

School-to-work pathways

Rasika Ranasinghe
Emerick Chew
Genevieve Knight
Gitta Siekmann
National Centre for Vocational Education Research



Publisher's note

To accompany this report, an interactive data visualisation, *Visualising school-to-work pathways using LSAY*, presents the school-to-work pathways of young Australians aged 16 to 25, and can be accessed from <https://www.ncver.edu.au/research-and-statistics/school-to-work-pathways>.

The views and opinions expressed in this document are those of NCVER and do not necessarily reflect the views of the Australian Government, or state and territory governments. Any interpretation of data is the responsibility of the author/project team.

To find other material of interest, search VOCEDplus (the UNESCO/NCVER international database <http://www.voced.edu.au>) using the following keywords: Employment; Higher education; Outcomes; Outcomes of education and training; Participation; Pathways; Students; Transition from education and training to employment; Vocational education and training; Youth; Youth transitions.

© Commonwealth of Australia, 2019



With the exception of the Commonwealth Coat of Arms, the Department's logo, any material protected by a trade mark and where otherwise noted all material presented in this document is provided under a Creative Commons Attribution 3.0 Australia <http://creativecommons.org/licenses/by/3.0/au> licence.

The details of the relevant licence conditions are available on the Creative Commons website (accessible using the links provided) as is the full legal code for the CC BY 3.0 AU licence <http://creativecommons.org/licenses/by/3.0/legalcode>.

The Creative Commons licence conditions do not apply to all logos, graphic design, artwork and photographs. Requests and enquiries concerning other reproduction and rights should be directed to the National Centre for Vocational Education Research (NCVER).

This document should be attributed as Ransinghe, R, Chew, E, Knight, G & Siekmann, G 2019, *School-to-work pathways*, NCVER, Adelaide.

This work has been produced by NCVER on behalf of the Australian Government and state and territory governments, with funding provided through the Australian Government Department of Employment, Skills and Family Business.

COVER IMAGE: GETTY IMAGES

ISBN 978-1-925717-35-8

TD/TNC 136.06

Published by NCVER, ABN 87 007 967 311

Level 5, 60 Light Square, Adelaide SA 5000

PO Box 8288 Station Arcade, Adelaide SA 5000, Australia

Phone +61 8 8230 8400 Email ncver@ncver.edu.au

Web <https://www.ncver.edu.au> <https://www.lsay.edu.au>

Follow us:  <https://twitter.com/ncver>  <https://www.linkedin.com/company/ncver>



About the research

School-to-work pathways

Rasika Ranasinghe, Emerick Chew, Genevieve Knight and Gitta Siekmann, NCVER

It is well established that a successful transition to the labour market has long-term social and economic implications for both individuals and society. However, the journey from school to the world of work is not straightforward and needs to be better understood.

Based on data from the 2006 cohort of the Longitudinal Surveys of Australian Youth (LSAY Y06), this research explores the school-to-work transitions of Australian youth aged 16 to 25 years. The study uses sequence analysis in combination with cluster analysis to summarise complex longitudinal data in a meaningful way and to investigate transitions in their entirety as 'pathways'.

This study captures the richness of the transition experience, both visually and analytically. Identifying the five key types of pathways taken by young people on their journey from school to work, this research describes these pathways and the implications of their evolution for labour market destinations over the 10 years from 2006 to 2016, when the cohort was aged 25 years.

Key messages

- Young people experience diverse and individualised school-to-work pathways. While the majority of young people in the study sample followed a generally simple higher education-to-work pathway or entered full-time work relatively early, some experienced complex post-school pathways, with frequent switching between higher education and vocational education and training (VET) activities, episodes of part-time work and repeatedly moving in and out of the labour market. The five key pathways followed by youth aged 16 to 25 years revealed by the analysis are:
 - Pathway 1: Higher education and work
 - Pathway 2: Early entry to full-time work
 - Pathway 3: Mix of higher education and VET
 - Pathway 4: Mixed and repeatedly disengaged
 - Pathway 5: Mostly working part-time.
- The factors that influenced specific pathways include studying VET subjects at school, individual school achievement and socioeconomic issues.
- The occupational outcomes at age 25 years differed for the various pathways.
- VET was involved in several pathways and emerges as an important avenue in school-to-work transitions which culminate in work at age 25 years.
- In Pathway 2, VET provided a direct and early route to work, resulting in 97.4% of these young people being in work at age 25 years – the highest proportion of any of the pathways – and they worked full-time for the longest, 69.8 months on average during the 10 years. Almost half had undertaken apprenticeships/traineeships, with the highest occupation group being technical and trades. This pathway was characterised by more males.
- Females who undertook VET had more often followed Pathways 3 and 5 and were mostly in work at the age of 25 years (91.7% and 90.2% respectively).

Contents



Tables and figures	5
Executive summary	6
Profiling the pathways	6
The role of VET in youth transitions	8
Implications and extensions	9
Introduction	10
Background	10
Data and method	11
Describing young people's pathways	16
Visualising the pathways	17
What factors contribute to different pathways?	33
Likelihood of following a given pathway relative to Pathway 3 (mix of higher education and VET)	36
Pathways associated with occupational outcomes at age 25 years	38
References	40
Appendix A: data and variable definitions	42
Variable derivation	42
Variable definitions	45
Appendix B: detailed methodology	54
What are sequences?	54
The Chi-square distance measure	54
Cluster analysis	55

Tables and figures

Tables

1	Profile of selected variables in 2006 LSAY and the research sample (proportions %)	15
2	Average number of months spent in each activity between ages 16 and 25 years by pathway	28
3	Number of transitions between age 16 and 25 years by pathway proportion of individuals (%)	29
4	Selected descriptive statistics of the sample: socioeconomic and demographic characteristics by pathway (sample proportions %)	30
5	Highest qualification, occupation and labour force status at age 25 years by pathway (sample proportions %)	31
6	All VET qualifications and apprenticeships/traineeships by age 25 years by pathway (sample proportions %)	32
A1	Derivation of variables	42
A2	Description and labels of characteristic variables	45
A3	Selected socioeconomic characteristics of the sample by pathway (with standard errors)	48
A4	Highest qualification, occupation and labour force status at age 25 years by pathway (with standard errors)	49
A5	Tests for proportions (in tables 4 and 5)	50
A6	Wald tests for independent variables (in table 7)	51
A7	Wald tests for independent variables (in table 8)	52
A8	Model fit tests for multinomial logistic model	53

Figures

1	Pathway 1: Higher education and work (60% of the sample)	20
2	Pathway 2: Early entry to full-time work (23% of the sample)	22
3	Pathway 3: Mix of higher education and VET (8% of the sample)	24
4	Pathway 4: Mixed and repeatedly disengaged (5% of the sample)	26
5	Pathway 5: Mostly working part-time (4% of the sample)	28
6	Occupation at age 25 years (% of pathway)	39
A1	Summary activity patterns for the entire study sample (n = 3186)	47
B1	Example of a sequence	54



Executive summary

The passage into work is a critical phase in young people's lives, with long-term implications for the future labour market and for social outcomes. An evolving labour market adds to the complexity of trajectories, further confounding youth transitions and highlighting the importance of understanding transitions as a process.

In this context, the current study has the following objectives:

- identifying the different types of pathways followed by young people in their journey from school to work, with the aim of developing a topology of transitions
- using this information to obtain a better understanding of the characteristics of young people in different pathways.

The analytical approach used in the study combines sequence analysis and cluster analysis in order to identify similarities between activity patterns. It enables the study of labour market transitions as a sequence of activities and exploits the longitudinal nature of the data by using a series of graphical representations, these providing a direct visual insight into the patterns of transition within each pathway. To accompany this report, an interactive data visualisation, *Visualising school-to-work pathways using LSAY*, presents the school-to-work pathways of young Australians aged 16 to 25, and can be accessed from <https://www.ncver.edu.au/research-and-statistics/school-to-work-pathways>.

This study is based on data from the 2006 cohort of the Longitudinal Surveys of Australian Youth (LSAY Y06). LSAY is a nationally representative survey that tracks 15-year-old students as they move from school to further education or other destinations until they are 25 years of age. The survey captures detailed information on education activities and employment, as well as socioeconomic and demographic characteristics, making it ideally suited for this analysis. The study sample is limited to the 3186 individuals who completed each annual survey until 2016. While analytical and data constraints limit the ability to generalise the results of this study, the initial LSAY sample is representative of the youth population of Australia and thus is useful in providing important insights into their transition pathways.

Profiling the pathways

The five pathways taken by young Australians between the ages of 16 to 25 years as they transitioned from school to work are:

Pathway 1: Higher education and work

This represents the largest group (60% of the sample) and encompasses an extended period of post-school higher education, followed by employment.

- This is a relatively simple pathway and is basically an academic track, whereby students enrol in university upon leaving school and have a prolonged higher education period before transitioning into employment.
- This pathway contains the highest proportion of youth from metropolitan areas, who have the highest socioeconomic status and who completed Year 12.

- This pathway also has the lowest proportion of youth with an Indigenous background or who were married or had children early, and the fewest number of young people who undertook vocational subjects during secondary school.

Pathway 2: Early entry to full-time work

This is an ‘express pathway’ to employment and includes apprenticeships and traineeships. This pathway comprises a relatively short spell (14.3 months on average) of post-school education or training, leading to full-time work (23% of the sample). For many respondents, however, it is likely that training jointly in combination with full-time work extends beyond early post-school years, as part of an apprenticeship or traineeship. Young people in this pathway have the fastest entry to employment and also spend the longest time in work.

- This is a predominantly male pathway, with a high proportion undertaking vocational subjects in secondary school; almost half had undertaken apprenticeship/traineeships by the age of 25 years.
- This pathway contains the highest proportion of young people who were married by the age of 25 years, and the highest proportion in technical and trades occupations at the age of 25 years.

Pathway 3: Mix of higher education and VET

This pathway comprises an extended period of higher education and VET activity, eventually leading to more stable employment or further VET activity (8% of the sample).

- Youth in this pathway have a relatively complex trajectory, with frequent switching between university and VET activities.
- This pathway is predominantly comprised of females, with a large number engaged in VET activities after the age of 20 years; they also spend the highest average number of months (35.2) in post-school VET activities and hold the most VET qualifications by age 25 years. At this age, 26.8% held a bachelor’s degree as their highest qualification, while a further 25.6% held an advanced diploma/diploma qualification, and 15.4% held a certificate IV.
- The highest proportions of these young people are working as professionals (20.1%) and community and personal service workers (22.8%), and in clerical and administrative occupations (16.5%) at age 25 years.

Pathway 4: Mixed and repeatedly disengaged

This pathway is characterised by multiple and repeated labour market movements and disengagement, indicating tenuous labour market attachment (5% of the sample).

- This represents the most complex pathway and contains the highest proportion of young people experiencing more than 10 transitions between the ages of 16 and 25 years.
- Young people in this pathway spend the highest average number of months disengaged from the labour market (16.2 months) or unemployed (41.2 months), with 53.1% not working at age 25 years.
- This pathway has the highest proportions of vulnerable youth, indicated by the higher incidence of teenage marriages or parenting, disability, early school leavers and youth from the lowest socioeconomic status (SES).

Pathway 5: Mostly working part-time

This represents the smallest group (4% of the sample), a group characterised by relatively early entry to the labour market and mostly employed part-time over the 10 years.

- Youth in this pathway spend the most time in part-time employment between the ages of 16 to 25 years.
- They hold the least qualifications of all the pathways (with the highest share, at 17.9%, holding a certificate III), and 50.9% have no post-school qualifications at the age of 25 years. They also spend the least amount of time in post-school education.
- At the age of 25 years, young people in this pathway are primarily in community and personal services (26.8%), sales (18.8%) and clerical and administrative occupations (12.5%).

The role of VET in youth transitions

The modelling suggests that a number of factors influence pathway choice, with school education and achievement playing key roles. The modelling shows that studying a vocational subject at school age is a significant positive factor in all non-academic pathways, in particular raising the probability of the more employment-oriented Pathway 2 by 13 percentage points. Attaining less than the top school maths and reading achievement by age 15 years also raises the probability of Pathway 2.

Personal backgrounds are also shown by the modelling to play a role, with a less advantaged socioeconomic background raising the probability of Pathway 2, while an overseas background lowers the probability of Pathway 2.

Males are more likely to follow Pathway 2, with a probability of 14 percentage points higher, and have lower chances of taking Pathway 1, Pathway 3, and Pathway 4.

Further modelling relating to Pathway 3 enables greater understanding of the factors determining which type of VET pathway is followed. The modelling confirmed that the relative likelihood of following Pathway 2 which had the highest rate of apprenticeships/traineeships was mostly increased by studying vocational studies in school but was not associated with school maths or reading test achievement. However, being male increased the chance of Pathway 2 by 2.9 times and having the most disadvantaged socioeconomic background increased the likelihood of this pathway by 1.6 times.

The transition pathways uncovered in this study confirm that, for some young people, by providing training opportunities for them, VET is an important means for facilitating pathways to the labour market. This is particularly true for young people in Pathway 2, where VET gave early entry to employment (mostly males), and for those in Pathway 3, who had an extended period of mixed VET and higher education activities (mostly females). Pathway 2, whereby VET provided a direct route to employment, resulted in 97.4% in work at the age of 25 years, the highest for any pathway. Note that in this pathway VET training started within school and extended beyond early post-school years but was mostly in combination with full-time work as part of an apprenticeship or traineeship.

Implications and extensions

This research, which summarises complex information relating to 10-year transition pathways, demonstrates that young people's post-school pathways are diverse, individualised and complex, underscoring the importance of understanding youth transitions as a process. The research also revealed the significant avenue offered by vocational education and training for those young people exploring alternative non-academic pathways.

Introduction

The journey into work is an important phase in young people's lives, with long-term implications not only for the individual, but also for society and the economy.

Following global trends, school-to-work transitions among young Australians have changed considerably over the last few decades. Despite Australia having escaped the worst of the Global Financial Crisis (GFC), labour market issues associated with the GFC, such as youth unemployment and those who are not in education, employment or training (NEET), persist. In 2017, nearly 14% of youth aged 15–24 years were unemployed, compared with just under 6% for the rest of the population (ABS 2017). In this context, there is renewed interest in developing an understanding of the nature of youth transition pathways.

Using data from the 2006 cohort of the Longitudinal Surveys of Australian Youth, this study aims to untangle the complex transition pathways of school leavers as they enter tertiary education and the labour market over a 10-year period. This is a useful exercise, because, as the Organisation for Economic Co-operation and Development (OECD 2000, p. 149) noted, the school-to-work transition is a process, and not merely a single event at a point in time:

More and more, it seems, periods of work and learning are being spread throughout life, rather than being concentrated in distinct and separated periods: education in the period up to the late teens and early 20s; work after that. A mingling of work with education is also being observed increasingly during the period of initial education.

To explore this process we use *sequence analysis*, an approach allowing the study of labour market transitions as a sequence of activities. Sequence analysis methods consider multiple transitions, as well as their complexity and timing, enabling a fuller comprehension of the entire pattern. The underlying notion is that a person can engage in a number of different activities over a period of time. While there are variations in individual sequences, it is possible to discern patterns in sequences, with groups of similar sets of sequences subsequently categorised into *pathways*. Such an approach allows more features to be incorporated when describing young peoples' transition to employment.

This study has the following objectives:

- identifying the different types of pathways followed by young people in their journey from school to work, with the aim of developing a topology of transitions
- using this information to obtain a better understanding of the characteristics of young people in different pathways.

Background

A number of international empirical studies have incorporated sequence analysis in investigations of youth transitions: McVicar and Anyadike-Danes (2002) for Northern Ireland; Aassve et al. (2007) who selectively followed work family transitions for young British women; Brzinsky-Fay (2007), who compared school-to-work transitions across 10 European countries, and Quintini and Manfredi (2009), who made comparisons between 19 European countries and the United States. School-to-work transitions were explored in Spain by Corrales-Herrero and Rodriguez-Prado (2012). Dorsett and Luccino (2014), Schoon and

Lyons-Amos (2016) and Anders and Dorsett (2017) explored UK data on this topic, while Albæk et al. (2015) and Lorentzen et al. (2018) analysed the transition patterns of young people in Denmark, Finland, Norway and Sweden, with Brzinsky-Fay and Solga (2016) exploring patterns in Germany. A common emerging theme of these findings is that entrenched disadvantages can push young people into long-term unemployment or disengagement and social exclusion. Specific risk factors identified included low education attainment, early marriage or child bearing, and a disadvantaged socioeconomic background.

