8-1-2003

A 'causal' estimate of the effect of schooling on full-time employment among young Australians

Chris Ryan
Australian National University

Longitudinal Surveys of Australian Youth

Research Report 35

A 'Causal' Estimate of the Effect of Schooling on Full-time Employment among Young Australians

Chris Ryan
Social Policy Evaluation, Analysis and Research Centre
Research School of Social Sciences
Australian National University

August 2003

This report forms part of the Longitudinal Surveys of Australian Youth:
a research program that is jointly managed by ACER and the
Commonwealth Department of Education, Science and Training (DEST).

The project has been funded by the DEST LSAY Analysis Grants Scheme.
The Scheme aims to widen the use of LSAY data amongst researchers and encourage new approaches to using the data to address policy issues.

Australian Council *for* Educational Research

ACER

Published 2003 by
The Australian Council for Educational Research Ltd
19 Prospect Hill Road, Camberwell, Victoria, 3124, Australia.

ISSN 1440-3455

ISBN 0 86431 721 2

CONTENTS

TABLES

FIGURES

EXECUTIVE SUMMARY

This paper exploits a policy change that occurred in South Australia in the mid-1980s to generate a 'causal' estimate of the effect of schooling on full-time employment outcomes.

The *Early Years of School* policy changed the way that an identifiable subset of students progressed through junior primary school, causing them to obtain an additional year of schooling for any completed Grade or level compared with their predecessors. The policy affected individuals born in specific months of the year (most of those born in October to February inclusive and some of those born in July to September inclusive). Those born at other times (March to June inclusive) were unaffected by it and provide a natural comparison group for assessing its impact.

The impact of the policy change on the age-grade structure of student cohorts in South Australia is captured in two waves of longitudinal data. The Youth in Transition 1975 cohort captures the environment before the policy change, the Longitudinal Surveys of Australian Youth 1995 cohort the post-policy change environment. Analysis of the observed background characteristics of individuals affected by the policy change and the identified 'comparison' group indicated that the major substantive change between them was the increased education obtained by the affected group because of the policy change.

Based on the analysis of the impact of this policy change, it appears that an additional year of junior primary school generates an increase in the probability of being employed full-time of about 11 percentage points in school leavers' first year out from school. In addition, there was an offsetting effect of about 8 percentage points on that probability because individuals were older when they left school, presumably because wages are institutionally age-based in the Australian youth labour market. These estimates apply only to those students who leave school and do not undertake full-time post-school studies immediately. The results provide evidence that the process of schooling itself has a considerable effect on labour market outcomes. The results also indicate that education reform can have complex outcomes, all of which need to be identified and analysed.

The policy implications of these results are not that all school students should undergo another year of school to improve their likelihood of obtaining full-time employment. The policy change analysed here increased the schooling of a group whose initial experiences may not have provided adequate grounding for their subsequent studies, and rectification of this situation improved their later full-time employment outcomes. The results provide encouragement for other policies designed to redress the disadvantage of students early in their schooling – interventions in junior primary school can have a substantial impact on the employment outcomes of individuals, but all of the effects of education reforms need to be identified carefully to assess their potential impact.

The policy induced 'natural experiment' exploited and the methodological approaches adopted in this paper are well suited to analyses of other phenomena thought to be influenced by education. Potential topics are identified in the conclusion. They include:

- the effect of schooling on other labour market outcomes; and
- the effect on educational achievement and attainment outcomes of slowing the progress of students through school.

Where possible, technical details about the data and methodology used here have been placed in Appendices to the report. Following the introduction, Chapter 2 contains a preliminary analysis of the data and the issues addressed here intended to motivate much of the remainder of the paper. Readers interested only in a broad discussion of the issues and the empirical results might read the introduction and Chapter 2, along with Chapters 6 and 7, which contain the empirical results and concluding remarks, respectively. Those readers who want to place the empirical results and the methodology employed in the context of the existing literature might also read Chapter 3, which summarises the literature on the association between schooling and labour market outcomes and how the causal effect of schooling can be isolated. Those interested in the policy change that forms the basis of the 'natural' experiment used here should read Chapter 4, where the *Early Years of School* policy change is described. Readers interested in the detail of the data and methodology used in this paper should read the descriptions contained in Chapter 5 and Appendices 2, 3 and 4.

1. INTRODUCTION

It is clear from countless international and Australian studies that more-educated people experience better labour market outcomes than those with less education. These better outcomes include lower unemployment rates, higher wages and employment in higher status occupations. What is less clear is the contribution of their education to those better outcomes. Are the better outcomes 'caused' by the education or do they reflect other factors, like innate ability, associated with both the outcomes and educational attainment?

Economists have used a number of approaches to identify the effect of schooling on labour market outcomes independent of the effect of ability and other factors. One has been to exploit the existence of some phenomenon that 'causes' a group of individuals to obtain a different level of schooling from their peers and predecessors to estimate how that different schooling influences their labour market outcomes. This approach is pursued in this paper to estimate the effect of an additional year of schooling on the full-time employment outcomes achieved by young Australians soon after they leave secondary school. These outcomes are estimated over the group of school leavers who do not go on to full-time study after leaving school.

In this case, the extra schooling undertaken by a subset of students was the outcome of a change of policy by one Australian State government. The *Early Years of School* policy influenced the way that some students in South Australia progressed through junior primary school from the mid-1980s. The policy affected individuals whose birthdays occurred in specific months of the year. Individuals born at other times were unaffected by the policy and provide a natural comparison group for assessing its impact on schooling levels and consequent labour market outcomes. Moreover, since ability, motivation and other factors like tastes for education are unlikely to be closely associated with birth date, the resulting estimated schooling effects can be given a causal interpretation.

The use of the term 'cause' has a limited application in this analysis. It should not be interpreted as implying that schooling in some way generated additional full-time employment. Rather it means that when employers selected people to fill available jobs, the additional schooling some young people had undertaken relative to their peers had improved their skills and caused them to be selected for those full-time jobs. In addition, the higher average productivity of school leavers may have meant they displaced other workers.

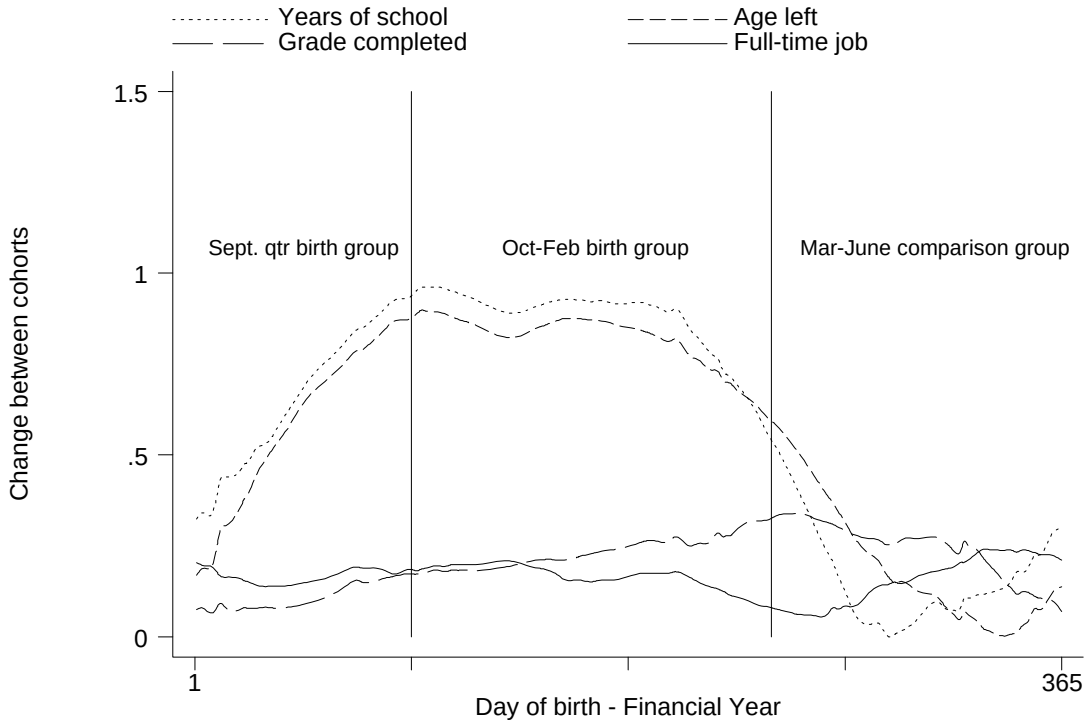
The remainder of the paper is structured as follows. The next Chapter contains a preliminary analysis of the data intended to motivate much of the remainder of the paper. The following Chapter contains a summary of the literature on the association between schooling and labour market outcomes, how the causal effect of schooling might be isolated and the implications of that literature for this study. Chapter 4 contains a description of the *Early Years of School* policy change that was implemented in South Australia in the mid-1980s and how it is exploited in this study. The following Chapter contains a description of the data and methodology used in this paper. Chapter 6 contains the empirical results and Chapter 7 some concluding remarks. Chapter 7 also identifies other potential applications where the policy change used here can be exploited to identify the causal impact of schooling on other phenomena schooling is thought to influence.

2. SETTING THE SCENE

This Chapter contains summary discussion of patterns in completed schooling and the full-time employment outcomes of school leavers in South Australia in the 1990s. It is intended to motivate the remainder of the paper. Information is provided about two cohorts of school leavers – one that left school in the early 1990s and another that left school in the late 1990s. The first group pre-dated the *Early Years of School* policy change (described in Chapter 4), while a subset of the second was affected by it. Only individuals from those cohorts who did not proceed to full-time post school studies are included in the analysis. The full-time employment outcomes of individuals are described only for their first year after leaving school.

Figure 1 shows how the schooling completed by individuals in South Australia changed between the two cohorts. It shows the (smoothed) changes between the cohorts in three dimensions of the schooling completed by individuals born at different times of the year.[1] These three dimensions are years of completed schooling, the age at which individuals left school and the highest grade they completed. The birthday of individuals in Figure 1 are presented on an Australian financial year basis – that is, the horizontal axis commences from 1 July and proceeds to 30 June. The various markers on the horizontal axis therefore reflect the end of the September, December, March and June quarters. The two vertical lines in the figure split the year into three birth groups: those born in the September quarter; those born between October and February (inclusive); and those born between March and June. The schooling of the first two groups was targeted directly by the *Early Years of School* policy change, but not that of the third group.

Figure 1 Change in average schooling and full-time employment rates by birthday between the early and late 1990s in South Australia



In general, individuals in the later cohort completed more years of schooling and left school later than those in the earlier one. The effect was most pronounced for those born in the first two birth groups. The additional year of junior primary school required by the *Early Years of School* policy for those in the second birth group and some of the first is apparent in the figure. The schooling of those in the third group, not influenced directly by the policy, changed very little. The average grade that individuals completed also increased between the cohorts, but the change was substantially smaller than and followed a slightly different pattern from that of years of completed schooling and the age at which students left school.

Figure 1 also contains the change between cohorts in the (smoothed) proportion of individuals born at different times of the year who worked full-time in their first year after leaving school.[2] Individuals in the second cohort enjoyed higher full-time employment outcomes than those in the earlier cohort, who completed their schooling either just before or during the recession of the early 1990s. There is no obvious relationship between the change in full-time employment outcomes between the two cohorts and the changes in the three dimensions of the education individuals completed.

Figure 1 does not point to any positive effect of schooling on full-time employment. The largest increases in the proportion employed full-time do not match those birth groups where the education measures increased the most. It is possible that such an effect has been hidden by other intervening factors. The remaining empirical Chapters of this paper address a series of questions designed to elicit whether there is any relationship between completed schooling and full-time employment outcomes. These questions include:

- Does the differential change in the various education dimensions between the cohorts between those born at different time of the year reflect the impact of the *Early Years of School* policy?
- With whom should the groups affected by the policy change be compared?
- How comparable are the data across the two cohorts?
- Did individual characteristics other than their education change between the cohorts?

One thing that may have changed between the cohorts is the proportion of individuals proceeding to further education and training after leaving school. Table 1 contains a summary of the destinations of South Australian school leavers in the two cohorts. The figures do show a small increase in the proportion of school leavers who undertook full-time study after leaving school. This change is much smaller than the change in the proportion that obtained full-time employment, which increased by almost 6 percentage points between the two cohorts. The proportion neither studying nor working fell by over 7 percentage points. This suggests that changes in the proportion undertaking further studies after leaving school are not likely to have had a major impact on the results presented later in the paper. Checks that changes in this proportion have not influenced the empirical results in this paper are described in Chapter 6, after the presentation of those results.

Table 1 Activities undertaken by individuals in their first year after leaving school in South Australia

Activity	Early 90s (per cent)[a]	Late 90s (per cent)[a]
Full-time work	19.7	25.6
Full-time study	45.8	47.0
Part-time work, not full-time work or study	17.7	19.6
Part-time study, no work	2.5	0.6
Not employed, not studying, seeking work	8.5	4.7
Not employed, not studying, not seeking work	6.0	2.6

(a) Figures may not sum to 100 because of rounding error.

Fundamentally, all of the empirical analysis in the remainder of this paper is designed to allow two questions to be answered. First, did the additional education undertaken by individuals in the second cohort who were born between July and February contribute to better full-time employment outcomes? Second, what does this reveal about the effect of the various dimensions of completed schooling on full-time employment? First, however, the literature on how the effect of schooling on labour market outcomes has and should be measured is reviewed briefly.

3. THE EFFECT OF EDUCATION ON INDIVIDUAL LABOUR MARKET OUTCOMES

There are a number of dimensions of the education that individuals receive that may affect their subsequent labour market outcomes. Two obvious factors are its quantity and its quality. The quantity of schooling individuals undertake may be measured in a number of ways – as years of completed schooling or as the grade or level individuals completed before they left school. However, grade completed may also contain a quality dimension – completion of a similar number of years of schooling by individuals who reached different grades by virtue of grade repetition or through school entrance arrangements that required only some to undergo a 'foundation' year may mean that their 'years of schooling' consist of different 'quality'.

A third dimension of the potential impact the education that individuals receive may have on their initial labour market outcomes in Australia is the age at which they leave school. While age on leaving may affect individual maturity and responsibility positively, it may also have a negative impact in Australia where the wages paid to many young workers are directly related to their age. These wages are set in industrial 'awards' that typically specify that young workers receive a particular percentage of the relevant adult wage. For example, in the hospitality industry, these percentages of the adult wages paid to 17, 18 and 19 year olds are 70, 80 and 90 per cent for respectively. In the metal industry, the rates for juniors who are not apprentices are 47, 58, 68 and 83 per cent of adult wages for 16, 17, 18 and 19 year olds respectively. In the retail industry, these rates for the same ages are 50, 55, 68 and 80, while 20 year olds receive 90 per cent of the adult rate. For full-time workers in 2003, these different proportions in all three industries translate to differences in wages of over $50 per week between juniors of neighbouring ages.[3] For otherwise similar individuals, with common levels of education, employers therefore face incentives to employ younger workers where that is possible.

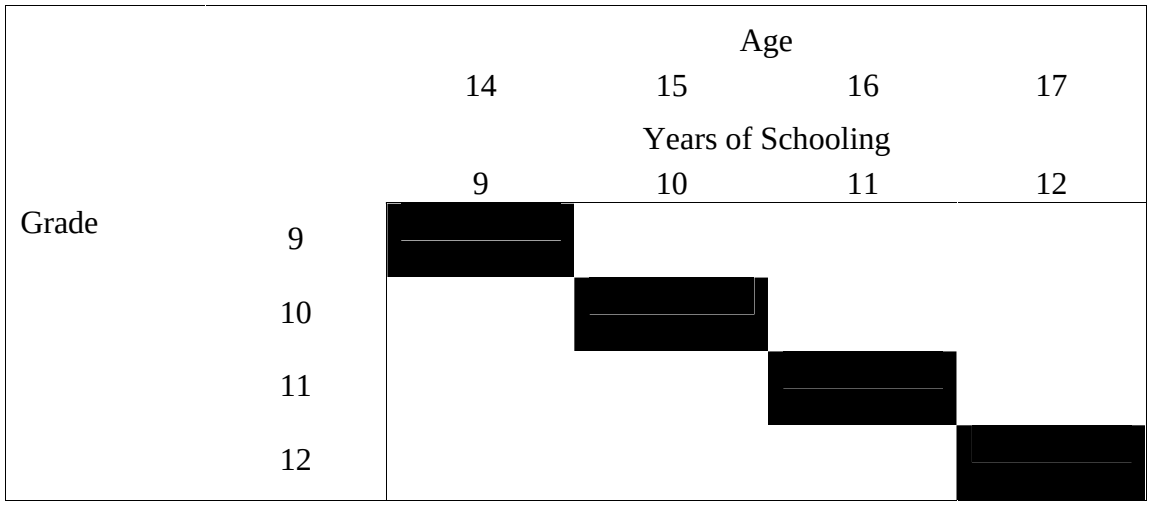
The separate effects of these various dimensions of education on individual outcomes can rarely be estimated, because the three elements – years of schooling, grade completed and age on leaving – are typically so closely related. Individuals who choose to progress from one grade to the next also undertake an additional year of schooling and consequently are at least a year older when they leave school.

This situation is characterised in Figure 2, which describes the progression of individuals through mid to upper secondary school in terms of the grades they attend, their years of schooling and their (assumed) age. Most people move diagonally through their schooling as characterised in such a diagram, unless they repeat or skip a grade. Moreover, most students are likely to move through the same cells – most are 14 years old in Year 9, 15 in Year 10 – meaning there is little that can be used to identify the effects of these three dimensions of schooling on the subsequent outcomes of individuals.

Consequently, most empirical studies rely on just one dimension to reflect the schooling of individuals in order to estimate its effect on labour market outcomes. One of the contributions of this study is to use variations in the design of the surveys used to collect the data used here and the policy change described in the next Chapter to estimate the separate effects of years of schooling, grade completed and age on leaving on individual labour market outcomes. Just how this is done is described in Chapter 5. Before that description, however, it is necessary to describe two related strands of the literature on the impact of education on labour market outcomes: first the literature on the type and

magnitudes of the estimated effects; and second, how any causal education effect might be identified. The first strand of the literature itself has two elements, one relating to the general effect of education on individual outcomes and a second related to the effect of interventions like the policy change described in the next Chapter. This change was similar in intent to early childhood or junior school enrichment programs. The following sections of this Chapter, therefore, deal with the literatures on general education effects, the effect of early childhood and junior school enrichment programs, and how such effects should be estimated, respectively.

Figure 2 Hypothetical student progression through mid to upper secondary school



A sample of the Australian literature on broad education effects

Australian and international studies have consistently found that individuals with higher levels of education enjoy superior labour market outcomes over those with less education. In some studies, the representation of an individual's education is provided by their highest completed qualification. In others, a continuous 'years of schooling' variable is constructed from the education and training qualifications individuals report they have completed or from information on their age when they left school.

Studies that use the first representation, such as Vella and Gregory (1996), Miller and Mulvey (1996, 1997), Preston (1997), Budd and Madden (1999), Marks and Fleming (1998a) and Borland, Dawkins, Johnson and Williams (2000) have consistently found that individuals with higher educational qualifications earn higher wages than individuals with lower levels of education. Other studies that use the years of schooling representation, such as Chang and Miller (1996) and Miller, Mulvey and Norris (1997) have found similarly that individuals with more years have higher wages.

These findings are replicated in studies that analyse other labour market outcomes. For example, Harris (1996), Lamb (1996), Marks and Fleming (1998b), Le and Miller (1999a), Kalb (2000), Stromback and Dockery (2000), Borland *et al*. (2000), Lamb, Dwyer and Wyn (2000) and Ryan (2002) analysed aspects of either unemployment incidence and duration or full-time employment rates and found that individuals with higher levels of education experienced better outcomes than other individuals. Moreover,

Lamb and McKenzie (2001) found that differences in employment and unemployment outcomes between Year 12 completers and those who do not finish school widen as individuals age.

Lamb (1996), Lamb *et al*. (2000) and Le and Miller (1999b) found that individuals with higher levels of education tended to obtain jobs in 'better' or higher status occupations than individuals with lower levels of education. Vella and Karmel (1999) found the same superior occupational outcomes for those with more education among two cohorts of twenty-one year old Australians. However, other aspects of their findings point to the issue at the heart of this paper. Vella and Karmel (1999) set out to determine what effect the expansion of education over the 1980s had had on the occupational outcomes achieved by twenty-one year olds in 1982 and 1991, since members of the latter group typically had higher education levels. They found that the occupational outcomes of the two groups at age twenty-one were the same. Therefore, the educational qualifications possessed by individuals may have had little impact on their knowledge, skills, competencies or attitudes. Rather, the importance of educational qualifications may have been in where they placed individuals on the distribution of education, that is, through their position relative to others in their cohort.

Literature on the impact of early childhood education programs

The *Early Years of School* policy change is described in more detail in the next Chapter. It was a policy change that required a subset of school students in South Australia to spend an additional year in junior primary compared with their predecessors. That is, it was an additional year of schooling for students at the outset of their school life. Therefore its impact might share similarities with other early childhood education and junior school enrichment programs (Raban 2000 reviews this literature). Unlike many of those programs, however, this policy was not directed towards students who were educationally or socially disadvantaged.

