

Siamak Baradaran

sia@kth.se

***Bilaga till rapporten "Tillvägagångssätt för skattning av körkortsmo-
dell"***

*I detta projekt använder vi ett mycket omfattande datasätt vilket gör att det är många saker som behöver justeras. I scriptet nedan i STA-miljö, dels tar vi fram de data som behövs, och dels rensar vi för eventuella problem. De rader som börjar med * är så kallade comment och innehåller förklaringar till vad som görs.*

***** 2003 *****

* vehicle data preparation

***** change here 1 ****

*odbc load, exec("select * from dbo.vy_bildata2003") dsn("P0374_VTI_KTH_Bilnehav") clear*

rename lopnr LopnrN

merge m:1 LopnrN using "\\Mfso01\p0374_barsia\$\ownership\Data\Basic data\key.dta"

rename LopnrG lopnr

keep lopnr chassi_id

rename lopnr id

tostring chassi_id, replace

destring chassi_id, replace

** dropping duplicates(id)*

```

duplicates tag chassi_id, generate(dups)
drop if (dups > 0)
drop dups
* generate a variable to later find out which veh (1 or 2) has the highest chassi_id, need this for
sort id chassi_id
generate delta = id - id[_n-1]
generate d1 = 1 if (delta == 0)
generate order = 1 if (d1 != 1)
replace order = (order[_n-1]+1) if (d1 == 1 & order != 1)
* change number "x" below on the line "order > x" to the number of vehicles you want to include
ps. this changes the dimension of matrix
drop if (order > 5)
generate int order2=1
replace order2=(order2[_n-1]+1) if (id==id[_n-1])
drop delta d1 order
* reshaping to matrix form/wide form
reshape wide chassi_id , i(id) j(order2)
generate int veh_nr=0
replace veh_nr=1 if
((chassi_id1!=.) & (chassi_id2==.) & (chassi_id3==.) & (chassi_id4==.) & (chassi_id5==.))
replace veh_nr=2 if
((chassi_id1!=.) & (chassi_id2!=.) & (chassi_id3==.) & (chassi_id4==.) & (chassi_id5==.))
replace veh_nr=3 if
((chassi_id1!=.) & (chassi_id2!=.) & (chassi_id3!=.) & (chassi_id4==.) & (chassi_id5==.))
replace veh_nr=4 if
((chassi_id1!=.) & (chassi_id2!=.) & (chassi_id3!=.) & (chassi_id4!=.) & (chassi_id5==.))
replace veh_nr=5 if
((chassi_id1!=.) & (chassi_id2!=.) & (chassi_id3!=.) & (chassi_id4!=.) & (chassi_id5!=.))
keep id veh_nr
save "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta", replace

```

** opening barnover18 database and identify parents*

*odbc load, exec("select * from dbo.VTI2003_barnover18") dsn("P0374_VTI_KTH_Bilinnhav")
clear*

keep lopnr lopnr barn

rename lopnr id_parent

rename lopnr barn id_child

sort id_child id_parent

generate delta = id_child - id_child[_n-1]

generate d1 = 1 if (delta == 0)

generate order = 1 if (d1 != 1)

replace order = (order[_n-1]+1) if (d1 == 1 & order != 1)

generate int order2=1

replace order2=(order2[_n-1]+1) if (id_child==id_child[_n-1])

drop delta d1 order

reshape wide id_parent , i(id_child) j(order2)

generate parents_nr=0

replace parents_nr=1 if ((id_parent1!=.)&(id_parent2==.))

replace parents_nr=2 if ((id_parent1!=.)&(id_parent2!=.))

gen int p1_veh=.

gen int p1_inc=.

gen str p1_SAMS=""

gen int p2_veh=.

gen int p2_inc=.

