Year 12 Completion and Retention in Australia in the 1990s

Chris Ryan
The Australian National University

Louise Watson
University of Canberra

Abstract
This paper analyses high school completion rates, known as ‘apparent retention rates’ in Australia, from 1989 to 2002. Unlike most measures of educational attainment or participation in Australia, the apparent retention rate was lower at the end of the 1990s than it had been in the early 1990s. We adjust ‘official’ retention rates directly and with the aid of regression equation parameters to remedy a number of well-known deficiencies in their measurement, notably Year 12 repetition and migration. The path followed by our adjusted retention measure during the 1990s departs substantially from that of the ‘official’ estimates. We conclude that the apparent retention rate was an especially poor measure of national school completion in the early 1990s, when it peaked. Unlike those ‘official’ estimates, the adjusted measure of Year 12 retention was no lower in the late 1990s than it had been in the early 1990s.

1. Introduction
With one prominent exception, indicators of educational participation and attainment increased in Australia over the 1990s (Australian Bureau of Statistics – ABS – 2002). More people undertake university and vocational education courses of study and formal training now than in the past. Each new yearly estimate indicates that the proportion of the population who possess a university qualification has increased. These indicators are informative about changes in the ‘quality’ of the Australian labour force over time.

In contrast, the official estimate of the proportion of commencing secondary school students who proceed to undertake the final year of school in Australia, Year 12, was lower in 2002 than it was a decade earlier. This would be of less significance if apparent retention rates had not been used as a key indicator of school system performance in Australia. Indeed, the apparent retention rate has been the specific target of government policy.

Address for correspondence: Chris Ryan, Social Policy Evaluation, Analysis and Research Centre, Research School of Social Sciences, Australian National University, Canberra ACT 0200, Australia. Email: Chris.Ryan@anu.edu.au

The authors are grateful to Tue Gørgens and two referees for their comments on earlier drafts of this paper. All errors are the responsibility of the authors.

1 Australian State and Commonwealth Ministers of Education agreed in the late 1980s to a national retention rate target of 65 per cent by the early 1990s (Australian Education Council, 1991, p. 8).

© The Centre for Labour Market Research, 2004
The purpose of this paper is to analyse whether the pattern apparent in these ‘apparent retention rates’ reflects problems in their measurement or the behaviour of school students.

The Year 12 apparent retention rate is estimated by the ABS as the number of students in Year 12 in a given calendar year divided by the number of students who were in the first Year of secondary school when that grade cohort commenced secondary school. The students observed in Year 12 need not necessarily have been in the specific cohort when it commenced secondary school. Official publications have repeatedly pointed to a list of factors that affect the measurement of apparent retention rates. In its publication, the ABS notes that: ‘Care should be exercised in the interpretation of apparent retention rates as the method of calculation does not take into account a range of factors. At the Australia level these include students repeating a year of education, migration and other net changes to the school population.’ (ABS Schools Australia 2003, p. 36). The warning goes on to identify further factors that cloud comparisons across jurisdictions, but may not influence the national estimate. These include different age-grade structures, the prevalence of part-time study and inter-state migration.

The aim of this paper is to estimate the impact of as many of the confounding influences identified above on apparent retention with a view to assessing their importance. We argue that it is possible to adjust apparent retention rates in a manner that deals with these issues. International and inter-state migration can be handled through adjustments to the estimated rates that build in changes in the school-age population. These adjustments are described in Section 3. The impact of grade repetition, age-grade structures and part-time school study can be estimated from available ABS data via regression analysis. Our regression methodology designed to do this is described in Section 4. We discuss our results in Section 5 and use them to re-assess what happened to Australian school completion in the 1990s in Section 6. The policy implications of the results are identified in the conclusion. Before we begin our re-analysis of the official apparent retention estimates, however, in the next Section we use information on Year 12 ‘retention’ from longitudinal data of cohorts of young Australians from the early and late 1990s. This analysis provides some benchmark with which the official estimates might be compared.

2. Estimates of School ‘Retention’ from Australian Longitudinal Data
We begin by using data from three Longitudinal Surveys of Australian Youth (LSAY) cohorts to estimate school ‘retention’ rates for the early 1990s and 1998 and 2001. The data are drawn from the Youth in Transition 1975 birth cohort (YIT 75) and the Longitudinal Surveys of Australian Youth Year 9 cohorts of 1995 and 1998 (LSAY 95 and LSAY 98). These data are collected under a program managed jointly by the Australian Council for Educational Research (ACER) and the Department of Education, Science and Training.
The benefits of using such longitudinal data for these purposes should be emphasised. It is possible to measure retention to Year 12 by individuals observed in the school system some years earlier, typically before they reached the minimum school leaving age. If measured carefully, retention estimates based on these data are not subject to influence by migration or other sources of population change, part-time studies or grade repetition. They will be affected by panel attrition, however, which may be concentrated among early school leavers. We describe the weights we use to deal with attrition below.

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). Members of the cohort were drawn from Years 8, 9 and 10 when first surveyed in 1989. The LSAY 95 and 98 cohorts were grade-based panels. Students were in Year 9 in 1995 or 1998, but varied by age, again depending on differences in the institutional features of the Australian jurisdictions and the reference date for age in the survey (the start of October).

For all cohorts, we analyse retention from the beginning of the panel to Year 12 by analysing whether individuals were in Year 12 in the first possible year, given the grade they were in when first surveyed. For YIT 75, this means analysing Year 12 participation in 1991 for those in Year 10 in 1989, in 1992 for those in Year 9 and in 1993 for those in Year 8. The estimate is of the proportion of respondents from 1991, 1992 and 1993 who were in Year 12 in the first year they could possibly have reached that grade. For LSAY 95 and 98 we analyse Year 12 participation in 1998 and 2001 respectively. Our measure differs from the official apparent retention rate in that the numerator and the denominator are measured over the same group of people.