International evaluations of the impact of early childhood education interventions have focussed on four main issues: the type of effects such programs have; the magnitude of those effects; the duration or persistence of the effects; and the financial savings to government of the programs. The programs have typically been of two types: high cost demonstration programs with evaluation methodologies built into them, involving control groups for comparison with the groups receiving the program intervention; and lower cost, large scale public programs without identified comparison groups.

Studies that evaluate the first type of program indicate that their positive effects range from improving measured IQ, performance in school achievement tests, higher school completion and college participation rates, lower teenage pregnancy rates, better mental health outcomes, lower arrest rates, increased employment probabilities and higher levels of earnings (Currie 2001, Karoly, Greenwood, Everingham, Hoube, Kilburn, Rydell, Sanders and Chiesa 1998). Moreover, these effects can be large and are persistent, since some of the outcome measures relate to individuals in their late twenties (Karoly *et al.* 1998). That the effects of such programs might be large is consistent with research about the role of positive stimulation and other environmental factors on early childhood brain development (McCain and Mustard 1999).

Evaluations of larger scale public programs are more difficult to assess in the absence of appropriate comparison groups (see Currie and Thomas 1995 for an attempt to construct

one in such circumstances). The effects appear to be smaller (like the per capita costs of the programs – Barnett 1998), but may act to reduce socio-economic status related disparities in school achievement. The effects may not persist, however, in the face of later, intervening factors such as low school quality (for example, Lee and Loeb 1995 and Currie and Thomas 2000, but see also Barnett 2002). Nevertheless, commentators such as Currie (2001) estimate the cost savings from such programs to warrant their expansion.

There are not comparable studies or programs in Australia. There are related literatures on the effect school-based programs can have on the development of literacy and numeracy skills among school entrants and on the social and academic development of boys (see de Lemos 2002 for a review of the literacy literature; Lingard, Martino, Mills and Bahr 2002 and Alloway, Freebody, Gilbert and Muspratt 2002 for interventions targeted at boys). Even here, however, de Lemos (2002) observed that there have been few 'systematic evaluations of specific teaching approaches or interventions on student outcomes in Australia' (2002: 13).

One exception is Ainley, Fleming and McGregor (2003) who assessed the impact of literacy programs undertaken in Catholic primary schools in Victoria on students in junior primary school. Schools were able to implement one of a menu of literacy programs. The Victorian Catholic Education Office had developed one of the programs, while the others had been developed elsewhere. Ainley *et al*. (2003) found that children in schools that adopted the local program experienced greater growth in literacy performance in Years 1 and 2 than children who did not undertake that program. Moreover, these performance differentials remained apparent at the end of Year 3. In addition, the literacy performance of a second cohort of Year 1 students from all Catholic schools who had experienced the new literacy programs only in their Preparatory Year was about one quarter of a year more advanced than the earlier cohort who did not start the programs until Year 1. More generally, however, Foley, Goldfeld, McLoughlin, Nagorcka, Oberklaid and Wake (2000) have observed that 'it appears that few Australian early childhood programs have been studied using rigorous research methods' (2000: 30).

How to deal with the endogeneity of schooling

While many studies have found that individuals with higher levels of education earn higher wages than those with less education, Card observes that:

> 'social scientists have been cautious to draw strong inferences about the causal effect of schooling. In the absence of experimental evidence, it is very difficult to know whether the higher earnings observed for better-educated workers are *caused* by their higher education, or whether individuals with greater earning capacity have chosen to acquire more schooling' (Card 1999, p. 1802).

Essentially, the problem social scientists face is that while the level or years of schooling an individual completes may be known, it reflects the outcome of a decision-making process in which the factors that may be paramount are not captured in the data available to social scientists. These factors may include individual ability, tastes for education, motivation levels and financial constraints. Since these factors may themselves influence subsequent individual labour market outcomes, their effects will be bundled up with the observed effect of schooling.

Researchers have followed a number of strategies to obtain more precise estimates of the effect of schooling on labour market outcomes. The first was to supplement observed schooling with measures that reflect some aspects of ability, such as IQ test scores, in regression equations designed to explain particular outcomes. Inclusion of such variables typically lowers the estimated schooling effects. Angrist and Krueger (1999) argue that such an approach biases down the schooling parameter, since the IQ scores individuals obtain in such tests reflect the outcome of their schooling. Moreover, Cawley, Heckman and Vytlacil (2001) argue that schooling and test scores are so strongly correlated it is typically not possible to identify their separate effects on labour market outcomes, unless very strong assumptions are imposed on the form of the relationships between these variables.

A second approach used by researchers to estimate the effect of schooling on labour market outcomes has been to compare the outcomes achieved by individuals assumed to have similar ability levels but who obtain different amounts of schooling. These comparisons have most typically been made between identical twins, such as Ashenfelter and Krueger (1994), Ashenfelter and Rouse (1998), Behrman, Rosenzweig and Taubman (1994) and Miller, Mulvey and Martin (1995), but other comparisons have also been exploited. For example, Rummery, Vella, and Verbeek (1999) identify the effect of schooling on wages by comparing individuals who had similar rankings in terms of their position in the distribution of schooling within their jurisdiction, but who differed in the absolute amount of schooling they had completed.[4] The approach of the latter paper remains relatively untested, while the 'equal abilities' assumption of the twins studies remains contentious, with conflicting evidence on the issue presented in different studies.[5] Essentially, the problem with twins studies is just the broader one already described faced by social scientists – how confident can we be that the observed differences in schooling between twins are generated by random factors and not important unobserved factors that also influence wages?

In contrast, the third approach to estimating the effect of schooling on labour market outcomes exploits some source of systematic variation in the schooling obtained by individuals. The hope is that the identified source of variation is exogenous in that it is unrelated to individual characteristics and does not affect labour market outcomes other than through the variation in schooling it induces. That is, this exogenous variation is assumed to be independent of the unobserved factors that drive other aspects of the education choices of individuals. The source of the exogenous variation is often described as a 'natural experiment' in such studies. Rosenzweig and Wolpin (2000) distinguish between 'natural natural experiments' and 'man-made' natural experiments. The man-made experiments include social experiments in which institutional arrangements and policy parameters are varied. The 'natural natural experiments' include biological and climate mechanisms that are plausibly independent of ability and preferences, such as birth date, sibling gender, the identical twin studies described earlier and climatic conditions.

Examples of the man-made natural experiments used to identify the effect of education on wages include: changes to minimum schooling legislation (Harmon and Walker 1995, Callan and Harmon 1999, Levin and Plug 1999); broad reforms to national education systems (Brunello and Miniaci 1999 and Vieira 1999); proximity to a university campus (Card 1995); school construction in a developing country that expanded the supply of secondary school places (Duflo 2001); and the role of World War II in disrupting the education of German and Austrian men (Ichino and Winter-Ebmer 1999). In many of

these studies the 'exogeneity' of the external shock is often doubtful, as is its isolation as the only potential explanation for observed differences in schooling and wage outcomes.

Rosenzweig and Wolpin (2000) show that while 'natural natural experiments' can allow identification of the effect of schooling on labour market phenomena, the validity of the instrument alone cannot guarantee that the estimate will be unbiased. Other aspects of the estimated equations need to be specified correctly for that to occur. Moreover, where the effect of schooling on labour market outcomes may vary between groups, 'natural natural experiments' will identify the effect of schooling only for that group affected by the natural experiment.

Rosenzweig and Wolpin (2000) make these points in discussion of a number of studies, but they can be demonstrated in relation to the study of Angrist and Krueger (1991). Angrist and Krueger (1991) used the quarter of birth of individuals, in conjunction with United States school commencement and minimum leaving legislation, as an instrument for their schooling. They found that individuals whose birthdays occurred after the cut-off for school commencement for a grade cohort reached the minimum school leaving age having completed fewer years of schooling than those born in the same year whose birth date preceded the designated cut-off. These quarter of birth effects should only affect the wages individuals received through the variation in schooling they induced, since ability is independent of birth date. Rosenzweig and Wolpin (2000) conclude that the quarter of birth instruments do indeed identify the effect of schooling on wages.[6] They argue, however, that the effect is only estimated consistently if the estimation procedure controls for the labour market experience of individuals, rather than the age of individuals as in Angrist and Krueger (1991). In addition, Rosenzweig and Wolpin (2000) note that the schooling effect estimated is for a specific group within the population – those who leave school at the legislated minimum leaving age. That is, the instrumental variable estimate of schooling on wages in the Angrist and Krueger (1991) paper was the effect of an additional year of schooling for the earliest (and on average, probably the least able, since ability and schooling appear to be positively correlated) group of school leavers. Finally, Rosenzweig and Wolpin (2000) argue that if schooling levels affect other 'intermediate' labour market phenomena that in turn affect wages, such as hours worked (either full-time/part-time status or hours over a career), any estimated schooling effect will also reflect these phenomena as well as the initial schooling effect.

Implications for this study

This study also makes use of a natural experiment to assess the impact of schooling on a specific labour market outcome. The experiment lies somewhere between Rosenzweig and Wolpin's (2000) man-made and 'natural natural experiment' distinction. The experiment involves analysis of the impact of a policy change that affected the progression through junior primary school of a group of students. What distinguished this group was when in the year they commenced school, which in turn reflected their birth dates. Since date of birth is likely to be independent of ability and preferences about schooling, the policy change itself should not have had an independent effect on the labour market outcomes examined here.

The analysis of Rosenzweig and Wolpin (2000) has a number of implications for this study. First, the group whose schooling effect is measured from the policy change needs to be identified carefully. It turns out that in this application, the estimated effect is for a

subset of those affected by the policy change only. This is discussed after the policy change is described in the next Chapter. Second, the schooling effect can only be estimated over a group with common levels of labour market experience, for example in the first year after they leave school. For the data used here, this involves measuring the labour market outcomes achieved by individuals in different calendar years. Third, the labour market outcome measure needs to be determined carefully, with the interplay of the effect of the policy change and labour market institutions in mind. The policy change affected the years of schooling obtained by some individuals. Since the analysis will be undertaken for individuals with a common level of labour market experience, the policy change also implies those affected will be older than they would have otherwise been. As described previously, in the Australian youth labour market, wages are institutionally age-related. Individuals affected by the policy change may receive increased wages for no reason other than their increased age. Consequently, this paper addresses one labour market phenomenon where any effect of age is less direct. This paper focuses on the effect of additional schooling arising from the policy change on the full-time employment outcomes of young Australians who do not engage in immediate post-school education.

4. THE *EARLY YEARS OF SCHOOL* POLICY

The *Early Years of School* policy was announced in 1984 in South Australia, with implementation to start in 1985 and be completed by February 1986. The elements of the policy and its rationale were set out in Education Department of South Australia (1983), which drew on analysis in the final report of the Committee of Enquiry into Education in South Australia (Keeves Enquiry, 1982). The major objective of the policy was to provide a better foundation for children's subsequent educational achievement by extending and enriching their junior primary education (that is, levels below Year 3).

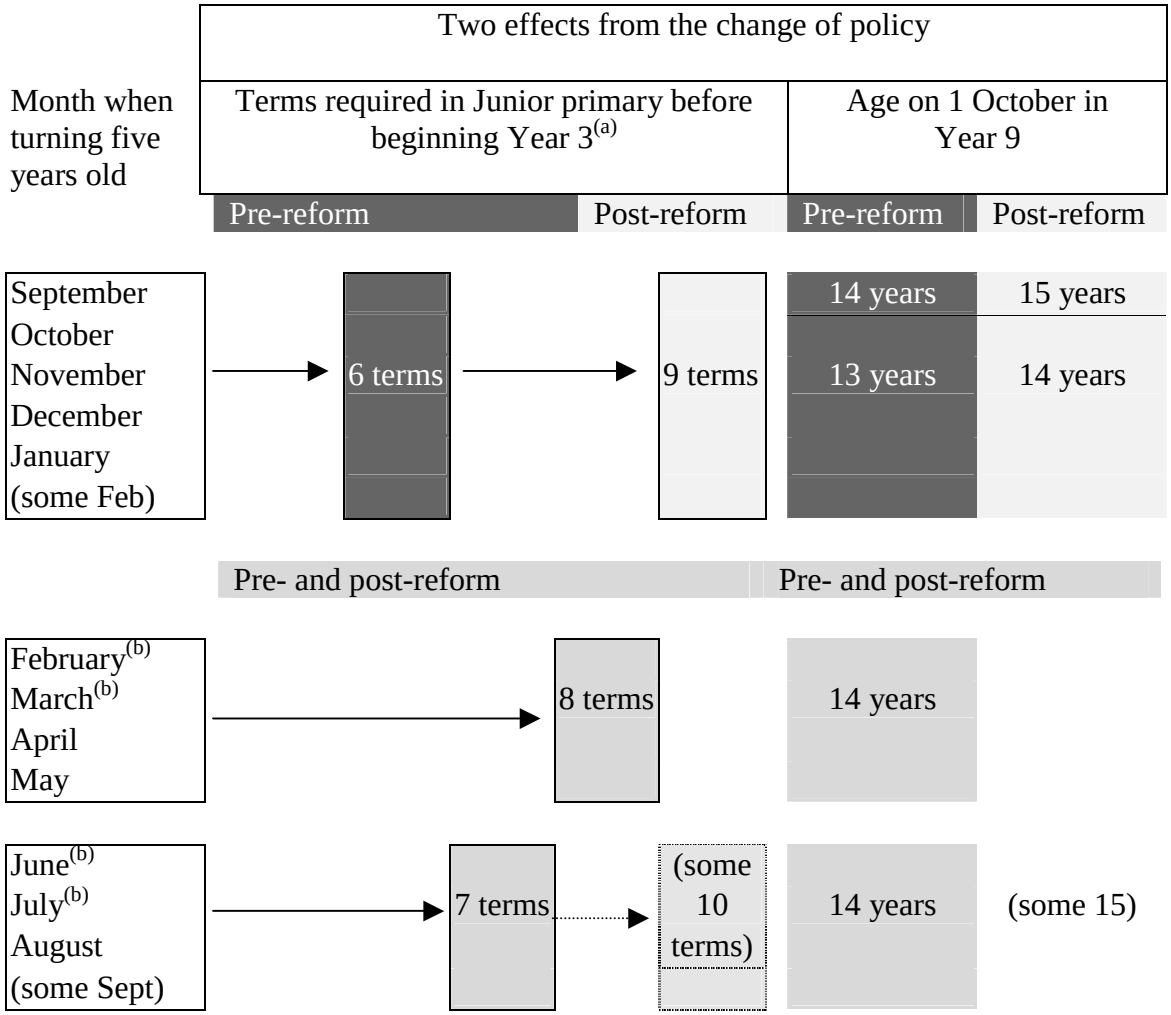
The new policy dictated that students 'enrolling in government schools have between seven and ten terms in junior primary classes, that is: reception, year 1 and year 2, rather than six to eight currently prevailing, such policy to be fully implemented by 1986' (Minister of Education, 14/11/1984, quoted in Director General of Education South Australia 1984: 10). Implementation of the policy began in the 1985 school year, when the school year consisted of three terms. Its effect was to make the pre-Year 1 entrance level, called Reception in South Australia, closer to a full year of schooling for most students rather than something less than that.

South Australia has a 'continuous admission' policy for 5 year olds (see Trethewey 1997 for a description of the history of this policy). It involves regular (not less than once a term) admission of recently turned five year olds into individual schools over the school year.[7] The way it operated prior to the *Early Years of School* policy meant that only those children who entered towards the end of the school year moved into Year 1 in the following year. Those five year olds who began towards the start of the school year moved directly into Year 2, having compressed Reception and Year 1 into just one year of school. Consequently, the Keeves Committee found that 'for 40 per cent of the students entering the South Australian school system the Reception Grade serves no useful purpose, and for a further 14 per cent it does little more than familiarize children with school for up to one term' (1982: 91).

The Keeves Committee had received proposals to add a further year of secondary education to the South Australian school system. Students going from school to university were younger than their Eastern State counterparts and were considered immature, a factor that was purportedly reflected in low levels of achievement in subjects such as Chemistry and Physics (1982: 88). Rather than accepting such proposals to address the perceived immaturity of school leavers, the Committee proposed that children spend more time in junior primary school, the changes effectively implemented under the *Early Years of School* policy. The Committee acknowledged that these changes would take a long time to take effect (13 years), but that increased use of grade repetition could be made at the junior primary and primary levels to achieve the same result (1982: 94).[8]

The nature of the *Early Years of School* policy reform is summarised in Figure 3 (and the formal policies set out in Appendix 1).[9] Figure 3 shows how children who turned five at different times of the year progressed through to Year 3 both before and after the policy change. It also shows how this affected their age at 1 October in Year 9, a reference point of significance in the data used in this research and described below.

Figure 3 Changed junior primary school arrangements from the implementation of the *Early Years of School* policy



(a) South Australia introduced a four-year term soon after the policy change. The timing of the terms requires some amendment to the diagram – with four terms, June probably belongs to the middle group of months and September wholly to the third group.

(b) Those who turned 5 after the commencement of a term may have enrolled during it if their school had more than one intake per term.

Those who turned five after the commencement of the last term in any year and before early February of the following year commenced school at the beginning of the school year in early February. Prior to the reform, they typically completed just two years of schooling (six terms) before entering Year 3. After the reform, they completed three years (nine terms) before entering Year 3. Since they were now a year older when they reached Year 3, they were a year older in all subsequent grades than their predecessors.

Other children were largely unaffected by the change. Children who turned five after the commencement of the school year and before the beginning of the final term (from mid-February to mid-September) entered school during the year and typically spent more than two years at school before they entered Year 3.[10] For most of these students the *Early Years of School* policy did not affect the way they progressed through school. It is possible that some of those who commenced school at the beginning of third term may

have undertaken 10 terms (one term plus three full years of school) before commencing Year 3.

Figure 3 is indicative only.[11] All policies provide a set of guidelines for the treatment of 'typical' individuals and a set of clauses that cover exceptions. Nevertheless, it appears consistent with the data contained in the Keeves Committee's Final Report that showed there was some 'bunching' of students commencing at the start of the school year (about 40 per cent of the grade cohort).

The *Early Years of School* policy also changed the age composition of grade cohorts. They became older. Figure 3 indicates how the age distribution of grade cohorts changed as a result of the policy change. It reports the age of students in October of Year 9. The focus is on their age at October 1, since this is the reference period for student ages in the data used in this paper. The change in the composition of the cohorts meant that the reported ages of students in Year 9 went from being approximately two-thirds aged fourteen with the balance aged thirteen years prior to the change to predominantly fourteen years after it, but with a sizeable group aged fifteen years.

The implementation of the *Early Years of School* policy took place over 1985 and 1986. It affected entrants to Year 3 from 1985. That particular cohort reached secondary school (which commences in Year 8 in South Australia) in 1990 and Year 9 in 1991. Figure 1 of this paper provides evidence that it had the intended impact on the schooling obtained by those born between October and February.[12]

The groups whose schooling was affected by the policy change

The policy change affected students who commenced school at particular times of the year. Given the continuous enrolment policy for five year olds, this translates into affecting individuals born at specific times of the year. Essentially, individuals born between October and part of February undertook an additional year of junior primary compared to their predecessors. Some born in July, August and September may also have done so.

There are, therefore, three groups of individuals to focus on in the analysis designed to determine the impact of the policy change. These groups are:

- those born between October and February, inclusive – who were most affected by the policy change
- those born between July and September – who may have been affected by the policy change
- those born between March and June –the comparison group, who were not affected by the policy change.

These groups are referred to throughout the remainder of this paper as the October – February and July – September (or September quarter) birth groups and the March – June comparison group, respectively.

5. DATA AND METHODOLOGY

Data used in this study

This paper exploits data from two Longitudinal Surveys of Australian Youth (LSAY) cohorts to assess the impact of the *Early Years of School* policy change. The Youth in Transition 1975 birth cohort (YIT 75) and the Longitudinal Surveys of Australian Youth Year 9 cohort (LSAY 95) fall either side of the policy change and should capture its impact.

The cohorts affected by the policy reached secondary school in South Australia after 1990 and Year 12 from 1994 onwards. The grade cohorts in YIT 75 reached Year 12 from 1991 through 1993. The LSAY 95 cohort started school after the policy change took effect and reached Year 12 in 1998. South Australian children in Year 9 in the LSAY 95 cohort should be older for their grade level than those in the YIT 75 cohort. Hereafter, the YIT 75 cohort is generally referred to as the 'early 90s' cohort and the LSAY 95 cohort as the 'late 90s', which reflects the period when their members completed school and entered the labour market.

