

# LMDG Dataset Documentation

Dale T. Mortensen Centre, Aarhus University —  
Version 03/04/2025

Henning Bunzel

## 1 ESR\_SE\_CVR, Yearly data with relations between SENR and CVRNR from ESR

**DSECONAU Updated:** It is an *LMDG* dataset.

**ECONAU Updated:** 03-04-2022. Special ESR version to ECONAU, last full year 2021. Delivered to LMDG project 702728.

### 1.1 Description

*ESR\_SE\_CVR* is a yearly dataset defining the relation between *SENR* and *CVRNR* constructed from *ESR\_SE\_CVR\_EVENTS* by creating an observation for each year in the interval 1980-2021, where there is an active Rel spell, i.e. an association between *SENR* and *CVRNR*.

*CVR* was created in 1999, but associations existing before 1999 have been added to *ESR*. Therefore it is possible to create yearly datasets from 1980.

There exist start- end dates for both *SENR* and Rel spells. It is assumed that Rel spells are contained within a SE spell, but this is not always the case. Therefore some editing of the dates have taken place, see section Detailed Description.

It seems to be a design of the dataset that a Rel spell is not made inactive. In the current version of the Yearly Rel spells, *ESR\_SE\_CVR*, the Rel spells are truncated to *SLUT\_AAR*, see section Detailed Description. The original *REL\_SLUT\_DATO* is kept unchanged for documentation.

The job *\_1\_esr\_se\_cvr* in subdirectory *dan* uses dataset *YEARS\_TRUNC* in directory *esr\_se\_cvr\_events\_2022\data* to create *ESR\_SE\_CVR*.

The job *\_1\_dups\_aar\_senr\_cvrnr* in subdirectory *test* checks if the keys (*AAR*, *SENR*) and (*AAR*, *SENR*, *CVRNR*) are unique.

The key (*AAR*, *SENR*) is not unique, it has 816 duplicates. The key (*AAR*, *SENR*, *CVRNR*) has 6 duplicates.

The duplicates are caused by short gaps within a year for the Rel spells (*REL\_START\_AAR*, *REL\_SLUT\_AAR*).

The job *\_2\_delete\_dups\_aar\_se\_cvr* in subdirectory *dan* concatenates these spells. The resulting dataset has 438 fewer observations.

The job *\_2\_dups\_aar\_senr\_cvrnr* in subdirectory *test* tests that the keys now are unique.

After *ESR\_SE\_CVR* has been updated *SPELL\_FIRM\_ID* must be updated. The jobs are in directory *spell\programs\firm\_id\before2008*

The job *\_1\_test\_match\_secvr\_conesr* tests for a match for all observations in datasets *RAW\_EMPLOMENTMENT\_CONESR\_RAS* and *ESR\_SE\_CVR* using (*AAR*, *SENR*).

If there is not a match for all observations it can be caused by either missing *SENR* in *ESR\_SE\_CVR* or errors in the Rel spell dates.

The job *\_1\_test\_match\_semr\_only* tests for a match only using *SENR*.

If there is a match for all observations the error is in Rel spell dates. If the Rel spell is singular the dates are not needed and the range is set to [1980,2021] by the job *\_3\_change\_rel\_dates\_conesr* in subdirectory *esr\_se\_cvr\_2022\dan*.

The job *freq\_edits* in subdirectory *analyse* shows the deletes and edits done on *ESR\_SE\_CVR*, see table 3.

The job *se\_cvr\_count* in subdirectory *analyse* produces descriptive tables for the dataset, see table 1 and 2.

#### **Remarks on the 2022 dataset *ESR\_SE\_CVR***

- The dataset has 21,347,392 observations.
- The key is (*AAR*, *SENR*, *CVRNR*).
- The dataset has 1,916,331 unique *SENR*.
- The dataset has 1,862,278 unique *CVRNR*.
- The dataset has 1,197,170 unique keys (*SENR*, *CVRNR*)

The *SENR* is assigned when an economic activity is registered for paying income tax, value added tax, or custom. The economic unit can change *CVRNR*. Therefore the Rel spell dates are needed to keep track of which *CVRNR* is assigned to a *SENR* in a given year. Table 1 shows how many *CVRNR* a *SENR* has been assigned to over the years. The table shows that 97.5% of the *SENR* has only been assigned one *CVRNR*. For these observations the Rel spell dates are not needed. The table 1 shows how many *SENR* a *CVRNR* is associated with. It is truncated at 30. The largest number is 472. The detailed list can be seen on the fseserver, xls file *se\_cvr\_count* in directory *econau\esr\esr\_se\_cvr\_2022\dokumenter*.

