Changes to the HILDA data made between Release 5.1 (March 2007) and Release 6.0 (January 2008)

Changes to the Master File

- Sampling strata variable (*ahhstrat*) created from Wave 1 major statistical region variable (*ahhmsr*) and serpentine ordering of CDs across Australia. The same stratification is given to new household members when they join the sample at a later wave. The previous name given to this stratification variable on the master file (*ahhmsr*) has a different specification on other files (i.e. it would not be applicable if person was not enumerated in Wave 1).
- [In-confidence master file only]. Minor corrections (about 30 cases) to most detailed levels of geography variables as now using the new updated national address file (GNAF). Minor changes to SEIFA raw scores, which are matched by geography.
- [In-confidence master file only]. Date of birth updated for 94 cases. This is usually a single correction to day of birth or month of birth or year of birth. Or the date of birth was unknown to the person answering the Household Form at a prior wave and has been supplied in the current wave.

Changes to Wave 5 files

Cleaning

- *ebncoth2* F16 "Current pensions and benefits – other (specify)". Removed 13 cases as contained other income, not pensions or benefits.
- Interviewer observations. 23 cases corrected in Z2 – Z10 due to skip issue. CPQ and NPQ codes at Z2 altered from Wave 7 to prevent this reoccurring.
- *ehgage* Age at 30 June 2005. Due to revisions in date of birth, 45 ages changed (usually by plus or minus 1 year). Age changes effected for prior waves.

Corrections in calculations

- *eoifrnta* Rental income – 15 cases that were imputed set to zero. The "Already reported (as part of business income)" code in the questionnaire was one generally used to indicate no answer so these cases were inappropriately set to missing and imputed. As already reported, rental income set to zero.
- *ejhljii2* Three cases corrected to ANZSIC 'no further information' codes
- *eehtuj eehtjb eehto* History variables for employment activities (time working, unemployed or other activity) since leaving full time education. The algorithm was updated which more accurately calculates the calendar periods to be added to previously known activities. The algorithm was revised for individuals who were not interviewed for one or more waves, resulting in an increase in the number of unable to be determined cases.
- *emrn* History variable for number of marriages. Algorithm corrected from not asked (-1) to zero marriages for never married individuals (who are not asked the number of marriages).
- *eedhists* and *eedhigh* For continuing persons no direct question was made of school year attended in current year (a new question has been added for Wave 7). For each year continuing persons haven't left school and are enrolled, the history variable 'Highest year of school completed/currently attending' (*_edhists*) is incremented by one year until Year 12 is achieved. This caused some movements between 'Year 11 and below' and 'Year 12' in *_edhigh*.

New variables

- *ehgdli1* to *ehgdli14* Date last interviewed (HF)
- *ehgwli1* to *ehgwli14* Wave last interviewed (HF)
- *ehgdli* date last interviewed (PQ)
- *ehgwsli* Weeks since last interviewed (for continuing persons)
- One-digit version of occupation and industry variables
- Replicate weights 31 thru 45

Deleted variables

- History variables: Previous job

Documentation corrections

- Household form marked up questionnaire (both u&c). At page 2, said DOB at June 30 2004, corrected to June 30 2005.
- In "Derived Variable Coding Framework e51c.pdf" page D80 it incorrectly states that orderly (*epnorder*) is reversed in the construction of *epnconsc* Personality scale Conscientiousness. Orderly is not reversed in the construction of this scale.

Changes to Wave 4 files

Cleaning

- 38 cases corrected in Z2 – Z10 (interviewer observations) due to skip issue. CPQ and NPQ codes at Z2 altered from wave 7 to prevent this reoccurring.

Corrections in calculations

- *doifrnta* Rental income – 14 cases that were imputed set to zero. The "Already reported (as part of business income)" code in the questionnaire was one generally used to indicate no answer so these cases were inappropriately set to missing and imputed. As already reported, rental income set to zero.
- Corrected family number and income unit numbers for household '24822' as relationship had been updated.
- *dehtuj dehtjb dehto* History variables for employment activity since leaving full time education. The algorithm was updated which more accurately calculates the calendar periods to be added to previously known activities. The algorithm was revised for individuals who were not interviewed for one or more waves, resulting in an increase in the number of 'unable to be determined' cases.
- *dmrn* History variable for number of marriages. Algorithm corrected from not asked (-1) to zero marriages for never married individuals (who are not asked the number of marriages).
- *dedhists* and *dedhigh* For continuing persons no direct question was made of school year attended in current year (a new question has been added for Wave 7). For each year continuing persons haven't left school and are enrolled, the history variable 'Highest year of school completed/currently attending' (_edhists) is incremented by one year until Year 12 is achieved. This caused some movements between 'Year 11 and below' and 'Year 12' in _edhigh.