In Australia, some earlier LSAY studies have focused on youth transitions and include:

- Buddelmeyer and Marks (2010), who used LSAY data from the 1995 cohort to analyse the annual transitions of 15 to 25-year olds who had completed a post-school qualification or training. They found that an individual's previous year's labour market state has the most significant implication for their current state.
- Fitzpatrick et al. (2011) used the 1995 cohort of LSAY to examine the average time taken to find employment after leaving education and found that those with a post-school qualification had the fastest entry to employment, while those who did not complete school took longer.

Two studies used sequence analysis for labour market transitions based on data from the Household, Income and Labour Dynamics in Australia (HILDA) Survey¹:

- Yu et al. (2012) examined individual vocational pathways for a sample of working-age individuals (age 18–64 years) in Australia using HILDA data. They discussed the patterns of occupational transitions for selected industries by different characteristics such as gender and age group, including those aged 18–24 years, but the overall analysis was not focused on youth transitions.
- Fry and Boulton (2013) used sequence analysis for labour market transitions in Australia that included a youth focus. Using HILDA data, they provided a descriptive analysis of the transition pathways in Australia they considered were likely to evolve over the life cycle. Their analysis by age group, including youths (15 to 24-year-olds), identified five pathways. Three of these were associated with increasing education levels and transitions to work; one was associated with moving in and out of work; and one was dominated by young women withdrawing from the labour force to raise children. Their findings indicated that many transitions in and out of work is a dominant pathway for youths and represented about 52% of all youths in their study.

Data and method

The 2006 cohort of LSAY (LSAY Y06) is used to produce the sequence analyses. As noted earlier, LSAY is a nationally representative survey that tracks 15-year-old students who are

1 A household-based panel study that collects valuable information about economic and personal wellbeing, labour market dynamics and family life. Started in 2001, the HILDA Survey follows the lives of more than 17 000 Australians each year. It collects information on many aspects of life in Australia, including household and family relationships, income and employment, and health and education. Participants are followed over the course of their lifetime
<<https://melbourneinstitute.unimelb.edu.au/hilda>>.

in school at their first interview, as they move from school to other destinations until they are 25 years of age. The survey collects information on education activities (school, transitions from school, post-school education and training), employment (job history, job search and mobility), living arrangements and health, as well as socioeconomic and demographic information. This makes it well suited for the analysis of the school-to-work pathways of young people.

LSAY Y06 consists of 14 170 students from across Australia in 2006. The sequence analysis required all 10 years of LSAY surveys to have been completed; the sample therefore consists of the 3186 individuals who completed each survey to 2016. Since all respondents were in school in 2006 when they were first interviewed, the analysis period allows us to follow 10 years of observed transitions.

Deriving monthly activity status

We have selected the LSAY monthly education and labour market activities of respondents for each year in the survey between 2007 and 2016. This information offers some advantages, in that the youth labour market often displays high volatility, meaning that monthly activity data can provide a finer level of analysis for describing youth transitions and their complexity. However, the monthly data include some further recall errors, which must be accommodated. The errors occur because the (monthly) data are based on individuals' recall of the events but are collected in the annual interview.² The schooling and other post-school education and training activities are used to construct the transitions, but while some of these are recorded at the annual survey others are only within the monthly recall data. Therefore, a number of additional, related, variables on schooling and post-school education/training activities were used in conjunction with the annual interview data to reconstruct monthly activities over the 10-year period. Further descriptive information about the youths is drawn from the annual interviews, some of which are collected only in the first-year interview or in selected years. Details on the derivation of variables are given in appendix A.³

Sequence methods are not tolerant of missing cases for variables (item non-response), although these are relatively few in the LSAY sample we use. A tailored inference approach was used to address these; appendix A indicates how this was addressed for each variable and the category to which missing cases were allocated.

Each young person in the sample had 120 months (12 months x 10 years) of activities. For analytical purposes, the activities were defined to be education in instances of multiple states, and activities were defined to be mutually exclusive. For example, if a young person was reported as engaged in both vocational education and employed for a given month, the monthly activity status would be classified as VET. It is also important to note that those identified as being at 'school' may have been taking part in VET as part of their studies (see table 4). There is the possibility of some weaknesses in this identification because, as with all surveys, there may be some oversimplification due to survey errors in recording at the

2 Each year, the survey collects detailed recall information on individual employment histories based on monthly employment calendars. The calendar data contain detailed information on what happened between the annual interviews, including full-time or part-time employment status and any periods of unemployment. Table A1 in appendix A describes how these are used to form activity variables.

3 See tables A1 and A2 in appendix A for details of derived variables and variable definitions.

interview what is a person's main status if they are undertaking more than one status jointly at a time point, (a chief example is apprenticeship/traineeship which inherently combines joint full-time work and training statuses), and as already discussed there can be further survey error in this activity identification arising from recall for the monthly information gathered about the past year activities. For Pathway 2, where the share with apprenticeships was quite high, and full-time employment was quite high, but training was briefer than most apprenticeships that result in technical and trades related occupations, survey error may explain this slightly odd combination of results (i.e. during the interview, due to the joint status, some period of their apprenticeship training was described as employment). Despite these aspects of survey error which cannot be redressed, the results still reflect a reasonable simplified description of the transitions recorded for these young people.

This classification format was selected to simplify the number of monthly activity states and the resultant sequence combinations. Using the information available, seven monthly activity states were derived:

- school
- university
- VET (includes apprenticeships and traineeships)⁴
- employed – full-time
- employed – part-time
- unemployed
- not in the labour force (NILF)/not in education, employment or training (NEET).⁵

Sequence and clustering analysis approach to youth transitions

Sequence analysis provides a method to capture a series of activity transitions over a time period. It allows for the utilisation of longitudinal data, enabling the identification of dominant patterns. The main task of the clustering in the analysis is to reduce complexity by comparing, sorting and grouping sequences. This results in groups of transition types that can be used for further analyses (see appendix B for methodological details).

The specific sequence and clustering analysis approach utilised in this study consists of the following steps:

- Derive the sequence of activities for each individual and quantify the dissimilarities (using the 'distance') between each possible pair of sequences.
- Use the 'distance' between sequences to classify them into clusters of pathways (cluster analysis).

4 Note that the sample size of apprenticeships/traineeships is too small (1.9% of the sample) to form a separate category. It is therefore combined with VET to form one vocational education category for analytical purposes.

5 Note that individuals who are NEET could be unemployed (and hence out of work, looking for work and so still in the labour force but not in education or training) or Not in the labour force (NILF) where they are not looking for work. Those who are NILF and not in education or training at a given month are considered *inactive* NEET, referred to as NILF/NEET in this study.

Note on data and methodological limitations

Several limitations to this methodology are discussed in the sequence analysis literature. Those relevant to this study mostly relate to the inability of existing methods to accommodate sample weights – either in the sequence analysis or the clustering process – to make the findings representative of the general population of 16 to 25-year-old Australians. Additionally, the sequence analysis methods used here require complete information on monthly activities over the entire study period. Thus, the study sample is limited to those who completed all 10 waves of the survey, resulting in the loss of information for the cases who failed to complete the survey in these subsequent years. In addition, any incomplete answers to the survey questions used as descriptive variables needed to be inferred.⁶ These aspects mean that this LSAY analysis may not remain representative of the general population of 16 to 25-year old Australians from 2006.⁷

Table 1 shows the initial Y06 cohort and the research sample who completed all surveys until 2016, for selected variables. The research sample had a higher tendency to be in the highest achievement quartile in mathematics and reading, based on their PISA⁸ scores at age 15 years. The loss of students from the lower achievement quartiles indicates that the research sample used in the sequence analysis tends to describe higher-achieving school students from LSAY Y06 (PISA), higher numbers from higher socioeconomic groups, fewer Indigenous young people, higher numbers from metropolitan locations, and more who undertook vocational studies in school than were in the first survey of LSAY Y06. The results should be interpreted in this context.

6 See table A2 in appendix A for how missing (item non-response i.e. non-response for a question) are treated within each variable description.

7 The potential biases due to attrition are an acknowledged limitation of sequence analysis. Most sequence analysis studies base the discussion of results only on the profile of the sample (Studer, Struffolino & Fasang 2018; Yu et al. 2012). As discussed in detail by Yu et al. (2012), sequence analysis methodology is inherently exploratory, without any statistical tests or prior hypotheses on how sequence patterns are generated. See, for example, Blanchard (2016) and Studer, Struffolino & Fasang (2018) for detailed a discussion on the progress of social sequence analysis and evolving methodological developments.

8 The Programme for International Student Assessment (PISA) is a worldwide assessment by the OECD and measures the performance of 15-year-old school pupils in mathematics, science and reading.

Table 1 Profile of selected variables in 2006 LSAY and the research sample (proportions %)

	2006 LSAY	Research sample*
Male	50.8	47.4
Indigenous	7.6	3.3
Overseas background	39.8	39.8
Language other than English	8.7	7.4
Metropolitan location	67.6	71.2
Socioeconomic status SES quartile		
Highest quartile	25.8	33.6
Third	24.8	28.2
Second	24.7	21.8
Lowest quartile	24.7	16.4
Vocational studies in school	16.5	25.3
Age 15 mathematics quartile (PISA)		
Highest quartile	25.0	42.1
Third	24.9	28.5
Second	25.1	20.1
Lowest quartile	25.1	9.2
Age 15 reading quartile (PISA)		
Highest quartile	24.9	41.8
Third	25.1	29.8
Second	24.9	19.3
Lowest quartile	25.1	9.1
Sample number	14,170	3,186

Notes: * The sample consists of Y06 respondents who participated in the survey until 2016.

See appendix A, table A2 for variable derivations

Source: LSAY 2006.



Describing young people's pathways

Pathway categories

The sequence and cluster analysis identified five pathways, as described by the following:

- **Pathway 1: Higher education and work**

The majority of the young people in this pathway have an extended period of higher education, followed by employment (60% of the young people).

- **Pathway 2: Early entry to full-time work**

About one-quarter of the young people (23%) followed an 'express pathway' to employment, distinguished by a short spell of post-school education or training (mostly VET) leading to full-time work approximately one year after leaving school. For many respondents, however, it is likely that training extended beyond the early post-school years; that is, in combination with full-time work as part of an apprenticeship or traineeship.

- **Pathway 3: Mix of higher education and VET**

This pathway comprises an extended period of higher education and VET activity, combined with short and intermittent episodes of employment, eventually leading to employment or further VET activity (8% of the sample).

- **Pathway 4: Mixed and repeatedly disengaged**

Young people in Pathway 4 undertook repeated labour market changes, with periods in the gap status of 'not in the labour force' or 'not in education, employment nor training'. While only a small proportion of the sample (5%) fell into this category, it indicates tenuous labour market attachment.

- **Pathway 5: Mostly working part-time**

The 4% of the sample followed in this pathway, entering the labour market relatively early and are mostly employed part-time.

It is noted that the pathways that emerged from this analysis are based on the sample of respondents who participated in *all* surveys of LSAY between 2006 and 2016 (3186 young people). As discussed earlier, the research sample used in the sequence analysis however tends to describe the higher-achieving school students from LSAY Y06 (PISA), with more from higher socioeconomic groups (SES), fewer Indigenous young people, more from metropolitan locations, and more who undertook vocational studies in school than those who participated in the first survey. This profile is partly reflected in the pathways identified, whereby the largest group of these young people in the sample follow a higher education-to-work pathway (Pathway 1).

Building a profile of pathways

Describing the characteristics of young people and their activity patterns for each of these pathways provides a useful account of their educational and labour market transitions between the ages 16 and 25 years. The section outlining each pathway provides a descriptive analysis of the pathways, with these descriptions summarising information drawn from the following analyses:

- Table 2 presents the average number of months spent in each activity, according to pathway, and demonstrates how time spent on activities is distributed within pathways.
- Table 3 shows the distribution of the number of transitions (defined as a shift from one activity to another) in each pathway, which is a proxy indicator of the complexity of pathways. For example, if a large proportion of individuals within a pathway experience only a few transitions over the 10-year follow-up period, it will be a relatively simple pathway compared with one where a high proportion experience a large number of transitions. For example, table 3 shows that 30% in Pathway 2 experienced up to five transitions over 10 years, making it a relatively simple pathway. However, Pathway 2 also included the combination of training and full-time work that extended beyond early post-school years as part of an apprenticeship or traineeship. In contrast, Pathway 4 included 32.1% of young people with more than 15 transitions across the 10-year period.
- Tables 4 and 5 provide additional socioeconomic and demographic characteristics for the individuals in each pathway.

In addition to the observed individual and family background characteristics, table 4 also includes initial education variables such as mathematics and reading achievement levels, based on the initial survey PISA data⁹ available for each young person, and their participation in vocational subjects at school.

Table 5 provides the highest qualification, employment and occupational outcomes of individuals in the sample at the age of 25 years, which helps to illustrate their labour market destinations. Table 6 provides additional information on all VET qualifications and apprenticeships/traineeships acquired by individuals by the age of 25 years and includes any multiple qualifications obtained during the 10-year period. This adds information to that of their highest qualification, showing how many and which type of credentials the young people had acquired by the age of 25 years.

Visualising the pathways

The most appreciated aspect of sequence analysis is the capacity to provide a visual analysis, which is shown in figures 1 to 5 where these transitions are graphically represented. There are three components to each pathway visualisation: the sequence index plots (panel a); chronographs (panel b); and modal plots (panel c). Figures 2 to 6 illustrate how each pathway can be shown using three different formats.¹⁰ Each section

⁹ As part of PISA 2006, students were assessed in mathematical and reading literacy. Details on the assessments and reporting are available at <<http://www.oecd.org/pisa/pisaproducts/42025182.pdf>>.

¹⁰ Summary visualisations for the entire sample are available in appendix A, figure A1, and give an overview of the monthly activity patterns over the 10-year period.

provides an explanation of the features of each pathway, which are described separately within each pathway figure.

- **Panel (a): Individual activity sequences (sequence index plot)**

Sequence index plots illustrate the activity sequences by using colour-coded horizontal stacked bars to indicate how young people moved between states over time. Individuals are numbered along the vertical axis, with time shown on the horizontal axis. This format provides a longitudinal perspective of the sequences and allows the pathway to be observed in its entirety.

For example, in figure 1 (panel a), most young people in Pathway 1 remained in school (yellow) until 2009 (aged approximately 18 years old). Then there is a brief gap before starting university education (pink). Some will complete their undergraduate education within three to five years and enter the labour market, while others will continue with further, postgraduate, education. It is possible here to observe those who have periods of unemployment or not in the labour force/not in education, employment nor training (orange).

- **Panel (b): Monthly proportion of activities (chronograph)**

Chronographs present the proportion of young people participating in each activity for each month. The vertical axis indicates the proportion of individuals, while the horizontal axis shows the time. They do not provide any information on the longitudinal nature of the data or the duration of activities.

By way of illustration, almost everyone in Pathway 1 was in school (yellow) in 2007 (figure 1, panel b). The share of individuals in school gradually falls over the following two years. At the beginning of 2009 those still in school fall to just under 80% of the entire sample. At the end of the 10-year period, in 2016, around 70% are in full-time employment (green) and about 10% in part-time employment (purple).

- **Panel (c): Most frequent (modal) activity for each month**

Modal plots provide yet another perspective on the transition pathways by illustrating the most frequent activity undertaken during each month, and the proportion of individuals participating in that activity. They are useful in identifying the dominant activity at different times. In other words, this format depicts the most popular activity undertaken by young people each month.

Figure 1, panel c, shows that those in Pathway 1 spent the most time in school (yellow) between 2007 and 2009. After a brief period (approximately two months) of unemployment or not in the labour force/not in education, employment or training, they begin university education (pink), which remains the most frequent activity until 2015, after which full-time employment (green) takes over.

Pathway 1: Higher education and work

Key characteristics – 60% of young people

- This is a relatively simple pathway that represents an academic track, whereby students enrol in university upon leaving school and have a prolonged higher education period before transitioning into employment.
- This pathway contains the highest proportion of youth from metropolitan areas, the highest SES quartile, the highest proportion with an overseas background and the highest proportion completing Year 12.
- This pathway had the lowest proportions of youth with an Indigenous background, of those who were married or had children early, and of youths who undertook vocational subjects during secondary school.

The 'higher education and work' pathway is the largest group, in that it contains 60% of the sample's individuals. In this pathway, most young people make a relatively smooth transition from school to higher education and remain in education for an extended period before entering full-time employment (see figure 1, panel a). They spent an average of 55.4 months at university, making it their dominant activity. They also spent the most months in school (21.5 months) compared with individuals in other pathways (table 2).