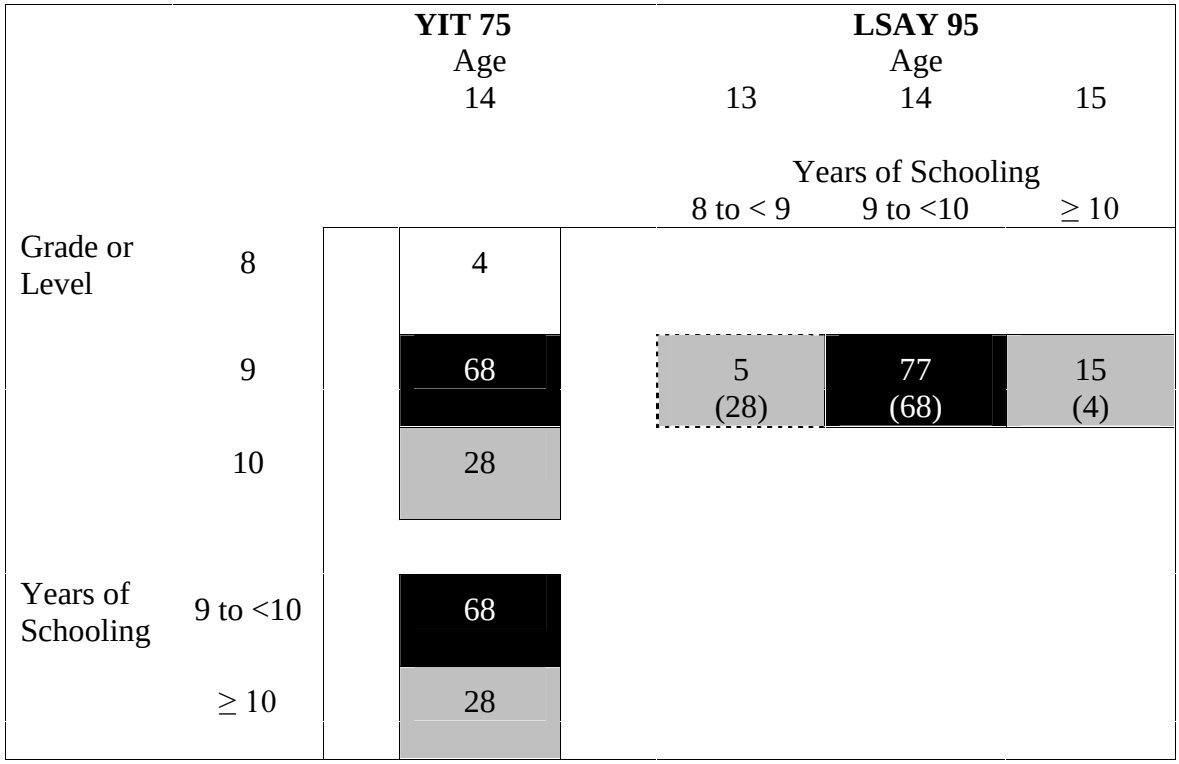
Analysis of data from the late 90s cohort in Appendix 2 indicates that the policy change is reflected in the age-grade structure of individuals in the data in exactly the way it would be anticipated to be, once aspects of the survey design of the LSAY 95 cohort are taken into account.

The major difference in the design of the two surveys is summarised in Figure 4. The first collection was an age-based sample of young Australians, the second a grade-based one. Consequently, individuals aged fourteen years (as of 1 October in the year they were surveyed) were distributed across grades or levels in South Australia as follows: 4 per cent were in Year 8; 68 per cent in Year 9; and 28 per cent in Year 10. For the second cohort, individuals in Year 9 were distributed across single years of age (on 1 October) as follows: 5 per cent were aged thirteen; 77 per cent were aged fourteen; and 15 per cent were aged fifteen. Had the *Early Years of School* policy not changed the age-grade structure in South Australia, these proportions would have been approximately 28, 68 and 4 per cent respectively, that is, the (reverse order) proportions from the earlier cohort.

Methodology

Both 'man-made' and 'natural natural experiments' can be exploited in one of two ways to provide 'causal' estimates of the effect of education on some other phenomenon. First, they allow estimation of difference in differences effects. These involve the comparison of changes in the mean values of key education and labour market outcome variables of those affected by the policy change with those of a natural control group. This type of estimator is summarised in Appendix 3 and described in more detail in Angrist and Krueger (1999), Meyer (1995) and Heckman, Lalonde and Smith (1999). This report makes limited use of such a difference in differences estimator.

Figure 4 Age, grade and years of schooling relationships in the two surveys



Analysis contained in Appendix 4 was designed to identify whether the data support difference in differences effects analysis. This analysis involved tests of whether the average characteristics of individuals in the three birth groups identified in the previous Chapter are similar in the two cohorts or change in a similar way between the cohorts. It suggests that in general they do but that there are three potential problems for the difference in differences estimator for the application analysed in this paper.

The first is that a selection process of individuals into different grade cohorts after the policy changes appears to have taken place for the July – September birth group. This weakens the case for arguing the policy change had an 'exogenous' effect on this group, so the results for this group need to be interpreted with care. The second is that the labour market conditions that faced members of the October – February birth group in the first cohort may have been better on average than those in the comparison group. The October – February birth group made up the bulk of the 28 per cent of students (see Figure 4) in Year 10 (in 1989) from the early 90s cohort, while the March – June comparison group were overwhelmingly in Year 9. With the timing of the recession of the early 1990s, on average members from the Year 10 grade cohort faced a state unemployment rate that was about 1.5 percentage points lower when they left school (one year earlier) than the comparison group. Early leavers from a common grade from all birth groups in the late 90s cohort faced the same labour market conditions. The third problem is that the number of observations available from the early 90s cohort is quite small (see Table 4.6). The implications of this are discussed in conjunction with the results.

The second way natural experiments can be used is to allow both tests of the exogeneity of schooling and the estimation of the effect of schooling through instrumental variable (IV) estimation where data on affected and unaffected individuals are available. Both

involve a two-stage regression approach. The first stage involves estimation of an equation that explains the amount of schooling obtained by individuals. This equation includes among its explanatory variables the instrument that reflects the natural experiment. The second stage involves estimation of an equation designed to explain the determinants of the labour market outcome of principal interest. In order to be valid, the policy-based instrument should affect the labour market outcome variable of principal interest only through its impact on individual schooling. The test for the exogeneity of schooling involves the addition to the second equation of either the predicted schooling levels for individuals from the first equation or its estimated error term.[13] The exogeneity test is based on whether the predicted variable or error term is significant in the second stage equation. If it is, estimation of the second stage equation should proceed by instrumental variables. In this case, the predicted schooling levels for individuals from the first equation are substituted in place of observed schooling in the second equation. The instrument that reflects the policy change is not included in this equation. If the instrument for education is valid, the predicted value of education from the first stage equation is independent of the error term of the second equation, allowing consistent estimation of the effect of education on the labour market outcome variable. In addition, regression analysis can deal with aspects of the problems identified with the difference in differences estimates referred to in the previous paragraph.[14] The results of such an estimation strategy for isolating the causal impact of schooling on the full-time employment outcomes of individuals are also reported in this paper.[15]

Individuals affected by the policy need not have undertaken more years of schooling than their predecessors. They may have simply substituted the additional year of junior primary schooling for one of secondary school, which is exactly what anyone who left at the minimum school leaving age of fifteen years would have done. If the grade distribution of school leavers affected by the change remained unchanged from their predecessors, then all affected individuals would have undertaken one more year of schooling than their predecessors. The evidence from Figure 1 is that this is exactly what happened. Hence, in general, the results of the instrumental variables estimation exploiting this policy change should be interpreted as measuring the effect of an additional year of schooling for a group of about average ability. Individuals who undertake post-school education and training, an above average ability group, are excluded from the full-time employment application. Therefore, for the specific application in this paper, the affected group are probably a below average ability group.

It should be noted that the *Early Years of School* policy change is not the only exogenous source of variation on the years of schooling or the relationships between it and grade completed and age left school. From Figure 4, the nature of the school entrance arrangements and the design of the surveys combined at the time of the initial interviews to produce variation across grades for a common age in YIT 75 and variation in age for a common grade in LSAY 95. That is, the relationships are less direct than those implied by Figure 3. One implication of Figure 4 is that if individuals born at various times of the year tend to complete similar grades, then there will be variation in their years of completed schooling and age on leaving. These different sources of variation, in addition to the effects of the policy change, are used in this study in an attempt to identify the separate effects of years of schooling, grade completed and age on leaving on individual labour market outcomes. It is possible that years of completed schooling and age on leaving will still be too closely correlated to distinguish their separate effects. In those circumstances, a substitute variable for the years of schooling variable will be utilised, one that captures directly the time spent in the Reception grade by individuals, estimated

from their birth month. In that case, the results will indicate the employment effect of a year of schooling on students undertaken at the commencement of their studies.

Not all students in mid-secondary school in South Australia need necessarily have commenced their schooling in that State, though here it is implicitly assumed that most did. Those born overseas are excluded from the analysis that follows. Those who started their schooling in other States cannot be separately identified, so all Australian-born students observed in South Australia in mid-secondary school are treated as though they commenced their schooling there. So long as migrating students are distributed randomly by birth date, this should have few implications for the analysis that follows.

As noted in Chapter 3, the schooling effects of interest can only be estimated over a group with common levels of labour market experience. The full-time employment outcomes of individuals in their first and second years in the labour market are analysed in this paper. For the data used here, this involves measuring the labour market outcomes achieved by individuals from the same data cohorts in different calendar years.

The dependent variables in such analyses take the value of one if individuals had a full-time job in their first year after completing school and zero if they did not. Since the dependent variable can take only two values, techniques that account for this limited range must be used to analyse such variables.[16] In the results reported later in this paper, whether individuals are employed full-time after they leave school is treated as being determined by:

- the state of the labour market when they leave school;
- the job information networks at their disposal;
- their age and gender;
- their work-related skills and school performance, as exhibited by both objective and subjective assessments of their school work; and
- the amount of schooling (years and level) they have completed.

The explanatory variables that capture these effects are described in the next Chapter and defined more fully in Appendix 5.

6. ANALYSIS OF THE IMPACT OF THE POLICY CHANGE ON FULL-TIME EMPLOYMENT

This Chapter contains an assessment of the impact of the *Early Years of School* policy on full-time employment, through its effect on the schooling of the individuals it affected. The experience of South Australian students from the YIT 75 and LSAY 95 cohorts is used to identify the impact of the policy. The two cohorts entered the labour market at two quite different points of the economic cycle. Most members of the first cohort entered the labour market during the recession of the early 1990s, while members of the second cohort entered it when full-time employment growth was stronger.

Table 2 contains the proportion employed full-time of the two birth groups affected by the policy change, along with the March – June comparison group in the two cohorts. The proportion is shown in both the first and second years out from school for those members of the groups who did not go on to full-time study after leaving school. Three patterns are evident in the table. Two are unsurprising: full-time employment outcomes are stronger in the second cohort than the first, which reflects the relative strength of the labour market; and the full-time employment rates are typically higher in the second than the first year out from school.[17] The third pattern evident in the table is that the outcomes for the two birth groups affected by the policy change improved relative to the comparison group between the two cohorts. The purpose of this Chapter is to analyse whether this different experience reflects the impact of the *Early Years of School* policy.

Table 2 Proportion employed full-time after leaving school in the two cohorts (among those not in further full-time education and training)

Cohort	First year	Second year
Early 90s		
October – February birth group	0.34	0.42
September quarter birth group	0.42	0.54
Comparison group	0.38	0.69
Late 90s		
October – February birth group	0.54	0.59
September quarter birth group	0.57	0.65
Comparison group	0.51	0.68
Change between cohorts[a]		
October – February birth group	0.21	0.18
September quarter birth group	0.15	0.11
Comparison group	0.13	-0.01

(a) Subject to rounding error

The impact of the policy on the schooling obtained by individuals

Difference in difference estimates

Individuals in the October – February birth group obtained an increased amount of schooling relative to the March – June comparison group as a result of the policy change. The estimates appear in Table 2. The average years of schooling of members of the October – February birth group increased by 0.798 of a school year relative to the comparison group. The change in schooling for the September quarter group was smaller, 0.494 of a year. As already highlighted in Figure 1, the policy change had the expected effect in these data on the average years of schooling of the two groups of individuals it was intended to influence. The relative grade or level of school completed by the various birth groups did not change between the cohorts.

Table 3 also provides evidence about age at which individuals left school, which was also affected by the policy change. The increased years of school completed by individuals in the two birth groups affected by the policy were translated quite directly into an increase in the age at which members of those groups left school relative to the comparison group. The average age on leaving school increased from less than 17 years to greater than 17 years for both groups. As noted earlier, it is conceivable that both years of schooling and individuals' ages when they seek work are important in explaining whether they obtain full-time employment when they leave school. The direction of a separate age-related effect is not clear, however, once the effect of the additional schooling has been taken into account. It may be positive if age is associated with maturity, which in turn has a positive effect on motivation, reliability and productivity, for example. It may also be negative, if it simply requires employers to pay older individuals the higher wages specified in industrial agreements. This issue is discussed further in the regression results described below.

Table 3 Change in mean values of years of schooling and employment outcomes between the cohorts: policy and comparison groups (among those not in further full-time education and training)

	Oct-Feb birth group	Comparison group	Difference in difference	T test
Years of schooling completed	0.878	0.080	0.798	4.50
Grade or Level of school completed	0.168	0.245	-0.077	-0.56
Age left school	0.707	0.020	0.687	4.64
Full-time job first year after leaving school	0.208	0.126	0.082	0.87
Full-time job second year after leaving school	0.177	-0.008	0.185	1.60
	Sept qtr birth group	Comparison group	Difference in difference	T test
Years of schooling completed	0.574	0.080	0.494	2.50
Grade or Level of school completed	-0.016	0.245	-0.262	-1.77
Age left school	0.509	0.020	0.489	2.74
Full-time job first year after leaving school	0.148	0.126	0.022	0.19
Full-time job second year after leaving school	0.110	-0.008	0.118	0.85

Regression estimates

Regression estimates are presented in Table 6.1 of Appendix 6 that capture the effect of the *Early Years of School* policy change on the three dimensions of education discussed in earlier Chapters: years of schooling, grade or level completed and age left school. The impact of the policy change is measured through two variables. The first captures the typical time individuals spent in the Reception grade, determined largely by when in the year they were born, and hence, commenced school. The second variable is the grade in which they turned fifteen years old, the minimum school leaving age. This grade would have fallen by one year for those in the late 90s cohort required to undertake a full year of Reception (the October – February birth group) and those among the September quarter who also undertook an additional year of Reception. These parameters can be used to estimate the effect of the policy on changes in the three dimensions of education. The total education effects of the policy change can be estimated by the parameters on the additional Reception year minus the parameter on the variable reflecting the grade when individuals turned 15 (the former increased by one unit, the second decreased by one for those who undertook an additional year of Reception). The estimates are consistent with the direction of the effects presented above: those who undertook the additional year of Reception completed more eventual years of school and were older when they left (an additional half a year in both cases). The grade they had completed when they left was little changed, however.

Most of the other explanatory variables that appear in Table 6.1 are those that are used in the equation designed to explain full-time employment outcomes in the next section. They include many of the individual influences that other studies have found to have an impact on school completion in Australia, such as gender, school achievement, school type, and father's occupation. Most of the variables are estimated here to have similar effects to those found in other studies, but in some cases the estimated effects depart from those found previously. For example, attendance at an independent school does not have an estimated positive effect on years of schooling in these data. This outcome simply reflects the nature of the individuals analysed here. Since attendance at an independent school is positively associated with post-school study (see Marks, Fleming, Long and McMillan 2000), more individuals from such school have been excluded from the sample utilised in this analysis. The result here indicates that among the group who do not undertake further full-time study when they leave school, attendance at an independent school has no positive impact on years of schooling or the age at which students left school. It retains its 'usual' positive effect on the grade of completed schooling.

The other effect of note in Table 6.1 is that associated with the State unemployment rate. Its effect is estimated to be negative on all three dimensions of education. This needs to be interpreted with care. The variable actually reflects the State unemployment rate in the year following the one when individuals left school. This timing makes it difficult to assign too much meaning to the result. Rather than capturing any inducement effect of students out of school with improved labour market conditions, the results effectively ascribe some of the increased education in the second cohort to the generally lower levels of unemployment operating at that time.

The impact of the policy on the full-time employment outcomes achieved by individuals

Difference in difference estimates

The changes in full-time employment outcomes between the two cohorts for the two groups affected by the policy change are also compared with those of the comparison group in Table 3. The differences between the changes for the two groups affected by the policy change are positive, but not significantly different from zero for either the first or second year out from school.

As indicated in Chapter 4, the policy change may have affected the different dimensions of education that in turn have offsetting effects on the probability of individuals obtaining a full-time job. The difference in differences estimator cannot distinguish those separate effects, while regression estimation may. These regression results are now described.

Regression estimates

The broad determinants of whether individuals obtain full-time employment after leaving school were set out in an earlier Chapter. The explanatory variables used in the regression equations that explain full-time employment outcomes among those who do not go onto further full-time education included:

- the aggregate state unemployment rate in the year after individuals left school and a metropolitan/non-metropolitan indicator
- social background indicators including father's occupation, whether their father was born overseas in an English-speaking or non-English speaking country;
- a gender indicator variable;
- the average of individual numeracy and literacy scores from tests undertaken when individuals were in mid-secondary school, as well as individuals own assessment of their performance relative to their peers at their school; and
- their years of completed schooling (or years of Reception), the age they left school and the grade they completed prior to leaving.

Since each of the last set of variables is potentially endogenous, the policy change analysed here is used to test this possibility and, if necessary, used as an instrument for completed schooling. As anticipated in an earlier Chapter, this involved a two-stage regression approach. The first stage involved estimation of an equation that explains the amount of education obtained by individuals, described in the previous section. This equation included among its explanatory variables the variables that capture the effect of the policy change, as well as all of the other explanatory variables used in the second stage estimation of the determinants of the full-time employment outcomes achieved by individuals. The estimated residual from that equation was added to a second equation that explains whether individuals obtained full-time jobs in their first year after completing school to test whether the relevant education variable was endogenous. If so, the predicted education levels of individuals from the first equation are substituted in place of observed one in the second equation. The variables that capture the policy change are not included in this second equation.[18]

In this application, the test of the exogeneity of the various education variables did not reject the possibility that they were exogenous. The z-value on the years of schooling equation error term was 1.42, the age individuals left 0.47 and the grade completed 0.31 respectively when they were included separately in the full-time employment equation along with the education variables. Therefore, there is no strict need to correct for the endogeneity of the education variables in the full-time employment equation. Does this outcome suggest that the earlier emphasis on the endogeneity of education choices in such labour outcome equations is misplaced? It seems more likely that the conditioning inherent in the sample analysed here has already taken account of many unobservable factors that influence education choices. That is, since the sample contains only individuals in one jurisdiction who did not proceed to full-time study after leaving school, it may consist of a group who are relatively homogeneous in terms of their ability, motivation and tastes for education.

The results of the preferred probit regression equation designed to explain whether individuals were employed full-time are presented in Table 6.2 of Appendix 6. The estimates of the effect of the education variables on full-time employment are summarised in Table 3. One feature to note is that in the preferred equation, the years of schooling variable has been replaced by the years of Reception undertaken by individuals. It did not prove possible to identify separately the effects of years of schooling from the age at which individuals left school (the variables were too closely correlated). It was possible to identify the separate effects of the years of Reception undertaken by individuals from that of the age at which they left school. Since the policy affected both the time individuals spent in Reception (by design) and the age they left school (Table 3), the complete impact of the policy is reflected in these two parameters. The parameters convey separate information about the impact of education on individual outcomes. The parameter on the years of Reception variable captures the employment effect of a better scholastic grounding provided early in the careers of students. The age when individuals left school is more likely to pick up the effect of the higher wages associated with later school leaving.[19]

The estimates presented in Table 4 are the estimated 'marginal effects' of the various education variables on the probability an individual obtained a full-time job in the relevant year after leaving school. The 'marginal effects' show the impact of a one-unit change in the explanatory variable on the full-time employment probability. These were calculated at the actual values taken by the explanatory variables for individuals and the estimates for all individuals averaged.

The results suggest that an additional year of Reception increases the probability individuals are employed full-time after leaving school by about 11 percentage points in their first year out of school. At the same time, any increase in the age at which individuals leave school, holding grade completed and years of Reception constant, leads to a reduction in the probability that individuals are employed full-time after leaving school. This age left school effect is around 8 percentage points. That is, the policy change resulted in offsetting effects on the full-time employment outcomes of individuals. On the one hand, it had a positive effect on the foundation experience of school for individuals and the skills they developed, but on the other, employers were less prepared to employ them because they had to pay a wage premium because they were older.

Table 4 Estimated effect of an additional year of schooling on the probability individuals work full-time (among those not in further full-time education and training)

	Effect on probability of working full-time			
	First year		Second year	
Variable	**Effect**	**Effect/Std Error**	**Effect**	**Effect/Std Error**
Years of Reception	0.109	1.99	0.048	0.71
Age left school	-0.078	- 2.42	-0.060	- 1.43
Grade completed	0.057	1.58	0.025	0.58

Individual characteristics also had an impact on whether individuals were employed full-time in their first year after leaving school. Males were more likely than females to be employed full-time, as were those with more positive assessments of their own school performance relative to their peers. Some characteristics of individuals' fathers appear to be important – specifically their birthplace and occupation. These effects may represent the operation of information networks, with Australian born fathers and those working as managers possessing better networks to identify available jobs for their children than others.