This approach to the analysis of the YIT 75 data differs from that adopted in a series of studies conducted under the auspices of ACER. For example, Long, Carpenter, and Hayden (1999), Marks, Fleming, Long and McMillan (2000) and Fullarton, Walker, Ainley and Hillman (2003) analysed participation in Year 12 by age 19 from the YIT 75 data, though the participation measure used Marks, et al. (2000) for LSAY 95 and by Fullarton, et al. (2003) for LSAY 98 is the same as that used here. We prefer to use a conceptually similar measure of Year 12 participation for both cohorts, one that in this case is closer to that of a retention rate. That is, our participation measure is the proportion of students in Year 12 in the first year they could possibly have reached that grade, given the grade they were in when first surveyed, a narrower definition than students who had reached Year 12 by age 19.

There is another source of difference between our estimates and those in the ACER reports. It relates to the way the data from these cohorts are weighted to account for panel attrition.
In LSAY 95 and LSAY 98 the attrition weights utilised in ACER studies and provided with the data are designed to ensure the sample matches the original distribution of school achievement (performance in literacy and numeracy tests) and gender observed when the samples were first drawn (Marks and Long, 2000). The re-weighting procedure corrects for high attrition among males and respondents from low achievement quartiles by giving greater weight to respondents with those characteristics who continue in the panel in any year.

The weighting schema used for the Youth in Transition series, on the other hand, differed from the LSAY re-weighting procedure in two important respects. First, the weights were based on school attainment (the level or grade completed) rather than school achievement. Second, the school attainment benchmarks were based on information external to the survey data. The benchmarks are based on published ABS school apparent retention data. The approach used is described in more detail in Appendix D of Williams (1987). The YIT re-weighting procedure aims to ensure that the distribution of the weighted achieved sample across attainment levels from Year 9 to Year 12 matches the ABS population proportions of individuals finishing school after Years 9, 10, 11 and 12 by gender, in each Australian jurisdiction and by school type (independent, Catholic and government schools).

Consequently, if we used the available weights for the YIT 75 cohort, we would simply replicate the retention patterns in the ABS data of the early 1990s, including the potentially important impact of Year 12 repetition. Rather than do that, we calculate attrition weights for the YIT 75 cohort like those used for LSAY 95 and LSAY 98, that is ones designed to match the original distribution of school achievement and gender observed when the YIT sample was first drawn. This allows us to observe the actual differences in retention in the data, rather than impose the patterns evident in the ABS data at the time. The resulting national ‘retention’ estimates appear in table 1, along with those published by ACER and the ABS apparent retention rates for the relevant years.

The estimates prepared for this paper show growth in national ‘retention’ between the first cohort and the later cohorts. These results contrast with the results presented in the ACER reports, where Year 12 completion is estimated to be slightly higher in YIT 75 than LSAY 95 but about the same as LSAY 98. This difference in results reflects the alternative approaches to measurement of Year 12 participation in YIT 75 (measured in 1994 by ACER, but in the first relevant year – 1991, 1992 or 1993 – by us).

The results from this analysis of longitudinal data contrast with the picture from national apparent retention rates in table 1. School completion was no
lower in the later cohorts than the cohort from the early 1990s. In fact, it appears that it was higher when measured across the same group of individuals. Of note for the analysis that follows, Year 12 repetition in the earliest cohort was higher (about double) than in the later cohorts.4

Table 1 Estimated National ‘Retention’ Rates

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Current estimates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUST</td>
<td>70.4</td>
<td>75.9</td>
<td>78.6</td>
<td>5.5</td>
<td>6.43</td>
<td>8.2</td>
<td>9.65</td>
</tr>
<tr>
<td>ACER estimates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUST</td>
<td>78</td>
<td>76</td>
<td>79</td>
<td>-2</td>
<td>-1.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Published ABS apparent retention rates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUST</td>
<td>75.0</td>
<td>71.6</td>
<td>73.4</td>
<td>-3.4</td>
<td>-1.6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


3. Official Apparent Retention Rates and Our Initial Adjustments

The ‘official’ Year 12 apparent retention rates are estimates produced by the Australian Bureau of Statistics (ABS) of the proportion of any cohort who commence secondary school and proceed to Year 12 in the minimum possible number of years in a jurisdiction. Secondary school commences in Year 7 in New South Wales, Victoria, Tasmania and the Australian Capital Territory. It commences in Year 8 in other jurisdictions (Queensland, South Australia, Western Australian and the Northern Territory). Consequently, the calendar year in which the denominator is measured varies between jurisdictions for any Year 12 retention figure in a calendar year.

Most of the data used subsequently in this paper are taken from the National Schools Statistics Collection (NSSC), which is published by the ABS in Schools Australia (Cat. No. 4221.0) (see Appendix 1 for the sources of the data used here). While the NSSC includes information on government and non-government schools and staff, the analysis in this paper mostly uses information on students. Specifically, the focus is on full-time student numbers in each Australian State and Territory cross-classified by age and year or level of education.

There are two particular limitations in the NSSC data for the analysis undertaken below. First, the data do not allow identification of the prevalence of Year 12 repetition. Second, the data relate primarily to full-time students. While some data on part-time students have been published since 1995 by the ABS, it is only available by year or level (not by age) and

4 The estimates were 4.7 per cent in the first cohort compared with 2.3 and 0.9 per cent in the later cohorts.
earlier unpublished data are considered too unreliable for release. Since both of these factors are forms of school participation that individuals may choose and their incidence has reportedly changed over time, their mismeasurement in these data might affect the validity of the inferences drawn here. Attempts to estimate both of these phenomena from the published data are described below. Other data utilised in this paper are drawn primarily from ABS sources. Appendix 1 contains the definitions of these variables and their sources.