As evident from table 3, Pathway 1 is the simplest, with one-quarter (26.8%) of individuals having five or fewer transition shifts between the ages 16 and 25 years. Just over one-half (52%) had between six and 10 transition shifts, and only 3.1% had more than 15 transitions over the 10-year period. Overall then, the young people in Pathway 1 followed a mostly uniform and simple trajectory.

Approximately three-quarters of individuals following Pathway 1 were from metropolitan areas (76.8%). They also tended to be from a higher SES background, with 43.4% belonging to the highest SES quartile. Compared with the other pathways, this group has the highest proportion of youth identified from overseas backgrounds. In a study comparing education attainment and choice of post-school studies by country of origin, Parasnis and Swan (2017), using the 2003 cohort of the LSAY, found similar outcomes. This artefact is also identified in 2016 Australian census data.¹¹

Consistent with the higher education focus of Pathway 1, there is a higher proportion of those aged 15 years who are high achievers in mathematics and reading (table 4) and 69.4% had obtained a university bachelor's degree as their highest qualification by the time they were 25 years old (table 5). A further 10.6% had a postgraduate qualification. Only 10.8% reported having no qualification. About one-half of them were in professional occupations (50.1%) by age 25 years. Only 3.8% had participated in an apprenticeship/traineeship (table 6).

Table 5 shows that by the age of 25 years 92.4% of young people in this pathway were employed, while 5.3% were not in the labour force. This higher education pathway has slightly fewer in work at age 25 years than does Pathway 2.

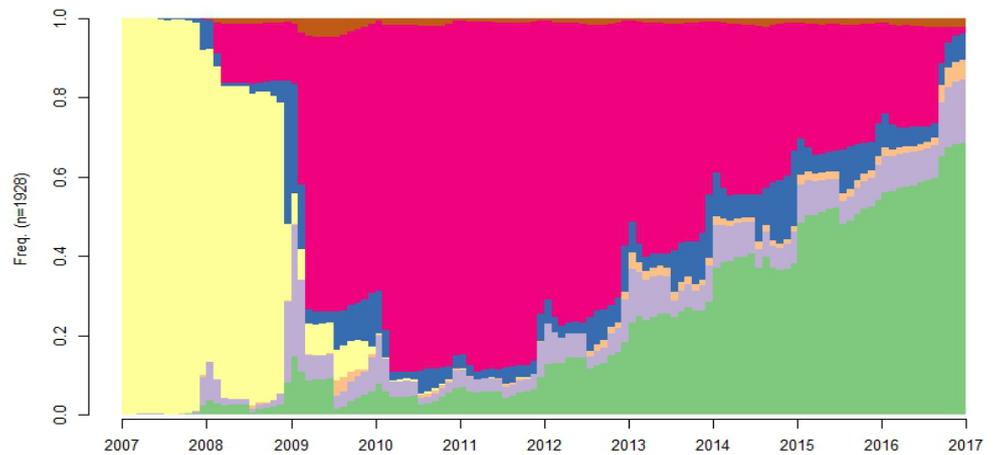
11 <<https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/2071.0-2016-Main%20Features-Educational%20Qualifications%20Data%20Summary%20-65>>.

Figure 1 Pathway 1: Higher education and work (60% of the sample)

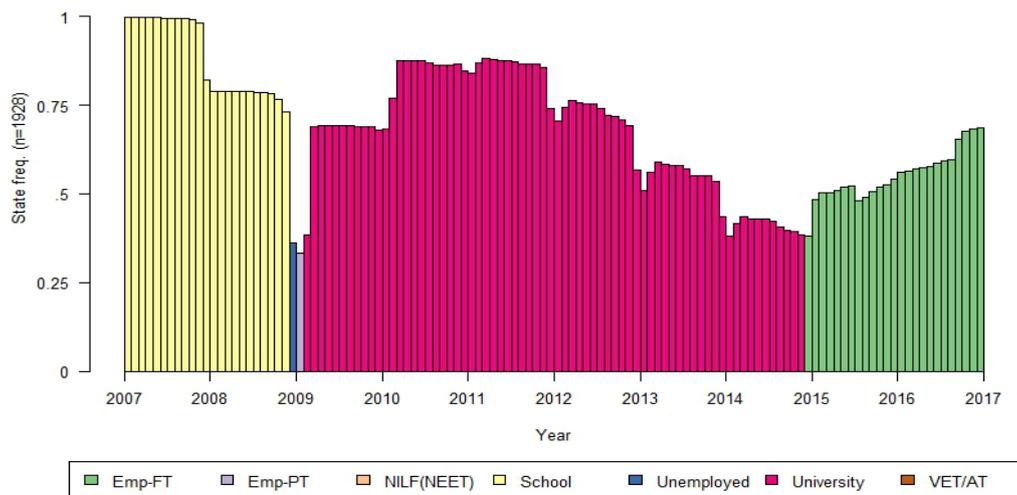
(a) Individual activity sequences



(b) Monthly proportion of activities



(c) Most frequent (modal) activity for each month



Notes: Emp-FT = full-time employment; Emp-PT = part-time employment; NILF = not in the labour force; NEET = not in education, employment nor training; VET/AT = vocational education or training/apprenticeship/traineeship.

Pathway 2: Early entry to full-time work

Key characteristics – 23% of young people

- Young people in this pathway had the fastest entry to employment.
- Predominantly male and comprising 44.4% who had studied vocational studies at school, they also spent the longest time in employment (69.8 months in full-time employment).
- The pathway contains the highest proportion of young people who were in work by age 25 years and married by the age of 25 years, and had the highest proportion in technical and trades occupations.

This pathway represents 23% of the sample. In the ‘early entry to full-time work’ pathway, young people spent 69.8 months on average in full-time employment, a figure signifying the highest number of months spent in any activity across all pathways.

The data indicate that the young people spent a relatively short time (14.3 months) in post-school VET or apprentice/trainee activity before moving onto full-time employment; however, in investigating this pathway, it is important to note that apprenticeships/traineeships were undertaken by 47.3% (table 6). For many respondents, it is likely that training jointly in combination with full-time work extends beyond early post-school years in both training and full-time work, in an apprenticeship or traineeship, and the activity sequences in figure 2 likely under-report the extent of their ongoing training (see data and methods section).

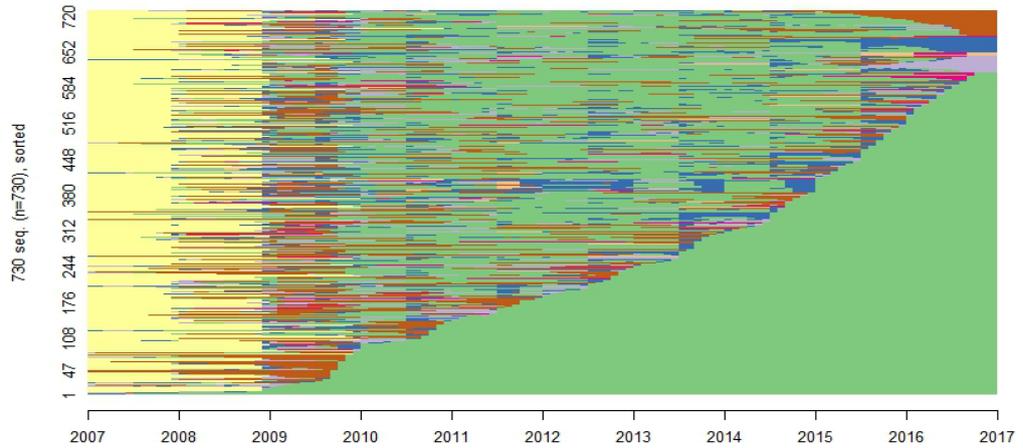
Youth in this pathway experienced a relatively simple trajectory, with 30% having one to five activity transitions, and a further 39.2% with six to 10 transition shifts during the 10-year period (table 3). As illustrated in table 4, this pathway is predominantly male (64.1%), with a relatively high share of early school leavers (21.5%) and those who belonged to the lowest mathematics and reading achievement quartiles at age 15 years (41% and 44%, respectively). Young people following this pathway also tended to establish families at a relatively young age, with 15.6% married between the ages 15 and 19, and 13.6% with dependent children by the time they were 25 years old.

Table 5 indicates that individuals in Pathway 2 were mostly VET-qualified by age 25 years, with certificate III (25.9%), certificate IV (13.7%) and advanced diploma/diploma (13.2%) their main three categories of highest qualification. Yet, a further 29.2% of young people reported having no qualification. Table 6 shows all credentials gained, revealing that 47.3% of young people in this pathway had taken up an apprenticeship or traineeship by age 25 years; 49% had certificate III and/or IV level qualifications.

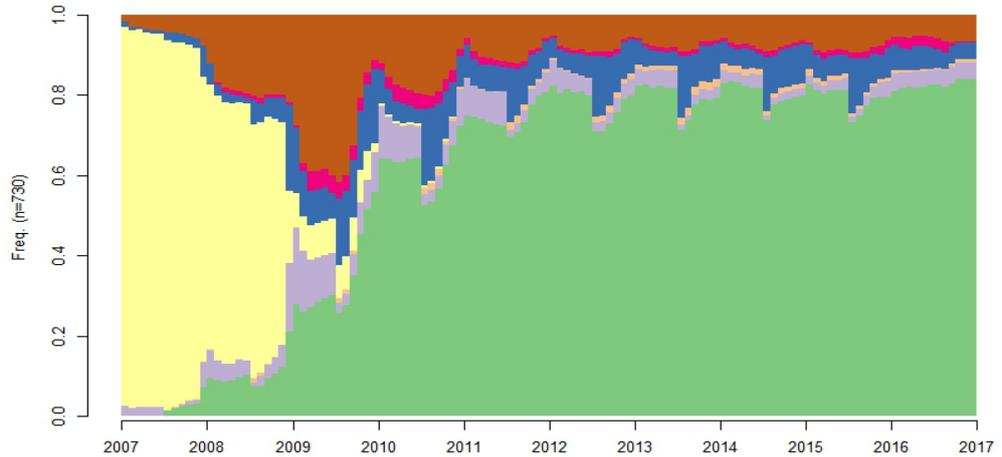
A large proportion of young people were employed in VET-related occupations at age 25 years. The highest occupational share is technical and trades related (29.3%). This was followed by clerical and administrative (16%) and community and personal service workers (11.4%) (table 5). Only 1.2% of young people in Pathway 2 remained unemployed at age 25 years, while another 1.4% were not in the labour force. Pathway 2 had the highest employment at age 25 years (97.4%) of all the pathways.

Figure 2 Pathway 2: Early entry to full-time work (23% of the sample)

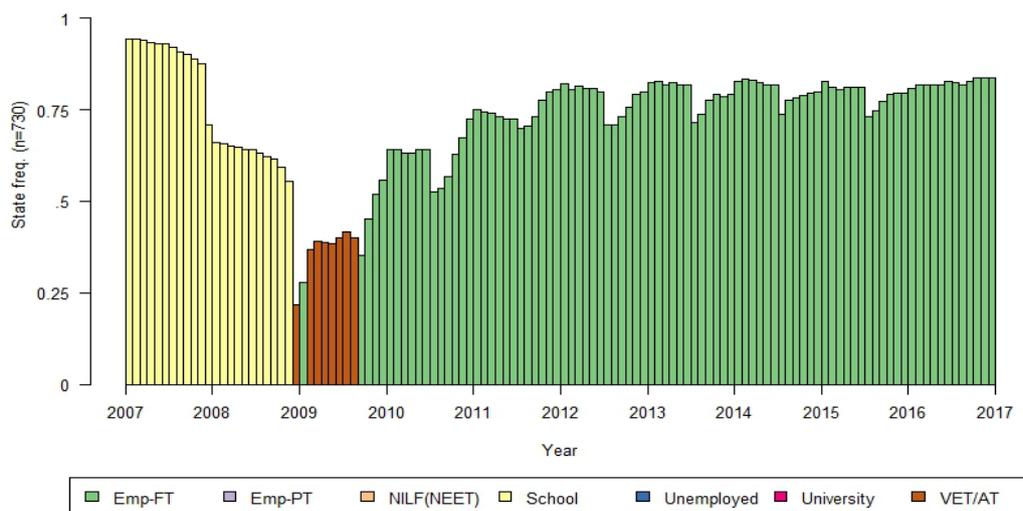
(a) Individual activity sequences



(b) Monthly proportion of activities



(c) Most frequent (modal) activity for each month



Notes: Emp-FT = full-time employment; Emp-PT = part-time employment; NILF = not in the labour force; NEET = not in education, employment nor training; VET/AT = vocational education or training/apprenticeship/traineeship.

Pathway 3: Mix of higher education and VET

Key characteristics – 8% of young people

- Youth in this pathway tended to experience a relatively complex trajectory, with frequent switching between university and VET activities.
- Predominantly female, a large share of this group engaged in VET activities after the age of 20 years. They also spent the highest number of months in post-school VET activities of all of the pathways and held the highest number of VET qualifications.
- This pathway supported the highest proportion of young people in the clerical and administrative work occupation group at age 25 years.

The ‘mix of higher education and VET pathway’ represents 8% of the sample, with the group spending the most months in VET activities (35.2 months on average). Young people in this pathway followed a somewhat complex path, experiencing a higher number of status shifts along the way. Nearly half (46.1%) had between six and 10 transition shifts, while a further 35.8% had between 11 and 15 transitions during the 10-year period (table 3). The sequence index plot (figure 3, panel a) indicates that a large share participated in some form of post-school VET activity. There is also some evidence that those who were in higher education went on to undertake VET activities at a later stage.

This pathway has more females (61.4%) and just over one-third of individuals had an overseas background (35%). The majority are from a metropolitan location (68.9%). Just under one-quarter (24%) belong to the highest SES quartile, and 17.3% from the lowest SES quartile, but mostly the spread was roughly even across the SES quartiles. Approximately one-third in this pathway were in the lowest mathematics (34.3%) and reading (29.5%) achievement quartiles at age 15 years. While 14.6% did not complete Year 12, 33.1% studied a vocational subject in secondary school, which again highlights VET as a means of providing alternative pathways (table 4).

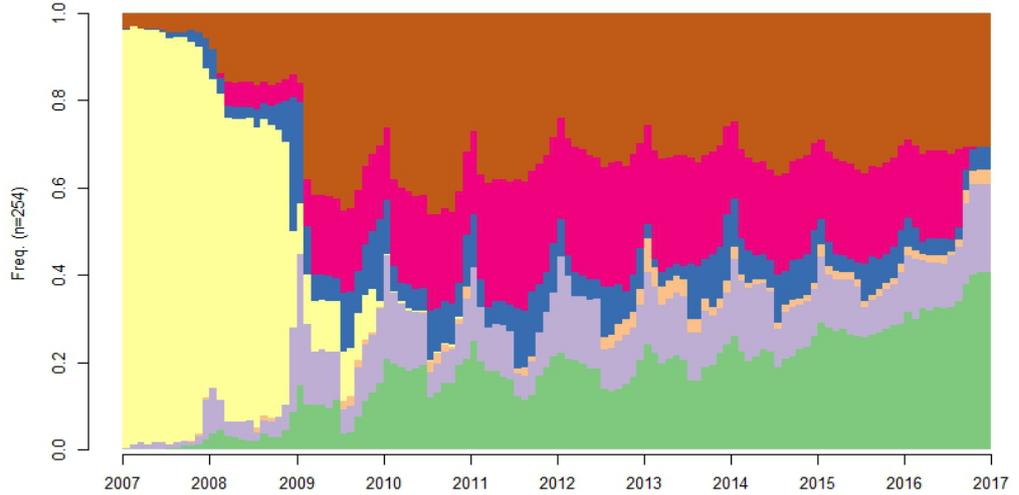
Those in Pathway 3 had a mix of tertiary education, with either a bachelor’s degree (26.8%), an advanced diploma or diploma (25.6%), or a higher-level certificate, such as certificate III or IV, as the highest qualification by age 25 years (table 5). The VET qualifications held by young people by age 25 years in this pathway were 37% advanced diploma and 30.7% certificate III (table 6). For 22.1% in this pathway, an apprenticeship/traineeship had been completed by age 25 years (table 6). A small share of young people in this pathway remained unemployed (3.5%) or out of the labour force (4.8%) at age 25 years (table 5). Table 5 also illustrates that, in terms of occupations, the highest share was in community and personal services (22.8%), followed by professionals (20.1%) and clerical and administrative workers (16.5%).