Finally and unsurprisingly, the state of the labour market has an impact on the proportion that finds full-time employment. Each percentage point increase in the state unemployment rate reduced the probability of finding a job by almost 6 percentage points. In terms of the average labour market conditions facing the two cohorts, the improved labour market experienced by the late 90s cohort probably increased their full-time employment outcomes by about 12 percentage points compared to the early 90s cohort.

The predictive performance of the equation appears satisfactory, correctly predicting the observed employment outcomes of over 60 per cent of those who did or did not get full-time jobs.[20] The equation and the estimated magnitudes of the effects appear to survive a number of specification and robustness tests. First, the equation survived tests for heteroskedasticity associated with the included explanatory variables.[21] Second, the equation survived specification tests of the probit model's capacity to capture key features of the observed data (Pagan 2002).[22]

Third, it is possible that the policy change had a differential impact on the two birth groups affected by it. Exclusion of the July – September birth group resulted in a slightly more positive point estimate of the marginal effect on the years of Reception variable (13.5 percentage points) and a less negative age left school effect (-6.8 percentage points). The parameters were estimated less precisely with the reduction in observations, however and were significant only at the 10 per cent level.

Fourth, when the equation was estimated by GMM to account for any endogeneity of the age left school and grade completed variables the estimated effects of the education variables on the probability of being employed full-time all increased. The variables that appear in Table 6.1 but were excluded from the full-time employment equation were used as instruments in the GMM estimation. The GMM estimates were measured with less

precision, such that age left school and grade completed variables were not significant at the 10 per cent level.

Another potential problem for the analysis is that any change in the proportion of individuals pursuing post-school education in the year after completion of their schooling may induce a 'selection' bias in the results. This just means that any comparison of the outcomes of those not going on to full-time post-school education may be affected by some change in the characteristics of that group over time that a researcher cannot observe. For example, if post-school participation in education is related to 'ability' and the proportion participating over time increases, the average ability of those who do not participate may fall. Since ability is difficult to observe, analysis that fails to deal with the possibility that average ability has fallen may wrongly attribute the impact of the change in average ability on an outcome variable to another factor correlated with ability, such as completed schooling.

Selection bias does not appear to be a problem here, however. This assessment is based on estimation of a bivariate probit equation that accounted for selectivity. The 'first' equation was designed to explain which individuals studied full-time in the year after completing school. This outcome was positively associated with school achievement, student self-rating of their school performance, completion of higher-level school grades, being female, being from a non-English language background, living in a metropolitan area, being from a high socio-economic status (SES) background and having attended a school with students from higher average SES backgrounds. The correlation of the error term from that equation with the error term of the full-time employment outcome equation was not significant. The parameter estimates of the full-time employment equation changed little when it was estimated this way. The effects of living in a metropolitan area and of school achievement on full-time employment were smaller, while that of the grade completed was slightly larger. These results do not suggest selection biases have influenced the estimated parameters in any substantial way.

While Table 4 contains estimates of the effect of an additional year of schooling on full-time employment outcomes in individuals' second year after leaving school, these estimates are quite speculative. They are based on a similar specification to that used for the first year outcome. The results suggest that the education related variables had little effect on the second year full-time employment outcomes of individuals in the labour market, though the smaller number of observations used in these results should be noted (536 for the second year results compared with 775 for the first year results).

The second year results reported in Tables 4 and 6.3 ignore the actual first year outcomes of individuals, which appear to be very important in determining their subsequent employment outcomes. Just over 80 per cent of those employed full-time in their first year out from school were employed full-time in the second year. Just fewer than 40 per cent of those not employed full-time in their first year out from school were employed full-time in the second year. It is not clear quite how these first year outcomes should be incorporated in the analysis. One approach is to estimate a sequential process where the set of variables in Table 6.2 determine the first year outcomes, with the second year outcomes determined by the actual first year outcome and a subset of the first year explanatory variables, with correlation allowed between the error terms of the two equations (Greene 2002: E17 – 32). Alternatively, the second year equation could be estimated separately for those with and without full-time jobs in the first year, again with correlation allowed between the error terms of the first and second year employment

equations (Stewart and Swaffield 1999). With the first approach, the first year outcome effect dominates all other variables. That is, once whether individuals had a full-time job in their first year out from school is taken into account, none of the other explanatory variables are significant. With the second approach, few of the variables in the second equation were significant. Grade completed may have had a negative effect on retaining full-time employment in the second year among those who had a full-time job in the first year, but a positive effect among those who did not have a full-time job in their first year out from school. In general, these attempts to incorporate a role for the first year outcomes were not satisfactory and were not pursued. Nevertheless, it appears that any second year years of Reception effect works only through its impact on first year outcomes, rather than having any separate effect.

These effects have been estimated with a relatively small number of observations (see Table 4.6). In general, the various surveys in the Longitudinal Surveys of Australian Youth series are designed to be nationally representative of school students when first surveyed, not necessarily representative of school students in specific jurisdictions such as South Australia. In addition, the effects are based on the measured impact of one specific South Australian policy change, from which it may not be possible to generalise. In this case the policy change exploited here targeted individuals born at specific times of the year. Its impact is not related to any aspects of the survey designs of the two cohorts used here. Nor were the group of students affected by the change selected on the basis of any specifically South Australian measure or phenomenon. Moreover, they simply undertook an additional year of school that was already (theoretically) part of the South Australian education system and existed in other Australian jurisdictions. The impact of the policy change appears capable of generalisation, therefore, and the available data can be used to support the analysis undertaken. It would clearly have been better to have available more data for the specific application analysed here, especially from the early 90s cohort. Nevertheless, the relatively small number of observations used in this application has not prevented the identification of statistically significant relationships between the education variables and the probability of full-time employment in individuals' first years out from school.

Discussion

The results presented indicate that the additional schooling undertaken during their junior primary years of school among those affected by the policy had a positive impact on their employment outcomes. However, another impact of the policy, that they were older when they left school, had a negative impact on individual employment outcomes.

One question worth asking is why might an additional introductory year of schooling have a positive impact on whether individuals obtain a full-time job when they leave school. The additional schooling may have influenced their employability in one of three conceivable ways. It may have influenced: their skills, knowledge or competencies; their attitudes; or other characteristics, such as their levels of maturity.

In general, these phenomena are difficult to measure. Evidence presented elsewhere in the paper (Appendix 4) indicates that the additional years of Reception undertaken by some students may have influenced their school achievement. For example, evidence in Table 4.3 indicates numeracy and, possibly, literacy levels for the entire October –

February birth group increased by up to a quarter of a standard deviation of performance compared with the March – June comparison group. While the estimated numeracy effect for the October – February group who did not go onto full-time post-school study was half that of the entire group and was not significantly different from zero (Table 4.7), it seems plausible that their improved employment outcomes may have been related to their improved school performance.

Attitudes and other characteristics are difficult to measure. There is weak evidence from the tables in Appendix 4 that members of the October – February birth group may have had more positive attitudes towards school, with more of them planning to undertake Year 12 when first interviewed, though this did not actually translate into substantially higher Year 12 completion. Identification of what 'intermediate' factors were influenced by their additional schooling and, in turn, improved the full-time employment outcomes, may be a valuable direction for further research.

Nevertheless, the results are consistent with those from the early childhood intervention literature that suggest that early childhood programs can have substantial, long-lasting impacts. Programs that enrich one's early school experience can influence the labour market outcomes achieved later in the lives of individuals. Such impacts appear to be mirrored in the results presented here.

Implications of the results for policy

The policy implications of these results are not that all Australian school students should undergo another year of school to improve their likelihood of obtaining full-time employment. The research has shown that such policies may have offsetting effects – a positive educational effect and a negative one by virtue of the wage implications of the age at which individuals leave school. These implications presumably apply to other policies that influence the age at which students leave school. For example, increases in minimum school leaving ages have a similar impact on the age of leaving school to those described here – they will increase that age by one year for some students. Based on the results presented here, it might be expected that such an increase may result in a negative effect on the full-time employment outcomes of affected individuals. Unless the additional year of schooling imparts improved skills, attitudes or other characteristics to these individuals , the net effect of such changes may be detrimental for those the reforms are intended to benefit.

The results presented here do provide encouragement for other policies designed to redress the disadvantage of students early in their schooling – an additional year of junior primary school can have a substantial impact on the employment outcomes achieved by individuals subject to such initiatives. Other junior school enrichment programs that improve the skills of individuals, without directly affecting the amount of schooling they undertake, such as the literacy programs analysed in Ainley *et al*. (2002), would appear to have the potential to provide substantial benefits to their participants.

7. CONCLUSION AND OTHER APPLICATIONS

This paper has exploited a policy change that occurred in South Australia in the mid-1980s and affected the years of schooling obtained by an identifiable subset of students to generate a 'causal' estimate of the effect of schooling on the full-time employment outcomes achieved by those students. It appears that an additional year of junior schooling generates an increase in the probability of being employed full-time of about 11 percentage points in school leavers' first year out from school. In addition, there is an offsetting effect of about 8 percentage points on that probability because individuals were older when they left school. These estimates apply only to those students who leave school and do not undertake full-time post-school studies immediately. It is evidence that the process of schooling itself has a considerable effect on the labour market outcomes achieved by individuals.

The policy implications of these results are not that all school students should undergo another year of school to improve their likelihood of obtaining full-time employment. The policy change analysed here increased the introductory schooling of a group who undertook too little of it compared to their peers and improved their full-time employment outcomes. This outcome provides encouragement for other policies designed to redress the disadvantage of students early in their schooling – an additional year of junior primary school can have a substantial impact on the employment outcomes achieved by individuals subject to such initiatives. This statement assumes that the outcomes of the effect of the *Early Years of School* policy change can be generalised. Since it involved no more than an additional year of school for some individuals based on their date of birth, no change in the formal school structure and, therefore (presumably), little development of new curriculum, such an assumption seems reasonable here.

In some senses the application here was a difficult test of whether the *Early Years of School* policy can be exploited to provide a causal estimate of the effect of schooling on some phenomenon. Analysing only those individuals who did not proceed to full-time study after leaving school cut the available sample size from the early 90s cohort substantially. Other potential applications, such as some of those described below, need not reduce the sample size so much. As demonstrated here, even where difference in difference estimates lacked precision because of the small sample size from the early 90s cohort or were not informative for other reasons, it was still possible to generate statistically significant regression-based estimates.

The approach adopted in this paper is well suited to analysis of a range of other topics, in terms of the effect of additional schooling on both the educational and labour market outcomes achieved by individuals. Other topics that could be analysed in the future using the approach adopted here include:

- the effect of schooling on other labour market outcomes (wages and unemployment incidence), again among school leavers not engaged in full-time post-school education;
- the effect of additional schooling on participation in further education and training. Does it improve student 'readiness', as the *Early Years of School* policy was intended to do, or do individuals substitute additional years of school for further study? ;

- the effect on educational achievement and attainment outcomes of slowing down the progress of students (especially boys) through school so that they are more mature during secondary school;
- the effect of additional schooling on individual social and political engagement and involvement in cultural activities;
- assessing the extent to which students choose their education in a strategic way as implied by screening or signalling theories of the role of schooling. This would involve comparing school completion in South Australia between the two cohorts analysed here with developments in other jurisdictions; and
- identifying whether it is attainment level or years of schooling that are more important in explaining youth labour market outcomes.

The *Early Years of School* policy is not the only Australian policy change that can be exploited in this way. For example, the school entry changes implemented in Queensland in the mid-1980s affected the age of an identifiable subset of students in grade cohorts in that State (those born in January and February). Once more, the effect of age and maturity on educational and labour market outcomes could be analysed through the impact of that policy change. This could also inform policy debate about the merits of slowing down the progress of students through school, for example. Other policy changes that affected school entry arrangements in various jurisdictions in the late 1970s and early 1980s should also be captured between other cohorts in the various Longitudinal Surveys of Australian Youth series.

None of this discussion should be taken as suggesting such analyses are straightforward. Identifying the right comparison group, eliminating differences in survey design, labour market conditions and other factors and assessing whether the groups being compared are similar other than for the effect of the policy can be complex, as this paper bears testimony. Nevertheless, possession of 'causal' estimates of the effect of some phenomenon on another is invaluable: the evidence from this paper is that the additional schooling undertaken by some individuals in South Australia improved their full-time employment outcomes. There was no signalling or qualification effect associated with this additional schooling, since the group influenced by the policy and the comparison group reached the same average grade or level in both years. The improved outcomes for those with more schooling must have arisen because their probability of selection for the available positions by employers increased. This is important evidence that schooling adds to the skills and competencies of students. Education reform can affect more than one characteristic of individuals, however. In this case, the increased age at which individuals left school offset part of the positive schooling effect. Clearly, education reforms can have a complex set of impacts on those affected by them. Proper analysis of such reforms requires identification of all of their various elements.

REFERENCES

Ainley, J., Fleming, M. and McGregor, M. (2002), *Three Years On: Literacy Advance in the Early and Middle Primary Years*, Catholic Education Commission of Victoria, Melbourne.

Alloway, N., Freebody, P., Gilbert, P. and Muspratt, S. (2002), *Boys, Literacy and Schooling: Expanding the Repertoires of Practice*, Curriculum Corporation, Melbourne.

Angrist, J.D. and Krueger, A.B. (1991), 'Does compulsory school attendance affect schooling and earnings?' *Quarterly Journal of Economics*, 106, 979 – 1014

— (1999), 'Empirical strategies in labor economics', Chapter 23 in Ashenfelter, O. and Card, D. (eds) *Handbook of Labour Economics*, Volume 3A, Elsevier Science, Holland.

Ashenfelter, O. and Krueger, A. (1994), 'Estimates of the economic return to schooling from a new sample of twins', *American Economic Review*, 84, 1157- 73.

Ashenfelter, O. and Rouse, C. (1998), 'Income, schooling and ability: evidence from a new sample of twins', *Quarterly Journal Economics*, 114, 1157 – 74.

Australian Industrial Relations Commission (1998), *Junior Rates Inquiry: Issues Paper*, AIRC, Sydney.

Barnett, S.W. (1998), Long-term cognitive and academic effects of early childhood education on children in poverty, *Preventive Medicine*, 27, 204 – 207.

— (2002), *The Battle over Head Start: What the Research Shows*, National Institute for Early Education Research Working Paper, available at http://nieer.org.

Behrman, J.R., Rosenzweig, M.R. and Taubman, P. (1994), 'Endowments and the allocation of schooling in the family and in the marriage market: the twins experiment', *Journal of Political Economy*, 102, 1131 – 74.

Borland, J., Dawkins, P., Johnson, D. and Williams, R. (2000), *Returns to Investment in higher education*, the Melbourne Economics of Higher Education Research Program Report No 1, mimeo, University of Melbourne.

Bound, J., Jaeger, D.A. and Baker, R. (1995), 'Problems with instrumental variables when the correlation between the instruments and the endogenous explanatory variables is weak', *Journal of the American Statistical Association*, 90, 443 – 450.

Brunello, G. and Miniaci, R. (1999), 'The economic returns to schooling for Italian men. An evaluation based on instrumental variables', *Labour Economics*, 6, 509-519.

Budd, B and Madden, G. (1999), *Determinants of Australian Full-Time Earnings*, Centre for Labour Market Research Discussion Paper 99/7, Murdoch University.

Callan, T. and Harmon, C. (1999), 'The economic returns to schooling in Ireland', *Labour Economics*, 6, 543 – 50.

Card, D. (1995), 'Using geographic variation in college proximity to estimate the return to schooling. In Christofides, L.N., Grant E.K. and Swidinsky, R. (eds) *Aspects of labour market behaviour: Essays in Honour of John Vanderkamp*, University of Toronto Press, Toronto, 201 – 202.

— (1999), 'The causal effect of education on earnings', Chapter 30 in Ashenfelter, O. and Card, D. (eds) *Handbook of Labour Economics*, Volume 3A, Elsevier Science, Holland.

Cawley, J., Heckman, J. and Vytlacil, E. (2001), 'Three observations on wages and measured cognitive ability', *Labour Economics*, 8, 419-442.

Chang, C-A and Miller, P.W. (1996), 'The inter-industry wage structure: evidence from the 1991 Australian Census', *Australian Bulletin of Labour*, 22, 28 – 48

Committee of Enquiry into Education in South Australia (Keeves Enquiry) (1981), *Education and Change in South Australia, First Report*, Government Printer, South Australia.

— (1982), *Education and Change in South Australia, Final Report*, Government Printer, South Australia.

Currie, J. (2001), 'Early childhood education programs', *Journal of Economic Perspectives*, 15, 213 – 238.

Currie, J. and Thomas, D. (1995), 'Does Head Start make a difference', *American Economic Review*, 85, 341 – 364.

— (2000), 'School quality and the longer-term effects of Head Start', *Journal of Human Resources*, 35, 755 – 774.

de Lemos, M. (2002), *Closing the Gap between Research and Practice: Foundations for the Acquisition of Literacy*, Australian Council for Educational Research, Camberwell.

Director General of Education, South Australia, (1984), *The Annual Report of the Director General of Education, South Australia*, Government Printer, South Australia.

Duflo, E. (2001), 'Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment', *American Economic Review*, 91, 795-813.

Education Department of South Australia (1983), *The Early Years of School*, Policy Development Paper No. 1, Government Printer, South Australia.

— (1984), *Education Gazette, South Australia*, No. 34, Volume 12, Week Ending 23 November 1984, Government Printer, South Australia.

— *Calendar*, various years, Government Printer, South Australia.

Everitt, B.S. and Dunn, G. (1991), *Applied Multivariate Data Analysis*, Edward Arnold, London

Foley, D., Goldfeld, S., McLoughlin, J., Nagorcka, J., Oberklaid F. and Wake M (2000), *A Review of the Early Childhood Literature*, Department of Family and Community Services, Canberra.

Greene, W. (1997), *Econometric Analysis*, Third Edition, Prentice-Hall, New Jersey.

— (2002), *LIMDEP*, Version 8, Econometric Software Inc., New York.

Harmon, C. and Walker, I. (1995), 'Estimates of the economic return to schooling for the United Kingdom', *American Economic Review*, 85, 1278 – 86.

Harris, M. (1996), 'Modelling the probability of youth unemployment in Australia', *Economic Record*, 72, 118-129.

Hausman, J.A. (1978), 'Specification tests in econometrics', *Econometrica*, 46, 1251 – 71.

Heckman, J.J., Lalonde, R.J. and Smith, J.A. (1999), 'The economics and econometrics of active labor market programs', Chapter 31 in Ashenfelter, O. and Card, D. (eds) *Handbook of Labour Economics*, Volume 3A, Elsevier Science, Holland.

Ichino, A. and Winter-Ebmer, R. (1999), 'Lower and upper bounds of returns to schooling: an exercise in IV estimation with different instruments', *European Economic Review*, 43, 889-901.

Kalb, G. (2000), *The Dynamics of Labour Market State and Benefit Receipt*, Australian Bureau of Statistics Occasional Paper, (Cat. No. 6293.0.00.004).

Karoly, L.A., Greenwood, P.W., Everingham, S.S., Hoube, J., Kilburn, M.R., Rydell, C.P., Sanders, M. and Chiesa J. (1998), *Investing in our Children: What We Know and Don't Know about the Costs and Benefits of Early Childhood Intervention*, RAND Corporation, Santa Monica.

Lamb, S. (1996), *Completing School in Australia: Trends in the 1990s*, Longitudinal Surveys of Australian Youth Research Report No 1, Australian Council for Educational Research, Melbourne.

Lamb, S. and McKenzie, P. (2001), *Patterns of Success and Failure in the Transition from School to Work in Australia*, Longitudinal Surveys of Australian Youth Research Report No 18, Australian Council for Educational Research, Melbourne.

Lamb, S., Dwyer, P. and Wyn, J. (2000), *Non-completion of school in Australia: the Changing Patterns of Participation and Outcomes*, Longitudinal Surveys of Australian Youth Research Report No 16, Australian Council for Educational Research, Melbourne.

Le, A.T. and Miller, P. (1999a), *A Risk Index Approach to Unemployment – An Application Using the Survey of Employment and Unemployment Patterns*, Australian Bureau of Statistics Occasional Paper, (Cat. No. 6293.0.00.001).

— (1999b), *Job Quality and Churning of the Pool of the Unemployed*, Australian Bureau of Statistics Occasional Paper, (Cat. No. 6293.0.00.003).

Lee, V. and Loeb, S. (1995), ' Where do Head Start attendees end up? One reason why pre-school effects fade out', *Educational Evaluation and Policy Analysis*, 17, 62 – 82.

Levin, J. and Plug, E.J.S. (1999), 'Instrumenting education and the returns to schooling in the Netherlands', *Labour Economics*, 6, 521 –34.

Lingard, B., Martino, W., Mills, M. and Bahr M. (2002), *Addressing the Educational Needs of Boys*, Report to the Department of Education Science and Training, Canberra.