**Table 1** Number of CVRNR assigned to a SENR during the years 1980 to 2021

| COUNTS | FREQUENCY |
|--------|-----------|
| 1      | 1944908   |
| 2      | 1083      |
| 3      | 10        |

**Table 2** Number of SENR assigned to a CVRNR during the years 1980 to 2021

| COUNTS | FREQUENCY |
|--------|-----------|
| 1      | 1,814,919 |
| 2      | 41,849    |
| 3      | 5,564     |
| 4      | 1,206     |
| 5      | 420       |
| ≥5     | 239,400   |

## 1.2 Detailed Description

The job *\_4\_esr\_se\_cvr\_events\_year* in directory *esr\_se\_cvr\_events\_2022\dan* creates two datasets, *YEARS* and *YEARS\_TRUNC* in directory *esr\_se\_cvr\_events\_2022\data* and documents the edits and deletes performed, see table 3

Apparently, Rel spells are never closed. In the TRUNC dataset Rel spells are truncated to *SLUT\_AAR* of the SE spell.

The job *esr\_se\_cvr* in directory *dan* inputs the dataset *YEARS\_REL\_SPELL\_TRUNC* and creates the dataset *ESR\_SE\_CVR* for the period 1980-2021.

All SE spells and Rel spells are truncated to the period [1980,2021].

Table 3 shows that almost all Rel spells are truncated to *SLUT\_AAR*.

**Table 3** Number of observations, deleted and edited

| Condition                                  | Code | Action                                           | Number of observations |
|--------------------------------------------|------|--------------------------------------------------|------------------------|
| <i>REL_SLUT_AAR</i> < 1980                 | 1    | deleted                                          |                        |
| <i>cvmr</i> missing                        | 2    | deleted                                          | 323,772                |
| <i>START_AAR</i> > <i>LAST_YEAR</i>        | 3    | deleted                                          | 18,428                 |
| <i>SLUT_AAR</i> < 1980                     | 4    | deleted                                          | 41                     |
| <i>REL_START_AAR</i> > <i>LAST_YEAR</i>    | 5    | deleted                                          | 272                    |
| <i>REL_START_AAR</i> > <i>SLUT_AAR</i>     | 6    | deleted                                          | 7,014                  |
| <i>REL_SLUT_AAR</i> < 1980                 | 7    | deleted                                          | 0                      |
| <i>REL_SLUT_AAR</i> < <i>START_AAR</i>     | 8    | deleted                                          | 9                      |
| <i>REL_SLUT_AAR</i> > <i>SLUT_AAR</i>      | 10   | <i>REL_SLUT_AAR</i>                              | 1,916,246              |
| <i>REL_SLUT_AAR</i> > <i>LAST_YEAR</i>     | 11   | <i>REL_SLUT_AAR</i> is assigned <i>LAST_YEAR</i> |                        |
| Both code 2 and code 10                    | 12   | <i>REL_SLUT_AAR</i>                              | 30,129                 |
| SE spell dates missing                     | 13   | assigned Rel spell                               | 30,129                 |
| <i>START_AAR</i> < 1980                    | 14   | <i>START_AAR</i> is assigned 1980                |                        |
| <i>SLUT_AAR</i> > <i>LAST_YEAR</i>         | 15   | <i>SLUT_AAR</i> is assigned <i>LAST_YEAR</i>     |                        |
| <i>REL_START_AAR</i> < 1980                | 17   | <i>REL_START_AAR</i> is assigned 1980            |                        |
| Rel spell is not contained with a SE spell | 90   | Rel dates are set to [1980,2021]                 | 85                     |

# ECONAU Variables in ESR\_SE\_CVR

Henning Bunzel

## 1 AAR, Year for the reference period of the observation

År for reference perioden for observation

**Period: 2000-2021**

### *1.1 Beskrivelse*

Året for reference perioden for observation

### *1.2 Description*

Year for the reference period of the observation

## 2 CVRNR, Encrypted legal unit identifier. Data break 1994

CVRNR, krypteret nøgle for juridisk enhed. Databrud 1994

**FSE Name: ARBGNR, ARBGNR8, JURNR, CVR\_NR**

---

Henning Bunzel at Department of Economics and Business Economics, Aarhus University, e-mail:  
hbunzel@econ.au.dk

## 2.1 *Beskrivelse*

*CVRNR* er et entydigt identifikationsnummer, som alle juridiske enheder tildeles i *CVR*, *Centrale Virksomheds Register*. Nummeret skal bruges til at identificere virksomheden entydigt over for alle offentlige myndigheder, og det bruges også i meget vid udstrækning til at identificere virksomheden over for andre virksomheder.