Figure 3 Pathway 3: Mix of higher education and VET (8% of the sample)

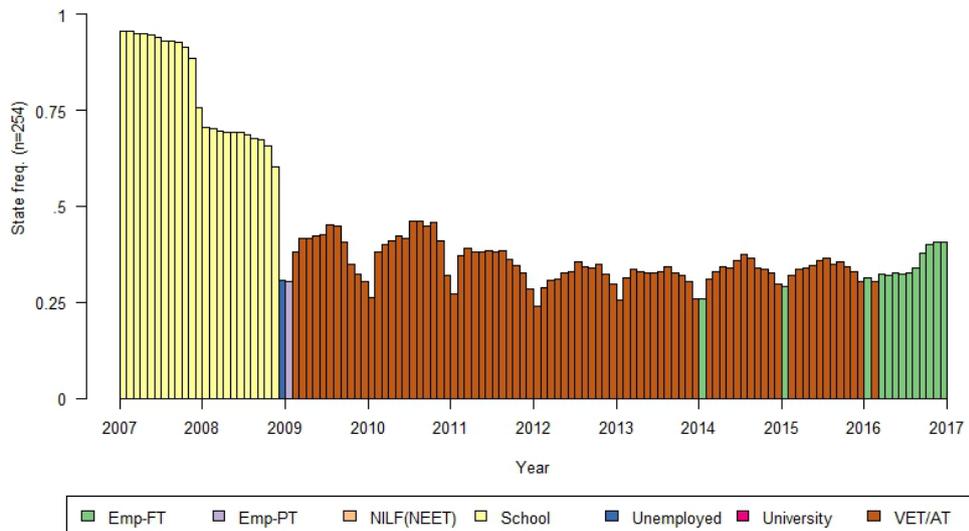
(a) Individual activity sequences



(b) Monthly proportion of activities



(c) Most frequent (modal) activity for each month



Notes: Emp-FT = full-time employment; Emp-PT = part-time employment; NILF = not in the labour force; NEET = not in education, employment nor training; VET/AT = vocational education or training/apprenticeship/traineeship.

Pathway 4: Mixed and repeatedly disengaged

Key characteristics – 5% of young people

- Pathway 4 represents the most complex pathway, containing the highest proportion of young people experiencing more than 10 transitions between the ages 16 and 25.
- Young people in this female-dominant pathway spent the most months disengaged from the labour market or unemployed. Just over one-half of youth remained unemployed at age 25 years.
- This pathway had the highest proportion of vulnerable youth, characterised by the highest incidence of teenage marriages or parenting, disability, early school leavers and youth from the lowest SES quartile.
- However, while this pathway had a *high* share with no qualifications (35.8%), it was not the pathway with the *highest* share with no qualifications (Pathway 5: Mostly working part time, with 50.9%). This figure is slightly at odds with the vulnerable youth characterisation.

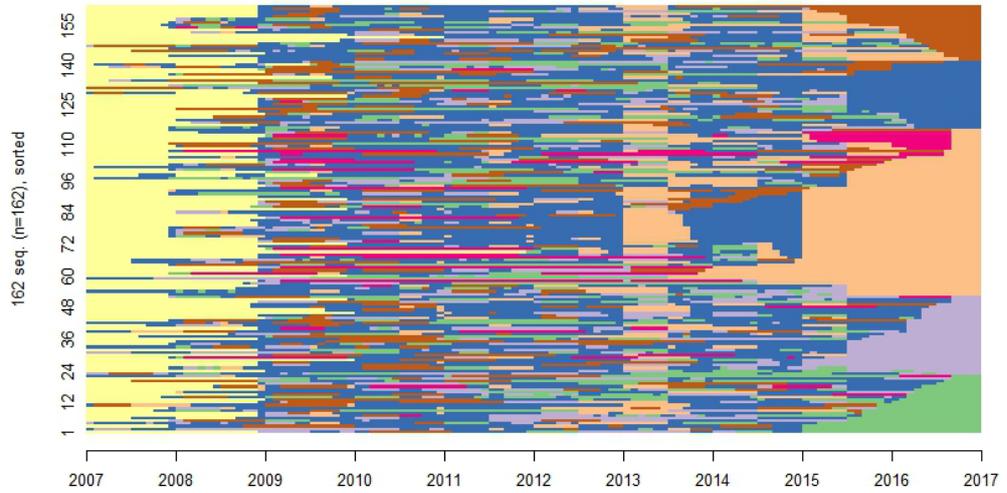
Of all of the pathways, young people in the ‘mixed and repeatedly disengaged’ pathway spent, on average, the most months unemployed (41.2 months) and out of the labour force (16.2 months), as well as the least months in school (17.9 months). It is clear from the large number of transition shifts that this is a particularly complex pathway. Nearly one-third had more than 15 transitions, while 38.3% had 11 to 15 transitions between the ages of 16 and 25 years. While this pathway represents only 5% of the sample, the high number of episodes in unemployment or NILF/NEET reveal low labour market attachment (see figure 4).

As shown in table 4, this pathway also has more females, and has a high incidence of teenage marriages (27.2%) and teenage parents (11.1%), as well as disability (10.5%). It is likely that caring duties and disability are factors in their labour market (dis)engagement. Just over one in three young people in Pathway 4 belong to the lowest SES quartile (34.6%), and 28.4% did not complete Year 12. In general, these characteristics are consistent with the literature on young people who are NEET or at risk of following precarious transitions pathways (Stanwick, Forrest & Skujins 2017; Furlong 2006).

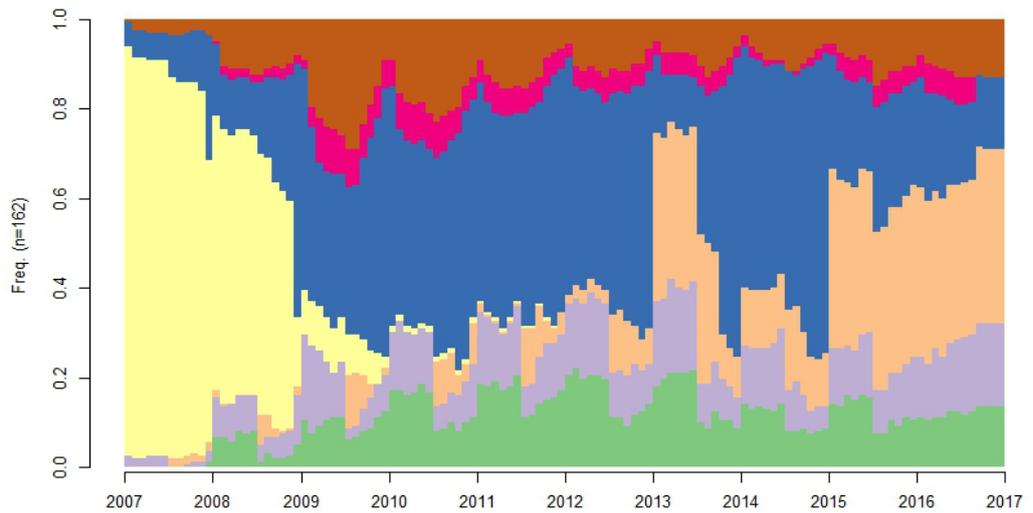
Approximately half in this pathway were in the lowest mathematics (53.7%) and reading (48.2%) achievement quartiles at age 15 years (table 4). By age 25 years, 35.8% of youth following this pathway had no qualification, and 36.4% were not in the labour force (table 5). If they held a qualification, most had a vocational-level qualification, primarily certificate III (24.1%) or advanced diploma/diploma (12.4%) as their highest qualification. Table 6 shows that 31.5% had completed one or more certificate III level qualifications by age 25 years, and 16.1% had completed an apprenticeship/traineeship. Only 46.9% in this pathway were employed at age 25 years. Occupations were in community and personal services (10.5%) or as sales workers (9.3%) or labourers (9.3%) (table 5).

Figure 4 Pathway 4: Mixed and repeatedly disengaged (5% of the sample)

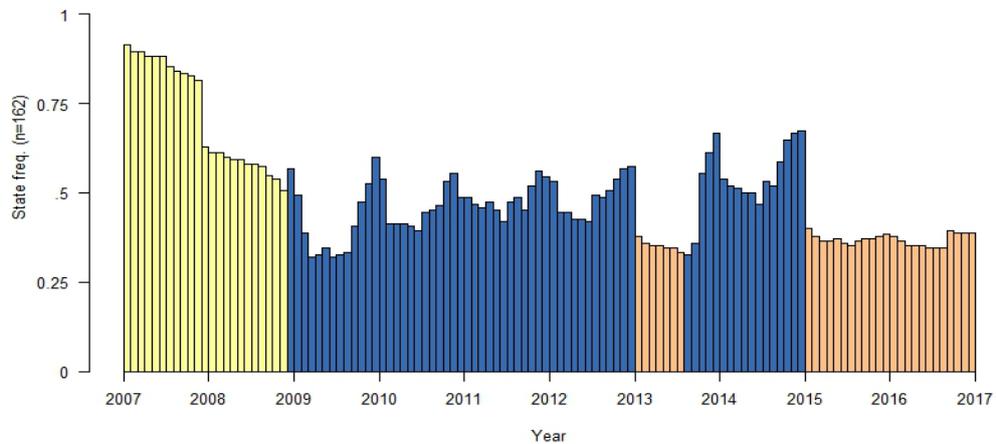
(a) Individual activity sequences



(b) Monthly proportion of activities



(c) Most frequent activity (modal) for each month



■ Emp-FT
 ■ Emp-PT
 ■ NILF(NEET)
 ■ School
 ■ Unemployed
 ■ University
 ■ VET/AT

Notes: Emp-FT = full-time employment; Emp-PT = part-time employment; NILF = not in the labour force; NEET = not in education, employment nor training; VET/AT = vocational education or training/apprenticeship/traineeship.

Pathway 5: Mostly working part-time

Key characteristics – 4% of young people

- The youth in this pathway spent the most time in part-time employment between the ages 16 to 25 years.
- One in two had no post-school qualification at the age of 25 years. They had also spent the least amount of time in post-school education.
- At age 25 years, young people in this pathway were primarily in community and personal services, clerical and administrative, and sales occupations.

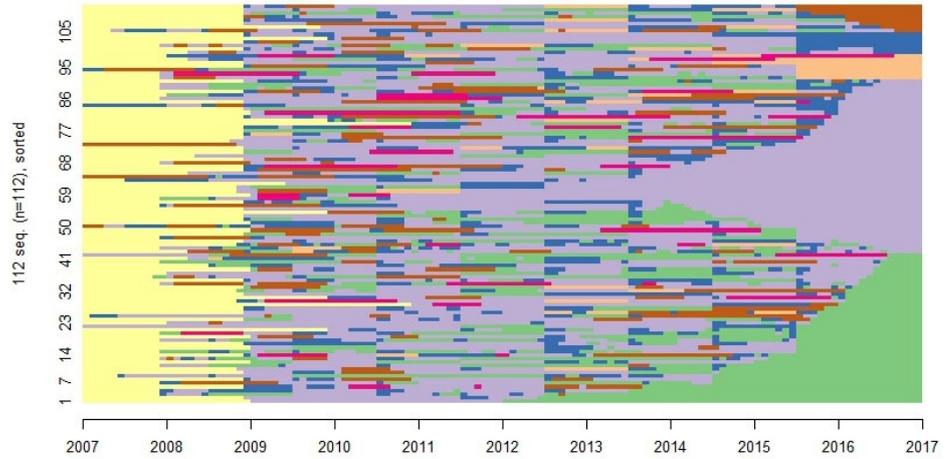
Young people in the ‘mostly working part-time’ pathway represent 4% of the sample and spent an average of 47.7 months in part-time employment. They also had a combined average of 13.2 months in post-school education (university and VET). Pathway 5 is therefore comprised of young people who were relatively less qualified and employed part-time. Their pathway is complex, with a large share experiencing 11 to 15 (37.5%) transition shifts, with 24.1% experiencing more than 15. As illustrated in the sequence index plot (figure 5, panel a), the changes in activities are primarily shifts between part-time employment and unemployment or education activities.

Table 4 provides some background characteristics of the individuals in Pathway 5. The share of males is 42.9%, indicating that slightly more females follow this pathway. Approximately one in 10 individuals in this pathway has an Indigenous background (10.7%). The observed incidence of teen marriages (19.6%) and early school leaving (20.5%) in this group is relatively high compared with other pathways. Members of this pathway are spread across different socioeconomic backgrounds, with 20.5% from the highest SES quartile and 33% from the lowest SES quartile.

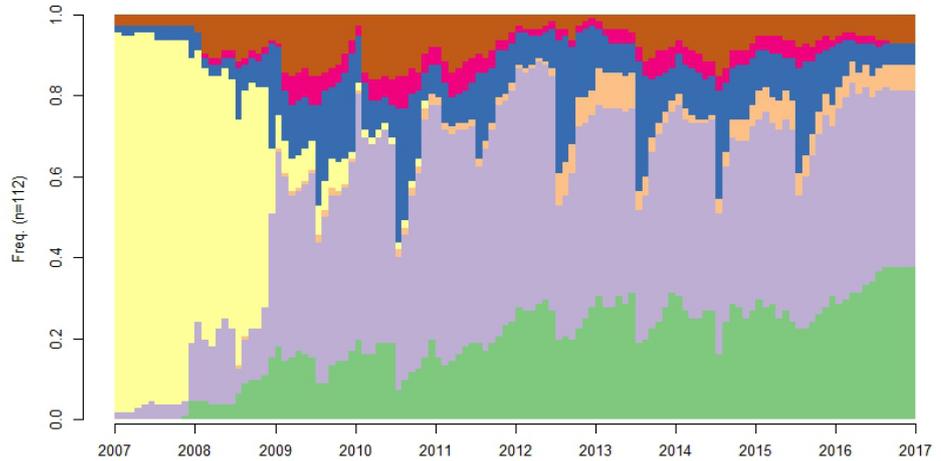
Just over one in two young people (51.8%) in this pathway belonged to the lowest mathematics achievement quartile and 42.9% were in the lowest reading achievement quartile at age 15 years (table 4). An interesting observation is that, even though 50.9% in this pathway did not complete a post-school qualification by the age of 25 years, they had maintained good contact with the labour market, with 90.2% employed (table 5). Certificate III (17.9%) and certificate IV (8.9%) were the most prevalent highest qualification held by the age of 25 years. In terms of all of the VET qualifications acquired by the age of 25 years, table 6 shows that 23.2% had a certificate III level qualification and 16.1% had undertaken an apprenticeship or traineeship. The occupation of their work at age 25 years indicates that they were mostly in community and personal services (26.8%), sales (18.8%) and clerical and administrative occupations (12.5%, table 5).

Figure 5 Pathway 5: Mostly working part-time (4% of the sample)

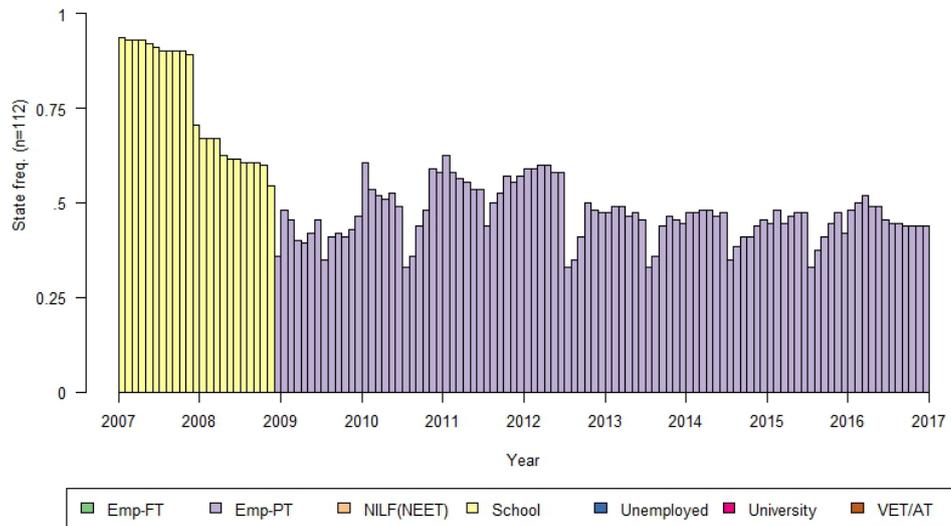
(a) Individual activity sequences



(b) Monthly proportion of activities



(c) Most frequent (modal) activity for each month



Notes: Emp-FT = full-time employment; Emp-PT = part-time employment; NILF = not in the labour force; NEET = not in education, employment nor training; VET/AT = vocational education or training/apprenticeship/traineeship.