gen str p2_SAMS=""

save "\\Mfso01\p0374_barsia\$\DL_model\data\parents_mtx.dta", replace

**merge to vehicle file to get number of cars available from each parent*

use "\\Mfso01\p0374_barsia\$\DL_model\data\parents_mtx.dta", clear

keep id_parent1

rename id_parent1 id

```

merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename id id_parent1
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta", replace
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent2
drop if(id_parent2==.)
rename id_parent2 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename id id_parent2
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta", replace
* put number of vehicles back in parents_matrix
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
merge m:m id_parent1 using "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta"
replace p1_veh=veh_nr if(_merge==3)
drop _merge veh_nr
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
merge m:m id_parent2 using "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta"
replace p2_veh=veh_nr if(_merge==3)
drop _merge veh_nr
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
* next fylling parents incomes
* individual data preparation
* reading from db and fixing the id's and deciding household head
*****
***** change here 2 ****

```

*odbc load, exec("select * from dbo.VTI2003 ") dsn("P0374_VTI_KTH_Bilnehav") clear*

duplicates drop Lopnr, force

*keep Lopnr Aterpnr Fodar Kon Antalbarn Bosams Sysstat Stud Arbsams Dispink Korkort
lopnrmak*

rename Lopnr id

rename Aterpnr recycled_id

rename Fodar birth_year

rename Kon sex

rename Antalbarn children_nr

rename Bosams home_sams

rename Sysstat emp_status

rename Stud student

rename Arbsams work_sams

rename Dispink income

rename Korkort drv_lic

rename lopnrmak id_partner

** reformating*

destring birth_year, replace

**destring home_sams, replace*

destring emp_stat, replace

**destring work_sams, replace*

destring drv_lic, replace

tostring children_nr, replace

destring children_nr, replace

***** change here 3 ****

save "\\Mfso01\p0374_barsia\$\DL_model\data\ind_03.dta", replace

```
*****
*****
```

```
*merging number of vehicles
```

```
merge 1:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
```

```
rename veh_nr own_veh
```

```
drop _merge
```

```
rename id own_id
```

```
rename id_partner id
```

```
merge m:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
```

```
rename veh_nr partner_veh
```

```
drop _merge
```

```
rename id id_partner
```

```
rename own_id id
```

```
*****
***** change here 3 *****
```

```
save "\\Mfso01\p0374_barsia$\DL_model\data\ind_03.dta", replace
```

```
*****
*****
```

```
keep id home_sams income
```

```
rename home_sams p_SAMS
```

```
rename income p_inc
```

```
save "\\Mfso01\p0374_barsia$\DL_model\data\p_data.dta", replace
```

```
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
```

```
keep id_parent1
```

```
rename id_parent1 id
```

```
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\p_data.dta"
```

```
rename id id_parent1
```

```
drop if(_merge!=3)
```

```

drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta", replace
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent2
drop if(id_parent2==.)
rename id_parent2 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\p_data.dta"
rename id id_parent2
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta", replace
* put income and SAMS-info back in parents_matrix
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
rename id_child id
merge m:m id_parent1 using "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta"
tostring p1_SAMS, replace
replace p1_SAMS=p_SAMS if(_merge==3)
replace p1_inc=p_inc if(_merge==3)
drop _merge p_SAMS p_inc
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
merge m:m id_parent2 using "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta"
tostring p2_SAMS, replace
replace p2_SAMS=p_SAMS if(_merge==3)
replace p2_inc=p_inc if(_merge==3)
drop _merge p_SAMS p_inc
*rename id_child id
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
*merge to all individuals

```

```

*****
***** change here 3 *****

use "\\Mfso01\p0374_barsia$\DL_model\data\ind_03.dta", clear

*****
*****

duplicates tag id, gen(dups)

drop if(dups!=0)

merge 1:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta"

drop _merge dups

*****
***** change here 3 *****

save "\\Mfso01\p0374_barsia$\DL_model\data\ind_03.dta", replace

*****
*****

***** 2004 *****

* vehicle data preparation

*****
*****

*****
***** change here 1 *****

odbc load, exec("select * from dbo.vy_bildata2004") dsn("P0374_VTI_KTH_Bilinnerhav") clear

*****
*****

rename lopnr LopnrN

merge m:1 LopnrN using "\\Mfso01\p0374_barsia$\ownership\Data\Basic data\key.dta"

rename LopnrG lopnr

keep lopnr chassi_id

rename lopnr id