Adjusted retention rate measures for jurisdictions are calculated and used in this paper. Three adjustments to the standard definition are made in this paper to produce estimates that are more consistent across jurisdictions and to minimise the effects of departures or additions on the cohort through either grade repetition or migration (both international and inter-state). That is, they deal with some of the limitations of apparent retention rates acknowledged by the ABS and described in Section 1.

- First, for all jurisdictions, retention rates are estimated by dividing the number of Year 12 students in calendar year $t$ with the number of Year 8 students in calendar year $t - 4$ (in jurisdictions with six years of secondary school, the ‘official’ estimate uses $t - 5$). This changes marginally the estimated retention rates for those jurisdictions where secondary school commences in Year 7, but leaves unchanged the estimates for jurisdictions where it commences in Year 8.

- Second, only those Year 8 students aged 12 to 15 years (inclusive) are counted in the denominator (these ages represented 99.8 per cent of Year 8 students in Australia in 1998) and only those Year 12 students aged 16 to 19 in the numerator (these ages represented 98.8 per cent of Year 12 students four years later in 2002).[^5]

- Third, the estimates are adjusted for changes in the population of the relevant age cohorts in jurisdictions over these years. That is, population growth in each Year 8 single year of age cohort over the years to Year 12 is used to adjust the retention rate estimate. In effect, the denominator is increased to match population growth over the intervening years.

Overall, the three adjustments lower estimated retention in 2002 by about four percentage points in Australia (see figure 3, which is described in more detail below). The first change is made simply to improve the comparability of the retention estimate across jurisdictions. As will become clearer below (see Section 6), the latter two changes minimise the contribution to our adjusted rates of phenomena that artificially inflated measured retention rates in the early 1990s.

---

[^5]: For 1991 and earlier years, the oldest age category in the ABS data was ‘19 years and older’. We used the observed 1992 share of 19 year olds among students aged ‘19 years’ or ‘20 years or older’ in jurisdictions to estimate those aged 19 years in Year 12 before 1992.
4. Regression Methodology for Further Retention Rate Adjustments

The adjustments set out in the previous Section deal with some of the problems with apparent retention estimates identified by the ABS. We use regression techniques to estimate the contribution of remaining problems, notably the effect of Year 12 repetition and part-time study on retention rates. The methodology we use to estimate the contribution of these phenomena to retention rates is described in this Section. Since we need a properly-specified econometric equation to distinguish their effect from other factors, this Section also contains a more general discussion of the determinants of Year 12 retention in Australia. Econometric estimation of school participation or retention rate behaviour over time by young people has not been particularly satisfactory in Australia. The studies have found surprisingly little role for many economic factors that might be expected to influence school participation decisions. Kenyon and Wooden (1996) consider that all of the relevant studies ‘are afflicted by data problems which give rise to severe multicollinearity’ (1996, p. 30).

The approach adopted here is to estimate an equation that explains the Year 12 retention rates of Australian States and Territories from 1989 to 2001. This period includes the final observations of the long period where school retention increased, its subsequent fall and ‘plateauing’ over the remainder of the 1990s. The aim is to identify whether economic or school structure factors (or some combination of both) were responsible for the retention patterns observed in Australia’s jurisdictions over the 1990s.

The estimated equation is of the form:

\[ R^*_j = \alpha_j + X'_j \beta + S'_j \gamma + \phi P^*_{12,j} + u^*_j \]  

(1)

where \( R^*_j \) is the population adjusted Year 12 retention rate in jurisdiction \( j \) in year \( t \) (described in the last section), \( X \) is a set of economic variables and \( S \) a set of variables reflecting the structure of schooling in Australian jurisdictions, \( P^*_{12,j} \) is a measure of ‘true’ Year 12 repetition and \( \alpha, \beta, \gamma \) and \( \phi \) are parameters or parameter vectors to be estimated. 6

The intercept is allowed to vary by jurisdiction, but other parameters are constant across jurisdictions and over time. The error term \( u^*_j \) is assumed to satisfy \( E[u^*_j] = 0 \) and \( Var[u^*_j] = \sigma^2_j \), that is, it has a zero mean and its variance is constant over time but allowed to vary across jurisdictions. 7

The explanatory variables include labour market and education system-related factors. The labour market variables are designed to reflect both

---

6 It is common to estimate the logit form of equation (1), that is with the dependent variable specified as \( \log[R^*_j/(1- R^*_j)] \) because this specification ensures that the predicted values will lie between zero and one. This approach is followed in many studies, such as Lang and Kropp (1986), Forster and Ryan (1989), and Karmel (1995). Results from both specifications are reported below, with the least squares estimates emphasised because they are marginally more straightforward to interpret.

