

Online Appendix Document 2

HILDA Re-education Project: Panel Sample: Sensitivity Analysis

1 Sample Selection

Treated sample: For any person in HILDA who ever reported *starting* a degree (determined by taking a person who switches from reporting not currently studying in one wave to currently studying in the next wave) and/or *completing* a degree, we select their first study event as a treatment observation if it satisfies three other conditions.

They are: (1) at least 21 years old in the starting year of study¹, (2) they were present in the two years before the start of study (in order to have information on their feature values), (3) there were not currently studying in any of the two years before the starting year of further study (to avoid reverse-causation issues), (4) they completed their further degree and (5) they were present in the survey and had a non-missing outcome 4 years after the start of study.

If a study event does not satisfy these conditions, we look to the next study event that satisfies these conditions or (if unavailable) delete the person from our sample completely. Conditions (3) and (5) together mean that we analyse a sample of individuals who started their degrees anytime between 2003 and 2015.

In our treated group, 1,814 individuals **started and completed a further educational degree**.

Control sample: These are those who had never started re-education throughout HILDA. From these control observations, we assign a time stamp to them for the year the control person theoretically started to study. We do this for every year from 2003 to 2019. This implies that never re-educated individuals can be duplicated and used

¹Note that we expanded the age range in this sensitivity analysis to ensure sufficient treatment observations for the estimation of the treatment outcome surfaces.

multiple times. For example, if a control individual is observed throughout the years 2001 to 2016, then they will be a control for the separate treated individuals that started re-education in 2003, in 2004, 2005 and up to 2017 i.e. the control individual will be duplicated 15 times.

There are 60,945 control observations i.e. individuals who never completed a further degree. However, as described above, these are non-unique observations in the sense that a control individual can be duplicated up to 15 times.

A Variable description

A.1 Outcome Variables

Weekly earnings from main job in fourth year after the individual started their re-education (*f4_wscmei*) records the weekly earnings from the main job for the individual in the fourth year after the individual started their re-education.

A.2 Treatment Variables: Re-education

Re-education completion based on both highest attainment and detailed qualifications (*redufl*) records whether the individual has completed re-education based on a comparison of the highest education attainment and the number of qualifications gained across waves 1 and 17. If either of these have gone up, *redufl* takes a value of 1 and 0 otherwise.

A.3 Input Variables

Characteristics in the Year Prior to Re-education Start

Demographics

Gender (*p1_hgsex*) records the gender of the individual. The value of 1 denotes males whereas the value 2 denotes females.

Age (*p1_hgage*) records the age of the individual in the year prior to re-education start.

Country of birth (*p1_anbcob*) records whether or not an individual was born in:

- Australia (value=1)

- English speaking countries (value=2)
- Non-English speaking countries (value=3)

Indigenous Status (*p1_anatsi*) records whether or not an individual is:

- Not indigenous (value=1)
- Aboriginal (value=2)
- Torres Islander (value=3)
- Both Aboriginal and Torres Islander (value=4)

Poor English speaking abilities (*p1_poeng*) records whether the individual has poor English speaking abilities in the year prior to re-education start.

State of residence (*p1_hhstate*) records the state of residence of the individual in the year prior to re-education start:

- NSW (value=1)
- VIC (value=2)
- QLD (value=3)
- SA (value=4)
- WA (value=5)
- TAS (value=6)
- NT (value=7)
- ACT (value=8)

Remoteness (*p1_hhsos*) records whether, in the year prior to re-education start, the individual lives in:

- A major city (value=0)
- An inner region (value=1)
- Outer and remote areas (value=2)
- migratory in nature (value=3)

Marital status (*p1_mrcurr*) records whether, in the year prior to re-education start, the individual was:

- Married (value=1)

- De facto (value=2)
- Separated (value=3)
- Divorced (value=4)
- Widowed (value=5)
- Single and never been married (value=6)

Household size (*p1_hhsize*) records the total number of individuals living in the same household as the individual (including the individual) in the year prior to re-education start.