Table 2 Average number of months spent in each activity between ages 16 and 25 years by pathway

Activity	Pathway 1 Higher education and work	Pathway 2 Early entry to full-time work	Pathway 3 Higher education and VET	Pathway 4 Mixed repeatedly disengaged	Pathway 5 Mostly working part-time
School	21.5 (0.2)	18.9 (0.3)	20.0 (0.5)	17.9 (0.8)	18.8 (0.9)
University	55.4 (0.4)	1.8 (0.2)	20.9 (1.3)	4.9 (0.9)	3.8 (0.8)
VET	1.7 (0.1)	14.3 (0.4)	35.2 (0.9)	13.6 (0.9)	9.4 (0.9)
Employed full-time	25.2 (0.4)	69.8 (0.7)	20.6 (1.0)	13.2 (1.2)	23.2 (1.9)
Employed part-time	7.5 (0.2)	5.5 (0.2)	13.1 (0.8)	12.9 (1.0)	47.7 (1.7)
Unemployed	7.3 (0.2)	8.8 (0.5)	8.5 (0.5)	41.2 (1.3)	13.7 (0.8)
NILF/NEET	1.4 (0.1)	0.9 (0.1)	1.6 (0.2)	16.2 (1.1)	3.5 (0.6)

Notes: NILF = not in the labour force; NEET = not in education, employment nor training. Standard errors are in parenthesis.

Source: LSAY 2006 (Y06).

Table 3 Number of transitions between age 16 and 25 years by pathway proportion of individuals (%)

Number of transitions	Pathway 1 Higher education and work %	Pathway 2 Early entry to full-time work %	Pathway 3 Higher education and VET %	Pathway 4 Mixed repeatedly disengaged %	Pathway 5 Mostly working part-time %
1 to 5	26.8	30.0	10.2	1.2	9.8
6 to 10	52.0	39.2	46.1	28.4	28.6
11 to 15	18.1	22.5	35.8	38.3	37.5
More than 15	3.1	8.4	7.9	32.1	24.1
Total	100	100	100	100	100

Notes: Chi² test for differences across pathways. Pearson chi² (12) = 426.8531 Pr = 0.000.

Source: LSAY 2006 (Y06).

Table 4 Selected descriptive statistics of the sample: socioeconomic and demographic characteristics by pathway (sample proportions %)

	Pathway 1	Pathway 2	Pathway 3	Pathway 4	Pathway 5
	Higher education and work	Early entry to full-time work	Higher education and VET	Mixed repeatedly disengaged	Mostly working part-time
	%	%	%	%	%
Male	43.3	64.1	38.6	38.3	42.9
Indigenous	1.7	4.9	4.7	7.4	10.7
Language other than English	10.2	2.7	5.5	4.3	0
Overseas background	45.8	29.9	35.0	29.6	26.8
Marital status					
Married at 15–19	4.4	15.6	8.3	27.2	19.6
Married at 20–25	40.4	51.0	47.6	32.7	34.8
Dependent children					
At age 15–19	0.3	0.6	0.4	11.1	0.9
At age 20–25	3.0	13.6	8.3	29.6	13.4
Metropolitan location	76.8	61.2	68.9	57.4	65.2
Socioeconomic status (SES)					
Highest quartile	43.4	17.7	24	15.4	20.5
Lowest quartile	10.1	25.6	17.3	34.6	33.0
Has a disability	3.3	3.3	6.7	10.5	7.1
Lives with parents (age 20–25)	80.9	80.7	82.3	72.2	84.8
Did not complete Year 12	0.9	21.5	14.6	28.4	20.5
Vocational studies in school	14.5	44.4	33.1	41.4	44.6
Age 15 mathematics (PISA)					
Highest quartile	34.0	9.5	15.8	10.5	12.5
Lowest quartile	13.9	41.1	34.3	53.7	51.8
Age 15 reading (PISA)					
Highest quartile	34.2	8.8	17.7	11.7	8.0
Lowest quartile	14.3	44.0	29.5	48.2	42.9
Government payments					
Youth allowance (YA)	62.4	34.0	58.7	75.9	57.1
Other payments (not YA)	37.5	27.4	43.3	67.9	38.4
Sample size	1928	730	254	162	112

Notes: Detailed table with standard errors is available in appendix A, table A4. Variable definitions are available in appendix A, table A3. Chi2 tests for differences across pathways indicate that all proportions except for living with parents (age 20–25) are significantly different from each other at the 5% level. See appendix A, table A5, for detailed tests.

Source: LSAY 2006 (Y06).

Table 5 Highest qualification, occupation and labour force status at age 25 years by pathway (sample proportions %)

	Pathway 1	Pathway 2	Pathway 3	Pathway 4	Pathway 5
	Higher education and work %	Early entry to full-time work %	Higher education and VET %	Mixed repeatedly disengaged %	Mostly working part-time %
Highest qualification (at age 25)					
Certificate I	0.0	3.7	0.8	2.5	1.8
Certificate II	0.2	5.8	3.2	6.2	7.1
Certificate III	1.4	25.9	11.0	24.1	17.9
Certificate IV	1.2	13.7	15.4	11.1	8.9
Certificate – unknown	0.4	7.1	2.4	1.9	1.8
Advanced diploma/diploma	1.7	13.2	25.6	12.4	6.3
Bachelor's degree	69.4	1.1	26.8	6.2	4.5
Postgrad. diploma/certificate	4.4	0.4	2.8	0.0	0.9
Postgraduate	10.6	0.0	0.8	0.0	0.0
None	10.8	29.2	11.4	35.8	50.9
Total	100	100	100	100	100
Occupation (at age 25)					
Managers	7.0	10.6	5.5	1.9	3.6
Professionals	50.1	7.1	20.1	3.1	6.3
Technicians and trades workers	4.1	29.3	8.7	4.3	8.9
Community and personal service workers	8.3	11.4	22.8	10.5	26.8
Clerical and administrative workers	11.2	16.0	16.5	4.3	12.5
Sales workers	6.1	7.5	7.9	9.3	18.8
Machinery operators and drivers	0.7	6.0	2.0	1.9	4.5
Labourers	2.2	6.0	5.9	9.3	8.0
Unknown/not classifiable	2.9	3.4	2.4	2.5	0.9
Not working (unemployed/NILF)	7.6	2.6	8.3	53.1	9.8
Total	100	100	100	100	100
Labour force status (at age 25)					
Employed	92.4	97.4	91.7	46.9	90.2
Unemployed	2.3	1.2	3.5	16.7	5.4
Not in the labour force	5.3	1.4	4.8	36.4	4.4
Total	100	100	100	100	100

Note: Proportions are by pathway. Detailed table with standard errors is available in appendix A, table A4. Pearson Chi2 tests for differences across pathways indicate that the proportion in each pathway for each of the variables is statistically different from each other at the 5% level. See appendix A, table A5 for detailed tests.

Source: LSAY 2006 (Y06).

Table 6 All VET qualifications and apprenticeships/traineeships by age 25 years by pathway (sample proportions %)

	Pathway 1	Pathway 2	Pathway 3	Pathway 4	Pathway 5
	Higher education and work %	Early entry to full-time work %	Higher education and VET %	Mixed repeatedly disengaged %	Mostly working part-time %
VET qualification					
Certificate I	0.2	4.9	2.0	3.1	3.6
Certificate II	0.7	9.0	9.8	9.9	9.8
Certificate III	3.6	33.0	30.7	31.5	23.2
Certificate IV	2.6	16.0	23.2	12.3	10.7
Certificate – unknown	0.5	7.7	2.8	1.9	1.8
Advanced diploma/diploma	3.3	13.6	37.0	12.3	9.8
Apprenticeships/traineeships	3.8	47.3	22.1	16.1	16.1

Note: Proportions are within pathway. Each row shows all of the qualifications obtained at that level; this is multiple response and column totals can add to more than 100%. Derived variables used, defined in appendix A, table A1.

Source: LSAY 2006 (Y06).

To accompany this report, an interactive data visualisation, *Visualising school-to-work pathways using LSAY*, presents the school-to-work pathways of young Australians aged 16 to 25, and can be accessed from <<https://www.ncver.edu.au/research-and-statistics/school-to-work-pathways>>.



What factors contribute to different pathways?

To better understand the transition pathways of Australian youth, this section investigates the role of socioeconomic and demographic factors in influencing these individual transition pathways. Given the extensive literature on the relationship between educational and labour market outcomes and individual socioeconomic background (Black & Devereux 2010; Ranasinghe 2015), it is of interest to explore the extent to which these factors affect young people's transition experiences.

Similar to other studies that use the output from sequence analysis and cluster analysis as the dependent variable in further explanatory analyses using regression models, the emphasis here is on identifying 'predictive markers' of transition pathways¹²; in this case, those observed in relation to an individual's (aged 15 years) future school-to-work trajectories. The aim is therefore to assess the link between starting conditions and future pathways, rather than to identify causal connections. In this context, the model includes a set of individual characteristics, observed at age 15 years, relating to socioeconomic and family background and education. Using the multinomial logistic regression model, we estimate the likelihood of belonging to one of these mutually exclusive pathways for a given set of background characteristics.

The framework of the logistic model makes it possible to estimate the percentage change in the probability of an individual with a given characteristic entering a specific pathway. Average marginal effects are obtained by averaging these estimates across all individuals in the sample. Table 7 presents these estimates from the multinomial logistic regression model, where Pathway 1 is the reference pathway.¹³ For example, the probability that a young person who studied vocational subjects in school then follows Pathway 1, the higher education and work pathway, is, on average, 20 percentage points lower than for an individual who did not take any vocational subjects in school. The meaning of a value close to zero for the marginal effect is that small changes in the value of the predictor variable are expected to be associated with almost no change in the pathway probability, although in some cases there are small but statistically meaningful differences. We focus attention on some selected marginal effects that are strongly statistically significant (marked with two stars).

12 See for example, Dorsett & Lucchino (2014), McVicar & Anyadike-Danes (2002) and Corrales-Herrero & Rodriguez-Prado (2017).

13 While there are different ways to present the results from a multinomial regression, in this instance they are presented as average marginal effects, following the approach in McVicar and Anyadike-Danes (2002) and Dorsett and Lucchino (2014), who also conduct regression analysis on the clusters based on sequence analysis.

Table 7 Age 15 years average marginal effects on future pathway outcomes: change in the probability of following a certain pathway for a given characteristic compared with the reference category

	Pathway 1	Pathway 2	Pathway 3	Pathway 4	Pathway 5
	Higher education and work	Early entry to full-time work	Higher education and VET	Mixed and repeatedly disengaged	Mostly working part-time
Male (ref: female)	-0.09** [0.02]	0.14** [0.02]	-0.03** [0.01]	-0.02** [0.01]	-0.01 [0.01]
Indigenous (ref: non-Indigenous)	-0.13** [0.05]	0.02 [0.04]	0.03 [0.03]	0.03 [0.02]	0.05* [0.02]
Overseas background (ref: non-overseas background)	0.11** [0.16]	-0.07** [0.01]	-0.01 [0.01]	-0.01 [0.01]	-0.01* [0.00]
Socioeconomic status (SES) (ref: top quartile)					
Second quartile	-0.05* [0.22]	0.04* [0.02]	0.02 [0.01]	-0.02 [0.01]	0.00 [0.01]
Third quartile	-0.18** [0.22]	0.12** [0.02]	0.03* [0.01]	0.01 [0.01]	0.01 [0.01]
Lowest quartile	-0.20** [0.02]	0.12** [0.02]	0.02 [0.01]	0.04* [0.01]	0.02* [0.01]
Metropolitan location (ref: non-metropolitan location)	0.04* [0.02]	-0.03* [0.02]	0.00 [0.01]	-0.01 [0.01]	0.00 [0.01]
Mathematics achievement PISA (ref: top quartile)					
Second quartile	-0.08** [0.02]	0.08** [0.02]	0.02 [0.02]	-0.01 [0.01]	-0.01 [0.01]
Third quartile	-0.12** [0.03]	0.11** [0.02]	0.01 [0.02]	0.01 [0.01]	-0.01 [0.01]
Lowest quartile	-0.25** [0.04]	0.16** [0.03]	0.03 [0.02]	0.04* [0.02]	0.02 [0.02]
Reading achievement PISA (ref: top quartile)					
Second quartile	-0.05* [0.03]	0.03 [0.02]	0.01 [0.02]	-0.01 [0.01]	0.02* [0.01]
Third quartile	-0.11** [0.03]	0.08** [0.03]	0.01 [0.02]	0.01 [0.02]	0.02* [0.01]
Lowest quartile	-0.16** [0.04]	0.10** [0.03]	0.01 [0.02]	0.02 [0.02]	0.03* [0.01]
Vocational studies in school (ref: no vocational studies in school)	-0.20** [0.02]	0.13** [0.02]	0.03* [0.01]	0.02* [0.01]	0.02** [0.01]
Sample size	1928	730	254	162	112

Note: ** significant at the 1% level (p<0.01); *significant at the 5% level (p<0.05).

Standard errors in parenthesis. See appendix A for details of variable definitions. Wald tests for overall significance of estimates indicate that all variables except for metropolitan location are significant at the 5% level. Wald tests for differences across pathways indicate that all variables except SES at the second quartile, metropolitan location and reading achievement at the second quartile are significantly different from each other at the 5% level. Detailed table available in appendix A, table A5. Model fit statistics are available in appendix A, table A8 The underlying multinomial logistic model reference pathway is Pathway 1. Multiple category variables (SES, PISA) are constructed as sets of 0/1 variables maintaining the reference category, so, for example, the third quartile SES is relative to the top quartile.

As observed in table 7, being a male is on average associated with a higher probability of following Pathway 2: Early entry to full-time work, and with a lower probability of following Pathways 1, 3 or 4.

The probability that a young person who studied vocational subjects in school subsequently follows Pathway 2 is, on average, 13 percentage points higher than for an individual who did not take any vocational subjects in school. The profile shows that Pathway 2 also had the highest share of apprentices/trainees (47.3%, table 6), and the highest share working at age 25 years as technician/tradesperson (29.3%, table 5).

Table 7 also shows that the probability that an Indigenous youth follows Pathway 1 is, on average, 13 percentage points lower than for non-Indigenous youth.¹⁴

In terms of the socioeconomic background, individuals from lower SES quartiles are associated with a lower probability of following the dominant Pathway 1: Higher education and work. For example, table 7 shows that the probability that a young person from the lowest SES quartile enters Pathway 1 is, on average, 20 percentage points lower than that for an individual from the highest SES quartile. On the other hand, coming from the third quartile or lowest socioeconomic background is associated with a 12-percentage point higher probability (than the highest SES individuals) of entering Pathway 2. These observations are consistent with existing overseas findings, of an association between socioeconomic background and the transition pathways experienced by young people, in particular indicating that those from an advantaged background tend to follow structured pathways between education and employment (Dorsett & Lucchino 2014; McVicar & Anyadike-Danes 2002).

Mathematics and reading achievements based on PISA scores are used as proxies for school attainment of individuals at the age of 15 years. It is well established in the literature that school performance is linked to the type of post-school transitions experienced by young people, particularly in relation to labour market outcomes (McLachlan, Gilfillan & Gordon 2013; Lee & Newhouse 2013). Having lower mathematics and reading achievement than those in the top quartile was associated with a lower probability of following Pathway 1: Higher education and work and with a higher probability of following a more employment-oriented pathway (Pathway 2).¹⁵

The results in table 7 indicate that there are alternative avenues for those who do not follow a traditional academic path. Studying a vocational subject while in school emerges as a statistically meaningful factor across all pathways (but not always large in the scale of effect on the pathway). Engaging in vocational studies at school was associated with a 20-percentage-point lower probability of following the dominant Pathway 1 but was found to be particularly associated with a 13-percentage-point rise in the probability of the Pathway 2: Early entry to full-time work. International studies on school-to-work transitions across several countries have consistently found that those with well-established VET and apprenticeship systems were more successful in facilitating a smooth transition to the labour market (Quintini & Manfredi 2009; Brzinsky-Fay & Solga 2016).

14 Indigenous background is associated with a higher probability of following Pathway 5: Mostly working part-time, at the lower 5% level, but it is not found to be statistically associated with Pathways 2, 3 or 4.

15 However, mathematics and reading achievement at age 15 years was not found to be statistically associated with Pathway 3, and limited evidence for statistical relationships were found in Pathways 4 and 5.