McCain, M.N. and Mustard, J.F. (1999), *Reversing the Real Brain Drain: Early Years Study Final Report*, The Canadian Institute for Advanced Research, Toronto.

Maddala, G.S. (1983), *Limited-dependent and qualitative variables in econometrics*, Cambridge University Press, Cambridge.

Marks, G.N. and Fleming, N (1998a), *Youth Earnings in Australia 1980 – 1994: A Comparison of Three Youth Cohorts*, Longitudinal Surveys of Australian Youth Research Report No. 8, Australian Council for Educational Research.

— (1998b), *Factors Influencing Youth Unemployment in Australia 1980 – 1994*, Longitudinal Surveys of Australian Youth Research Report No. 7, Australian Council for Educational Research.

Marks, G.N., and Long, M. (2000), *Weighting the 1995 Year 9 Cohort Sample for Differential Response Rates and Sample Attrition*, Longitudinal Surveys of Australian Youth Technical Paper No 15, Australian Council for Educational Research, Melbourne.

Marks, G.N., Fleming, N., Long, M. and McMillan, J. (2000), *Patterns of Participation in Year 12 and Higher Education in Australia: Trends and Issues*, Longitudinal Surveys of Australian Youth Research Report No 17, Australian Council for Educational Research, Melbourne.

Meyer, B.D. (1995), 'Natural and quasi-experiments in economics', *Journal of Business and Economic Statistics*, 13, 151 – 161.

Miller, P. and Mulvey, C. (1996), Unions, firm size and wages, *Economic Record*, 72, 138-153.

— (1997), Computer skills and wages, *Australian Economic Papers*, 36, 106 – 113.

Miller, P., Mulvey, C. and Martin, N. (1995), 'What do twins studies reveal about the economic returns to education? A comparison of Australian and U.S. findings', *American Economic Review*, 85, 586 – 99.

Miller, P., Mulvey, C. and Norris, K. (1997), 'Compensating differentials for risk of death in Australia', *Economic Record*, 73, 336-372.

Murphy, K.M. and Topel, R.H. (1985), 'Estimation and inference in two-step econometric models', *Journal of Business and Economic Statistics*, 3, 370 – 79.

Nakamura, A. and Nakamura, M. (1981), 'On the relationship among several specification error tests presented by Durbin, Wu and Hausman', *Econometrica*, 49, 1583 – 88.

Pagan, A.R. (2002), 'Learning about models and their fit to data', *International Economic Journal*, 16, 1 – 18.

Preston, A (1997), 'Where are we now with human capital theory in Australia', *Economic Record*, 73, 51-78.

Queensland Department of Education (1984), *Education Office Gazette*, Vol.86, No. 5, 23 March 1984, Government Printer, Queensland.

Raban, B. (2000), *Just the Beginning . . .* , Department of Education, Training and Youth Affairs, Canberra, available at:
http://www.dest.gov.au/research/fellowship/docs/Bridie_Raban/Just_the_Beginning.pdf .

Rosenzweig, M.R. and Wolpin, K.I. (2000), 'Natural "natural experiments" in economics', *Journal of Economic Literature*, 38, 827 – 874.

Rummery, S., Vella, F., and Verbeek, M. (1999), Estimating the returns to education for Australian youth via rank-order instrumental variables, *Labour Economics*, 6, 491 – 507.

Ryan, C. (2002), *Longer-term outcomes for individuals completing vocational education and training qualifications*, National Centre for Vocational Education Research, Adelaide.

Smith, R.J. and Blundell, R.W. (1986), 'An exogeneity test for a simultaneous equation tobit model with an application to labor supply', *Econometrica*, 54, 679 – 685.

Stewart, M.B. and Swaffield, J.K. (1999), 'Low pay dynamics and transition probabilities', *Economica*, 66, 23 – 42.

Strombach, T. and Dockery, A.M. (2000), *Labour Market Programs, Employment and Unemployment Hazards: an application using the 1994 – 1997 Survey of Employment and Unemployment Patterns*, Australian Bureau of Statistics Occasional Paper, (Cat. No. 6293.0.00.002).

Trethewey, L. (1997), 'Beginning school in the 1970s and 80s: the (re-) introduction of continuous admission in South Australia', *New Zealand Journal of Educational Studies*, 32, 175 – 189.

Vella, F. and Gregory, R.G. (1996), Selection bias and human capital investment: estimating the rates of return to education for young males, *Labour Economics*, 3, 197-219.

Vella, F. and Karmel, T. (1999), 'Evaluating the impact of educational expansion on the occupational structure of youth', *Australian Economic Papers*, 38, 310 – 27.

Vieira, J.A.C. (1999), 'Returns to education in Portugal', *Labour Economics*, 6, 535 – 41.

Williams, T. (1987), *Participation in Education*, ACER Research Monograph No. 30, Australian Council for Educational Research, Melbourne.

Wooldridge, J. (2002), *Econometric Analysis of Cross Section and Panel Data*, Massachusetts Institute of Technology Press, Massachusetts.

APPENDIX 1: SOUTH AUSTRALIAN SCHOOL ENTRANCE AND PROGRESSION POLICY

Policy prior to 1985

Children admitted at age five will be regarded initially as reception enrolments, except that such children admitted at the beginning of the year may be regarded as Year 1 enrolments.

Only in exceptional circumstances should children admitted at age five spend less time than two years in reception to Year 2. No child, at whatever age admitted, shall spend more than three years in reception to Year 2 without referral to a guidance officer.

(Quoted in Committee of Enquiry into Education in South Australia 1982: 91)

The Policy following implementation of the Early Years of School Policy

It is Education Department policy that children enrolling in government schools have between seven and ten terms in junior primary classes, that is: reception, year 1 and year 2. This policy will be fully introduced by 1986.

Depending on the date of admission, children will progress according to the following patterns:

- Children admitted at five years of age in February will have nine terms, that is, three years in junior primary classes.
- Children admitted at the beginning of second term will have eight terms.
- Children admitted at the beginning of third term will normally have seven or ten terms, depending on individual development, competence and maturity.

In exceptional cases the length of time that particular children will spend in junior primary classes and the age at which they commence school may be varied through discussions between parents and teachers and the principal with advice from guidance officers if appropriate. This might apply to children who have begun school close to six years of age or to children with special social, emotional, physical or intellectual needs. Decisions taking age and maturity into account with reference to departmental guidelines may favour proceeding to Year 3 after completing only six terms of junior primary education or staying longer than ten terms. (*Education Gazette*, South Australia, No. 34, Volume 12, Week Ending 23 November 1984: 1048)

Both before and after the implementation of the Early Years of School Policy, departmental policy further stated that:

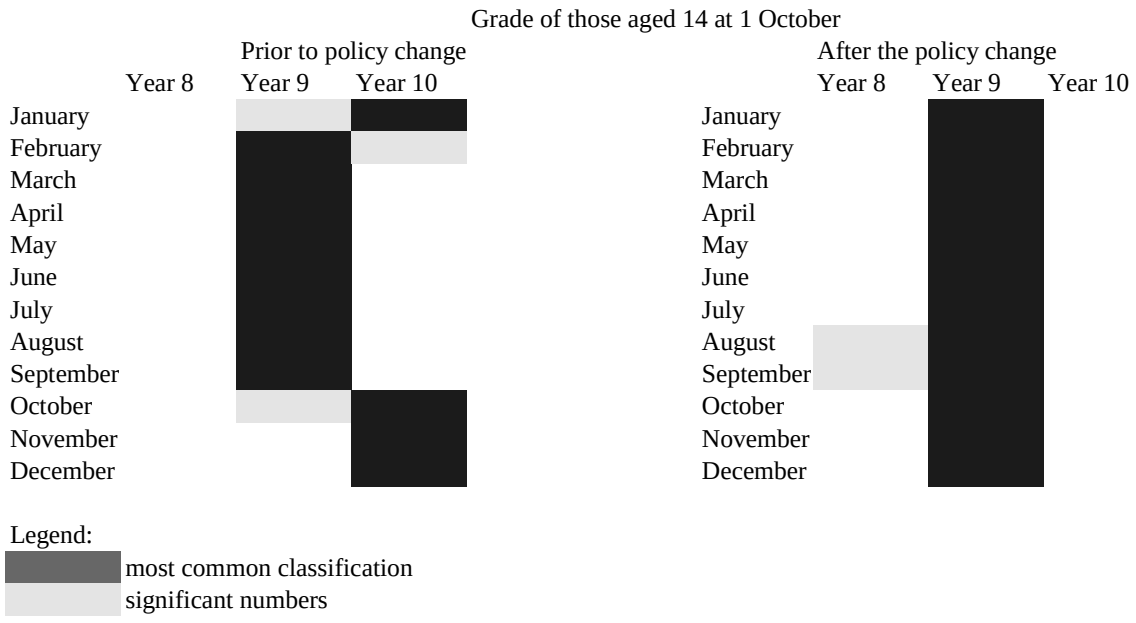
> Entry to Year 3 should occur only at the beginning of the school year.

APPENDIX 2: HOW THE *EARLY YEARS OF SCHOOL* POLICY IS CAPTURED IN LSAY DATA

The policy known as the *Early Years of School* policy was introduced in South Australia in the mid-1980s. It was implemented to slow the way an identifiable subset of students progressed through junior primary, so that they were older when they reached secondary school. Groups of individuals influenced by the policy change can be broadly identified, since it was directed at individuals born at specific times of the year.

It is possible to use data from two Longitudinal Surveys of Australian Youth (LSAY) cohorts to assess the impact of the policy change. The Youth in Transition 1975 birth cohort (YIT 75) and the Longitudinal Surveys of Australian Youth Year 9 cohort (LSAY 95) fall either side of the policy change and should reflect its impact. The reference date for the ages of individuals in these surveys is the beginning of October. That is, individuals needed to be 14 years old when surveyed in the YIT 75 survey at the beginning of October in 1989. Some individuals in the survey were about to turn 15 years over the remainder of 1989. This means that the individuals aged 14 in YIT 75 in South Australia were likely to have been in Year 9 if they had turned 14 after the start of the 1989 school year and Year 10 otherwise. This pre-policy age-grade structure, for a survey with an age reference point like YIT 75 is depicted in the left hand panel of Figure 2.1.

Figure 2.1 Effects of the policy change on age and grade structure in South Australia the YIT 75 survey by month of birth



The major effect of the policy change was to slow the progress through primary school of individuals who started school at the beginning of the school year. This means that 14-year-old individuals at October 1 who would have been in Year 10 prior to the policy change were likely to be in Year 9 after it. The age-grade structure following the policy change, for a survey with an age reference point like YIT 75, is depicted in the right hand

panel of Figure 2.1. Fourteen year olds (at 1 October) who were in Year 10 prior to the policy change (who were born between October and early February) were likely to be in Year 9 after it. In addition to affecting those individuals who started school at the beginning of the school year, the policy change envisaged that some individuals who started school at the beginning of term three might undertake a full three years of junior primary in addition to their initial term. Therefore, it seems likely that some whose birthday was in August and September and might formerly have been in Year 9, may have instead been in Year 8 following the policy change in a survey with an age reference point of early October like YIT 75. This possibility is also shown in the right hand panel of Figure 2.1.

As already described, the YIT 75 cohort was an age-based cohort. When surveyed in October 1989, the respondents were aged 14 years. The actual grades students were in across Australia when surveyed in 1989 reflected differences in the structure of the schooling systems, school commencement procedures across Australian jurisdictions and the timing of the survey (or at least, the reference date for age in the survey of the beginning of October). These are summarised in the left hand panel of Figure 2.2. The numbers in the various cells show the proportion of students born in particular months in each jurisdiction who were in the grade identified. That is, looking at the first number in the Year 9 column, 92 per cent of fourteen-year-old students in New South Wales (NSW), Victoria (VIC), Tasmania (TAS) and the Australian Capital Territory (ACT) whose month of birth was January, were in Year 9 in 1989.

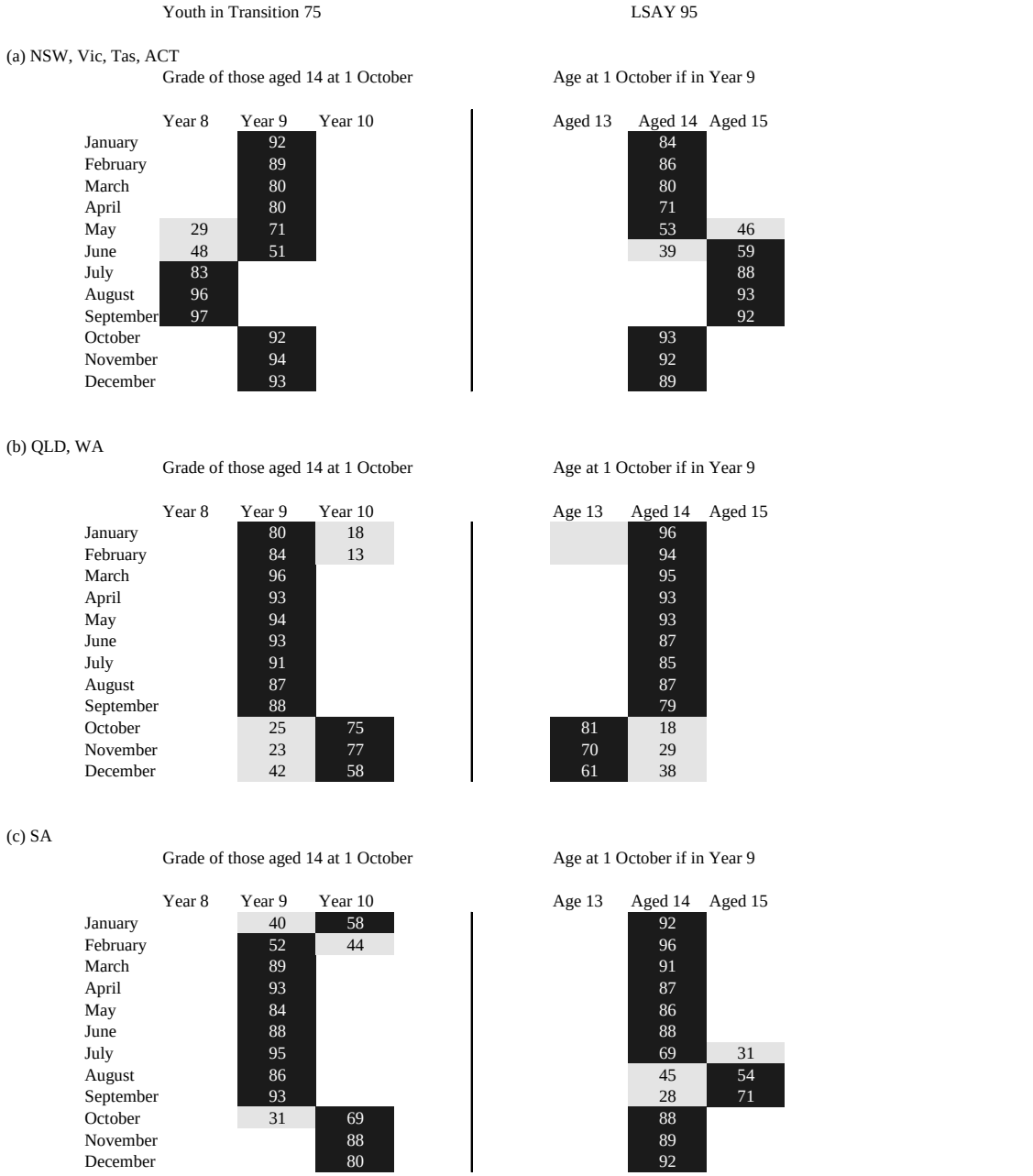
Students from jurisdictions with thirteen full years in their formal school structure (NSW, VIC, TAS and the ACT) were primarily in Year 9 if their birthday was in the first half of the year or if their fifteenth birthday was due in the December quarter. They were typically in Year 8 if they turned fourteen in the September quarter.[23,24] Students from jurisdictions with twelve full years in their formal structure (Queensland – QLD and Western Australia – WA) were primarily in Year 9 unless their fifteenth birthday was due in the December quarter, in which case they tended to be in Year 10.

The pattern was similar in SA to that of QLD and WA, except that some students whose birthdays fell in January or February were also in Year 10.[25] The grade structure for the YIT 75 in Figure 2.2 matches that foreshadowed by the left hand panel of Figure 2.1.

The cohorts affected by the policy reached secondary school in South Australia after 1990 and Year 12 from 1994 onwards. The late 90s (LSAY 95) cohort were a grade cohort that started school after the policy change and should, therefore, reflect its impact. South Australian children in Year 9 in the late 90s cohort should be older for their grade level than those in the early 90s (YIT 75) cohort.

The late 90s cohort was a grade-based panel. Students were in Year 9 in 1995, but varied by age, again depending on differences in the structure of the schooling systems, school commencement procedures across Australian jurisdictions and the timing of the survey (or at least, the reference date for age in the survey of the beginning of October). These patterns are shown in the right-hand panel of Figure 2.2. The numbers in the various cells show the proportion of students born in particular months in each jurisdiction in Year 9 who were of the age specified. That is, looking at the first number in the Aged 14 column, 84 per cent of Year 9 students in NSW, VIC, TAS and the ACT whose month of birth was January, were aged fourteen years in 1995.

Figure 2.2 Comparison of the two cohorts: age and grade structure by jurisdiction



Sample size - based on unweighted data

	NSW	VIC	QLD	SA	WA	TAS	NT	ACT	TOTAL
YIT 75	990	1028	774	645	833	653	286	436	5645
LSAY 95	2928	2779	2383	1636	1785	560	370	586	13027

Legend: most common classification; significant numbers

The diagram allows the identification of the effect of the different survey designs on the composition of the respondents by jurisdiction. For example, individuals from NSW, VIC, TAS and the ACT born between January to April or October to December are at similar stages of their schooling in both surveys – they were typically aged fourteen years and in Year 9 in both surveys. Differences in the survey designs did affect the stage of schooling for individuals born between May and August in those jurisdictions. The

fourteen year olds in the early 90s cohort were more likely to be in Year 8. Hence, those born in those months in the late 90s cohort who were in Year 9 tended to be 15 years old.

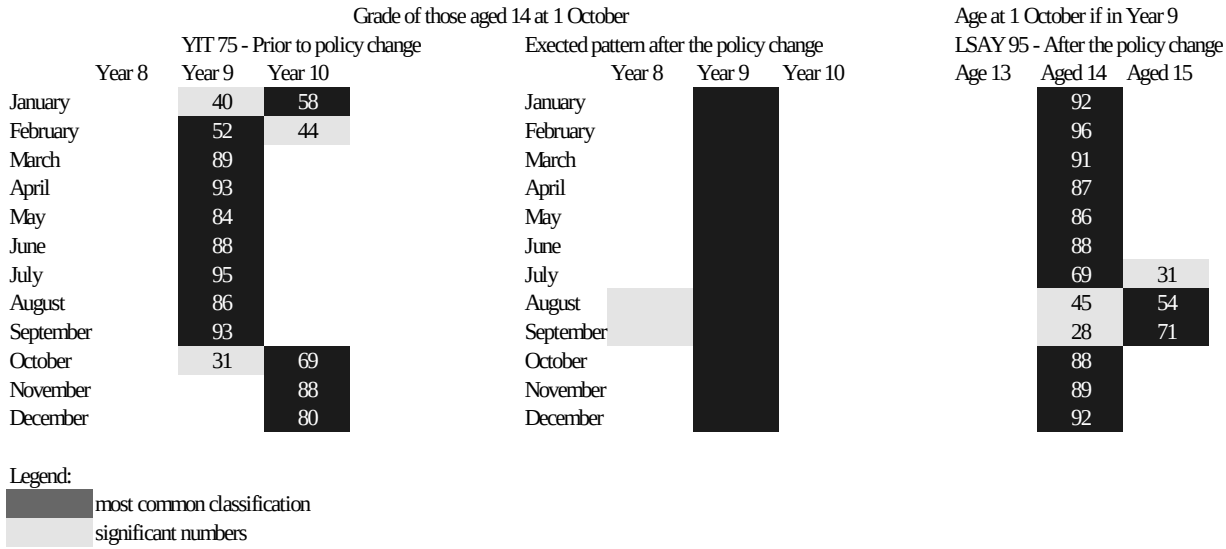
Essentially, for the top two groups, NSW, VIC, TAS and the ACT and QLD and WA, the right-hand panels are mirror images of the left-hand panels, with the impact of the difference in the survey design readily apparent. There are two exceptions to this mirror image pattern. In the NSW, VIC, TAS and the ACT group, individuals born in May or June tended to be more likely in the late 90s cohort to commence school with the following grade cohort rather than the cohort in which they would have been the youngest members.[26] Second, there are no students who turned thirteen years in January or February of Year 9 in the late 90s cohort, while there would have been some in the early 90s cohort. This reflects changed commencement arrangements in Queensland. Prior to 1986, students there were allowed to commence school provided they had turned five by the end of February. In 1986, they had to have turned five by the end of January in that year to commence school. From 1987, the year the late 90s cohort commenced school, they were required to have turned five in the preceding calendar year.[27]

While the ages of students in the late 90s cohort are largely a mirror image of the grade structure of fourteen year olds in the early 90s cohort for the two groups of jurisdictions described, that is not the case in South Australia. There, the pattern is different. In fact, the right hand side of Figure 2.2 is the mirror image pattern of the right hand side of Figure 2.1. That is, students in South Australia in the late 90s cohort are considerably older in Year 9 than would have been the case under the old school structure captured in the early 90s cohort. This is clearer in Figure 2.3, where the 'expected' grade distribution of fourteen years olds after the policy change (the right hand panel from Figure 2.1) is interposed between the two South Australian panels from Figure 2.2. Almost all students in Year 9 whose birthdays were from October to February were aged fourteen (rather than thirteen as would have been the case under the old arrangements) and many students whose birthdays were in the September quarter were aged fifteen (rather than fourteen).