Sexual orientation (*p1_lgtb*) records that the individuals sexual orientation is not heterosexual. The variable is constructed from the Sexual Identity question that is only asked in waves 12 and 16. We combine answers from both waves to create a binary indicator for the individual ever reporting a sexual identity that is not heterosexual, treating sexual orientation as a fixed trait for a given individual.

Parental Status

Number of dependents (*p1_totalkids*) records the number of children under 15 the individual had in the household in the year prior to re-education start.

Having children (*p1_anykid*) records the individual had any dependents in the household in the year prior to re-education start.

Children under 5 (*p1_kidu5*) records the individual had children under 5 in the household in the year prior to re-education start.

Age of youngest (*p1_rcyng*) records the age of the youngest children living with the respondent in the year prior to re-education start (including adult children).

Physical Health

Severity of health conditions (*p1_disdeg*) records whether, in the year prior to re-education start, the individual had:

- No health conditions (value=0)
- A mild condition (value=1)
- A moderate condition (value=2)
- A severe condition (value=3)

Labour Force Variables

Labour market status (*p1_lfs*) records whether the individual was:

- Employed (value=1)
- Unemployed (value=2)
- Not in the labour market (value=3)

Extent of working hour match with preferences (*p1_whpref*) records whether, in the year prior to re-education start, the match between the individuals total weekly working hours across all jobs and their preferred number of working hours made them:

- Underemployed by at least 4 hours a week (value=1)
- Roughly Matched: Preferred and Actual Hours Worked differ by less than 4 hours a week (value=2)
- Overemployed by at least 4 hours a week (value=3)

Employee type (*p1_emptype*) records whether, in the year prior to re-education start, the individual was:

- An employee (value=1)
- An employee of own business (value=2)
- Self Employed (value=3)
- Unpaid family worker (value=4)

Contract type (*p1_contype*) records whether, in the year prior to re-education start, the individual was:

- On a fixed term contract (value=1)
- On a casual contract (value=2)
- On a permanent contract (value=3)
- On other types of contracts (value=4)

Occupation (*p1_occ*) records whether, in the year prior to re-education start, the individual was working as:

- Armed forces (value=0)
- Legislators, Senior Officials and Managers (value=1)
- Professionals (value=2)

- Technicians and Associate Professionals (value=3)
- Clerks (value=4)
- Service Workers and Shop and Market Sales Workers (value=5)
- Skilled Agriculture and Fishery Workers (value=6)
- Craft and Related Trades Workers (value=7)
- Plant and Machine Operators and Assemblers (value=8)
- Elementary Occupations (value=9)

Union membership (*p1_union*) records whether the individual was a union member in the year prior to re-education start.

Real household income (*p1_rhdi*) records the real value of the individuals total household income indexed at 2012 price levels and adjusted for household size in the year prior to re-education start.

Partner labour force status (*p1_plfs*) records whether, in the year prior to re-education start, the individual:

- Had no partner or no resident partner (value=0)
- Had a partner who was employed (value=1)
- Had a partner who was unemployed (value=2)
- Had a partner who was not in the labour force (value=3)

Years in paid work (*p1_egtjb*) records the total number of years in paid work the individual has spent in the year prior to re-education start.

Percent finding as least as good a job (*p1_jbmggj*) records, for employees, the percentage that they will find as least as good a job as they currently have in their own estimation in the year prior to re-education start

Occupational scale (*p1_jbm06s*) records the Australian Socioeconomic Index 2006 ranking of the individuals occupation in the year prior to re-education start. It ranges from 0 to 100, with higher scores indicating higher occupational status.

Tenure with employer (*p1_jbempt*) records the total years spent with the current employer for the individual in the year prior to starting re-education.