New variables

- *dhgdli1* to *dhgdli14* Date last interviewed (HF)
- *dhgwli1* to *dhgwli14* Wave last interviewed (HF)
- *dhgdli* date last interviewed (Eperson)

- *dhgwsli* Weeks since last interviewed (for continuing persons)
- One-digit version of occupation and industry variables
- Replicate weights 31 thru 45

Deleted variables
- History variables: Previous job

Documentation corrections
- None

Changes to Wave 3 files

Cleaning
- Section Z: 30 cases corrected Z2 – Z10.
- Corrected family number and income units for 76571.

Corrections in calculations
- *coifrnta* Rental income – 14 cases that were imputed set to zero. The "Already reported (as part of business income)" code in the questionnaire was one generally used to indicate "no answer" so these cases were inappropriately set to missing and imputed. As already reported, rental income set to zero.
- *cehtuj cehtjb cehto* History variables for employment activity since leaving full time education. The algorithm was updated which more accurately calculates the calendar periods to be added to previously known activities. The algorithm was revised for individuals who were not interviewed for one or more waves, resulting in an increase in the number of 'unable to be determined' cases.
- *cmrn* History variable for number of marriages. Algorithm corrected from not asked (-1) to zero marriages for never married individuals (who are not asked the number of marriages).
- *cedhists* and *cedhigh* For continuing persons no direct question was made of school year attended in current year (a new question has been added for Wave 7). For each year continuing persons haven't left school and are enrolled, the history variable 'Highest year of school completed/currently attending' (_edhists) is incremented by one year until Year 12 is achieved. This caused some movements between 'Year 11 and below' and 'Year 12' in _edhigh.

New variables
- *chgdli1* to *chgdli14* Date last interviewed (HF)
- *chgwli1* to *chgwli14* Wave last interviewed (HF)
- *chgdli* date last interviewed (Eperson)
- *chgwsli* Weeks since last interviewed (for continuing persons)
- One-digit version of occupation and industry variables
- Replicate weights 31 thru 45

Deleted variables
- History variables: Previous job

Documentation corrections
- None

Changes to Wave 2 files

Cleaning

- Corrected family number and income units for households 52331 and 35712.

Corrections in calculations

- Rental income – 10 cases that were imputed set to zero. The "Already reported (as part of business income)" code in the questionnaire was one generally used to indicate "no answer" so these cases were inappropriately set to missing and imputed. As already reported, rental income set to zero.
- *bmrn* History variable for number of marriages. Algorithm corrected from not asked (-1) to zero marriages for never married individuals (who are not asked the number of marriages).
- Since Wave 2, wealth was first released. Superannuation for retirees (*bpwsuprt*) nine cases set to zero as had incorrectly been set to refused. Superannuation for non-retirees (*bpwsupwk*) 135 cases set to zero as had incorrectly been set to refused or don't know. Superannuation imputation flags at person level corrected. [Imputed variables and flags at household level had been constructed correctly.]
- *bedhists* and *bedhigh* For continuing persons no direct question was made of school year attended in current year (a new question has been added for Wave 7). For each year continuing persons haven't left school and are enrolled, the history variable 'Highest year of school completed/currently attending' (*_edhists*) is incremented by one year until Year 12 is achieved. This caused some movements between 'Year 11 and below' and 'Year 12' in *_edhigh*.

New variables

- *bhgdli1* to *bhgdli14* Date last interviewed (HF)
- *bhgwli1* to *bhgwli14* Wave last interviewed (HF)
- *bhgdli* date last interviewed (Eperson)
- *bhgwsli* Weeks since last interviewed (for continuing persons)
- One-digit version of occupation and industry variables
- Replicate weights 31 thru 45
- Household level joint and own credit card debt supplied for completeness of wealth model. The total credit card debt at household level supplied in previous releases is still provided.

Deleted variables

- History variables: Previous job

Documentation corrections

- *bw1dt1* to *bw1dt14* (date last interviewed) renamed to *bhgdli1* to *bhgdli* for conformity with later waves.
- *bpwsupr* renamed to *bpwsuprt*. Wealth: retireees superannuation. For consistency with household level wealth variables

Changes to Wave 1 files

Cleaning

- *aoifrnta* Rental income – nine cases that were imputed set to zero. The "Already reported (as part of business income)" code in the questionnaire was one generally used to indicate

"no answer" so these cases were inappropriately set to missing and imputed. As already reported, rental income set to zero.

- Corrected family number and income units for 010319 and 003910.
- Where it was unknown whether the respondent's mother or father was working when the respondent was aged 14, the occupation was set to not asked. Removes two father occupations and four mother occupations.
- Minor corrections to skip pathways for three variables (resets a handful of cases from no answer to not asked).

Corrections in calculations

- *ajbmday* Main job – days usually worked – code 5 "Other – specify days below" changed to code 8 "Other – specify days below" in accord with code values in subsequent waves.

New variables

- One-digit version of occupation and industry variables

Deleted variables

- None

Documentation corrections

- None