What Figure 2.3 shows is that the change in the age structures of grade cohorts in South Australia resulting from the implementation of the *Early Years of School* policy is captured between the two cohorts. Therefore, analysis of the effect of the policy can proceed.

Some subtleties need to be reflected in any analysis, however. The first is that only individuals from South Australia born between March and June appear to provide a suitable control group – they were a common age in the same grade in both surveys. Individuals born between those months will be described hereafter as the 'comparison' group. The second is that individuals born between October and February appear to provide a group to assess the impact of the policy on specific outcome measures (this group is described hereafter as the 'October – February birth group'). They are the same age in both surveys, but most are in Year 9 in the late 90s cohort as opposed to Year 10 in the early 90s cohort. The third is that any comparison of specific outcomes for individuals born between July and September between the two surveys will contain both survey design and policy effects (this group is described hereafter as the 'July – September or September quarter birth group'). A second age-based survey would have shown significant numbers of fourteen year olds in Year 8 for this group.

Figure 2.3 Comparison of the two cohorts in South Australia: age and grade structure and the effect of the *Early Years of School* policy



Instead, the late 90s cohort contains a substantial number of fifteen year olds in Year 9. For individuals born in those months, it appears it became much more common for them to add a full three years of junior primary to the initial term they would have undertaken in the year they turned five years old. For this group, the effect of the policy could only be identified through a comparison with another group subject to similar age-based survey design effects, but who were not affected by the policy change. The direction of any survey design effects that would have been apparent for individuals born in this period in South Australia is evident among those born in the September quarter in NSW, VIC, TAS and the ACT in Figure 2.2. Therefore, this group can potentially be used to identify the effect of any changes in education that was induced by the policy change in South Australia on any outcome measure, such as full-time employment rates among school leavers.

The estimators of the effect of the policy for both the July – September birth group and the October – February group are set out in Appendix 3. Both estimators require that the group affected by the policy change (whose schooling changed) and the comparison group be 'alike', other than that the first group's behaviour was influenced by the policy change. The outcome of the comparison of the background characteristics of individuals in the various groups is described in Appendix 4.

APPENDIX 3: THE DIFFERENCE IN DIFFERENCES ESTIMATOR

The effect of the *Early Years of School* policy on some outcome measure of interest (for example, full-time employment) can be described with the aid of the following simple characterisation. That is, measured full-time employment, *F*, in South Australia (jurisdiction *j*) in year *t* for individuals affected by the policy change (*h*) who obtain education level *e* can be portrayed as

(3.1) $F_{ehjt} = \alpha_j + P_{ehjt} + C_{ehjt} + L_{\boldsymbol{jt}} + D_t + S_{hjt} + u_{ehjt}$

where

α_j captures jurisdiction-specific factors that affect the full-time employment of young people with education level *e*, but are constant through time

P_{ehjt} measures the impact of the policy introduced in year τ that affects group *h*, but not group *f* so that $P_{efjt} = 0$. In this case, P_{ehjt} reflects the effect of an additional year of schooling for the group affected by the policy change.

C_{ehjt} reflects the personal characteristics of the group affected by the policy change

$L_{\boldsymbol{jt}}$ reflects other aspects of the youth labour market in the jurisdiction that may vary from year to year

D_t reflects factors that affect all jurisdictions in the same way, but which vary from year to year

S_{hjt} reflects measurement issues that may influence the observed outcome, in this case factors associated with changes in survey design that may affect groups differently, and

u_{ehjt} is an error term with a zero mean.

Equation (3.1) simply means that in any calendar year, full-time employment among the group affected by the policy change that obtained education level *e* reflects:

- historical patterns (α_j),
- the characteristics of the group (C_{ehjt}),
- trends in other factors common across jurisdictions (D_t),
- factors that influence the youth labour market in that jurisdiction in that year ($L_{\boldsymbol{jt}}$),
- differences in the estimated employment outcomes of groups that arise from the way the data were collected (S_{hjt}),
- the additional year of schooling they obtained because of the policy change (P_{ehjt}) and
- factors unique to that group in that jurisdiction in that year (u_{ehjt}).

The effect of the policy change in period τ can be isolated by comparing the pre- and post policy implementation full-time employment outcomes of the group affected by the policy change with the change in the full-time employment outcomes of the control group who were not affected. Based on Appendix 2, a suitable comparison appears to be between individuals in SA born between October and February (affected by the policy) and those born in March to June (not affected).

A second comparison can be made between those born between July and September in SA (affected by the policy) and those born in March to June (not affected). This estimate will also include the impact of differences in survey design between the early and late 90s cohort data collections. From this observed difference, it is necessary to extract the survey design effect. This can be done by subtracting the change in employment among individuals with the same education level from other jurisdictions with the same likely survey design effect.

In the first case, the expected effect of the policy change is given by the difference in difference (DD) estimator:

$$(3.2) \qquad E[(F_{ehjt} - F_{ehj\tau-1}) - (F_{efjt} - F_{efj\tau-1})] = P_{jt}$$

$$\text{if} \qquad (C_{ehjt} - C_{ehj\tau-1}) - (C_{efjt} - C_{efj\tau-1}) = 0 \qquad \text{and}$$

$$(S_{ehjt} - S_{ehj\tau-1}) - (S_{efjt} - S_{efj\tau-1}) = 0.$$ [1]

That is, if the survey design features and characteristics of the respective groups obtaining education level *e* are either the same in both periods, different but unchanged over time, or if they differ and change, they change in similar ways between the surveys, then the difference in the change in the proportion employed full-time between the two groups provides an estimate of the effect of an additional year of schooling on full-time employment. In this case, the groups were chosen for whom the design features were identical in the two surveys. Therefore, the validity of the test requires only that the change in the background characteristics of individuals in the two groups be the same.

The second case arises where survey design effects are important, so that a difference in the difference in differences (DDD) estimator is required. In this case,

$$(3.3) \qquad E\{[(F_{ehjt} - F_{ehj\tau-1}) - (F_{efjt} - F_{efj\tau-1})] - [(F_{ehkt} - F_{ehk\tau-1}) - (F_{efkt} - F_{efk\tau-1})]\} = P_{jt}$$

$$\text{if} \qquad (C_{ehjt} - C_{ehj\tau-1}) - (C_{efjt} - C_{efj\tau-1}) - (C_{ehkt} - C_{ehk\tau-1}) - (C_{efkt} - C_{efk\tau-1}) = 0, \text{ and}$$

$$(S_{ehjt} - S_{ehj\tau-1}) - (S_{efjt} - S_{efj\tau-1}) - (S_{ehkt} - S_{ehk\tau-1}) - (S_{efkt} - S_{efk\tau-1}) =$$
$$(S_{ehjt} - S_{ehj\tau-1}) - (S_{ehkt} - S_{ehk\tau-1}) = 0, \text{ since } (S_{efjt} - S_{efj\tau-1}) = (S_{efkt} - S_{efk\tau-1}) = 0.$$

That is, the difference in the proportion employed full-time between the respective groups in the jurisdictions can provide an estimate of the effect of an additional year of schooling on employment outcomes. It does this if the characteristics of the individuals with the same education level in the various groups are unchanged over time, or if they change in similar ways. In addition, the affect of the change in survey design between groups in jurisdiction *j* (SA) needs to be the same as in the other jurisdictions (NSW, VIC, TAS and the ACT). The 'control' jurisdictions were selected with that purpose in mind (see Figure 2.2 of Appendix 2). The magnitude of the survey design effect in those jurisdictions may be overstated for SA, however. From Figure 2.2, the survey design effect is apparent for almost all individuals in the comparison jurisdictions between the

[1] Since $E[((\alpha_j + P_{ehjt} + C_{ehjt} + L_{jt} + D_t + S_{hjt} + u_{ehjt}) - (\alpha_j + C_{ehj\tau-1} + L_{j\tau-1} + D_{\tau-1} + S_{hjt-1} + u_{ehj\tau-1})) -$
$(\alpha_j + C_{efjt} + L_{jt} + D_t + S_{fjt} + u_{efjt}) - (\alpha_j + C_{efj\tau-1} + L_{j\tau-1} + D_{\tau-1} + S_{fjt} + u_{efj\tau-1}))]$
$= P_{ehjt} + (C_{ehjt} - C_{ehj\tau-1}) - (C_{efjt} - C_{efj\tau-1}) + (S_{ehjt} - S_{ehj\tau-1}) - (S_{efjt} - S_{efj\tau-1}) + E[u_{ehjt} - u_{ehj\tau-1} + u_{efjt} - u_{efj\tau-1}]$
The residuals disappear because each term has an expected value of zero.

two cohorts. In contrast, only about one half of those born in the September quarter in SA were aged 15 years. While this might have little impact on any analysis of unchanging demographic characteristics, it may have more important implications for developmental phenomena such as literacy and numeracy performance. This issue is discussed further in Appendix 4.

It should be noted that this estimator does not require that both general influences on employment (the state of the economic cycle, for example) and influences specific to the youth labour market (structural changes in the availability of full-time jobs for young people) should be common across jurisdictions. The double differencing removes these effects from the estimated effect of the policy on employment.

A version of the second estimator (equation 3.3) may also be more appropriate for analysis of the October – February birth group for the specific issue analysed here. Analysis contained in Appendix 4 indicates that it is not reasonable to assume that the labour market conditions when individuals in the early 90s cohort left school were similar between that group and the comparison group, since they were one grade ahead of the comparison group and could have entered the labour market prior to the onset of the recession of the early 1990s. Since there is no reason to expect $L_{jt} - L_{j\tau-1} = L_{jt} - L_{j\tau-2}$, the resulting estimator may not be a good one of the impact of the policy. It may be possible to use groups in other jurisdictions, notably Queensland and Western Australia, who entered the labour market at different times to remove this effect from the estimated impact of the policy. This would require the assumption that the timing of the economic cycle and the magnitude of the changes in economic conditions were the same in those jurisdictions in those years as in South Australia. An alternative approach is to rely on regression-based estimates of the effect of the policy, which can incorporate the labour market conditions at the time individuals entered it more directly.

The estimators in equations (3.2) and (3.3) rely on assumptions that in other circumstances may not be met. Where comparisons are made of groups across different jurisdictions, these include (see Meyer 1995 for discussion on these and other issues):

- That the jurisdiction- specific effects on the youth labour market do not change over time. This is obviously more reasonable the shorter the time between the two years in which the comparisons are made;
- That there are no interaction effects between any of the separately identified influences on the youth labour market in equation (1); and
- That either other policies that affect the youth market are constant over time or that all jurisdictions adopt policies with similar effects other than the one under consideration.

In this case, since the comparisons are fundamentally between groups in the same jurisdiction, the assumptions seem more reasonable. To some extent, these various qualifications to the difference in differences estimator and the assumptions it relies on can be tested through regression based techniques that allow for the different characteristics of individuals and for interactions between jurisdiction effects and economic factors to be taken into account. Nevertheless, the simplicity of difference in differences estimators such as those in equations (3.2) and (3.3) make their use very attractive. The use of difference in differences estimators requires careful analysis to ensure it is appropriate for any specific application. This analysis for this paper is described in Appendix 4.

APPENDIX 4: THE GENERAL SCOPE FOR 'NATURAL EXPERIMENT' ANALYSIS

The description of the difference in differences estimators in Appendix 3 shows that the difference in the change in the proportion employed full-time between two groups in two cross-sectional surveys can provide an estimate of the effect of completed schooling arising from policy changes on full-time employment. It will do this if the characteristics of the respective groups obtaining a common education level are either:

- the same in both periods,
- different, but unchanged over time,
- or if they differ and change, they change in similar ways between the two surveys.

In addition, any change in the design of the surveys needs to affect both groups similarly to allow estimation of the impact of the policy. The first purpose of this Appendix is to establish that, in general, the two surveys contain enough observations of individuals affected by the policy change or in the comparison group to allow the analysis of phenomena of interest to proceed. The second purpose is to establish if any of the three conditions about the change in the characteristics of individuals in the groups affected by the policy change and the comparison groups are met.

As discussed in the body of the paper, the groups of individuals from the two surveys analysed from South Australia are:

- those born between October and February, inclusive – who were affected by the policy change
- those born between July and September – who may have been affected by the policy change
- those born between March and June – the comparison group, who were not affected by the policy change.

In addition, a further comparison of individuals born in the September quarter in NSW, VIC, TAS and the ACT with those born in the June quarter in those jurisdictions is used to account for the age-based survey design effects that influence any comparison of the September quarter birth group with the March – June comparison group.

There are two related issues that need to be addressed: whether the data appear to support analysis in general and whether they support it in a specific application where the sample may be defined narrowly. A general comparison of the available data between the two groups affected by the policy change and the comparison group is considered first. This analysis supports the use of the difference in differences estimator between at least October – February birth group and the comparison group. That is, the changes in the background characteristics of individuals in that group affected by the policy change and the comparison group appear to be common between the two cohorts.

This Appendix also contains an analysis of the data for the specific application of interest here – to estimate the effect of schooling on the full-time employment outcomes of individuals who do not undertake full-time post-school study in their first two years out

of school. Unfortunately, for this application, the difference in differences estimator has problems arising from the timing of school leaving between the groups affected by the policy and the control group in the first cohort and the timing of the economic cycle.

Background to the data

Data drawn from the individuals who responded to the fourth interviews conducted in the early 90s and late 90s surveys are analysed first in this Appendix. These were conducted in 1992 and 1998 respectively. They inevitably contain some attrition that reduces the original panel size, but these years are probably the first in which meaningful labour market outcomes of those affected by the policy change might be observed through comparison of the two surveys.

As already indicated in the body of the paper, the two surveys differed in their design features. Among these design differences are those that affected the way the samples were drawn and the proportion of total observations from each jurisdiction and between school types. Both surveys have weights contained in the data files that account for these design features and allow re-weighting of observations to be 'nationally representative'. ACER has also developed and utilised a re-weighting procedure that aims to account for sample attrition, so that remaining observations are re-weighted each year to ensure that the respondent sample remains broadly representative of the population from which it was drawn. Unfortunately, the approach used by ACER to estimate these attrition-related weights differs between the two surveys (see Williams 1987 for a description of the YIT 75 approach and Marks and Long 2000 for a description of the LSAY 95 one). This weighting issue could be resolved by:

- ignoring attrition, since it might reasonably be argued that attrition should not be related to birth date, which here determines the groups to which individuals belong
- using a common approach to estimate attrition-related weights for both surveys
- presenting results where similarly based weights are used for both surveys and comparing them with results which use weights that account for the survey designs, but not attrition. In this case the approach used for the late 90s cohort was used to generate weights for the early 90s sample.

Evidence for why attrition might reasonably be ignored in the analysis is presented in Table 4.1. It contains the survey response rates in each of the first five surveys after the initial interview in the two collections for the groups of interest for this analysis. Attrition appears to follow a broadly comparable pattern across the various groups in the two series, though it is lower in the late 90s cohort. More detailed analysis of these patterns suggests that there are not significant differences in attrition between the two birth groups affected by the policy change and the comparison group. Tests of differences in response rates between the two groups affected by the policy change and the comparison group in the various years were insignificant, other than for September quarter birth group and the comparison group in the last year of the late 90s series.

Table 4.1 Response rates by cohort, survey year and analysis group, South Australian respondents

	Early 90s (YIT 75)			Late 90s (LSAY 95)		
	Oct- Feb birth group	Sept qtr birth group	Comparison group	Oct- Feb birth group	Sept qtr birth group	Comparison group
2nd year	0.863	0.860	0.824	0.775	0.838	0.794
3rd year	0.775	0.848	0.803	0.806	0.855	0.817
4th year	0.735	0.774	0.700	0.770	0.828	0.778
5th year	0.594	0.646	0.588	0.720	0.760	0.713
6th year	0.614	0.622	0.597	0.645	0.720	0.648

Attrition obviously involves the loss of observations for analysis. This points to the issue of the size of the sample available for the analysis of substantive issues, which is now taken up.

Sample sizes

Table 4.2 contains the (unweighted) sample sizes of respondents to the fourth interviews conducted in the early 90s and late 90s surveys for the South Australian policy and comparison groups. The differences in cell sizes reflect the relative sample sizes of two collections and the number of months of the year accounted for by each of the groups (five, three and four respectively).

In aggregate, the sample sizes in Table 4.2 are adequate, but not particularly large, especially for the early 90s cohort. In reality, the sample sizes may well constrain the type of analyses that can be undertaken. For example, it is possible to analyse the impact of the policy change on Year 12 completion, since the denominator of that proportion is the total sample in any cell. More detailed analysis of the impact of the policy on the employment outcomes by either gender or by Year 12 completion becomes more problematic as the cell sizes shrink, especially if non-completers are classified by 'early school leaving' status (Year 10 or earlier). Such problems are exemplified in the application analysed in the body of the paper. The greater the disaggregation required for the analysis, the more problematic is the sample size. What this means is that the adequacy of the sample size for analysis in the way proposed here for specific research applications needs to be established in each application.

The analysis contained in the body of the paper indicates that this does not mean that the policy change cannot be used to identify the causal effect of additional education on particular outcomes of interest in cases where only a sub-set of the population is analysed. Such problems can be analysed successfully with regression techniques.

Table 4.2 Sample sizes in the fourth survey year by analysis group

	Oct- Feb birth group	Sept qtr birth group	Comparison group	Total
Early 90s (YIT 75)	183	127	163	473
Late 90s (LSAY 95)	439	289	393	1121
Total	622	416	556	1594

Note: based on unweighted observations.

Analysis of the characteristics of individuals

The characteristics of the various birth groups are compared in this section. Four different tests are described. The results of all of the tests generally support the use of difference in differences techniques to analyse the impact of the *Early Years of School* policy change to identify the impact of education on outcomes of interest.

These four tests involve:

- A comparison of changes in the mean values of individual background characteristics across surveys between the policy and comparison groups;
- A regression-based multivariate test of whether the changes in the mean values of background characteristics across surveys between the policy and comparison groups are comparable;
- A regression-based multivariate test of whether the mean values of individual background characteristics between the policy and comparison groups are equal in both surveys; and
- A multivariate test of whether the changes in the mean values of individual background characteristics across surveys between the policy and comparison groups in South Australia are similar to the changes in mean values of those characteristics for individuals in other states.

The outcomes of the first test appear in Tables 4.3 and 4.4. Table 4.3 contains the comparison between the October – February birth group and the comparison group of the change in the mean values between the surveys for a set of background characteristics. These variables are described in Appendix 5. Table 4.4 contains the same comparison between the September quarter birth group and the comparison group, as well as the further differencing estimate involving the comparison with individuals from other jurisdictions to remove any age-based survey design effects. The first column of both tables shows the change in the mean value of the relevant variable between the two surveys for the policy group. The second shows the change in the mean value of the variable for the comparison group. The third column shows the difference in the changes in the mean values (column one minus column two) and the fourth column its significance (values over 1.96 mean that the possibility that the changes in the mean values for that variable are the same between the groups should be rejected). The fifth column of Table 4.4 shows the comparison with the mean changes between groups in the other jurisdictions and the sixth column the significance of that triple difference.

By way of interpretation, from the first line of Table 4.3 the proportion of individuals in the October – February birth group whose father had a degree increased by 9.4 percentage points between the two surveys. In the comparison group, this proportion increased by 8.7 percentage points, which was 0.8 percentage points lower that the policy group. The ‘t’ statistic for this difference between the groups was 0.15, which means the difference is not significantly different from zero.

The key feature of both tables is that very few of the background variables (those above the first horizontal line) are individually significant. The joint test described below suggests that the changes in the mean values of the variables are similar in the two birth groups affected by the policy change to those of the comparison group.

The remaining variables in the tables include the age of individuals and other variables that might be both outcomes of the policy change and reflect aspects of the 'character' of individuals. These include the average number of years of schooling individuals in the group eventually undertook and changes in their numeracy and literacy skill levels.

From Table 4.3, individuals in the October – February birth group were slightly younger in the second survey than the first (not inconsistent with Figure 2.3), but they completed increased years of schooling, were older when they left and enjoyed greater improvement in their numeracy and literacy levels than the comparison group. On the face of it, these 'outcomes' suggest the policy change had positive educational outcomes for this group.