En økonomisk enhed kan have flek *CVRNR* eller skifte *CVRNR*. De fleste økonomiske enheder har samme *CVRNR* i hele levetiden. For bestemte virksomhedstyper kan *CVRNR* genbruges.

I ECONAU data er *JURNR* omdøbt til *CVRNR* selvom de ikke er et *CVRNR*. Perioden afhænger af datasættet. I nyere version af data sæt has FSE converteret *JURNR* og *SENR* til *CVRNR*.

I ECONAU findes *CVRNR* i følgende datasæt:

- *AKM* 1985-2002
- *BFL* alle år fra 2008
- *CONESR* 1980-2007
- *LOEN* alle år fra 2002
- *MOMS* 1990-2000
- *RAS* fra 2006
- *ESR\_SE.CVR* alle år fra 1989
- *UHDI*, *UHDM* alle år fra 1993

I ECONAU findes *CVRNR* i følgende datasæt:

- *BFL* alle år fra 2008
- *MOMS* 2001-2009
- *FIKS* alle år fra 2001
- *RAS* fra 2008
- *ESR\_SE.CVR* alle år fra 1989
- *UHDI*, *UHDM*, alle år fra 1993

**Reelt aktive firmaer** I *CVR* der danner grundlag for *ESR*, registreres alle firmaer, der er registreringspligtige i henhold til skatte- og afgiftslovene eller i henhold til selskabslovgivningen, uanset aktivitetsniveauet. Det betyder, at selv den mindste aktivitet, selvom den ikke er egentlig erhvervsmæssig, resulterer i en registrering af et firma. I firmastatistikken medtages kun reelt aktive firmaer. Et firma anses for at være reelt aktivt når det har ansatte, der leverer, hvad der samlet svarer til mindst et halvt fuldtidsårsværk.

## 2.2 *Description*

The *CVRNR* is a unique identifier given to all legal units in *CVR*, *Central Firm Register*. The identifier is used to identify the *firm* uniquely to all public authorities and it is also widely used to identify the *firm* to other *firms*.

The *CVRNR* is unchanged during the entire life cycle of the legal unit. *CVRNR* for closed *firms* may be reused for certain types of *firms*.

In ECONAU data *JURNR* is renamed to *CVRNR* even though they are not a *CVRNR*. The period depends on the dataset. In recent version of datasets FSE has mapped *JURNR* and *SENr* to *CVRNR*.

I ECONAU *SENr* is found in the following datasets:

- *AKM* 1985-2002
- *BFL all years* from 2008
- *CONESR* 1980-2007
- *LOEN all years* from 2002
- *MOMS* 1990-2000
- *RAS all years* from 2006
- *ESR\_SE\_CVR all years* from 1989
- *UHDI, UHDM all years* from 1993

I ECONAU *CVRNR* is found in the following datasets:

- *BFL all years* from 2008
- *LOEN all years* from 1999
- *MOMS* 2001-2009
- *FIKS all years* from 2001
- *RAS all years* from 2008
- *ESR\_SE\_CVR all years* from 1989
- *UHDI, UHDM, all years* from 1993

**Really active firms** In *CVR* which is the basis of *ESR*, all firms are registered which are subjects to registration according to the taxes laws or to the company legislation, regardless of the level of activity. It means that even the smallest activity results in a registration of a firm, even though it is not really professional. The firms statistics only includes really active firms. A firm is regarded as being really active when the employees produce what corresponds to at least half of full-time man-years.

### 3 SENR, A firm account number for tax and custom payments

Virksomhedskontonummer til skat- og toldbetalinger

#### 3.1 Beskrivelse

Et *SENr* er et ottecifret nummer, der blev indført i foråret 1985 af det daværende Statsskattedirektorat i forbindelse med etableringen af *SE, Stamregister for erhvervsdrivende*.

Nummeret blev anvendt som identifikation af juridiske enheder som for eksempel virksomheder og selvstændige erhvervsdrivende i forbindelse med afregning af kildeskat. Senere hen blev det også brugt til afregningen og registreringen af moms, AM-bidrag og A-skat, tilhørsforhold til barselsudligningsordning, periodeoplysninger og evt. produktionsenhedsnummer mv. Den indberetningspligtige har adgang til oplysninger i *eIR*, *Danmarks Statistiks E-Indkomst Register* som denne selv har indberettet, eller som er indberettet på den indberetningspligtiges vegne.

*CVR*, *Centrale Virksomheds Register* blev sat i kraft i oktober 1999. Generelt skal *CVRNR* anvendes til identifikation af virksomheden, idet det alene er *CVRNR*, der udgør virksomhedens entydige identifikation.