Parental information

Fathers country of birth (*p1_fcob*) records whether or not the individuals father was born in:

- Australia (value=1)
- English speaking countries (value=2)
- Non-English speaking countries or indigenous (value=3)

Mothers country of birth (*p1_mcob*) records whether or not the individuals mother was born in:

- Australia (value=1)
- English speaking countries (value=2)
- Non-English speaking countries or indigenous (value=3)

Fathers education records whether the individuals fathers highest education, as reported in 2005, was:

- None (value=1)
- Primary (value=2)
- Below secondary (value=3)
- Secondary (value=4)
- Post-secondary, non-university (value=5)
- Post-secondary, university (value=6)

Mothers education records whether the individuals mothers highest education, as reported in 2005, was:

- None (value=1)
- Primary (value=2)
- Below secondary (value=3)
- Secondary (value=4)
- Post-secondary, non-university (value=5)
- Post-secondary, university (value=6)

Father undertaken post-school qualification through employer or non-tertiary means (*p_fpsm*) records whether the individuals father had undertaken his highest qualification through employers or other channels other than tertiary education, as reported in 2005.

Mother undertaken post-school qualification through employer or non-tertiary means (*p_mpsm*) records whether the individuals mother had undertaken his highest qualification through employers or other channels other than tertiary education, as reported in 2005.

Fathers Employment at age 14 (*p1_femp*) records whether the individuals father was working or not when they were aged 14.

Mothers Employment at age 14 (*p1_memp*) records whether the individuals mother was working or not when they were aged 14.

Father substantially unemployed growing up (*p1_fsue*) records whether the individuals father had been unemployed or 6 months or more when they were aged 14.

Fathers Occupation (*p1_focc*) records whether at age 14 the individuals father was last known working as:

- Armed forces (value=0)
- Legislators, Senior Officials and Managers (value=1)
- Professionals (value=2)
- Technicians and Associate Professionals (value=3)
- Clerks (value=4)
- Service Workers and Shop and Market Sales Workers (value=5)
- Skilled Agriculture and Fishery Workers (value=6)
- Craft and Related Trades Workers (value=7)
- Plant and Machine Operators and Assemblers (value=8)
- Elementary Occupations (value=9)

Mothers Occupation (*p1_mocc*) records whether at age 14 the individuals mother last known working as:

- Armed forces (value=0)
- Legislators, Senior Officials and Managers (value=1)

- Professionals (value=2)
- Technicians and Associate Professionals (value=3)
- Clerks (value=4)
- Service Workers and Shop and Market Sales Workers (value=5)
- Skilled Agriculture and Fishery Workers (value=6)
- Craft and Related Trades Workers (value=7)
- Plant and Machine Operators and Assemblers (value=8)
- Elementary Occupations (value=9)

Income Support

On income support (*p1_onis*) records the individual was on income support in the year prior to starting re-education

On Newstart (*p1_onnsa*) records the individual was on Newstart Allowance in the year prior to starting re-education

On Age Pension (*p1_onap*) records the individual was on Age Pension in the year prior to starting re-education

On DSP (*p1_ondsp*) records the individual was on Disability Support Pension in the year prior to starting re-education

On Carer Payment (*p1_oncp*) records the individual was on Carer Payment in the year prior to starting re-education

On Widow Allowance/Wife Pension (*p1_onww*) records the individual was on Widow Allowance/Wife Pension in the year prior to starting re-education

On Youth Allowance (*p1_ony*) records the individual was on Youth Allowance in the year prior to starting re-education

On Mature Age Allowance (*p1_onma*) records the individual was on Mature Age Allowance in the year prior to starting re-education

On Mature Age Partner Allowance (*p1_onmap*) records the individual was on Mature Age Partner Allowance in the year prior to starting re-education

On Ab/Austudy (*p1_onsdy*) records the individual was on Ab/Austudy in the year prior to starting re-education

On Bereavement Allowance (*p1_onba*) records the individual was on Bereavement Allowance in the year prior to starting re-education

On Sickness Allowance/Special Benefits (*p1_onsab*) records the individual was on Sickness Allowance/Special Benefits in the year prior to starting re-education

On Partner Allowance (*p1_onpa*) records the individual was on Partner Allowance in the year prior to starting re-education

On Parenting Payments (*p1_onpp*) records the individual was on Parenting Payments in the year prior to starting re-education

Housing situation

Mortgage balance (*p1_hsmgowe*) records the amount still owing on the mortgage that the individual had in the year prior to re-education start. For those without a mortgage or not home owner, the mortgage balance is set to 0.