```



```

tostring chassi_id, replace
destring chassi_id, replace
* dropping duplicates(id)
duplicates tag chassi_id, generate(dups)
drop if (dups > 0)
drop dups
* generate a variable to later find out which veh (1 or 2) has the highest chassi_id, need this for
sort id chassi_id
generate delta = id - id[_n-1]
generate d1 = 1 if (delta == 0)
generate order = 1 if (d1 != 1)
replace order = (order[_n-1]+1) if (d1 == 1 & order != 1)
* change number "x" below on the line "order > x" to the number of vehicles you want to include
ps. this changes the dimension of matrix
drop if (order > 5)
generate int order2=1
replace order2=(order2[_n-1]+1) if (id==id[_n-1])
drop delta d1 order
* reshaping to matrix form/wide form
reshape wide chassi_id , i(id) j(order2)
generate int veh_nr=0
replace veh_nr=1 if
((chassi_id1!=.) & (chassi_id2==.) & (chassi_id3==.) & (chassi_id4==.) & (chassi_id5==.))
replace veh_nr=2 if
((chassi_id1!=.) & (chassi_id2!=.) & (chassi_id3==.) & (chassi_id4==.) & (chassi_id5==.))
replace veh_nr=3 if
((chassi_id1!=.) & (chassi_id2!=.) & (chassi_id3!=.) & (chassi_id4==.) & (chassi_id5==.))
replace veh_nr=4 if
((chassi_id1!=.) & (chassi_id2!=.) & (chassi_id3!=.) & (chassi_id4!=.) & (chassi_id5==.))

```

```

replace veh_nr=5 if
((chassi_id1!=.)&(chassi_id2!=.)&(chassi_id3!=.)&(chassi_id4!=.)&(chassi_id5!=.))
keep id veh_nr
save "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta", replace
* opening barnover18 database and identify parnets
odbc load, exec("select * from dbo.VTI2004_barnover18") dsn("P0374_VTI_KTH_Bilinnhav")
clear
keep lopnr lopnr_barn
rename lopnr id_parent
rename lopnr_barn id_child
sort id_child id_parent
generate delta = id_child - id_child[_n-1]
generate d1 = 1 if (delta == 0)
generate order = 1 if (d1 != 1)
replace order = (order[_n-1]+1) if (d1 == 1 & order != 1)
generate int order2=1
replace order2=(order2[_n-1]+1) if (id_child==id_child[_n-1])
drop delta d1 order
reshape wide id_parent , i(id_child) j(order2)
generate parents_nr=0
replace parents_nr=1 if ((id_parent1!=.)&(id_parent2==.))
replace parents_nr=2 if ((id_parent1!=.)&(id_parent2!=.))
gen int p1_veh=.
gen int p1_inc=.
gen str p1_SAMS=""
gen int p2_veh=.
gen int p2_inc=.
gen str p2_SAMS=""
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace

```

```

*merge to vehicle file to get number of cars available from each parent
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent1
rename id_parent1 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename id id_parent1
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta", replace
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent2
drop if(id_parent2==.)
rename id_parent2 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename id id_parent2
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta", replace
* put number of vehicles back in parents_matrix
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
merge m:m id_parent1 using "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta"
replace p1_veh=veh_nr if(_merge==3)
drop _merge veh_nr
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
merge m:m id_parent2 using "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta"
replace p2_veh=veh_nr if(_merge==3)
drop _merge veh_nr
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace

```

** next fylling parents incomes*

** individual data preparation*

** reading from db and fixing the id's and deciding household head*

***** *change here 2* *****

*odbc load, exec("select * from dbo.VTI2004 ") dsn("P0374_VTI_KTH_Bilnehav") clear*

duplicates drop Lopnr, force

*keep Lopnr Aterpnr Fodar Kon Antalbarn Bosams Sysstat Stud Arbsams Dispink Korkort
lopnrmak*