Nogle virksomheder har flere *SENR* hos *SKAT*. Der er ingen begrænsninger, og det er alene op til virksomheden, hvad den finder mest praktisk. Til eksempel kan der anvendes et til afregning af moms, og et andet til afregning af A-skat m.v.

I langt de fleste tilfælde har virksomheder kun et *SENR*, som så er identisk med *CVRNR*.

I nogle indberetninger bruger flere *CVR* enheder et fælles *SENR* afregningsnummer, såkaldt fællesafregning.

*ESR* er Danmarks Statistik register for virksomheder. Datasættet *ESR\_SE\_CVR* viser, hvilket *CVRNR* et *SENR* er tilknyttet på et givet tidspunkt.

I mange datasæt er *SENR* benyttet i indberetninger, men i de senere år har *DST* tilføjet *CVRNR* eller erstattet *SENR* med *CVRNR*. Det kan medføre flere observationer med samme nøgle (*AAR*, *CVRNR*).

I *ECONAU* findes *SENR* i følgende datasæt:

- *AKM* 1985-2002
- *BFL* fra 2008, alle år
- *CONESR* 1980-2007, alle år
- *LOEN* fra 2002
- *MOMS* 1990-2000
- *RAS* fra 2006
- *ESR\_SE\_CVR* fra 1989, alle år
- *UHDI*, *UHDM* fra 1993, alle år

I *ECONAU* findes *CVRNR* i følgende datasæt:

- *BFL* fra 2008, alle år
- *LOEN* fra 1999, alle år
- *MOMS* 2001-2009
- *FIKS* fra 2001, alle år
- *RAS* fra 2008
- *ESR\_SE\_CVR* fra 1989,
- *UHDI*, *UHDM*, fra 1993 alle år

### 3.2 Description

A *SENR* is an 8-digit number which was introduced in Spring 1985 by the then State Tax Directorate in connection with the creation of the *SE, Master register for traders*.

The number was used as identification of legal units as for instance firms and self-employed in connection with withholding of tax. Later on it was also used for registration and payment of VAT, Labour market contributions and income taxes, relation to contributions to maternity leave payments, periodical information and maybe production unit number etc. The reporting unit has access to information in *eIR, Danmarks Statistiks E-Indkomst Register* which has been reported by itself or on behalf of the reporting unit.

*CVR*, *Central Firm Register* was initialized in October 1999. Generally *CVRNR* is an identifier of a legal unit. *CVRNR* is a unique identifier of an economic unit, named firm or enterprise. Later, *OKNR, Identifier for Economic Units* was introduced as an identifier of the economic unit. An economic unit may have several *CVRNR*. In many datasets one of the *CVRNR* is selected as identifying the economic unit. Often it is the firm that decides to have several *CVRNR*, and several *SENR* associated each *CVRNR*. At any time one *SENR* can only be associated one *CVRNR* and some firms have more *SENR* at *SKAT*. The registration of *SENR* can be done for different purposes, for example one number can be used for withholding of VAT, and another for withholding of A-taxes mv.

In the vast majority of cases firms have only one *SENR* which then is identical to *CVRNR*.

In some data several *CVRNR* units use a common *SENR*, a so-called common settlement.

*ESR* is the Statistics Denmark register for firms. *ESR\_SE\_CVR* shows which *CVRNR* a *SENR* is associated at a given time.

In many datasets *SENR* is used in reports but in recent years *DST* has added *CVRNR* or replaced *SENR* with *CVRNR*. It may cause more observations with the same key (*AAR, CVRNR*.)

I ECONAU *SENR* is found in the following datasets:

- *AKM* 1985-2002
- *BFL* from 2008, **all years**
- *CONESR* 1980-2007, **all years**
- *LOEN* from 2002
- *MOMS* 1990-2000
- *RAS* from 2006
- *ESR\_SE\_CVR* from 1989, **all years**
- *UHDI, UHDM* from 1993 **all years**

I ECONAU *CVRNR* is found in the following datasets:

- *BFL* from 2008, **all years**
- *LOEN* from 1999, **all years**

- *MOMS* 2001-2009
- *FIKS* from 2001, **all years**
- *RAS* from 2008
- *ESR\_SE\_CVR* from 1989, **all years**
- *UHDI, UHDM*, from 1993 **all years**

#### **4 SLUT\_AAR, End year of a SENR**

Slut år for et SENR

#### **5 SLUT\_DATO, End date for a spell**

Slutdato for et forløb

**FSE Name:** ARB\_TIL\_DATO

**Period:** Period is not defined

#### **6 START\_AAR, Startyear of a SENR**

Start år for et SENR

#### **7 START\_DATO, Start date of a spell**

Startdato for et forløb

**Period:** Period is not yet defined