7 The standard errors were estimated taking account of the ‘clustering’ of observations by jurisdiction.
cyclical and structural factors likely to influence the employment opportunities for young people. The cyclical factors are captured through the change in annual state unemployment rates over the period.\textsuperscript{8} Like Karmel (1995), we allow for the cyclical effects to have an asymmetric effect on retention – deteriorations may add more to retention than improved conditions detract from it. The structural factors are captured through measures of the annual state proportions of fifteen to nineteen year olds who work full-time.\textsuperscript{9}

Differences in the education systems are captured with a series of variables. The first is the proportion of 17 year olds in a jurisdiction studying at a TAFE college in any year. This variable is interpreted as reflecting alternatives to school in jurisdictions, both educational and labour market alternatives, since apprentices and trainees are included in the TAFE student figures. The second is an estimate of the proportion of the relevant Year 8 cohort studying part-time in Year 12. The third variable is the proportion of Year 8 students of any grade cohort who are indigenous. This was highest in the Northern Territory in 1998 at about 25 per cent of the grade cohort, but was no more than 5 per cent in the other jurisdictions. The final variable is the proportion of grade cohorts who were aged twelve years in Year 8. This proportion is highest in jurisdictions without a pre-Year 1 year of primary school (Queensland and – at the time – Western Australia). The effect of the age structure of grade cohorts is identified here because these proportions fell with changes in school entrance arrangements in two jurisdictions in the estimation period: from over 50 per cent to less than 10 in South Australia and from 40 to about 30 per cent in Queensland.\textsuperscript{10}

There are a number of potential problems with the regression equation specified in equation (1). First, there is no official measure of Year 12 repetition, so it needs to be estimated. Since any estimate is bound to be wrong, the related problem of measurement error and its impact on the estimated parameters of equation (1) needs to be addressed. Second, some of the explanatory variables are potentially endogenous; notably the full-time teenage employment ratio and the TAFE participation rate. Third, estimates of part-time employment are not available for the entire estimation period. We now describe how we deal with each of these issues.

\textsuperscript{8} We use the aggregate rate rather than the youth unemployment rate because it is not influenced as directly by the retention rate as the youth rate may be and to limit the number of potentially endogenous variables we need to deal with in equation (1).

\textsuperscript{9} The observed retention behaviour for any cohort reflects decisions made over a number years by its members. Therefore, the economic conditions pertaining at the end of Years 9, 10 and 11 could affect retention, not just the conditions when the cohort reached Year 12. Consequently, for any particular cohort that reaches Year 12 in some year, the value included for any specific economic variable is its average value over the two prior financial years.

\textsuperscript{10} These policy changes are described in more detail in Ryan (2003). Other features of their educational systems may explain divergent patterns between jurisdictions and over time. Some jurisdictions introduced ‘Certificates of Education’ encompassing students’ senior school performance during the estimation period. Lamb (1996, 1998) reported that Year 12 repetition increased substantially in the late 1980s and early 1990s in South Australia and the Northern Territory following changes in certification requirements. The changed certification arrangements first inflated the jurisdictions’ retention rates then exaggerated their decline from 1992. Vickers and Lamb (2002) conclude that differences in educational structures between Australian jurisdictions explain divergent patterns of school retention in the late 1990s.
**Year 12 Repetition**

The retention rate estimates used here exclude those Year 12 students aged 20 years or older in the data. This adjustment will have removed some of the effects of Year 12 repetition. We adopt a two-stage strategy to estimate the impact of any residual Year 12 repetition on retention rates. In the first stage, we generate an estimate of Year 12 repetition in all jurisdictions in all years. It may be a poor estimate of ‘true’ Year 12 repetition, since there are no official ABS figures published on grade repetition. In the second stage we make assumptions about the way we have mismeasured the ‘true’ Year 12 repetition and use instrumental variable regression techniques to obtain better estimates of the extent of Year 12 repetition.

We adopt a very simple approach to estimating the first stage estimates of Year 12 repetition. Since repetition affects the age distribution of any grade, we endeavour to capture changes in the age distribution of cohorts as they move through school grades. We make assumptions about the average age of individuals in each single year of age group in Year 11 (specifically, the average age of fifteen-year-olds is 15.5 years, sixteen-year-olds 16.5 years and so on) and multiply those ages by the distribution of students across the ages in Year 11 to obtain an average age for the entire grade cohort. We then add one year to each age weight for estimation of the average age of the grade cohort in the following year in Year 12. Any change in the average age that exceeds unity we attribute to grade repetition, since it reflects a shift in the cohort’s age distribution.

In the second stage, we assume that our estimates of Year 12 repetition, \( P_{12,jt} \), vary by jurisdiction, \( j = 1, \ldots, J \), and mismeasure true Year 12 repetition, \( P^*_{12,jt} \), in the following way:

\[
P_{12,jt} = \lambda P^*_{12,jt} + e_{jt} \quad \leftrightarrow \quad P^*_{12,jt} = \frac{1}{\lambda} \left( P_{12,jt} - e_{jt} \right)
\]

with \( e_{jt} \sim (0, \sigma^2_e) \). That is, our estimated repetition departs from the true rate by some scaling factor and an error term. The scaling factor is assumed constant through time and across jurisdictions. If we substitute for \( P^*_{12,jt} \) in equation (1) using equation (2), we get

\[
R^*_{jt} = \alpha_j + X_j' \beta + S_j' \gamma + (\phi/\lambda) P_{12,jt} + (u_{jt} - (\phi/\lambda)e_{jt}) \quad \text{or} \quad (3)
\]

\[
R^*_{jt} = \alpha_j + X_j' \beta + S_j' \gamma + \phi P_{12,jt} + u'_{jt}
\]

where \( \phi = \phi/\lambda = 1/\lambda \), since \( \phi = 1 \) because each Year 12 repeater is included in the ‘official’ apparent retention estimate, and \( u'_{jt} \) is a composite error term. Because Year 12 repetition is measured with error, least squares estimation of equation (4) will produce biased and inconsistent estimates of the parameters (Greene, 1997, p. 437). Consistent estimates of the parameters may be obtained through estimation of equation (4) by instrumental variables. This requires the existence of a suitable instrument, one that is correlated with Year 12 repetition, but not Year 12 retention except through the repetition variable. Note that if such an instrument exists, equation (4) would provide us with an estimate of \( 1/\lambda \). This parameter
estimate would allow us to obtain a better estimate of ‘true’ retention from our estimated Year 12 repetition rate via equation (2). Our estimates in any year may be wrong, but on average, they will be of a similar magnitude to the true Year 12 repetition rate.