rename Lopnr id

rename Aterpnr recycled_id

rename Fodar birth_year

rename Kon sex

rename Antalbarn children_nr

rename Bosams home_sams

rename Sysstat emp_status

rename Stud student

rename Arbsams work_sams

rename Dispink income

rename Korkort drv_lic

rename lopnrmak id_partner

** reformating*

destring birth_year, replace

**destring home_sams, replace*

destring emp_stat, replace

**destring work_sams, replace*

destring drv_lic, replace

tostring children_nr, replace

destring children_nr, replace

***** *change here 3* *****

save "\\Mfso01\p0374_barsia\$\DL_model\data\ind_04.dta", replace

**merging number of vehicles*

merge 1:1 id using "\\Mfso01\p0374_barsia\$\DL_model\data\veh_mtx.dta"

rename veh_nr own_veh

drop _merge

rename id own_id

rename id_partner id

merge m:1 id using "\\Mfso01\p0374_barsia\$\DL_model\data\veh_mtx.dta"

rename veh_nr partner_veh

drop _merge

rename id id_partner

rename own_id id

***** *change here 3* *****

save "\\Mfso01\p0374_barsia\$\DL_model\data\ind_04.dta", replace

keep id home_sams income

rename home_sams p_SAMS

rename income p_inc

save "\\Mfso01\p0374_barsia\$\DL_model\data\p_data.dta", replace

use "\\Mfso01\p0374_barsia\$\DL_model\data\parents_mtx.dta", clear

keep id_parent1

```
rename id_parent1 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\p_data.dta"
rename id id_parent1
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta", replace
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent2
drop if(id_parent2==.)
rename id_parent2 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\p_data.dta"
rename id id_parent2
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta", replace
* put income and SAMS-info back in parents_matrix
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
rename id_child id
merge m:m id_parent1 using "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta"
tostring p1_SAMS, replace
replace p1_SAMS=p_SAMS if(_merge==3)
replace p1_inc=p_inc if(_merge==3)
drop _merge p_SAMS p_inc
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
merge m:m id_parent2 using "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta"
tostring p2_SAMS, replace
replace p2_SAMS=p_SAMS if(_merge==3)
replace p2_inc=p_inc if(_merge==3)
```

```

drop _merge p_SAMS p_inc

*rename id_child id

save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace

*merge to all individuals

*****
***** change here 3 ****

use "\\Mfso01\p0374_barsia$\DL_model\data\ind_04.dta", clear

*****
*****

duplicates tag id, gen(dups)

drop if(dups!=0)

merge 1:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta"

drop _merge dups

*****
***** change here 3 ****

save "\\Mfso01\p0374_barsia$\DL_model\data\ind_04.dta", replace

*****
*****

***** 2005 *****

* vehicle data preparation

*****
*****

*****
***** change here 1 ****

odbc load, exec("select * from dbo.vy_bildata2005") dsn("P0374_VTI_KTH_Bilinnehav") clear

*****
*****

rename lopnr LopnrN

merge m:1 LopnrN using "\\Mfso01\p0374_barsia$\ownership\Data\Basic data\key.dta"

```

```

rename LopnrG lopnr
keep lopnr chassi_id
rename lopnr id
tostring chassi_id, replace
destring chassi_id, replace
* dropping duplicates(id)
duplicates tag chassi_id, generate(dups)
drop if (dups > 0)
drop dups
* generate a variable to later find out which veh (1 or 2) has the highest chassi_id, need this for
sort id chassi_id
generate delta = id - id[_n-1]
generate d1 = 1 if (delta == 0)
generate order = 1 if (d1 != 1)
replace order = (order[_n-1]+1) if (d1 == 1 & order != 1)
* change number "x" below on the line "order > x" to the number of vehicles you want to include
ps. this changes the dimension of matrix
drop if (order > 5)
generate int order2=1
replace order2=(order2[_n-1]+1) if (id==id[_n-1])
drop delta d1 order
* reshaping to matrix form/wide form
reshape wide chassi_id , i(id) j(order2)
generate int veh_nr=0
replace veh_nr=1 if
((chassi_id1!=.) & (chassi_id2==.) & (chassi_id3==.) & (chassi_id4==.) & (chassi_id5==.))
replace veh_nr=2 if
((chassi_id1!=.) & (chassi_id2!=.) & (chassi_id3==.) & (chassi_id4==.) & (chassi_id5==.))
replace veh_nr=3 if
((chassi_id1!=.) & (chassi_id2!=.) & (chassi_id3!=.) & (chassi_id4==.) & (chassi_id5==.))

