## **ECONAU Data Set Documentation**

# Dale T. Mortensen Centre, Aarhus University — Version 21/01/2022

Henning Bunzel

#### 1 IDAP, IDA, Persons, labour market attachment and experience

**DSECONAU Updated:** 19072018, 1980-2016, yearly data set.

ECONAU Updated: 04032019.

### 1.1 Remarks on Update

**NYT 4-kvartal 2018** 2007-2016. Alvorlig fejl. **NYT 2-kvartal 2018** 2008-2016. Alvorlig fejl.

**Changes:** The last FSE update of DSECONAU data covers the years 2008-2016.

- ALDER\_NOV and KOEN are new variable in FSE and not included in ECONAU-IDAP. The variables are missing 1980 to 2007, because the ECONAU data sets are not updated for 1980 to 2007. The variables are available in RAS and PER-SONER.
- LEDDEL and LEDFULD is not continued after 2008. New variables are LEDIGHED\_NETTO and LEDIGHED\_BRUTTO. In ECONAU-IDAP the variable LEDIGHED\_NETTO is a concanetaion of (LEDDEL-LEDFULD) and LEDIGHED\_NETTO.
- ARLEDGR is not continued after 2008. In ECONAU-IDAP GRADAAR\_NETTO is a concatenation of ARLEDGR and ARLEDGR\_NETTO.
- *AKASSE\_ID* is a new variable and not available before 2008. It is not included in *ECONAU-IDAP*. *AKASSE\_KODE* is included in *IDAPALL* from *RAS*.

 $\begin{tabular}{ll} \textbf{Department of Economics and Business Economis, Aarhus University, e-mail: $hbunzel@econ.au.dk} \end{tabular}$ 

2 Henning Bunzel

IDAP has a number of wage variables aggregated over groups of jobs. After 2008
there exist a basic and an extended wage. The 1980 to 2007 variables are concatenated with basic wage variables.

• A large number of new descriptive variables have been added from 2018.

#### **Errors and Questions:**

- 1. 153.481 persons in *PERSONER* cannot be found in *IDAP*, 2006, 2008-2016.
- 2. 98.444 persons in *IDAP* cannot be found in *PERSONER*, 2008-2016.
- 3. The following values are not included in 2008-2013: LEDDEL, LEDFULD.
- 4. LEDIGHED starts in 2008.
- 5. *EJNOVSUM* has 4.968.640 missing values in 2004. It jumps from 559 DKK in 2007 to 1507 DKK in 2008.
- 6. *FORSKAT\_RAS* has app 2.5 mio missing values 1991-2003. In the other years missing is coded as 0. Code values 0 are recoded to missing.
- 7. ARLEDGR has a low value in 2008. Check?
- 8. The data set is unique on: (AAR, PNR)
- 9. RENAME: ATPXX to ATP\_BELOEB\_SATSB, FORSKAT\_RASto FORSIKRINGSKAT-EGORI\_KODE

#### Remarks.

1. The key(AAR,PNR) is unique.