In fact, we possess a suitable instrument, one that is correlated with Year 12 repetition but not directly with Year 12 retention as we measure it. The instrument is the proportion of the total Year 12 cohort that is older than 19 years of age in the ABS data. Such individuals are excluded from our calculation of retention, so the proportion is not correlated directly with our estimate of retention. However, a substantial proportion of such individuals will themselves be repeating Year 12 and it seems reasonable to anticipate that Year 12 repetition among those aged 20 years or older would be closely correlated with repetition among those aged 19 years or younger. Hence the proportion of the total Year 12 cohort that is older than 19 years of age in ABS data can be used as an instrument for our estimated repetition in equation (4).  

Endogeneity of Explanatory Variables

The potential endogeneity of any of the explanatory variables can also be accommodated through instrumental variable regression techniques. The potential endogeneity of two variables – the teenage full-time employment proportion and the proportion of 17 year olds studying at TAFE – was addressed in the analysis undertaken here.  

Part-time Study

To address the absence of official Year 12 part-time study figures over the entire estimation period, estimates were made of the proportion of the original Year 8 cohort studying Year 12 part-time in jurisdictions. There are two problems with the published part-time study figures. The first is that the ABS only began to publish part-time student numbers from 1995. Total part-time students in government schools by jurisdiction in 1994 were reported in Steering Committee for the Review of Commonwealth/State Service Provision (SCRCSSP, 2000). The second source of measurement error is that the ABS does not publish part-time student figures by age. Therefore, the age of those enrolled in Year 12 is not known.  

The resulting repetition estimates from the two-stage estimation process broadly match what little is known about it in Australia, where no official estimates are published. DEET (1994, p. 12) contains an estimate that at that time Year 12 repetition added about 4 percentage points to the national retention estimate in 1993. Morgan (1995) contains larger estimates of Year 12 repetition in Australian jurisdictions in the early 1990s, the magnitudes of which support earlier estimates in Russell (1993). Those estimates were based on figures provided by the various Boards of Study, who certify whether individuals complete their Year 12 studies at some required level.  

The teenage full-time employment proportion failed a Hausman (1978) exogeneity test, so we used lagged values as instruments for current values in estimation. In fact, for reasons outlined below, we used the square of the proportion in estimation and lagged values to instrument for current values. The proportion of 17 year-olds studying at TAFE variable survived a Hausman (1978) exogeneity test, so it is treated as an exogenous variable. The instrument used for the test was the proportion of the population aged 20 to 24 studying at TAFE in each jurisdiction in each year.  

It is known for some states. For example, about 20 per cent of part-time Year 12 students were aged 20 years or older in South Australia in both 2000 and 2001 (Department of Education, Training and Employment 2001, 2002).
assumptions about growth in full-time equivalent part-time students prior to 1994 were made and the resulting variables included in regression equations. Alternatively, the equation was estimated only over the period since the ABS began to publish the part-time student numbers. The part-time student variable was never significantly different from zero and generally of the ‘wrong’ sign if part-time students do detract from full-time Year 12 retention (it was positive). Consequently, the variable was excluded from the results reported.

5. Regression Results
The regression results reported in table 2.2 of Appendix 2 largely confirm the role of education system effects on Year 12 retention. Retention is lower the higher the proportion of seventeen year olds who study at TAFE and the higher the indigenous share of the commencing Year 8 cohort. The results imply that indigenous retention is about 30 per cent of that of non-indigenous Australians. That estimate seems low compared to estimates based on national figures in Long, Frigo and Batten (1999, p. 50) of just over 40 per cent. Our estimates are heavily dependent on retention patterns and the treatment of indigenous students in the ABS collection for the Northern Territory. Changes in the treatment of indigenous students in the Northern Territory affected apparent retention rates there from 1999.

Retention is higher among jurisdictions with ‘young’ Year 8 cohorts and in jurisdictions where Year 12 repetition is higher. There was no evidence from the regression results that growth in part-time Year 12 student numbers had detracted from Year 12 retention, so that variable was excluded from the results presented.

The results also confirmed the role of economic factors in explaining retention (both the long-term decline in full-time youth jobs and the role of cyclical effects through changes in the unemployment rate). Like findings in Larum and Beggs (1989) and Lewis and Koshy (1996), the results suggest that the decline in full-time job opportunities for teenagers was the main determinant of increased school retention. Like Karmel (1995), we find that general labour market conditions have an asymmetric effect on retention – deteriorations add to retention, but improved conditions do not lead to reversals in retention. In addition, there were significant remaining differences in retention outcomes between Australian jurisdictions.

Specification tests suggested that linear forms of the explanatory variables did not capture non-linearities in retention over the estimation period. These were captured adequately with a quadratic term for the proportion of teenagers working full-time. However, the parameter on that variable itself

14 The estimates in Morgan (1995) suggest that there were substantial numbers of part-time students in Year 12 in South Australia and some other jurisdictions prior to 1994.
15 The change involved reclassifying secondary-aged students in remote Aboriginal schools to an ‘ungraded’ category from 1995. This lowered the indigenous share of the 1995 Year 8 cohort by 10 percentage points compared with the 1994 cohort. Consequently, the apparent retention rate increased by 10 percentage in the Northern Territory between 1998 and 1999. This episode is likely to be very influential in the regression estimates. Also, the regression data are not weighted by the population, so they reflect retention by jurisdictions, rather than necessarily capturing national averages as in Long, et al. (1999).
varied between jurisdictions, being significantly larger in jurisdictions with older grade cohorts than younger ones.\textsuperscript{16} A separate parameter on that variable was estimated for two sets of jurisdictions: one for the NSW, VIC, TAS and the ACT group; the other consisting of the remaining jurisdictions. The results imply that the loss of teenage jobs had a greater impact on retention in those jurisdictions where students tended to be older for a given grade. The loss of teenage jobs in states with older grade cohorts may have had a larger impact on retention because it changed the opportunity cost of remaining in school for potential school leavers in such states by lowering the probability of employment.

