose of the cluster analysis is to make an initial contribution to addressing this problem in Australia by investigating the presence of "like" groups of the long-term and very long-term unemployed for which assistance could be tailored.

- The statement of the arguments for or against random assignment is overly simplified. There are of course many other factors that policy makers and administrators must take into account in deciding whether evaluation by random assignment is feasible or preferable. Indeed random assignment as an evaluation approach carries its own set of difficulties, as is discussed in some depth in pages18 and 23-26 of the introduction to this Thesis.

References

Burtless, G. (2002), "Can supply-side policies reduce unemployment: lessons from America", *Australian Journal of Labour Economics*, 5, 2, 115-142.

Department of Employment, Education, Training and Youth Affairs (DEETYA) (1997), *The net impact of labour market programs*, EMB Report 2/97,

Evaluation and Monitoring Branch, Analysis and Evaluation Division, Canberra.

Le, A. T. and Miller, P. W. (2001), "Is a risk index approach to unemployment possible?", Economic Record, 77, 236, 51-70.

Martin, J. P. (1998), "What works among active labour market policies: evidence from OECD countries' experiences", in Debelle, G. and Borland, J. *Unemployment and the Australian labour market*, Alken Press: Australia, 276-302.

Sloan, J. and Wooden, M. (1987), "Labour Market Programs", Freebairn, J., Porter, M. and Walsh, C. (eds) *Spending and taxing: Australian reform options*, Allen and Unwin: Sydney, 146-165.

[7]   Dockery, A. M. (2002), "The New Enterprise Incentive Scheme: an evaluation and a test of the Job Network", *Australian Journal of Labour Economics*, Vol. 5, No. 3, pp. 351-371.

ADDENDUM

On page 355 it is noted that the estimates of the coefficients on some variables in models 5.3 and 5.4 are not as robust as the estimates in model 5.1. I suggest that this "… will in part be due to smaller sample sizes." In fact, this is unlikely to be a contributing factor since the sample sizes are still large (around 9000 observations) and estimates for some other coefficients are indeed more robust than in model 5.1. Moreover, in the case of some variables higher degrees of significance are observed in the estimation of model 5.2 for which the sample size is less than 6,000 observations. It is more likely that the differences in results between these models reflect either differences in the effect of these characteristics on "employment" outcomes compared to "off-benefit" outcomes; that the missing observations are non-random; or a combination of these two factors.

1. **Thorsten Stromback**

2. **Elizabeth Webster**