The picture from Table 4.4 is somewhat different. The average age of individuals in the September quarter birth group increased by more than that of the comparison group (as expected, see Figure 2.3), as did their completed schooling. Their relative literacy and numeracy outcomes did not increase significantly, however. Compared to the experience of like groups in other states, the age of individuals in the September quarter birth group did not increase as much but their relative numeracy and literacy performance was comparable. This latter comparison appears somewhat unsatisfactory, however. From Figure 2.2, almost all individuals born in the September quarter were a year older in the late 90s cohort than the early 90s one in the comparison jurisdictions, while that was true of only half the individuals born in South Australia in that quarter. Hence the magnitude of the age-based survey design effects are much more pronounced in the comparison jurisdictions than in South Australia, so that the third 'differencing' in column five of Table 4.4 may impose a larger estimate of the survey design effect for South Australia than is warranted. It does appear though, that the numeracy and literacy performance of the September quarter birth group did not increase relative to the South Australian comparison group like that of the October – February birth group.

This may reflect a problem for the difference in differences estimator and its interpretation in this case. Respondents in the September quarter birth group are split fairly evenly between being aged fourteen and fifteen in the late 90s cohort. In the early 90s cohort, almost all of the fourteen year olds in that policy group were in Year 9. That is, while there was relatively little discretion in the Year of schooling fourteen year olds were in prior to the policy change, after it there was considerable discretion. In terms of the policy set out in Appendix 1, this decision was to depend on 'individual development, competence and maturity'. That is, students born in the same months were allocated to two school levels, with the selection based on ability and maturity considerations. If those entering the higher grade were more able, the allocation of the less able older children to the lower class may mask the beneficial effects of their slower progression through school and increased maturity on their numeracy and literacy outcomes.

Table 4.3 Change in mean values of background characteristics between the surveys: October – February birth group and the comparison group (fourth survey year)

	Oct-Feb birth group	Comparison group	Difference in difference	T test
Father with degree	0.095	0.087	0.008	0.15
Mother with degree	0.124	0.096	0.028	0.55
Both parents with degrees	0.071	0.055	0.016	0.41
Father's occupation - ASCO manager	0.044	-0.007	0.051	0.79
Father's occ - ASCO professional	0.057	0.035	0.022	0.43
Father's occ - ASCO associate professional	-0.012	-0.027	0.015	0.48
Father's occ - ASCO trade	-0.067	0.014	-0.081	-1.38
Father's occ - ASCO clerk	-0.058	-0.012	-0.045	-1.40
Father's occ - ASCO sales and personal service	0.023	0.002	0.020	0.54
Father's occ - ASCO machine operator	-0.050	0.005	-0.055	-1.36
Metropolitan region	-0.053	-0.126	0.074	0.99
Catholic school	0.092	0.031	0.060	1.07
Independent school	0.030	-0.004	0.034	0.67
Male	-0.071	-0.133	0.062	0.81
Planned to complete Year 12	0.033	-0.065	0.098	1.56
Self-assessed school performance-well above average	0.085	0.047	0.037	0.72
Self-assessed school performance - above average	0.030	-0.039	0.069	0.94
Self-assessed school performance - average	-0.013	0.050	-0.063	-0.82
Father born overseas, English-speaking	-0.012	0.025	-0.037	-0.74
Father born overseas, non-English-speaking	-0.023	-0.005	-0.018	-0.30
Mother born overseas, English-speaking	-0.003	0.025	-0.027	-0.56
Mother born overseas, non-English-speaking	-0.056	-0.023	-0.033	-0.63
Father's occupation - ANU 3	0.110	0.073	0.038	1.44
Age in first survey	-0.008	0.111	-0.119	-4.68
Years of schooling completed	0.535	-0.141	0.676	5.84
Grade of school completed	0.178	0.086	0.092	1.07
Age left school	0.446	-0.088	0.533	4.94
Numeracy (standardised)	0.341	-0.004	0.344	3.01
Literacy (standardised)	0.248	0.036	0.212	1.84

Table 4.4 Change in mean values of background characteristics between the surveys: September quarter birth group and the comparison group (fourth survey year)

	Sept qtr group	Compari-son group	Difference in difference	T test	Diff. in diff. in differences	T test
Father with degree	0.033	0.087	-0.054	-0.87	-0.081	-1.13
Mother with degree	0.049	0.096	-0.047	-0.81	-0.026	-0.37
Both parents with degrees	0.025	0.055	-0.030	-0.64	-0.014	-0.29
Father's occupation - ASCO manager	-0.051	-0.007	-0.044	-0.61	-0.101	-1.06
Father's occ - ASCO professional	-0.002	0.035	-0.037	-0.64	-0.031	-0.44
Father's occ - ASCO associate professional	-0.028	-0.027	-0.002	-0.04	-0.049	-1.04
Father's occ - ASCO trade	-0.058	0.014	-0.072	-1.10	-0.024	-0.27
Father's occ - ASCO clerk	-0.044	-0.012	-0.032	-0.95	-0.017	-0.36
Father's occ - ASCO sales and personal service	0.023	0.002	0.021	0.48	0.010	0.18
Father's occ - ASCO machine operator	0.035	0.005	0.030	0.74	0.022	0.30
Metropolitan region	-0.188	-0.126	-0.062	-0.75	0.030	0.26
Catholic school	-0.039	0.031	-0.071	-1.13	-0.088	-1.12
Independent school	-0.005	-0.004	-0.001	-0.01	-0.014	-0.21
Male	-0.067	-0.133	0.067	0.78	0.036	0.30
Planned to complete Year 12	-0.116	-0.065	-0.051	-0.74	0.029	0.28
Self-assessed school performance - well above average	0.068	0.047	0.021	0.34	-0.014	-0.21
Self-assessed school performance - above average	-0.123	-0.039	-0.084	-1.03	-0.046	-0.42
Self-assessed school performance - average	0.072	0.050	0.022	0.26	0.011	0.09
Father born overseas, English-speaking	-0.016	0.025	-0.041	-0.72	-0.052	-0.70
Father born overseas, non-English-speaking	-0.061	-0.005	-0.056	-0.85	-0.077	-0.97
Mother born overseas, English-speaking	-0.043	0.025	-0.068	-1.23	-0.087	-1.18
Mother born overseas, non-English-speaking	-0.043	-0.023	-0.020	-0.34	-0.050	-0.77
Father's occupation - ANU 3	0.025	0.073	-0.047	-1.61	-0.071	-2.26
Age in first survey	0.507	0.111	0.396	9.96	-0.367	-8.86
Years of schooling completed	0.317	-0.141	0.459	3.69	0.359	2.72
Grade of school completed	0.080	0.086	-0.006	-0.06	0.012	0.11
Age left school	0.225	-0.088	0.312	2.58	0.285	2.22
Numeracy (standardised)	0.044	-0.004	0.047	0.37	-0.137	-1.00
Literacy (standardised)	0.039	0.036	0.003	0.02	-0.130	-0.98

The numeracy and literacy scores of individuals (which reflect their ability, among other factors) in this group are consistent with the operation of some form of ability-based selection process. In the early 90s cohort, the fourteen year olds in this group in Year 9 had average standardised numeracy and literacy scores of 0.04 and 0.14 respectively. In the late 90s cohort, the average scores among Year 9 students in this group who were fourteen year old were 0.16 and 0.14. In contrast, the older fifteen-year-old Year 9 students in this group had average standardised numeracy and literacy scores of -0.20 and –0.19 respectively. That choices were being made about the grade cohorts those individuals entered makes any 'causal' interpretation of any variation in the education policy group two acquired because of the *Early Years of School* policy on other outcomes more difficult to sustain.

The second test reported here is one of whether the changes in the mean values of individual background characteristics across surveys between the groups affected by the policy change and the comparison group are comparable. It is a test based on the output of a logit regression equation. The same set of explanatory variables as those used in the top section of Tables 3.3 and 3.4 are used in the equation. The dependent variable is a binary one that took the value one if individuals were in the late 90s cohort and zero if they were in early 90s cohort. Since the average values of these explanatory variables do change between the surveys, the equation has some explanatory power. However, the test undertaken focuses on whether a set of interactions between these variables and the respective policy group indicators have any predictive power. The outcomes of separate (for each policy group) joint tests of the significance of those interactions (plus the relevant birth group indicator variables) suggest that they do not. These results are reported in the top section of Table 4.5. The test involves a likelihood ratio test of the change in the log likelihood function when the interaction terms were added to the logit equations. For both groups affected by the policy change, the test statistic is less than the critical value. This suggests that the test cannot reject the possibility that the changes in the background characteristics of individuals between the surveys were the same for the two birth groups affected by the policy change as the South Australian comparison group.

Table 4.5 Joint tests of the significance of the explanatory variables in regression equations

	October – February birth group versus the comparison group	September quarter birth group versus the comparison group
First regression likelihood ratio test		
Test statistic	32.9	21.6
Degrees of freedom	24	24
Probability value	0.893	0.395
Critical value	36.4	36.4
Second regression likelihood ratio test		
Test statistic	51.1	51.3
Degrees of freedom	47	47
Probability value	0.684	0.691
Critical value	63.7	63.7

The third test used here is similar to the one just described. It is a test of whether the mean values of individual background characteristics between the policy and comparison groups are equal in both surveys. It is also based on the output of a logit regression equation. The explanatory variables used in the equations for the last test were used in this test. This time the dependent variable was whether individuals belonged to the policy or the comparison groups and the equation was estimated separately for each policy group contrasted with the comparison group. This time, a set of interactions was constructed between an indicator for the LSAY 95 survey and the explanatory variables. The variables and the interactions could not explain whether individuals were in the policy or the comparison groups. Essentially, there was no regression. The results for this test are presented in the lower half of Table 4.5. Once more, the test is a likelihood ratio test and for both policy groups the test statistics was lower than the critical value. This suggests that the mean values for the background variables were similar in the policy and comparison groups across both surveys. This outcome is not surprising. It simply means that birth date of individuals is not associated with their background characteristics.

The final test described here compares the changes in the mean values of individual background characteristics across surveys between the policy and comparison groups in South Australia with the changes in mean values of those characteristics for individuals in other states. The change in the average values between the surveys for the set of background characteristics used in the two previous tests were calculated for individuals born in each quarter in each jurisdiction – thirty-two sets of changes in mean values (eight jurisdictions by four quarters). The *Euclidean distance* between each set of these changes and the remaining thirty-one was then calculated.[28] The average distance measures between each set of mean changes and the others were then calculated and the distribution of these average distance measures used to identify whether the estimated changes for the South Australian groups differed from those of the other jurisdictions. The 't' statistics for the March, June (roughly the comparison group), September (policy group two) and December quarters (the bulk of policy group one) were 0.15, -1.26, -0.77 and –0.28 respectively. These values suggest that the changes in background characteristics in South Australia were comparable to the changes that occurred in the other Australian jurisdictions between the surveys.

Taken together, the tests suggest that the changes in the observed characteristics between the policy groups and the comparison groups in South Australia were similar between the two surveys. Therefore, the results support the use of the difference in difference estimator to identify the effect of education on phenomena it is considered to influence. This support is stronger for the October – February birth group than for the September quarter group, where a selection process of individuals into different grade cohorts after the policy changes appears to have taken place.

Scope for natural experiment analysis of full-time employment outcomes

The tests and analysis reported above are now repeated for the specific application dealt with in the body of the report to ascertain whether the data support the proposed analysis.

Sample sizes for this application

Table 4.6 contains the (unweighted) sample sizes of respondents observed in their first year after completing their secondary schooling in the early and late 90s surveys for the South Australian policy and comparison groups. The table excludes all full-time post-secondary students. The differences in cell sizes reflect the relative sample sizes of the two surveys and the number of months of the year accounted for by each of the groups (five, three and four respectively).

The sample sizes reported in Table 4.6 are small for the early 90s cohort, but adequate for the late 90s cohort. More detailed analyses of the impact of the policy on the employment outcomes by either gender or by Year 12 completion are clearly out of the question with the available data. Even in the presence of large estimated effects, measurement precision may be quite low, with consequent large standard errors. What this means is that it may be necessary to rely on regression-based techniques to identify the causal effect of additional education on the full-time employment outcomes of this group.

Table 4.6 Sample sizes in the year after individuals who completed their schooling and did not proceed to post-school education by analysis group

	Oct- Feb birth group	Sept qtr birth group	Comparison group	Total
Early 90s (YIT 75)	86	45	69	200
Late 90s (LSAY 95)	231	132	212	575
Total	317	177	281	775

Note: based on unweighted observations.

Analysis of the background characteristics of individuals

The characteristics of education leavers among the two birth groups affected by the policy change and the comparison group are compared in this section. The first three tests described earlier in this Appendix are used again to compare the various groups. The outcomes of the tests are consistent with there being only modest differences in the background characteristics of individuals in the policy and comparison groups and, therefore, support the use of difference in differences techniques to analyse the impact of the *Early Years of School* policy change on individual full-time employment outcomes. These results are similar to those reported for the entire South Australian policy and comparison groups reported earlier.

The outcomes of the first test for the October – February birth group and the comparison group appear in Table 4.7 and the test for the September quarter birth group and the comparison group in Table 4.8. The tables contain the changes in the mean values for the various groups between the surveys for a set of background characteristics. The formats of the tables are identical to Tables 4.3 and 4.4. Once more, the key feature of the tables is that very few of the background variables (those above the first horizontal line) are individually significant.

Table 4.7 Change in mean values of background characteristics between the surveys: October – February birth group and the comparison group individuals who did not proceed to full-time study after school

	Oct-Feb birth group	Comparison group	Difference in difference	T test
Father with degree	0.000	0.034	-0.034	-0.51
Mother with degree	0.062	0.068	-0.006	-0.09
Both parents with degrees	0.034	0.028	0.006	0.13
Father's occupation - ASCO manager	-0.032	-0.040	0.008	0.09
Father's occ - ASCO professional	0.007	0.042	-0.035	-0.53
Father's occ - ASCO associate professional	-0.040	-0.050	0.010	0.22
Father's occ - ASCO trade	0.030	0.087	-0.057	-0.68
Father's occ - ASCO clerk	-0.048	-0.044	-0.004	-0.09
Father's occ - ASCO sales and personal service	0.037	0.031	0.006	0.11
Father's occ - ASCO machine operator	-0.081	-0.027	-0.055	-0.78
Metropolitan region	-0.079	-0.139	0.060	0.54
Catholic school	0.060	0.010	0.050	0.68
Independent school	-0.098	-0.018	-0.080	0.48
Male	-0.016	-0.173	0.157	1.40
Planned to complete Year 12	0.055	-0.096	0.151	1.49
Self-assessed school performance- well above average	0.102	0.048	0.055	0.86
Self-assessed school performance - above average	0.091	-0.053	0.143	1.38
Self-assessed school performance - average	-0.049	0.059	-0.108	-0.96
Father born overseas, English-speaking	-0.030	0.026	-0.056	-0.78
Father born overseas, non-English-speaking	-0.035	-0.026	-0.009	-0.12
Mother born overseas, English-speaking	-0.027	0.020	-0.047	-0.65
Mother born overseas, non-English-speaking	-0.056	-0.042	-0.014	-0.23
Father's occupation - ANU 3	0.055	0.070	-0.015	-0.44
Age in first survey	0.009	0.113	-0.105	-3.78
Years of schooling completed	0.878	0.080	0.798	4.50
Grade of school completed	0.168	0.245	-0.077	-0.56
Age left school	0.707	0.020	0.687	4.64
Numeracy (standardised)	0.235	0.121	0.114	0.71
Literacy (standardised)	-0.067	0.035	-0.102	-0.63

Table 4.8 Change in mean values of background characteristics between the surveys: September quarter birth group and comparison group individuals who did not proceed to full-time study after school

	Sept qtr group	Compari-son group	Difference in difference	T test	Diff. in diff. in differences	T test
Father with degree	-0.010	0.034	-0.044	-0.74	-0.094	-1.18
Mother with degree	0.025	0.068	-0.043	-0.71	0.004	0.06
Both parents with degrees	0.004	0.028	-0.024	-0.57	-0.007	-0.13
Father's occupation - ASCO manager	-0.100	-0.040	-0.059	-0.74	-0.125	-1.19
Father's occ - ASCO professional	0.047	0.042	0.004	0.08	0.037	0.47
Father's occ - ASCO associate professional	0.027	-0.050	0.076	2.07	-0.014	-0.27
Father's occ - ASCO trade	-0.137	0.087	-0.225	-3.04	-0.261	-2.62
Father's occ - ASCO clerk	-0.054	-0.044	-0.009	-0.27	-0.024	-0.47
Father's occ - ASCO sales and personal service	0.000	0.031	-0.031	-0.58	-0.093	-1.38
Father's occ - ASCO machine operator	0.041	-0.027	0.068	1.06	0.131	1.58
Metropolitan region	-0.341	-0.139	-0.203	-2.11	-0.293	-2.28
Catholic school	-0.163	0.010	-0.173	-2.50	-0.213	-2.39
Independent school	-0.022	0.007	-0.030	-0.60	-0.071	-0.99
Male	0.006	-0.173	0.179	1.85	0.294	2.27
Planned to complete Year 12	0.007	-0.096	0.103	1.17	0.100	0.83
Self-assessed school performance - well above average	0.033	0.048	-0.015	-0.26	-0.053	-0.74
Self-assessed school performance - above average	-0.001	-0.053	0.052	0.56	0.067	0.56
Self-assessed school performance - average	-0.052	0.059	-0.111	-1.15	-0.141	-1.09
Father born overseas, English-speaking	-0.010	0.026	-0.036	-0.60	0.013	0.16
Father born overseas, non-English-speaking	-0.119	-0.026	-0.094	-1.44	-0.191	-2.20
Mother born overseas, English-speaking	-0.026	0.020	-0.046	-0.80	-0.057	-0.71
Mother born overseas, non-English-speaking	-0.016	-0.042	0.027	0.49	-0.112	-1.57
Father's occupation - ANU 3	0.019	0.070	-0.051	-1.34	-0.06441	-1.58
Age in first survey	0.590	0.112	0.478	8.22	-0.270	-4.48
Years of schooling completed	0.574	0.080	0.494	2.50	0.446	2.12
Grade of school completed	-0.016	0.245	-0.262	-1.77	-0.165	-1.03
Age left school	0.509	0.020	0.489	2.74	0.549	2.87
Numeracy (standardised)	0.159	0.121	0.038	0.21	0.035	0.18
Literacy (standardised)	-0.235	0.035	-0.270	-1.47	-0.305	-1.57

From Table 4.7, individuals in the October – February birth group who did not go on to full-time study after leaving school increased their years of schooling between the first and second surveys. This increase was greater than the increase in years of schooling completed by the comparison group. Unlike the entire group, the increase in literacy and numeracy achievement between the surveys was less apparent among those who did not proceed to full-time post-school education after school.

In Table 4.8, more of the changes in background characteristics between surveys are significantly different between the September quarter policy group and the comparison group. The estimated changes in completed schooling and age left school are also smaller than between the October – February birth group and the comparison group.

The outcomes of the second and third tests are reported in Table 4.9. These repeat the tests described earlier and reported in Table 4.5. The test statistics for the October – February birth group were not statistically significant (they were lower than the critical values). This implies that: the changes in the background characteristics of individuals between the surveys did not differ between that policy group and the comparison group; and that the mean values for the background variables were similar in the October – February birth group and the comparison group across both surveys. In contrast, the statistics for both tests were statistically significant for the comparisons between the September quarter birth group and the comparison group. This suggests that regression analysis, which can control for these differences (and may act to reduce the residual variance of the estimated policy effect – Meyer 1995), may provide a better approach to the estimation of the effect of the policy change for this group.

While the tests for differences in the characteristics of individuals from the October – February birth group from those of the comparison group may not have been rejected, there may still be problems for the difference in differences estimator for this specific application. This occurs because the October – February birth group in the early 90s cohort was predominantly in Year 10 when first interviewed in 1989. In contrast, the March – June comparison group was in Year 9. Hence, for completion of a common grade, the October – February birth group left school one calendar year earlier than the comparison group. With the timing of the recession of the early 1990s, on average members from the Year 10 grade cohort faced a state unemployment rate that was about 1.5 percentage points lower when they left school (one year earlier) than the comparison group. Therefore, members of the October – February birth group in the first cohort may have faced better initial labour market conditions on average than those in the comparison group. In the late 90s cohort, members of the October – February birth group and the comparison group were in the same grade when first interviewed. Hence any comparison of changes in the employment outcomes of the two groups between the cohorts may confound the impact of the policy change with one arising from differential labour market conditions.

Conclusions on the scope for 'natural experiment' analysis

Taken together the tests suggest that, for general applications, the differences in the background characteristics of individuals in the policy and comparison groups between the surveys are modest and should not prevent analysis that utilises difference in difference techniques. This support is stronger for the October – February birth group than for the September quarter group, where a selection process of individuals into different grade cohorts after the policy changes appears to have taken place. For that

group, regression techniques that can deal with differences in the observed characteristics of individuals may be preferable.