```



```

replace veh_nr=4 if
((chassi_id1!=.)&(chassi_id2!=.)&(chassi_id3!=.)&(chassi_id4!=.)&(chassi_id5==.))
replace veh_nr=5 if
((chassi_id1!=.)&(chassi_id2!=.)&(chassi_id3!=.)&(chassi_id4!=.)&(chassi_id5!=.))
keep id veh_nr
save "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta", replace
* opening barnover18 database and identify parnets
odbc load, exec("select * from dbo.VTI2005_barnover18") dsn("P0374_VTI_KTH_Bilinnehav")
clear
keep lopnr lopnrbarn
rename lopnr id_parent
rename lopnrbarn id_child
sort id_child id_parent
generate delta = id_child - id_child[_n-1]
generate d1 = 1 if (delta == 0)
generate order = 1 if (d1 !=1)
replace order = (order[_n-1]+1) if (d1 == 1 & order != 1)
generate int order2=1
replace order2=(order2[_n-1]+1) if (id_child==id_child[_n-1])
drop delta d1 order
reshape wide id_parent , i(id_child) j(order2)
generate parents_nr=0
replace parents_nr=1 if ((id_parent1!=.)&(id_parent2==.))
replace parents_nr=2 if ((id_parent1!=.)&(id_parent2!=.))
gen int p1_veh=.
gen int p1_inc=.
gen str p1_SAMS=""
gen int p2_veh=.
gen int p2_inc=.

```

```
gen str p2_SAMS=""
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
*merge to vehicle file to get number of cars available from each parent
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent1
rename id_parent1 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename id id_parent1
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta", replace
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent2
drop if(id_parent2==.)
rename id_parent2 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename id id_parent2
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta", replace
* put number of vehicles back in parents_matrix
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
merge m:m id_parent1 using "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta"
replace p1_veh=veh_nr if(_merge==3)
drop _merge veh_nr
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
merge m:m id_parent2 using "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta"
replace p2_veh=veh_nr if(_merge==3)
```

```
drop _merge veh_nr
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
* next fylling parents incomes
* individual data preparation
* reading from db and fixing the id's and deciding household head
*****
***** change here 2 ****
odbc load, exec("select * from dbo.VTI2005 ") dsn("P0374_VTI_KTH_Bilnehav") clear
*****
*****
duplicates drop Lopnr, force
keep Lopnr Aterpnr Fodar Kon Antalbarn Bosams Sysstat Stud Arbsams Dispink Korkort
lopnrnak
rename Lopnr id
rename Aterpnr recycled_id
rename Fodar birth_year
rename Kon sex
rename Antalbarn children_nr
rename Bosams home_sams
rename Sysstat emp_status
rename Stud student
rename Arbsams work_sams
rename Dispink income
rename Korkort drv_lic
rename lopnrnak id_partner
* reformatting
destring birth_year, replace
*destring home_sams, replace
destring emp_stat, replace
```

```
*destring work_sams, replace
destring drv_lic, replace
tostring children_nr, replace
destring children_nr, replace
*****
***** change here 3 ****
save "\\Mfso01\p0374_barsia$\DL_model\data\ind_05.dta", replace
*****
*****
*merging number of vehicles
merge 1:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename veh_nr own_veh
drop _merge
rename id own_id
rename id_partner id
merge m:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename veh_nr partner_veh
drop _merge
rename id id_partner
rename own_id id
*****
***** change here 3 ****
save "\\Mfso01\p0374_barsia$\DL_model\data\ind_05.dta", replace
*****
*****

keep id home_sams income
rename home_sams p_SAMS
rename income p_inc
save "\\Mfso01\p0374_barsia$\DL_model\data\p_data.dta", replace
```

```
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent1
rename id_parent1 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\p_data.dta"
rename id id_parent1
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta", replace
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent2
drop if(id_parent2==.)
rename id_parent2 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\p_data.dta"
rename id id_parent2
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta", replace
* put income and SAMS-info back in parents_matrix
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
rename id_child id
merge m:m id_parent1 using "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta"
tostring p1_SAMS, replace
replace p1_SAMS=p_SAMS if(_merge==3)
replace p1_inc=p_inc if(_merge==3)
drop _merge p_SAMS p_inc
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
merge m:m id_parent2 using "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta"
tostring p2_SAMS, replace
```

```

replace p2_SAMS=p_SAMS if(_merge==3)
replace p2_inc=p_inc if(_merge==3)
drop _merge p_SAMS p_inc
*rename id_child id
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
*merge to all individuals
*****
***** change here 3 ****
use "\\Mfso01\p0374_barsia$\DL_model\data\ind_05.dta", clear
*****
*****
duplicates tag id, gen(dups)
drop if(dups!=0)
merge 1:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta"
drop _merge dups
*****
***** change here 3 ****
save "\\Mfso01\p0374_barsia$\DL_model\data\ind_05.dta", replace
*****
*****

***** 2006 *****
* vehicle data preparation
*****
*****
*****
***** change here 1 ****
odbc load, exec("select * from dbo.vy_bildata2006") dsn("P0374_VTI_KTH_Bilnehav") clear
*****
*****