Likelihood of following a given pathway relative to Pathway 3: Mix of higher education and VET

The results of the multinomial regression in terms of relative risk ratios is shown in table 8, with Pathway 3: Mix of higher education and VET as the base category. Pathway 3 contains individuals who had a high level of participation in VET activity during the 10-year period (table 6) and therefore is chosen as the reference category. Thus, the results from the regression modelling is interpreted as the estimated likelihood of belonging to Pathway 1, 2, 4 or 5, relative to Pathway 3.

The *relative risk ratios* shown give the proportionate change in the relative likelihood of belonging to a given pathway rather than the reference pathway (Pathway 3 in this instance), when the variable changes by one unit.¹⁶

A relative risk ratio must be greater than zero, and a value of 1.00 means that the likelihood is identical in the two groups. A value greater than 1.00 indicates that the likelihood is higher in the group concerned, relative to the reference group. Conversely, a value less than 1.00 implies that the likelihood is lower compared with the reference group. A zero relative risk ratio implies that there were no cases in one group and some cases in the reference group; however, this situation is unlikely to occur as covariates with no values in a given category are not used in the regression.

Table 8 shows that:

- Males are 2.9 times more likely to be in Pathway 2's early entry to full-time work than in Pathway 3, with its mix of VET and higher education activities.
- Young people from lower quartile 3 socioeconomic backgrounds are less likely to be in the higher education and work Pathway 1 than in Pathway 3. On the other hand, those from the lowest SES quartile are more likely to be in Pathway 2, the early entry to full-time work pathway, rather than in Pathway 3.
- Those in the lowest mathematics achievement quartile are less likely to be in Pathway 1, higher education and work, than in Pathway 3. However, mathematics or reading achievement at age 15 years do not emerge as key factors in other pathways.
- Studying vocational subjects in school lowers the chances of being in Pathway 1 as opposed to being in Pathway 3. It increases the likelihood of following Pathway 2, early entry to full-time work, but does not appear to be a significant factor in other pathways.

¹⁶ Another way to think about relative risk ratios is that they are simply ratios of two conditional probabilities.

Table 8 Multinomial logistic results (relative risk ratios) relative to Pathway 3

	Pathway 1	Pathway 2	Pathway 4	Pathway 5
	Higher education and work	Early entry to full-time work	Mixed and repeatedly disengaged	Mostly working part-time
Relative to Pathway 3: Mix of higher education and VET (n = 254)				
Male (ref: female)	1.12 [0.17]	2.91** [0.48]	1.01 [0.23]	1.21 [0.31]
Indigenous (ref: non-Indigenous)	0.52 [0.19]	0.87 [0.31]	1.29 [0.56]	2.06 [0.90]
Overseas background (ref: non-overseas background)	1.52* [0.22]	0.78 [0.13]	0.80 [0.18]	0.68 [0.18]
Socioeconomic status (SES) (ref: top quartile)				
Second quartile	0.73 [0.15]	1.07 [0.26]	0.49 [0.19]	0.91 [0.37]
Third quartile	0.44** [0.09]	1.35 [0.32]	0.89 [0.29]	0.98 [0.38]
Lowest quartile	0.49* [0.11]	1.61* [0.39]	1.63 [0.53]	1.69 [0.63]
Metropolitan location (ref: non-metropolitan location)	1.09 [0.17]	0.83 [0.14]	0.75 [0.17]	1.06 [0.27]
Mathematics achievement PISA (ref: top quartile)				
Second quartile	0.67 [0.15]	1.31 [0.35]	0.67 [0.28]	0.66 [0.29]
Third quartile	0.66 [0.17]	1.69 [0.49]	1.13 [0.49]	0.63 [0.29]
Lowest quartile	0.37* [0.11]	1.67 [0.55]	1.47 [0.70]	1.10 [0.55]
Reading achievement PISA (ref: top quartile)				
Second quartile	0.80 [0.18]	1.10 [0.29]	0.76 [0.31]	2.02 [0.95]
Third quartile	0.67 [0.17]	1.36 [0.39]	1.05 [0.44]	2.09 [1.04]
Lowest quartile	0.58 [0.17]	1.50 [0.49]	1.39 [0.64]	2.31 [1.28]
Vocational studies in school (ref: no vocational studies in school)	0.45** [0.07]	1.42* [0.22]	1.18 [0.25]	1.40 [0.33]
Sample size	1928	730	162	112

Note: Reference category: Pathway 3: Mix of higher education and VET.

** significant at the 1% level (p<0.01); *significant at the 5% level (p<0.05)

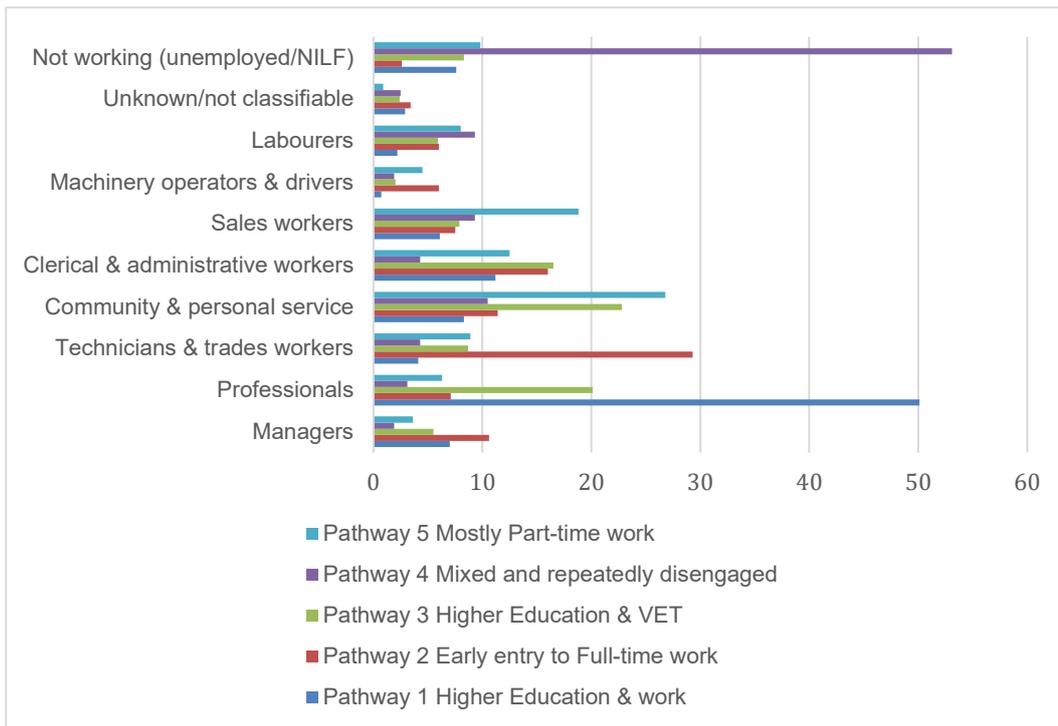
Standard errors in parenthesis. See appendix A for details of variable definitions. Wald tests for overall significance of estimates indicate that all variables except for metropolitan location are significant at the 5% level. Wald tests for differences across pathways indicate that all variables except SES at the second quartile, metropolitan location and reading achievement at the second quartile are significantly different from each other at the 5% level. Detailed table available in appendix A, table A6; model fit statistics are available in table A8.

Pathways associated with occupational outcomes at age 25 years

The occupations held at the age of 25 years and the pathways young people had followed are shown in figure 6. A key feature of figure 6 is that at age 25 years many occupations did not have particularly high shares from any specific pathway, with, at most, half of a pathway directed to a specific occupation. This suggests that, from the pathways defined here, the occupational outcomes can be many. The clearest occupational link is the high share of professional occupations arising from Pathway 1: Higher education and work (50.1%); however, interestingly, Pathway 3: Mix of HE and VET also featured for this occupation as it supported 20% professionals. While a reasonably high share (29.3%) of technicians and trades came from Pathway 2: Early entry to full-time work, some other occupations such as managers, clerical and administrative, and community and personal service workers had shares of over 10% from Pathway 2.

It is apparent that Pathway 4: Mixed and repeatedly disengaged is associated with the highest share of young people with poor work outcomes at the age of 25 years, with 53.1% unemployed or not in the labour force. The pathway sequences in figure 4 show that over time these young people experienced, beginning after school, high shares of both unemployment and out of the labour force, which preceded these poor work destinations at age 25 years. This pathway has the highest share of government payments (Youth Allowance 75.9%, and 67% other govt payments, table 4) and non-completion of school (28.4%). The poor work outcomes over time seem to culminate in the poor outcome at age 25 years, despite the group gaining some qualifications (64.2%, table 6). This perhaps reflects the findings of Buddelmeyer and Marks (2010), who found that the previous year's labour market state of an individual has the most significant implication for their current state.

Figure 6 Occupation at age 25 years (% of pathway)



Note: Underlying figures are column % of pathway, shown in table 5.

Source: LSAY 2006 Y06.



References

- Aassve, A, Billari, F & Piccarreta, R 2007, Strings of adulthood: a sequence analysis of young British women's work-family trajectories, *European Journal of Population*, vol.23, pp.369–388.
- ABS (Australian Bureau of Statistics) 2017, Labour Force Survey, cat.no.6202.0, ABS, Canberra.
- Albæk, K, Asplund, R, Barth, E, Lindahl, L, von Simson, K & Vanhala, P 2015, 'Youth unemployment and inactivity: a comparison of school-to-work transitions and labour market outcomes in four Nordic countries', *TemaNord*, 548, Denmark.
- Anders, J & Dorsett, R 2017, 'What young English people do once they reach school-leaving age: a cross-cohort comparison for the last 20 years', *Longitudinal and Life Course Studies*, vol.8, pp.75–103.
- Black, S E, Devereux, P 2010, *Recent developments in intergenerational mobility*, IZA discussion papers 4866, Institute for the Study of Labor (IZA), Bonn.
- Blanchard, P 2016, 'Mapping the field of sequence analysis', in G Ritschard & M Studer (eds), *Proceedings of the International Conference on Sequence Analysis and Related Methods*, Lausanne, June, 2016, pp.31–3.
- Brzinsky-Fay, C 2007, 'Lost in transition? Labour market sequences of school-leavers in Europe', *European Sociological Review*, vol.23, no.4, pp.409–22.
- Brzinsky-Fay, C & Solga, H 2016, 'Compressed, postponed, or disadvantaged? School-to-work transition patterns and early occupational attainment in West Germany', *Research in Social Stratification and Mobility*, vol.46, pp.21–36.
- Buddelmeyer, H & Marks, G 2010, *Annual transitions between labour market states for young Australians*, NCVER, Adelaide, <<https://www.ncver.edu.au/research-and-statistics/publications/all-publications/annual-transitions-between-labour-market-states-for-young-australians>>.
- Corrales-Herrero, H & Rodríguez-Prado, B 2012, 'Characterizing Spanish labour pathways of young people with vocational lower-secondary education', *Applied Economics*, vol.44, no.29, pp.3777–92.
- Dorsett, R & Lucchino, P 2014, 'Explaining patterns in the school-to-work transition: an analysis using optimal matching', *Advances in Life Course Research*, vol.22, pp.1–14.
- Fasang, A 2014, 'Sequence analysis in family life course research', European Consortium for Sociological Research Spring School on 'Family complexity and its implications of inequality', Turin, Italy.
- Fitzpatrick, D, Lester, L, Mavromaras, K, Richardson, S & Sun, Y 2011, *From education to employment: how long does it take?*, NCVER, Adelaide, <<https://www.ncver.edu.au/research-and-statistics/publications/all-publications/from-education-to-employment-how-long-does-it-take>>.
- Fry, J & Boulton, C 2013, *Prevalence of transition pathways in Australia*, Productivity Commission Staff Working Paper, Canberra.
- Furlong, A 2006, 'Not a very NEET solution: representing problematic labour market transitions among early school-leaver', *Work, Employment and Society*, vol.20, no.3, pp.553–69.
- Gabardinho, A, Ritschard, M, Müller, N S & Studer, M 2011, Analyzing and visualising state sequences in R with TraMineR, *Journal of Statistical Software*, vol. 40, pp.1–37.
- Lee, J & Newhouse, D 2013, *Cognitive skills and youth labor market outcomes*, Background paper for the *World development report 2013*, World Bank, Washington, DC.
- Lorentzen, T, Bäckman, O, Ilmakunnas, I, & Kauppinen, T 2018, Pathways to adulthood: sequences in the school-to-work transition in Finland, Norway and Sweden, Social Indicators Research, pp.1–21.
- McLachlan, R, Gilfillan, G & Gordon, J 2013, *Deep and persistent disadvantage in Australia*, Productivity Commission staff work paper, Australian Government, Canberra.
- McVicar, D & Anyadike-Danes, M 2002, 'Predicting successful and unsuccessful transitions from school to work using sequence methods', *Journal of the Royal Statistical Society, Series A*, vol.165, part 2, pp.317–34.
- Martin, P, Schoon, I & Ross, A 2008, 'Beyond transitions: applying optimal matching analysis to life course research', *International Journal of Social Research Methodology*, vol.11, no.3, pp.179–99.
- OECD (Organisation for Economic Co-operation and Development) 2000, *From initial education to working life: making transitions work*, Paris.

- Parasnis, J & Swan, J 2017, 'Differences in educational attainment by country of origin: evidence from Australia', Department of Economics Monash University, Discussion paper 05/17, Melbourne.
- Quintini, G & Manfredi, T 2009, *Going separate ways? school-to-work transitions in the United States and Europe*, OECD, Paris.
- Ranasinghe, R 2015, 'The transition of education across generations: evidence from Australia', *BE Journal of Economic Analysis and Policy*, vol.15, no.4, pp.1893–917.
- Schoon, I & Lyons-Amos, M 2016, 'Diverse pathways in becoming an adult: the role of structure, agency and context', *Research in Social Stratification and Mobility*, vol.46, pp.11–20.
- Stanwick, J, Forrest, C & Skujins, P 2017, *Who are the persistently NEET young people?*, NCVER, Adelaide, <<https://www.ncver.edu.au/research-and-statistics/publications/all-publications/who-are-the-persistently-neet-young-people>>.
- Strauss, T & von Maltitz, M 2017, 'Generalising Ward's method for use with Manhattan distances', *PLOS ONE*, vol.12, no.1: e0168288, <<https://doi.org/10.1371/journal.pone.0168288>>.
- Studer, M 2013, 'WeightedCluster library manual: a practical guide to creating typologies of trajectories in the social sciences with R', *LIVES Working Papers*, vol.24, viewed 2 June 2019, DOI: <<http://dx.doi.org/10.12682/lives.2296-1658.2013.24>>.
- Studer, M & Ritschard, G 2014, 'A comparative review of sequence dissimilarity measures', *LIVES Working Paper* 2014/33.
- Studer, M & Ritschard, G 2016, 'What matters in differences between life trajectories: a comparative review of sequence dissimilarity measures', *Journal of the Royal Statistical Society*, vol.179, part 2, pp.481–511.
- Studer, M, Struffolino, E & Fasang, A 2018, 'Estimating the relationship between time-varying covariates and trajectories: the sequence analysis multistate model procedure', *Sociological Methodology*, vol.48, no.1, pp.103–35.
- Yu, S, Bretherton, T, Schutz, J & Buchanan, J 2012, *Understanding the nature of vocations today: exploring labour market pathways*, NCVER, Adelaide, <<https://www.ncver.edu.au/research-and-statistics/publications/all-publications/understanding-the-nature-of-vocations-today-exploring-labour-market-pathways>>.

Appendix A: data and variable definitions

Variable derivation

This section presents the variables used in setting up the monthly data as required for sequence analysis, including any assumptions made. Where calendar data were not available, interview years and interview months have been used to derive the monthly variables for each category/state.

General assumptions:

- If a person is said to have stopped studying in a specified month, the person is classified to be studying for that month. This is because the exact day the person stopped studying is unknown.
- If a person is interviewed in the same year as the year they stopped studying, but the month they stopped study is unknown, the person is classified to be studying in January and the remaining months up to the interviewed month are classified to be unknown.