The equations survived RESET specification tests and the parameter estimates were robust to the inclusion of other variables subsequently excluded from the results presented. Alternative specifications involving other economic variables did not affect the parameters of the education system variables. The standard errors are ‘robust’ estimates to account for different error variances between the states and for any serial correlation in the errors.

6. Why did the ‘Official’ National Apparent Retention Rate Fall in the 1990s?

The regression equation described in earlier sections explained differences in retention between jurisdictions. All observations from all jurisdictions are accorded the same weight in the equation. Hence, the equation does not strictly ‘explain’ changes in the Australian apparent retention rate over time, since this would require greater weight being given to observations from the more populous than the less populous jurisdictions.

\textbf{Figure 1  Contribution of Explanatory Variables to Changes in National Retention – 1989 to 2001\textsuperscript{a}}

(a) The data are smoothed for presentational purposes.

\textsuperscript{16} Wald tests suggested the linear terms for the teenage full-time employment proportions could be excluded once the quadratic terms were included.
Nevertheless, it is possible to use the results to identify the factors that appear to have played the greatest role in shaping Australian retention since the late 1980s. We do this with the aid of a diagram, figure 1, which shows the contributions of the variables included in the equation to changes in retention from 1989 to 2001. The value taken by each variable in each year was multiplied by the estimated parameter from the regression equation. The change in the impact of each variable on retention, relative to its 1989 effect, is shown in figure 1, along with the actual change in retention and the aggregate predicted change.

The first point to note about figure 1 is that predicted retention, based on the regression equation, tracks changes in actual retention reasonably, though it fails to capture the increase in retention that occurred in 1990. Thereafter, predicted retention increases quickly, peaking in 1992 and 1993 before falling and levelling off from 1995 about 10 percentage points higher than its 1989 value.

Three factors seem to have contributed most to the peak in retention. The broad magnitude of retention rate changes appears to have been driven by the decline in available full-time jobs for teenagers. Overlaid on this effect in the early 1990s were two other factors. The first contribution to retention was provided by the recession (through the increase in the unemployment rate), which added about three percentage points to retention in 1991. In reality, the teenage full-time employment effect also contains a substantial recessionary effect – the largest falls in that proportion occurred in the recession of the early 1990s, with the proportion being relatively stable in other years. The other contributing factor to the growth in retention in the early 1990s was the effect of Year 12 repetition, which added close to three percentage points to retention in 1992 (relative to its impact on retention in 1989). The relative magnitude of these latter two effects fell after 1992 and the impetus to retention from the loss of full-time teenage jobs slowed from 1993. Consequently, retention fell after 1993. In addition, another effect of a structural nature detracted from retention after 1994. Grade cohorts aged somewhat (this ageing resulted from changes in school entrance arrangements in South Australian and Queensland, described in Ryan, 2003), which had the effect of lowering retention by about one and a half percentage points.

These factors seem to provide a reasonable explanation for the changes in our adjusted measure of retention between 1989 and 2002. They are not a complete explanation for the pattern observed in the ‘official’ measure of apparent retention, however. To do that, we need to return to how the ‘official’ rate is measured.

---

17 When the data were weighted according to population shares, the parameter estimates on the variables that measure full-time employment, unemployment rate changes, Year 12 repetition and the proportion aged twelve in Year 8 changed little.

18 On average, the predictions from the regression equation must equal the actual values. The estimates in figure 1 show retention growth relative to 1989 and suggest the predicted growth from 1989 was less than the actual, so the predicted values for 1989 must exceed the actual.
In figure 2 we combine three factors that influence the ‘official’ measure that do not reflect true school completion by commencing Year 8 cohorts: school age population growth; the addition of older students to the Year 12 class; and Year 12 repetition. Between 1989 and 2002 these factors made varying contributions to the ‘official’ apparent retention estimate. However, they all peaked in magnitude in either 1992 or 1993, and their sum peaked in 1992, adding about eight and a half percentage points to apparent retention in that year. By 1996, the trough in ‘official’ apparent retention in the 1990s, the contribution of these three factors had fallen by 5 percentage points, just less than the decline in ‘official’ apparent retention over that period.

Taken together these various effects suggest that the decline in ‘official’ national apparent retention of six percentage points between 1992 and 1996 was made up of four elements. The first was the decline in the number of Year 12 students aged 20 or older which must have included a substantial proportion of students repeating Year 12. This effect contributed about one percentage point to the decline in apparent retention. The second was the decline in Year 12 repetition among students aged 16 to 19 in Year 12. This effect contributed about 3 percentage points to the decline in apparent retention. The third effect, of another percentage point arose from a decrease in the growth rate of the senior school age population. The final effect of about one and a half percentage points arose from the ageing of grade cohorts (the fall in the share of the Year 8 student body aged 12 years).

One result of some note, given their importance in pushing retention rates up, is that changed labour market conditions had no net effect on retention
after 1992, because the effect of lower teenage full-time employment on retention was offset by better aggregate labour market conditions.\textsuperscript{19}

After we correct the figures to take account of these measurement issues, what does the path followed by Australian retention over the 1990s look like? Figure 3 shows this picture. It contains the ‘official’ estimate of apparent retention between 1989 and 2002, an estimate when the rate is adjusted only for population change and the older Year 12 students are removed, and a third estimate when the effect of Year 12 repetition is also removed (that is, these estimates include the earlier adjustment). No adjustment is made for the ageing of the grade cohort.