```

```

rename lopnr LopnrN
merge m:1 LopnrN using "\\Mfso01\p0374_barsia$\ownership\Data\Basic data\key.dta"
rename LopnrG lopnr
keep lopnr chassi_id
rename lopnr id
tostring chassi_id, replace
destring chassi_id, replace
* dropping duplicates(id)
duplicates tag chassi_id, generate(dups)
drop if (dups > 0)
drop dups
* generate a variable to later find out which veh (1 or 2) has the highest chassi_id, need this for
sort id chassi_id
generate delta = id - id[_n-1]
generate d1 = 1 if (delta == 0)
generate order = 1 if (d1 != 1)
replace order = (order[_n-1]+1) if (d1 == 1 & order != 1)
* change number "x" below on the line "order > x" to the number of vehicles you want to include
ps. this changes the dimension of matrix
drop if (order > 5)
generate int order2=1
replace order2=(order2[_n-1]+1) if (id==id[_n-1])
drop delta d1 order
* reshaping to matrix form/wide form
reshape wide chassi_id , i(id) j(order2)
generate int veh_nr=0
replace veh_nr=1 if
((chassi_id1!=.)&(chassi_id2==.)&(chassi_id3==.)&(chassi_id4==.)&(chassi_id5==.))

```

```

replace veh_nr=2 if
((chassi_id1!=.)&(chassi_id2!=.)&(chassi_id3==.)&(chassi_id4==.)&(chassi_id5==.))

replace veh_nr=3 if
((chassi_id1!=.)&(chassi_id2!=.)&(chassi_id3!=.)&(chassi_id4==.)&(chassi_id5==.))

replace veh_nr=4 if
((chassi_id1!=.)&(chassi_id2!=.)&(chassi_id3!=.)&(chassi_id4!=.)&(chassi_id5==.))

replace veh_nr=5 if
((chassi_id1!=.)&(chassi_id2!=.)&(chassi_id3!=.)&(chassi_id4!=.)&(chassi_id5!=.))

keep id veh_nr

save "\\Mfso01\p0374_barsia$DL_model\data\veh_mtx.dta", replace

* opening barnover18 database and identify parnets

odbc load, exec("select * from dbo.VTI2006_barnover18") dsn("P0374_VTI_KTH_Bilinnehav")
clear

keep lopnr lopnrbarn

rename lopnr id_parent

rename lopnrbarn id_child

sort id_child id_parent

generate delta = id_child - id_child[_n-1]

generate d1 = 1 if (delta == 0)

generate order = 1 if (d1 !=1)

replace order = (order[_n-1]+1) if (d1 == 1 & order != 1)

generate int order2=1

replace order2=(order2[_n-1]+1) if (id_child==id_child[_n-1])

drop delta d1 order

reshape wide id_parent , i(id_child) j(order2)

generate parents_nr=0

replace parents_nr=1 if ((id_parent1!=.)&(id_parent2==.))

replace parents_nr=2 if ((id_parent1!=.)&(id_parent2!=.))

gen int p1_veh=.

gen int p1_inc=.