Table A1 Derivation of variables

State	Status	Variables used	Filter variables
School	-	<ul style="list-style-type: none"> ▪ Still at school ▪ Month left school ▪ Year left school 	-
Employment	-	<ul style="list-style-type: none"> ▪ Months worked – FT/PT (calendar) ▪ No full-time work since last interview ▪ Currently work in a job/own business or farm ▪ Jobs away from ▪ Any other jobs since last interview 	-
University/VET	Post-school study	<ul style="list-style-type: none"> ▪ Started any study or training since leaving school ▪ Month began study ▪ Year began study ▪ Current Year 12 TAFE/short course/other module/undefined study ▪ Month stopped study ▪ Year stopped study ▪ Current qualification type ▪ Still at school ▪ Current study or training (not elsewhere reported) ▪ Month began apprenticeship/traineeship/study ▪ Year began apprenticeship/traineeship/study 	<ul style="list-style-type: none"> ▪ Qualification type ▪ Type of qualification (for study/training not elsewhere reported)
	Continuation of previous study	<ul style="list-style-type: none"> ▪ Confirmation of previous study ▪ Continuation of previous study ▪ Month stopped study ▪ Year stopped study 	<ul style="list-style-type: none"> ▪ Qualification type ▪ Type of qualification (for study/training not elsewhere reported)

			<ul style="list-style-type: none"> ▪ Qualification type (for changed courses)
	Deferred study	<ul style="list-style-type: none"> ▪ Confirmation of deferred studies ▪ Resumption of deferred studies ▪ Continuation of resumed studies ▪ Month stopped study ▪ Year stopped study ▪ Same course as previously reported deferred ▪ Month began changed study ▪ Year began changed study 	<ul style="list-style-type: none"> ▪ Qualification type ▪ Type of qualification (for study/training not elsewhere reported) ▪ Qualification type (for changed courses)
	Commencements of new study	<ul style="list-style-type: none"> ▪ Type of study or training ▪ Still studying current qualification ▪ Month began study ▪ Year began study ▪ Month stopped study ▪ Year stopped study 	<ul style="list-style-type: none"> ▪ Qualification type
	Changed course	<ul style="list-style-type: none"> ▪ Study completed, withdrawn, deferred or changed ▪ Currently doing changed qualification ▪ Month stopped study ▪ Year stopped study ▪ Month stopped changed qualification ▪ Year stopped changed qualification 	<ul style="list-style-type: none"> ▪ Qualification type (for changed courses)
	No new study	<ul style="list-style-type: none"> ▪ New study or training since last interview 	-
	Current study	<ul style="list-style-type: none"> ▪ Current study or training (not elsewhere reported) ▪ Month began apprenticeship/traineeship/study ▪ Year began apprenticeship/traineeship/study 	<ul style="list-style-type: none"> ▪ Type of qualification (for study/training not elsewhere reported)
Apprenticeship or traineeship	Post-school study	<ul style="list-style-type: none"> ▪ Still doing current apprenticeship/traineeship ▪ Month started apprenticeship/traineeship ▪ Year started apprenticeship/traineeship ▪ Month stopped apprenticeship/traineeship ▪ Year stopped apprenticeship/traineeship ▪ Current study or training (not elsewhere reported) ▪ Month began apprenticeship/traineeship/study ▪ Year began apprenticeship/traineeship/study 	-
	Continuation of previous studies	<ul style="list-style-type: none"> ▪ Confirmation of previous apprenticeship/traineeship ▪ Still doing current apprenticeship/traineeship ▪ Month stopped apprenticeship/traineeship ▪ Year stopped apprenticeship/traineeship 	-
	Commencements of new study	<ul style="list-style-type: none"> ▪ Type of study or training ▪ Still doing current apprenticeship/traineeship ▪ Month started apprenticeship/traineeship 	-

	No new study	<ul style="list-style-type: none"> ▪ Year started apprenticeship/traineeship ▪ Month stopped apprenticeship/traineeship ▪ Year stopped apprenticeship/traineeship
	Current study	<ul style="list-style-type: none"> ▪ New study or training since last interview ▪ Current study or training (not elsewhere reported) ▪ Month began apprenticeship/traineeship ▪ Year began apprenticeship/traineeship
Other education (short courses/online courses etc.)	Commencements of new study	<ul style="list-style-type: none"> ▪ Type of study or training ▪ Still doing current Year 12 (post-school) / short-course/other/module/undefined study ▪ Month began study ▪ Year began study ▪ Month finished study ▪ Year finished study

As LSAY does not have a separate set of questions for studying at university or participating in VET, the filter variables were used to distinguish survey participants who have been in either university and/or VET at a given month.

Survey participants who studied certificates I–IV, VET/TAFE diploma, VET/TAFE advanced diploma/associate degree or VET/TAFE graduate diploma/graduate certificate were classified to be in VET, while survey participants who studied university diploma, university advanced diploma/associate degree, bachelor’s degree (including honours), or university graduate diploma/graduate certificate were classified to be in university.

In the analysis, VET consists of a combination of TAFE, apprenticeships/traineeships, and other education (such as short courses or online courses).

Variable definitions

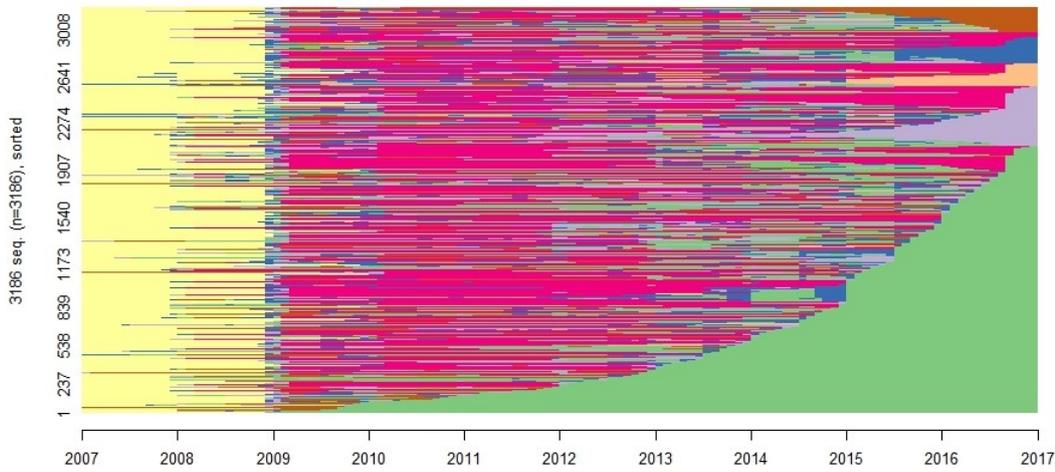
Table A2 Description and labels of characteristic variables

Socioeconomic and demographic characteristic	Description	Details	Value label
Male	Gender of the respondent	Dummy variable	1 – Respondent is male 0 – Respondent is female
Indigenous	Indigenous status of the respondent	Dummy variable	1 – Respondent is an Aboriginal or Torres Strait Islander 0 – Otherwise
Language other than English	Respondent who speaks a language other than English at home	Dummy variable	1 – Respondent speaks a language other than English at home 0 – Respondent speaks English at home NA, invalid, missing (n = 21) coded to 0
Overseas background	Respondent whose parent(s) were born in another country	Dummy variable from PISA 2006 index of overseas background.	1 – Respondent reported to be either <ul style="list-style-type: none"> ▪ First-generation students (those students born outside the country of assessment and whose parents were also born in another country) ▪ Second-generation students (those born in the country of assessment but whose parent(s) were born in another country) 0 – Respondent is a native (those students who had at least one parent born in the country) NA, invalid, missing (n = 39) coded to 0
Marital status	Marital status of the respondent	Dummy variable	1 – Respondent is married or de facto 0 – Respondent is not married NA, invalid, missing (n = 50) coded to 0
Dependent children	Respondent with at least one dependent children	Dummy variable	1 – Respondent has at least one dependent children 0 – Otherwise Don't knows coded to 0
Metropolitan location	Respondent attended school located in the metropolitan area at age 16	Dummy variable	1 – Respondent attends a school located in the metropolitan 0 – Otherwise

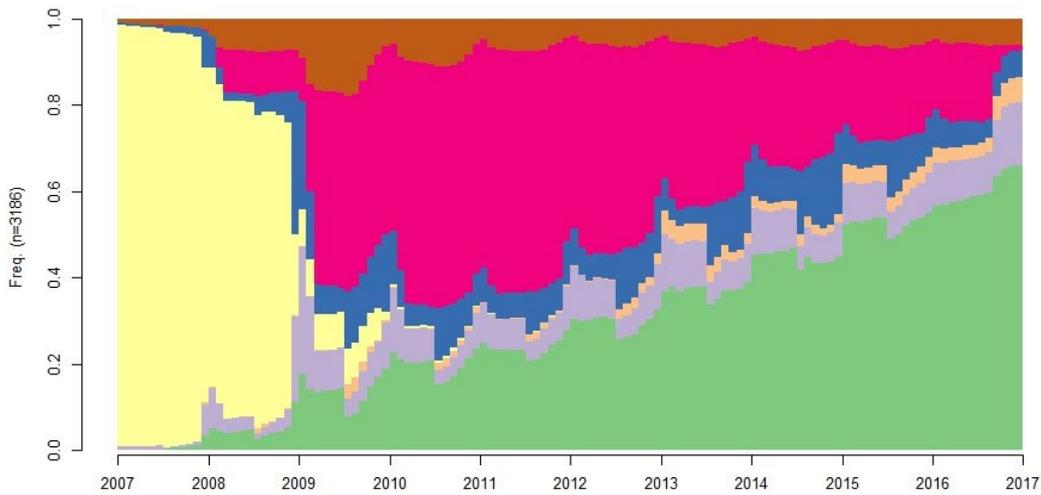
Socioeconomic status (SES)	Economic, social and cultural status of the respondent	Categorical variable. Has four quartiles and is derived from PISA 2006 index of economic, social and cultural status (ESCS). More information on the derivation of ESCS index is available at: < https://www.lsay.edu.au/_data/assets/pdf_file/0028/181486/LSAY_Y06UserGuideDataElementsA_2258.pdf >	1 – Highest quartile 2 – Second quartile 3 – Third quartile 4 – Lowest quartile NA, invalid, missing (n = 5) coded to 4 – Lowest quartile
Has a disability	Disability status of the respondent	Dummy variable	1 – Respondent has any disability which limits the amount or type of work they can do 0 – Otherwise NA, invalid, missing coded to 0
Lives with parents (age 20–25)	Respondent that lives with parents between the age of 20 to 25	Dummy variable	1 – Respondent lives with their parents between the age of 20 to 25 0 – Otherwise
Did not complete Year 12	Respondent who did not complete Year 12 studies	Dummy variable	1 – Respondent completed Year 12 0 – Respondent did not complete Year 12
Vocational studies in school	Respondent who undertook at least one vocational subject in school	Dummy variable	1 – Respondent undertook at least one vocational subject in school 0 – Otherwise
Mathematics and reading achievement (PISA)	Mathematics and reading achievement based on 2006 PISA assessment	Categorical variable with four quartiles based on PISA assessment mathematics and reading achievement Additional information available at: < http://www.oecd.org/pisa/pisaproducts/42025182.pdf >	1 – Highest quartile 2 – Second quartile 3 – Third quartile 4 – Lowest quartile NA, invalid, missing (n = 5) coded to 4 – Lowest quartile
Government payments	Respondent that has received either Youth Allowance or other payments (including parenting payment, sickness allowance, disability support pension, family tax benefit or any other government payment) from the government	Dummy variable Details of the Youth Allowance are available at: < https://www.humanservices.gov.au/individuals/services/centrelink/youth-allowance >. Government payment types are available at < https://www.humanservices.gov.au/individuals/services/centrelink >	1 – Respondent has received government allowance 0 – Otherwise

Figure A1 Summary activity patterns for the entire study sample (n = 3186)

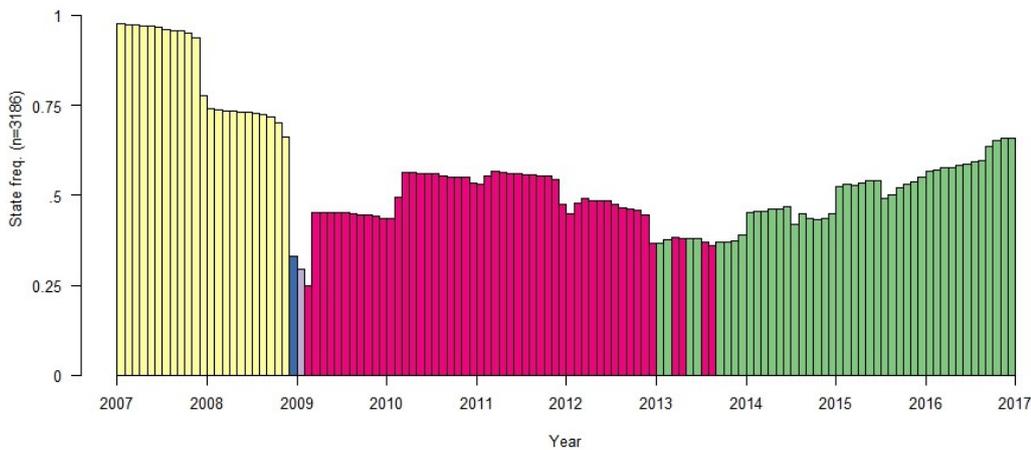
(a) Individual activity sequences



(b) Monthly proportion of activities



(c) Most frequent (modal) activity for each month



■ Emp-FT
 ■ Emp-PT
 ■ NILF(NEET)
 ■ School
 ■ Unemployed
 ■ University
 ■ VET/AT

Notes: Emp-FT = full-time employment; Emp-PT = part-time employment; NILF = not in the labour force; NEET = not in education, employment nor training; VET/AT = vocational education and training/apprenticeship/traineeship.

Table A3 Selected socioeconomic characteristics of the sample by pathway (with standard errors)

	Pathway 1		Pathway 2		Pathway 3		Pathway 4		Pathway 5	
	Higher education and work		Early entry to full-time work		Higher education and VET		Mixed and repeatedly disengaged		Mostly working part-time	
	%	<i>Std error</i>	%	<i>Std error</i>	%	<i>Std error</i>	%	<i>Std error</i>	%	<i>Std error</i>
Male	43.3	1.1	64.1	1.8	38.6	3.1	38.3	3.8	42.9	4.7
Indigenous	1.7	0.3	4.9	0.8	4.7	1.3	7.4	2.1	10.7	2.9
Language other than English	10.2	0.7	2.7	0.6	5.5	1.4	4.3	1.5	0	0
Overseas background	45.8	1.1	29.9	1.7	35	3.0	29.6	3.6	26.8	4.2
Marital status										
Married at 15–19	4.4	0.5	15.6	1.3	8.3	1.7	27.2	3.5	19.6	3.8
Married at 20–25	40.4	1.1	51	1.9	47.6	3.1	32.7	3.7	34.8	4.5
Dependent children										
At age 15–19	0.3	0.1	0.6	0.3	0.4	0.4	11.1	2.5	0.9	0.9
At age 20–25	3.0	0.4	13.6	1.3	8.3	1.7	29.6	3.6	13.4	3.2
Metropolitan location	76.8	1.0	61.2	1.8	68.9	2.9	57.4	3.9	65.2	4.5
Socioeconomic status (SES)										
Highest quartile	43.4	1.1	17.7	1.4	24	2.7	15.4	2.8	20.5	3.8
Lowest quartile	10.1	0.7	25.6	1.6	17.3	2.4	34.6	3.7	33	4.5
Has a disability	3.3	0.4	3.3	0.7	6.7	1.6	10.5	2.4	7.1	2.4
Lives with parents (age 20–25)*	80.9	0.9	80.7	1.5	82.3	2.4	72.2	3.5	84.8	3.4
Did not complete Year 12	0.9	0.2	21.5	1.5	14.6	2.2	28.4	3.6	20.5	3.8
Vocational studies in school	14.5	0.8	44.4	1.8	33.1	3.0	41.4	3.9	44.6	4.7
Mathematics achievement (PISA)										
Highest quartile	34.0	1.1	9.5	1.5	15.8	2.3	10.5	2.4	12.5	3.1
Lowest quartile	13.9	0.8	41.1	1.8	34.3	2.9	53.7	3.9	51.8	4.7
Reading achievement (PISA)										
Highest quartile	34.2	1.1	8.8	1.0	17.7	2.4	11.7	2.5	8.0	2.6
Lowest quartile	14.3	0.8	44.0	1.8	29.5	2.9	48.2	3.9	42.9	4.7
Government payments										
Youth allowance (YA)	62.4	1.1	34	1.8	58.7	3.1	75.9	3.4	57.1	4.7
Other payments (not YA)	37.5	1.1	27.4	1.7	43.3	3.1	67.9	3.7	38.4	4.6
Sample size	1928		730		254		162		112	

Notes: *lives with parents (age 20–25) is the only variable where the chi² test indicates that they are not significantly different across pathways at the 5% significance level.