Non home owners (*p1_renter*) records whether the individual was renting or not living in their own homes in the year prior to re-education start.

Prior Year Outcomes

Weekly income from all jobs (*p1_earning*) records the weekly earnings from all jobs for the individual in the year prior to the individual starting their re-education.

Weekly income from main job (*p1_wscmei*) records the weekly earnings from the main job for the individual in the year prior to the individual starting their re-education.

Weekly working hours (*p1_wkhr*) records the total number of hours the individual works in all jobs in a week on average in the year prior to the individual started their re-education. Working hours are set to 0 for those not working.

Real hourly wage (*p1_rlwage*) records the real hourly wage of the individual in the year prior to the individual starting their re-education, indexed at 2012 price levels. Hourly wages are set to 0 for those not working and set to missing for those reporting working more than 100 hours a week. All wages have then been adjusted up by \$1 to preserve sample size for the logarithm transformation.

Log hourly wage (*p1_lnwage*) records the log of *p1_rlwage*.

Mental health (*p1_ghmh*) records the transformed mental health scores from the aggregation of mental health items of the SF-36 Health Survey, as reported by the individual in the year prior to the individual started their re-education. It ranges from 0 to 100, with higher scores indicating better mental health.

Life satisfaction (*p1_losat*) records the life satisfaction score reported by the individual in the year prior to the individual started their re-education. It ranges from 0 to 10, with higher scores indicating higher life satisfaction.

Delta variables

For all the variables described in the preceding section titled Characteristics in the Year Prior to Re-education Start, we create a further set of change or delta variables. Specifically, each delta variable is the subtracting of the value of a given characteristic in the two years prior to starting re-education from the value of this characteristic in the year prior to re-education start.

All delta variables are denoted by the *d_* prefix.

Education-related variables

Level of re-education completed: Bachelor and above (*bachab*) records whether the individual had completed re-education at bachelor and above. The variable is set to 0 for the control group and missing for those who had completed certificates.

Level of re-education completed: Below Bachelor (*bbach*) records whether the individual had completed re-education that is below bachelor level. The variable is set to 0 for the control group and missing for those who had completed a bachelor or higher qualification.

Main field of study: technical degree (*techdeg*) records whether the individuals main field of study was a technical degree. The variable is set to 0 for the control group and missing for those whose main field of study was a qualitative degree. Technical degrees include:

- Natural and physical sciences
- Information technology
- Engineering and related technologies
- Architecture and building
- Agriculture, environment and related studies
- Medicine
- Nursing
- Other health-related (e.g. Pharmacy, Dental studies, Rehabilitation therapies, Optical science, Veterinary studies)
- Management and commerce (e.g. Accounting, Business, Sales and marketing, Banking and finance, Office studies)

- Law

Main field of study: qualitative degree (*qualdeg*) records whether the individuals main field of study was a qualitative degree. The variable is set to 0 for the control group and missing for those whose main field of study was a technical degree. Qualitative degrees include:

- Education
- Society and culture (e.g. Economics, Political science, Social work, History, Psychology, Languages, Religion, Sport)
- Creative arts
- Food, hospitality and personal services
- Other

Study duration (*fsddur*) records the total number of waves an individual had spent studying from the start of their first study event counted in our sample.

Starting Study intensity (*csftsd*) records whether the individual was studying full time or not when they started their re-education.

Finishing Study intensity (*fsftsd*) records whether the individual was studying full time or not when they completed their re-education.

Other variables

Number of waves in HILDA (*numwave*) records the number of waves in which the respondent has submitted a valid response for the HILDA survey.

A.4 Variables that are not included in the model

The unique person identifier (*xwaveid*)

Wave started re-education (*icswave*)

Wave completed re-education (*ifswave*)

Control group indicator (*control*)

Started but did not complete re-education between 2003-2017 (*ncomp*)

Starting year of re-education imputed (*impute*) is a binary indicator for individuals for which we observe their re-education completion but they never reported ever starting re-education and so we had to impute a starting wave for these individuals.

Started re-educaton in wave 2018/19 (*lateststart*) is an indicator for those individuals who had started their re-education in 2018 or 2019.