For the specific application in this paper, analysis of the initial labour market outcomes of individuals who do not proceed to full-time post-school education, there appear to be problems with the use of difference in difference estimators. For the September quarter group, there are significant differences in the background characteristics of individuals compared to the control group. For the October – February birth group, the difference in labour market conditions between the two surveys may have differed from that of the comparison group, which would confound difference in difference estimators. A better approach, therefore, is to rely on regression based analysis.

Table 4.9 Joint tests of the significance of the explanatory variables in regression equations for individuals who did not proceed to full-time study after school

	October – February birth group versus the comparison group	September quarter birth group versus the comparison group
First regression likelihood ratio test		
Test statistic	31.6	38.9
Degrees of freedom	24	24
Probability value	0.863	0.972
Critical value	36.6	36.4
Second regression likelihood ratio test		
Test statistic	53.6	74.4
Degrees of freedom	47	47
Probability value	0.763	0.993
Critical value	63.7	63.7

APPENDIX 5: DESCRIPTION OF THE SAMPLES AND VARIABLE DEFINITIONS

The samples analysed

The data on individuals are drawn from two samples, the *Youth in Transition* 1975 birth cohort and the *Longitudinal Surveys of Australian Youth* 1995 Year 9 cohort. The YIT survey commenced with classroom-based literacy and numeracy tests in 1989 when the cohort was aged fourteen. Individuals also completed a questionnaire that collected background information at that time. They were followed by mail survey in subsequent years. The LSAY 95 cohort also commenced with classroom-based literacy and numeracy tests undertaken in 1995 when the modal age of the cohort was fourteen years. Individuals also completed a questionnaire that collected family background information. Individuals were initially followed by mail survey, but the data collection methodology moved to phone surveys from 1997.

Individuals in those surveys were excluded from the analysis undertaken in this report for a variety of reasons. In the analysis contained in the first part of Appendix 4, anyone who did not complete a questionnaire in the fourth survey year (1992 or 1998) was excluded. In addition, individuals who did not provide information on critical variables such as their date of birth were excluded.

For the analysis contained in the body of the paper, individuals had to have responded to the survey in the year after they had left school. If not, they were excluded from the analysis, as were those individuals who reported that they had left school but were studying full-time at another educational institution when surveyed. Individuals born overseas were also excluded, because they may not have commenced their schooling in Australia. The year of arrival in Australia of those born overseas was not collected in the YIT 75 survey. Once more, individuals who did not provide information on variables used in the analysis such as their date of birth or family background were excluded.

Variable descriptions and summary statistics

The following table, Table 5.1 contains descriptions of the variables used in this paper and summary statistics for the sample of individuals from South Australia analysed in assessing the impact of schooling on full-time employment outcomes. Summary statistics are provided for the two cohorts separately. Many of the variables are indicator variables that show the proportion of the cohort with the specific characteristic. For example, 8.7 per cent of the 'early 90s' cohort had a father (but not a mother) who had completed a university degree. One variable requires description. The self–assessed school performance variable was constructed as follows: individual assessments were assigned a value representing the proportion of the cohort they considered their schoolwork to be superior to: for example, those who indicated they were 'well above average' were assigned a value reflecting the proportion of students who assessed their work as less than 'well above average' (in fact, the mid-point between 1 and this proportion). This variable was then regressed on individuals' actual numeracy and literacy scores and the residual of that regression was used to reflect their self-assessment of their school performance. The resulting estimate is uncorrelated with actual achievement and probably reflects a dimension of self-confidence or self-efficacy.

Table 5.1 Variable descriptions and summary statistics

Variable	Definition	Early 90s		Late 90s	
		Mean	Std Dev.	Mean	Std Dev.
Father with degree[a]	Father had completed a university degree.	0.087		0.097	
Mother with degree[a]	Mother had completed a university degree.	0.055		0.111	
Both parents with degrees[a]	Both parents had completed a university degree.	0.025		0.050	
Father's occ - ASCO manager[a]	Father's occupation was as a manager in the ABS' First Edition ASCO classification.	0.245		0.196	
Father's occ - ASCO professional[a]	Father's occupation was as a professional in the First Edition ASCO classification.	0.071		0.100	
Father's occ - ASCO associate professional[a]	Father's occupation was as an associate professional in the First Edition ASCO classification.	0.067		0.039	
Father's occ - ASCO trade[a]	Father's occupation was as a tradesman in the First Edition ASCO classification.	0.176		0.184	
Father's occ - ASCO clerk[a]	Father's occupation was as a clerk in the First Edition ASCO classification.	0.080		0.028	
Father's occ - ASCO sales and personal service[a]	Father's occupation was as a sales and personal service worker in the First Edition ASCO classification.	0.050		0.076	
Father's occ - ASCO machine operator[a]	Father's occupation was as a machine operator in the First Edition ASCO classification.	0.124		0.091	
Father's occ – ASCO labourer[a]	Father's occupation was as a labourer in the First Edition ASCO classification.	0.109		0.134	
Father's occ – residual category[a]	Father's occupation not reported – includes those not employed or deceased	0.023		0.045	
Father's occupation – ANU 3[a]	Father's occupation based on the ANU 3 occupational scale, with values assigned on the basis of First edition ASCO minor group occupations.	0.317	0.21	0.359	0.20
Metropolitan region[b]	Attended a school in a metropolitan region when first surveyed.	0.673		0.514	
Catholic school[b]	Attended a Catholic school when first surveyed.	0.128		0.120	
Independent school[b]	Attended an Independent school when first surveyed.	0.086		0.099	
Self-assessed school performance – well above average[c]	Individual indicated that compared to their peers in their class they were 'well above average'.	0.045		0.113	
Self-assessed school performance – above average[c]	Individual indicated that compared to others in their class they were 'above average'.	0.279		0.300	
Self-assessed school performance – average[c]	Individual indicated that compared to others in their class they were 'average'.	0.530		0.514	

Table 5.1 Variable descriptions and summary statistics (continued)

Variable	Definition	Early 90s		Late 90s	
		Mean	Std Dev.	Mean	Std Dev.
Self-assessed school performance	Residual from equation explaining self-assessment with literacy and numeracy scores.	-0.054		0.010	
Father born o/s, English-speaking [a]	Father born overseas in a predominantly English-speaking country	0.112		0.105	
Father born o/s, non-English-speaking [a]	Father born overseas in a non-English-speaking country	0.159		0.107	
Mother born o/s, English-speaking [a]	Mother born overseas in a predominantly English-speaking country	0.126		0.115	
Mother born o/s, non-English-speaking [a]	Mother born overseas in a non-English-speaking country	0.102		0.061	
Male [b]		0.550		0.483	
Age in first survey [b]	Age at October 1 in the year when first surveyed.	14.5	0.29	14.7	0.42
Age left school	Individuals age in years and months when they left school	16.9	1.05	17.3	0.97
Years of schooling completed	Full years of schooling completed by individuals. Individuals were assigned partial years for Reception in accordance with their birth date. Where individuals left school during the course of the school year, that schooling was not included in this variable.	11.9	1.25	12.4	0.89
Grade completed at school		11.4	0.97	11.4	0.86
Years of Reception	Years of pre-Year 1 studies.	0.2	0.22	0.8	0.33
Numeracy (standardised) [c]	Standardised numeracy score from test undertaken in the first survey year.	-0.337	1.06	-0.186	0.93
Literacy (standardised) [c]	Standardised literacy score from test undertaken in the first survey year.	-0.164	0.96	-0.244	1.06
Achievement (standardised) [c]	Standardised average of literacy and numeracy scores from tests undertaken in the first survey year.	-0.283	1.01	-0.235	0.97
Employed full-time	Individual worked no less than 35 hours per week in their main job in October of the relevant year in YIT 75 or at the time of the survey in LSAY 95.	0.370		0.535	
Unemployment rate	South Australian unemployment rate in the year following individual's departure from school, taken from the ABS *Labour Force Survey*.	10.6	0.95	8.6	0.64
Number of observations		200		575	

Notes:

a. based on responses provided in 1991 in YIT 75 and 1995 in LSAY 95

b. based on school and individual characteristics in 1989 in YIT 75 and 1995 in LSAY 95

c. based on responses provided by individuals in 1990 in YIT 75 and 1995 in LSAY 95

APPENDIX 6: TABLES OF RESULTS

Table 6.1 Least squares education equation results

	Years of schooling		Age left school		Grade completed	
	Coef.	Coef. / Std. Error	Coef.	Coef. / Std. Error	Coef.	Coef ./ Std. Error
Constant	6.680	5.09	19.288	25.66	11.162	20.62
Male	0.033	0.55	0.019	0.28	-0.109	-1.81
Metropolitan region	0.033	0.49	0.040	0.54	-0.146	-2.18
Self-assessed school performance	0.380	3.07	0.471	3.44	0.580	4.43
Father born o/s, English-speaking	-0.052	-0.52	-0.027	-0.24	0.009	0.09
Father born o/s, non-English-speaking	0.040	0.37	0.019	0.16	0.117	1.13
School achievement	0.081	2.34	0.084	2.27	0.108	3.26
Unemployment rate	-0.148	-4.06	-0.144	-3.54	-0.098	-2.61
Father's occ - ASCO manager	0.386	4.31	0.388	3.90	0.411	4.52
Father's occ - ASCO professional	0.235	1.75	0.226	1.65	0.286	2.26
Father's occ - ASCO associate professional	0.219	1.15	0.169	0.85	0.294	1.66
Father's occ - ASCO trade	0.239	2.61	0.176	1.68	0.288	2.99
Father's occ - ASCO clerk	0.206	1.50	0.244	1.63	0.395	2.45
Father's occ - ASCO sales and personal service	0.262	1.97	0.217	1.47	0.350	2.92
Father's occ - ASCO machine operator	0.487	4.03	0.523	3.72	0.509	4.17
Catholic school	0.247	2.36	0.289	2.61	0.291	3.33
Independent school	0.113	1.18	0.160	1.59	0.227	2.40
Grade turned fifteen years	0.600	4.77	-0.130	-2.09	0.103	2.74
Years of Reception	1.086	7.74	0.409	3.18	-0.083	-0.71
Observations		775		775		775
Deg. Fr.		756		756		756
Mean Dep. Var.		12.2		17.2		11.4
Std Dev		1.0		1.0		0.9
Res. Sum Sq.		551.6		663.5		530.0
Std Dev		0.9		0.9		0.8
R-squared		0.34		0.16		0.15
F[18, 474]		21.9		8.1		7.2
Prob value		0.0		0.0		0.0
Breusch- Pagan statistic		131.5		104.6		137.4

Table 6.2 Full-time employment probit results: first year after leaving school (for those not in further full-time education and training)

Variables	Coefficient	Coefficient / Std. Error	Marginal effect	Marginal eff / Std. Error
Constant	2.887	2.83		
Male	0.412	4.30	0.151	4.36
Metropolitan region	-0.192	-1.94	-0.071	-1.94
Self-assessed school performance	0.518	2.62	0.188	2.62
Father born o/s, English-speaking	-0.374	-2.36	-0.134	-2.46
Father born o/s, non-English-speaking	-0.385	-2.58	-0.139	-2.69
School achievement	-0.061	-1.24	-0.022	-1.24
Unemployment rate	-0.160	-3.43	-0.058	-3.43
Father's occ - ASCO manager	0.421	2.91	0.153	2.97
Father's occ - ASCO not a labourer	0.265	2.26	0.096	2.27
Years of Reception	0.299	1.99	0.109	1.99
Grade completed	0.156	1.57	0.057	1.58
Age left school	-0.214	-2.42	-0.078	-2.42
Number of observations		775		
Log likelihood function		-492.7		
Restricted log likelihood		-536.7		
Chi squared		87.9		
Degrees of freedom		12		
Prob[Chi Sqd > value]		0		
McFadden's R^2		0.07		
		Predicted (0.5 cut-off)		
Actual	Not Full-time	Full-time	Total	% correct
Not employed full-time	245	145	390	62.8
Employed full-time	138	247	385	64.2
Total	383	392	775	63.5

Table 6.3 Full-time employment probit results: second year after leaving school (for those not in further full-time education and training)

Variables	Coefficient	Coefficient / Std. Error	Marginal effect	Marginal eff / Std. Error
Constant	2.887	1.88		
Male	0.451	3.84	0.160	3.90
Metropolitan region	-0.266	-2.20	-0.094	-2.22
Self-assessed school performance	0.581	2.43	0.204	2.43
Father born o/s, English-speaking	-0.264	-1.44	-0.095	-1.42
Father born o/s, non-English-speaking	-0.329	-1.82	-0.119	-1.80
School achievement	-0.034	-0.54	-0.012	-0.54
Unemployment rate	-0.106	-1.74	-0.037	-1.74
Father's occ - ASCO manager	0.450	2.51	0.151	2.69
Father's occ - ASCO not a labourer	0.465	3.24	0.163	3.28
Years of Reception	0.136	0.71	0.048	0.71
Grade completed	0.070	0.57	0.025	0.58
Age left school	-0.170	-1.43	-0.060	-1.43
Number of observations		536		
Log likelihood function		-330.0		
Restricted log likelihood		-359.5		
Chi squared		59.1		
Degrees of freedom		12		
Prob[Chi Sqd > value]		0		
McFadden's R^2		0.08		
		Predicted (0.6 cut-off)		
Actual	Not Full-time	Full-time	Total	% correct
Not employed full-time	125	76	201	62.2
Employed full-time	99	236	335	70.4
Total	224	312	536	67.4

ENDNOTES

[1] The means of these variables conditional on date of birth for the two cohorts were estimated using the *ksm* smoothing program of STATA, utilising the *lowess* option and a bandwidth of 0.3. Differences in the conditional means were then estimated, which appear in Figure 1. Variations in the bandwidth did not alter the shape of the lines from those presented in Figure 1.

[2] It has a different scale from the other three lines in the diagram, since it can be no greater than one.

[3] Source: Australian Industrial Relations Commission (1998:19, 20). These various proportions are unchanged in the relevant current industrial awards.

[4] In fact, in the Rummery *et al*. (1999) study, the ranking was of the error terms of the first stage schooling equation of individuals in different Australian states. This identification strategy requires that the distribution of ability be similar across jurisdictions (or any other way the population might be partitioned) but that the distributions of schooling should differ.

[5] Ashenfelter and Rouse (1998) report that schooling differences between identical twins are uncorrelated with birth order and with their spouses' education. Behrman *et al*. (1994) report a positive correlation between schooling differences between identical twins and with differences in their spouses' education. Behrman *et al*. (1994) also report a strong positive correlation between the birth weights of 'identical' twins and subsequent differences in schooling between them.

[6] Bound, Jaeger and Baker (1995) found that the instruments used in some specifications estimated in the Angrist and Krueger (1991) paper had very little explanatory power. These instruments were predominantly jurisdiction by quarter of birth dummy variables. Bound *et al.* (1995) argued that these weak instruments asymptotically biased the instrumental variables estimates of the return towards the least squares ones.

[7] Of the other Australian jurisdictions, only the Northern Territory now has a similar policy. It operates only for children turning five in the first half of the year. Primary school in South Australia consists of Reception plus Years 1 to 7; Secondary school Years 8 to 12.

[8] The Committee found that the average age of South Australian school students in Year 7 had fallen relative to that in other States (by half a year) between 1964 and 1979 not because of changes in the way Reception operated, but from changes in grade repetition policies in lower and middle primary school levels (1982: 88).

[9] The current policy has been updated to reflect the move from a three to four term school year in South Australia.

[10] Hence, prior to the policy change they entered the grade cohort that followed those in the first group in Figure 3. After the change, they were part of the same cohort.

[11] Soon after the introduction of the *Early Years of School* policy the structure of the school year changed in South Australia which requires some amendment to the diagram. Prior to the introduction of the four-term year in 1987, the three school terms typically commenced in early February, late May and mid-September respectively. With four terms, the terms commenced in late January, late April, mid to late July and mid-October (South Australian Department of Education *Calendar*, various years).

[12] Other aspects of the education system, including educational curriculum policy changed between cohorts. First, the South Australian Certificate of Education was introduced in 1992, which affected the Year 12 cohort of 1993, a subset of the early 1990s cohort likely to consist of only the September quarter and March – June comparison groups. It was the certificate available to all groups in the late 90s cohort. Second, resources were committed to the development of additional curriculum to support the *Early Year of School Policy*. If it supported studies where little curriculum existed applicable for the group that formerly did three years of school in two calendar years, this curriculum simply filled a gap that some curriculum must have been used for. If it was of a better 'quality' than existing curriculum, it seems most likely that it would have been used for the entire grade cohort and, hence, affected both the birth groups affected by the policy and the comparison group.

[13] The test was popularised by Hausman (1978). Nakamura and Nakamura (1981) identify its earlier origins and demonstrated the equivalence of including either the predicted value from the first stage least squares regression or the residual in the second stage.

[14] Specifically, the exogeneity of the years of schooling can be tested explicitly, measures of labour market conditions can be included in the regression equation and regression estimation removes sources of residual variance that may help identify significant relationships in the presence of small samples.

[15] This paragraph describes instrumental variables estimation where the second stage equation is estimated by least squares. Where the second stage equation has a discrete dependent variable, as here, the test procedure is identical to that specified in the text (see Smith and Blundell 1986). However, the estimation procedure differs. The second equation can be estimated by Generalised Method of Moments (GMM) (see Greene 1997 for a description of GMM estimation). This approach is followed in this paper.

[16] Such techniques are described in Maddala (1983), Greene (1997) and Wooldridge (2002), for example.

[17] That the proportion employed full-time in the second year after leaving school is higher in the early 90s cohort for the September quarter birth group and, particularly, the March – June comparison group may reflect a reversal of the advantageous conditions faced by the October – February birth group on leaving school. By two years after they left school, the late leavers faced a post-recession labour market. The early leavers from the October – February birth group may have still faced a recessed labour market two years after leaving school.

[18] In fact, where the education variables are endogenous, the entire process (both stages) is estimated by Generalized Method of Moments (GMM - see Greene 1997).

[19] The age left school variable may also pick the effect of grade repetition on the probability individuals obtain a full-time job after leaving school. If the students who repeat a grade are typically of lower 'ability' than others in the sample who do not repeat a grade, the age left school variable, holding grade completed constant, may capture a negative ability effect. When an additional indicator variable was included in the employment equation that reflected whether individuals were older than 'normal' when they left school, conditional on the grade they completed, it was insignificant (a 'z' value of 0.24) and lowered the magnitude of the age left school effect by only a small amount (from 0.078 to 0.070). Therefore, the results suggest that the estimated age left school effect captures the impact of the policy change on employment outcomes rather than any separate grade repetition effect.

[20] The 'out of sample' predictive performance of the model was tested in the following way: the sample was split randomly into ten approximately equal groups. Each 10 per cent sample was alternately excluded from the estimation and the resulting parameters used to predict the employment outcome of the excluded 10 per cent. If the predicted probability was greater than or equal to 0.5, the prediction was set equal to 1 (predicted to be full-time employed); otherwise it was set equal to zero (not employed). The average correct prediction rate of the 10 'out of sample' sets of predictions was 60.5 per cent, with the worst correct prediction rate 55 per cent. The average of 60.5 per cent was higher than the prediction rate if individuals had been simply assigned randomly to one state or the other, or if everyone was assigned to 'employed' or 'not employed', which all would provide a success rate of 50 per cent in these data.

[21] The LM statistic against a form of multiplicative heteroskedasticity was 15.6, less than the critical $\chi^2_{(12)}$ value of 21.0. Where 'Robust' standard errors are used, the estimated age left school effect is not quite significant at the 5 per cent level ($p = 0.0504$) and the years of Reception effect is not quite significant at the 10 per cent level ($p = 0.1082$). Following Murphy and Topel (1985), the reported standard errors take account of the inclusion of the constructed self-assessed school performance variable.

[22] These are described in Pagan (2002) and involve tests of whether the mean and variance derived from the probit model depart from those of the data. In this case they were not. The second stage of Pagan's approach is to compare diagrams of the conditional mean and variance of the data with those imposed by the probit model, where the conditioning variable is the 'index' estimated by the probit model, $X'\beta$. In this case, the conditional mean and variance of the probit model tracked non-parametric estimates of the conditional mean and variance from the sample very closely.

[23] Unless otherwise identified, the age-grade pattern for each group in Figure 2.2 is also evident in the jurisdictions that make up the groups.

[24] The YIT 70 cohort exhibits similar age-grade structures across jurisdictions to those shown for the YIT 75 cohort in Figure 2.2.

[25] In fact, the pattern in South Australia was somewhat similar to that of Queensland, but not WA. In Queensland, some of those born in January and February were in Year 10 like South Australia.

[26] School commencement procedures in most of these jurisdictions now require that individuals turn five by the end of April to commence school at the start of Term 1 in any year.

[27] See Queensland Department of Education (1984).

[28] For the changes in the mean values of j variables for group i, denoted by $\mathbf{X_{ij}}$, this distance measure is equal to $[\Sigma_j (\mathbf{X_{ij}} - \mathbf{X_{kj}})^2]^{1/2}$ (see Everitt and Dunn 1991: 67). The more dissimilar are the changes in mean values in the vector of variables, the larger is the distance measure between groups i and k.