Source: LSAY 2006.

Table A4 Highest qualification, occupation and labour force status at age 25 years by pathway (with standard errors)

	Pathway 1		Pathway 2		Pathway 3		Pathway 4		Pathway 5	
	Higher education and work		Early entry to full-time work		Higher education and VET		Mixed and repeatedly disengaged		Mostly working part-time	
	%	Std error	%	Std error	%	Std error	%	Std error	%	Std error
Highest qualification (at age 25)										
Certificate I	0.0	-	3.7	0.7	0.8	0.6	2.5	1.2	1.8	1.3
Certificate II	0.2	0.1	5.8	0.9	3.2	1.1	6.2	1.9	7.1	2.4
Certificate III	1.4	0.3	25.9	1.6	11.0	2.0	24.1	3.4	17.9	3.6
Certificate IV	1.2	0.2	13.7	1.3	15.4	2.3	11.1	2.5	8.9	2.7
Certificate – unknown	0.4	0.1	7.1	1.0	2.4	1.0	1.9	1.1	1.8	1.3
Advanced diploma/diploma	1.7	0.3	13.2	1.3	25.6	2.7	12.4	2.6	6.3	2.3
Bachelor's degree	69.4	1.1	1.1	0.4	26.8	2.8	6.2	1.9	4.5	2.0
Postgraduate diploma/certificate	4.4	0.5	0.4	0.2	2.8	1.0	0.0	-	0.9	0.9
Postgraduate	10.6	0.7	0.0	-	0.8	0.6	0.0	-	0.0	-
None	10.8	0.7	29.2	1.7	11.4	2.0	35.8	3.8	50.9	4.7
Total	100		100		100		100		100	
Occupation (at age 25)										
Managers	7.0	0.6	10.6	1.1	5.5	1.4	1.9	1.1	3.6	1.8
Professionals	50.1	1.1	7.1	1.0	20.1	2.5	3.1	1.4	6.3	2.3
Technicians & trades workers	4.1	0.4	29.3	1.7	8.7	1.8	4.3	1.6	8.9	2.7
Community & personal service workers	8.3	0.6	11.4	1.2	22.8	2.6	10.5	2.4	26.8	4.2
Clerical & administrative workers	11.2	0.7	16.0	1.4	16.5	2.3	4.3	1.6	12.5	3.1
Sales workers	6.1	0.5	7.5	1.0	7.9	1.7	9.3	2.3	18.8	3.7
Machinery operators & drivers	0.7	0.2	6.0	0.9	2.0	0.9	1.9	1.1	4.5	2.0
Labourers	2.2	0.3	6.0	0.9	5.9	1.5	9.3	2.3	8.0	2.6
Unknown/not classifiable	2.9	0.4	3.4	0.7	2.4	1.0	2.5	1.2	0.9	0.9
Not working (unemployed/NILF)	7.6	1.0	2.6	0.6	8.3	1.7	53.1	3.9	9.8	2.8
Total	100		100		100		100		100	
Labour force status (at age 25)										
Employed	92.4	0.1	97.4	0.1	91.7	0.2	46.9	0.4	90.2	0.3
Unemployed	2.3	0.0	1.2	0.0	3.5	0.1	16.7	0.3	5.4	0.2
Not in the labour force	5.3	0.1	1.4	0.0	4.8	0.1	36.4	0.4	4.4	0.2
Total	100		100		100		100		100	

Table A5 Tests for proportions (in tables 4 and 5)

Variable	Chi2(df)	p
Male	chi2(4) = 108.94	0.000
Indigenous	chi2(4) = 50.87	0.000
Language other than English	Fisher's exact	0.000
Overseas background	chi2(4) = 50.87	0.000
Marital status		
Married at 15–19	chi2(4) = 169.92	0.000
Married at 20–25	chi2(4) = 36.63	0.000
Dependent children		
At age 15–19	chi2(4) = 197.77	0.000
At age 20–25	chi2(4) = 213.21	0.000
Metropolitan location	chi2(4) = 82.71	0.000
Socioeconomic status (SES)	chi2(12) = 309.83	0.000
Has a disability	chi2(4) = 58.68	0.000
Lives with parents (age 20–25)	chi2(4) = 9.10	0.059
Did not complete Year 12	chi2(4) = 401.95	0.000
Vocational studies in school	chi2(4) = 311.81	0.000
Age 15 mathematics PISA	chi2(12) = 468.50	0.000
Age 15 reading PISA	chi2(12) = 473.32	0.000
Youth allowance (YA)	chi2(4) = 202.73	0.000
Other payments (not YA)	chi2(4) = 94.92	0.000
Highest qualification (at age 25)	chi2(36) = 2.2e+03	0.000
Highest occupation (at age 25)	chi2(36) = 1.4e+03	0.000
Labour force status (at age 25)	chi2(8) = 434.71	0.000
Sample size	3186	

Source: LSAY 2006.

Table A6: Wald tests for independent variables (in table 7)

Variable	Chi2(df)	Prob>chi2
Male <i>(ref: female)</i>	chi2 (4) = 90.37	0.000
Indigenous <i>(ref: non-Indigenous)</i>	chi2 (4) = 15.35	0.004
Overseas background <i>(ref: non-overseas background)</i>	chi2 (4) = 48.76	0.000
Socioeconomic status (SES) <i>(ref: top quartile)</i>	chi2 (12) = 124.73	0.000
Second quartile	chi2 (4) = 9.40	0.052
Third quartile	chi2 (4) = 61.69	0.000
Lowest quartile	chi2 (4) = 74.439	0.000
Metropolitan location <i>(ref: non-metropolitan location)</i>	chi2 (4) = 8.57	0.073
Mathematics achievement PISA <i>(ref: top quartile)</i>	chi2 (12) = 69.91	0.000
Second quartile	chi2 (4) = 16.09	0.003
Third quartile	chi2 (4) = 25.70	0.000
Lowest quartile	chi2 (4) = 56.10	0.000
Reading achievement PISA <i>(ref: top quartile)</i>	chi2 (12) = 32.08	0.001
Second quartile	chi2 (4) = 7.75	0.101
Third quartile	chi2 (4) = 19.23	0.001
Lowest quartile	chi2 (4) = 24.92	0.000
Vocational studies in school <i>(ref: no vocational studies in school)</i>	chi2 (4) = 124.48	0.000
Sample size	3186	

Note: Wald test Hypothesis Ho: all coefficients associated with given variable(s) are 0.

Source: LSAY 2006.

Table A7 Wald tests for independent variables (in table 8)

Variable	Chi2(df)	Prob>chi2
Male (ref: female)	chi2 (4) = 90.37	0.000
Indigenous (ref: non-Indigenous)	chi2 (4) = 15.35	0.004
Overseas background (ref: non-overseas background)	chi2 (4) = 48.76	0.000
Socioeconomic status (SES) (ref: top quartile)	chi2 (12) = 124.73	0.000
Second quartile	chi2 (4) = 9.40	0.052
Third quartile	chi2 (4) = 61.70	0.000
Lowest quartile	chi2 (4) = 74.44	0.000
Metropolitan location (ref: non-metropolitan location)	chi2 (4) = 8.57	0.073
Mathematics achievement PISA (ref: top quartile)	chi2 (12) = 69.91	0.000
Second quartile	chi2 (4) = 16.09	0.003
Third quartile	chi2 (4) = 25.70	0.000
Lowest quartile	chi2 (4) = 56.10	0.000
Reading achievement PISA (ref: top quartile)	chi2 (12) = 32.08	0.001
Second quartile	chi2 (4) = 7.75	0.101
Third quartile	chi2 (4) = 19.23	0.001
Lowest quartile	chi2 (4) = 24.92	0.000
Vocational studies in school (ref: no vocational studies in school)	chi2 (4) = 124.48	0.000
Sample size	3186	

Note: Wald test Hypothesis Ho: all coefficients associated with given variable(s) are 0.
Source: LSAY 2006.

Table A8 Model fit tests for multinomial logistic model

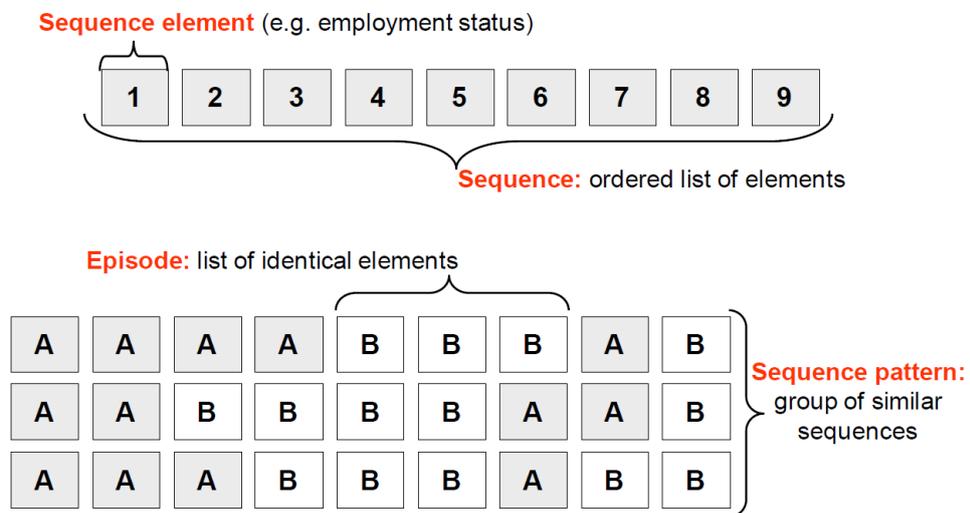
Log-likelihood	
Model	-3003.828
Intercept-only	-3544.018
Chi-square	
Deviance (df = 3126)	6007.655
LR (df = 56)	1080.38
p-value	0.00
R2	
McFadden	0.152
McFadden (adjusted)	0.135
Cox-Snell/ML	0.288
Cragg-Uhler/Nagelkerke	0.322
Count	0.661
Count (adjusted)	0.141
IC	
AIC	6127.655
AIC divided by N	1.923
BIC (df = 60)	6491.646

Appendix B: detailed methodology

What are sequences?

Sequences are ordered listings of elements (see figure B1). An element can be a certain status (for example, employment or marital status), an event (for example, a dance step or music note), or an object (for example, base pair of DNA). These elements are either ordered based on time (employment status in a given month, for example) or in a specific manner (such as a list of numbers, or a comic strip where the sequence of drawings arranged in interrelated panels form a narrative).

Figure B1 Example of a sequence



Source: Fasang (2014).

The Chi-square distance measure

The Chi-square distance measure quantifies differences in state distributions (Studer & Ritschard 2014, 2015). The aim of the measure is to focus on the time spent in each state within the sequences that is appropriate for the data and the analytical purposes of this study. Moreover, this approach retains the contemporaneity of the sequences and is ideally suited to the LSAY data, which contains a cohort of individuals of the same age who are likely to make transitions at similar times.

The Chi-square distance weights the squared differences for each state by the inverse of the overall proportion of time spent in the state and is defined as

$$d_x^2(x, y) = \sum_{j=1}^m \frac{(p_{jx} - p_{jy})^2}{p_j}, \quad (1)$$

where p_{jx} is the proportion of time spent in state j in sequence x , p_{jy} is the proportion of time spent in state j in sequence y , and p_j is the overall proportion of time spent in state j (Studer & Ritschard 2016). However, this measure is sensitive to the time spent in the states but insensitive to the order and exact timing of the states. This lack of sensitivity to

ordering is a considerable limitation, since the ordering of the states reflects the internal dynamics of a trajectory. Studer and Ritschard (2014, 2015) propose a solution to overcome this issue, whereby the distribution is evaluated in successive (potentially overlapping) periods or sub-sequences. Then the period-dependent Chi-square distance is

$$d_{x,k}^2(x, y) = \sum_{k=i}^K \sum_{j=1}^m \frac{(p_{jx_k} - p_{jy_k})^2}{p_{jk}}, \quad (2)$$

where the sub-sequence of x over period k is x_k and the overall proportion of time in state j in the k th interval is p_{jk} . The advantage of this measure is that it takes into account the timing as well as the duration of the transitions.¹⁷

Implementation in R

The **TraMineR** package in R statistical software was used for sequence analysis and the computation of pairwise dissimilarity between sequences. Gabadinho et al. (2011) was used to guide this part of the sequence analysis.

Below is the code used to compute the distance matrix in R:

```
chiDist <- seqdist(stwseq1, method = "CHI2", step =120)
```

- `seqdist` computes pairwise dissimilarities between sequences or dissimilarity from a reference sequence.
- `stwseq1` contains the sequences of 3186 individuals.
- `Method = "CHI2"` specifies the chi-square distance measure to be used to compute the pairwise dissimilarities.
- `'step = 120'` (or alternatively, `step = length(stwseq1)`) indicates that the chosen $K = 120$. Here, K is chosen to be equal to the length of the sequence so that the chi-square distance measure gains sensitivity to the timing aspect of sequences while maintaining some sensitivity to differences in durations (Studer & Ritschard 2015).

Cluster analysis

Utilising the evaluated distance matrix, hierarchical agglomerative cluster analysis was used to group similar sequences together.

There are two main approaches to hierarchical clustering, one being agglomerative clustering, whereby clusters are progressively grouped into larger ones; the other is divisive clustering, whereby clusters are divided into smaller ones.

In the analysis, Ward's hierarchical agglomerative clustering method was used. It is the most common approach in the sequence analysis literature (Martin, Schoon & Ross 2008).

As with all agglomerative clustering algorithms, each individual starts off as their own cluster. The closest pair of clusters is then merged at each iteration based on a distance measure of how close two clusters are.

¹⁷ See Studer & Ritschard (2016) for additional details.

Ward's method in agglomerative clustering merges two clusters A and B, which minimises the increase in the error sum-of-squares (ESS). More formally, the ESS of a cluster A is given by,

$$ESS_A = \sum_{i=1}^{n_a} (\mathbf{a}_i - \bar{\mathbf{a}})' (\mathbf{a}_i - \bar{\mathbf{a}}),$$

where \mathbf{a}_i represents the i th observation vector in cluster A, $\bar{\mathbf{a}} = \frac{1}{n_a} \sum_{i=1}^{n_a} \mathbf{a}_i$ is the mean vector of cluster A and n_a is the number of individuals in cluster A. The ESS of cluster B is calculated similarly (Strauss & von Maltitz 2017).

In each iteration, Ward's algorithm merges two clusters, cluster A and cluster B, which minimises the increase in the ESS, defined as,

$$\begin{aligned} I_{AB} &= ESS_{AB} - (ESS_A + ESS_B) \\ &= \frac{n_A n_B}{n_A + n_B} (\bar{\mathbf{a}} - \bar{\mathbf{b}})' (\bar{\mathbf{a}} - \bar{\mathbf{b}}), \end{aligned}$$

where ESS_{AB} is the error sum-of-squares of the newly combined cluster AB, $\bar{\mathbf{a}}$ and $\bar{\mathbf{b}}$ are the mean vectors of cluster A and cluster B respectively, and n_a , n_b represent the number of individuals in cluster A and cluster B respectively (Strauss & Maltitz 2017).

As hierarchical agglomerative clustering methods ultimately stop when all individuals are grouped into one single cluster, it does not present any meaning contextually. The second part of cluster analysis therefore involves deciding on the appropriate number of clusters, which involves striking a balance between low variation within clusters and sufficient variation between clusters while being analytically meaningful.

There are several statistical measures to assess the quality of the clusters obtained. Studer (2013) compiled them in the **WeightedCluster** library available in **R**. In the current analysis, all statistical measures that were available in the **WeightedCluster** library were used for comparison (including, but not limited to, Hubert's Gamma, Average Silhouette Width and Calinski-Harabasz index). A 'majority rule' approach was adopted to determine the number of clusters. That is, the number of clusters that the majority of the statistical measures recommended was chosen. As such, five clusters emerged as the representative pathways in the study sample, which were also meaningful within a socioeconomic context.



National Centre for Vocational Education Research

Level 5, 60 Light Square, Adelaide, SA 5000
PO Box 8288 Station Arcade, Adelaide SA 5000, Australia

Phone +61 8 8230 8400 **Email** ncver@ncver.edu.au

Web [<https://www.ncver.edu.au>](https://www.ncver.edu.au) [<https://www.isay.edu.au>](https://www.isay.edu.au)

Follow us:  [<https://twitter.com/ncver>](https://twitter.com/ncver)  [<https://www.linkedin.com/company/ncver>](https://www.linkedin.com/company/ncver)