**Figure 3 ‘Official’ and Adjusted National Retention Rate Estimates 1989 to 2002**

The picture from the population, cohort and repetition adjusted measure is somewhat different from that provided by the ‘official’ apparent retention estimate. The peak in retention during the early 1990s recession is less pronounced, the subsequent decline smaller and the adjusted retention estimates exceed those of 1992 and 1993 from 1999. That is, unlike the pattern in the ‘official’ estimate, the measure of adjusted retention is higher at the end of the estimation period than it was during the recession of the

\textsuperscript{19}The parameters on the two labour market variables in table 2.2 in Appendix 2 (the changes in unemployment rates and the proportion of 15 to 19 year olds employed full-time), in conjunction with the changes in the values of those variables, are consistent with about a forty percentage point increase in retention between the late 1970s and the 1990s. This is the magnitude of the actual change in retention over the period. However, the parameters assign most of this change to the two recessions of the early 1980s and 1990s and do not pick up as much of the growth in retention that occurred between the two recessions.
early 1990s. This picture is consistent with the Year 12 ‘retention’ or ‘completion’ estimates from longitudinal data presented in Section 2 and supports the proposition that it was no lower in the late 1990s than it had been in the early 1990s.

7. Conclusion
After we adjust official apparent retention rates for a number of acknowledged problems in their measurement, we obtain estimates that are more consistent with both other measures of school completion and of broader education participation and attainment. Our estimates suggest that the apparent retention rate was an especially poor measure of national school completion in the early 1990s. Its peak was exaggerated by Year 12 repetition, abnormally high numbers of older students and relatively strong population growth among 15 to 19 year olds. Removal of these important influences on retention in the 1990s substantially changes the interpretation of the path followed by retention nationally. The peak in retention during the early 1990s recession is less pronounced in the adjusted estimates and the subsequent decline smaller. Unlike the ‘official’ estimates, adjusted Year 12 retention in the late 1990s was no lower than it had been in the early 1990s.

Concern that retention fell during the 1990s in Australia therefore appears misplaced to us. Underlying levels of school completion did not fall. However, the main reason adjusted Year 12 retention was higher in 2001 than it was in 1989 had little to do with the school system: it arose from the loss of available full-time jobs for young people. This factor contributed about 12 percentage points to the increase in retention over the period analysed. Since the key factor driving school retention over this period was the state of the youth labour market, the status of retention rates as an indicator of school system ‘performance’ also seems misplaced.

While this paper has shown it is possible to adjust the National Schools Statistics Collection data to produce retention estimates that are less subject to measurement error over time, it is always preferable to start with data that are already more comparable. Efforts to increase the usefulness of the National Schools Statistics Collection via a unique student identifier would make such comparisons more straightforward and allow issues such as the impact of grade repetition and part-time studies on retention and participation to be addressed more directly than was possible in this paper.
Appendix 1 Data Definitions and Sources

Data used in the paper involve:

- full-time student numbers in each Australian State and Territory cross classified by their age and year or level of education, which is published by Australian Bureau of Statistics in *Schools Australia* (Cat. No. 4221.0). The reference period for the data in this publication is 1 July of the relevant calendar year;

- data on part-time students by grade or level are also drawn from *Schools Australia*;

- monthly total full-time employment, total unemployment rates, 15 to 19 year old full-time employment and 15 to 19 year old civilian population estimates in the various jurisdictions from the *Labour Force Australia* (Cat. No. 6201.0);

- annual estimates of TAFE participation by age and jurisdiction in September of each year from the *Participation in Education* (Cat. No. 6272.0). This publication ceased after 1999. For 2000 and 2001, growth in TAFE student numbers by age and jurisdiction published by the National Centre for Vocational Education Research, *Australian Vocational Education and Training Statistics: In detail*, were used to scale the 1999 ABS TAFE participation estimates;

- State population by single year of age estimates in July of each year, taken from *Population by Age and Sex* (Cat. No. 3201.0).

Appendix 2 Data Descriptives and Regression Results

Table 2.1 Means of Retention Rate Equation Variables (1989 to 2001)

<table>
<thead>
<tr>
<th></th>
<th>NSW</th>
<th>VIC</th>
<th>QLD</th>
<th>SA</th>
<th>WA</th>
<th>TAS</th>
<th>NT</th>
<th>ACT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retention rate</td>
<td>0.658</td>
<td>0.748</td>
<td>0.778</td>
<td>0.735</td>
<td>0.706</td>
<td>0.582</td>
<td>0.475</td>
<td>0.912</td>
</tr>
<tr>
<td>Adjusted retention rate</td>
<td>0.639</td>
<td>0.717</td>
<td>0.719</td>
<td>0.710</td>
<td>0.657</td>
<td>0.570</td>
<td>0.479</td>
<td>0.837</td>
</tr>
<tr>
<td>Logit (Adjusted retention rate)</td>
<td>0.249</td>
<td>0.404</td>
<td>0.409</td>
<td>0.389</td>
<td>0.282</td>
<td>0.123</td>
<td>-0.036</td>
<td>0.710</td>
</tr>
<tr>
<td>Proportion aged 12 in Year 8</td>
<td>0.024</td>
<td>0.017</td>
<td>0.366</td>
<td>0.316</td>
<td>0.422</td>
<td>0.020</td>
<td>0.223</td>
<td>0.015</td>
</tr>
<tr>
<td>TAFE proportion</td>
<td>0.102</td>
<td>0.052</td>
<td>0.057</td>
<td>0.076</td>
<td>0.138</td>
<td>0.084</td>
<td>0.114</td>
<td>0.026</td>
</tr>
<tr>
<td>Year 12 Repetition</td>
<td>0.003</td>
<td>0.014</td>
<td>0.009</td>
<td>0.020</td>
<td>0.019</td>
<td>0.026</td>
<td>0.005</td>
<td>0.026</td>
</tr>
<tr>
<td>Unemployment rate (increases)</td>
<td>0.4</td>
<td>0.5</td>
<td>0.3</td>
<td>0.4</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>Unemployment rate (decreases)</td>
<td>-0.5</td>
<td>-0.5</td>
<td>-0.4</td>
<td>-0.5</td>
<td>-0.3</td>
<td>-0.6</td>
<td>-0.3</td>
<td></td>
</tr>
<tr>
<td>Indigenous share of Year 8 students</td>
<td>0.019</td>
<td>0.004</td>
<td>0.038</td>
<td>0.019</td>
<td>0.044</td>
<td>0.031</td>
<td>0.313</td>
<td>0.007</td>
</tr>
<tr>
<td>Full-time employment share of 15-19 population</td>
<td>0.216</td>
<td>0.181</td>
<td>0.233</td>
<td>0.206</td>
<td>0.267</td>
<td>0.233</td>
<td>0.223</td>
<td>0.146</td>
</tr>
<tr>
<td></td>
<td>Adjusted Retention Rate</td>
<td>Logit of Retention Rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-------------------------</td>
<td>-------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Coeff.</td>
<td>Std. Err.</td>
<td>t-ratio</td>
<td>Coeff.</td>
<td>Std. Err.</td>
<td>t-ratio</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 12 Repeats</td>
<td>1.577</td>
<td>0.44</td>
<td>3.55</td>
<td>7.080</td>
<td>2.22</td>
<td>3.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full-time employment share squared</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(NSW, VIC, TAS, ACT)</td>
<td>-2.277</td>
<td>0.17</td>
<td>-13.46</td>
<td>-9.980</td>
<td>0.83</td>
<td>-12.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full-time employment share squared</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(QLD, SA, WA, NT)</td>
<td>-1.248</td>
<td>0.24</td>
<td>-5.28</td>
<td>-6.43</td>
<td>1.11</td>
<td>-5.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion aged 17 at TAFE</td>
<td>-0.190</td>
<td>0.09</td>
<td>-2.10</td>
<td>-0.910</td>
<td>0.38</td>
<td>-2.41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Increase in the unemployment rate</td>
<td>0.015</td>
<td>0.01</td>
<td>2.67</td>
<td>0.082</td>
<td>0.03</td>
<td>2.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decrease in the unemployment rate</td>
<td>0.003</td>
<td>0.01</td>
<td>0.41</td>
<td>0.002</td>
<td>0.03</td>
<td>0.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indigenous share of Year 8 cohort</td>
<td>-0.717</td>
<td>0.28</td>
<td>-2.52</td>
<td>-3.070</td>
<td>1.16</td>
<td>-2.64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Twelve year old share of Year 8 cohort</td>
<td>0.339</td>
<td>0.05</td>
<td>6.61</td>
<td>1.774</td>
<td>0.25</td>
<td>6.98</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

State Effects

NSW  0.769  0.01  53.14  0.755  0.04  17.56
VIC  0.779  0.01  58.68  0.767  0.04  19.29
QLD  0.686  0.02  38.36  0.677  0.08  8.91
SA   0.653  0.01  45.47  0.651  0.07  8.81
WA   0.626  0.02  30.53  0.615  0.13  4.89
TAS  0.681  0.02  32.35  0.687  0.09  7.83
NT   0.690  0.09  7.82   0.702  0.29  2.41
ACT  0.847  0.02  56.00  0.845  0.04  23.10

Marginal Effects

Year 12 Repeats  1.577  0.44  3.55  1.575  0.49  3.19
Full-time employment share squared |       |           |         |        |           |         |
(NSW, VIC, TAS, ACT)                | -0.970 | 0.07      | -13.46  | -0.946 | 0.08      | -12.06  |
Full-time employment share squared (QLD, SA, WA, NT) |       |           |         |        |           |         |
-Proportion aged 17 at TAFE         | -0.580 | 0.11      | -5.28   | -0.665 | 0.12      | -5.78   |
-Increase in the unemployment rate  | 0.015  | 0.01      | 2.67    | 0.018  | 0.01      | 2.90    |
-Decrease in the unemployment rate  | 0.003  | 0.01      | 0.41    | 0.001  | 0.01      | 0.09    |
-Indigenous share of Year 8 cohort  | -0.717 | 0.28      | -2.52   | -0.683 | 0.26      | -2.64   |
-Twelve year old share of Year 8 cohort | 0.339  | 0.05      | 6.61    | 0.394  | 0.06      | 6.98    |

Observations  104
Dependent Variable: Mean 0.666 0.739
Std dev. 0.115 0.550
Residual Sum of squares 0.078 1.807
Std dev. 0.030 0.143
R-squared 0.943 0.942
F[14, 89] 97.2 95.4
Probability value 0.00 0.00
Log likelihood 226.6 63.2
Restricted(b=0) Log likelihood 77.6 -84.9
Estd. Autocorrelation of e(i,t) 0.29 0.29

References

— Schools Australia, Cat. No. 4221.0, various years, Canberra.
— Labour Force Australia, Cat. No. 6201.0, various years, Canberra.
— Participation in Education, Cat. No. 6272.0, various years, Canberra.
— Population by Age and Sex, Cat. No. 3201.0, various years, Canberra.