```



```
gen str p1_SAMS=""
gen int p2_veh=.
gen int p2_inc=.
gen str p2_SAMS=""
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
*merge to vehicle file to get number of cars available from each parent
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent1
rename id_parent1 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename id id_parent1
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta", replace
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent2
drop if(id_parent2==.)
rename id_parent2 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename id id_parent2
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta", replace
* put number of vehicles back in parents_matrix
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
merge m:m id_parent1 using "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta"
replace p1_veh=veh_nr if(_merge==3)
drop _merge veh_nr
```

```
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
merge m:m id_parent2 using "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta"
replace p2_veh=veh_nr if(_merge==3)
drop _merge veh_nr
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
* next fylling parents incomes
* individual data preparation
* reading from db and fixing the id's and deciding household head
*****
***** change here 2 *****

odbc load, exec("select * from dbo.VTI_2006MS ") dsn("P0374_VTI_KTH_Bilinnhav") clear
*****
*****

rename LopnrMak mak
duplicates drop lopnr, force
keep lopnr Fodar Kon AntalBarn BoSAMS sysstat stud ArbSAMS Dispink korkort mak
rename lopnr id
rename Fodar birth_year
rename Kon sex
rename AntalBarn children_nr
rename BoSAMS home_sams
rename sysstat emp_status
rename stud student
rename ArbSAMS work_sams
rename Dispink income
rename korkort drv_lic
rename mak id_partner
* reformating
```

```
destring birth_year, replace
*destring home_sams, replace
destring emp_stat, replace
*destring work_sams, replace
destring drv_lic, replace
tostring children_nr, replace
destring children_nr, replace
*****
***** change here 3 ****
save "\\Mfso01\p0374_barsia$\DL_model\data\ind_06.dta", replace
*****
*****
*merging number of vehicles
merge 1:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename veh_nr own_veh
drop _merge
rename id own_id
rename id_partner id
merge m:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename veh_nr partner_veh
drop _merge
rename id id_partner
rename own_id id
*****
***** change here 3 ****
save "\\Mfso01\p0374_barsia$\DL_model\data\ind_06.dta", replace
*****
*****
keep id home_sams income
```

```

rename home_sams p_SAMS
rename income p_inc
save "\\Mfso01\p0374_barsia$\DL_model\data\p_data.dta", replace
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent1
rename id_parent1 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\p_data.dta"
rename id id_parent1
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta", replace
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent2
drop if(id_parent2==.)
rename id_parent2 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\p_data.dta"
rename id id_parent2
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta", replace
* put income and SAMS-info back in parents_matrix
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
rename id_child id
merge m:m id_parent1 using "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta"
tostring p1_SAMS, replace
replace p1_SAMS=p_SAMS if(_merge==3)

* ta bort detta sedan

```

```

destring p_inc, replace
*****

replace p1_inc=p_inc if(_merge==3)

drop _merge p_SAMS p_inc

save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace

merge m:m id_parent2 using "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta"

tostring p2_SAMS, replace

replace p2_SAMS=p_SAMS if(_merge==3)

* ta bort detta sedan

destring p_inc, replace
*****

replace p2_inc=p_inc if(_merge==3)

drop _merge p_SAMS p_inc

*rename id_child id

save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace

*merege to all individuals

*****
***** change here 3 ****

use "\\Mfso01\p0374_barsia$\DL_model\data\ind_06.dta", clear

*****
*****

duplicates tag id, gen(dups)

drop if(dups!=0)

merge 1:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta"

drop _merge dups

*****
***** change here 3 ****

save "\\Mfso01\p0374_barsia$\DL_model\data\ind_06.dta", replace

```


***** 2007 *****

** vehicle data preparation*

***** *change here 1* *****

*odbc load, exec("select * from dbo.vy_bildata2007") dsn("P0374_VTI_KTH_Bilinnehav") clear*

rename lopnr LopnrN

merge m:1 LopnrN using "\\Mfso01\p0374_barsia\$\ownership\Data\Basic data\key.dta"

rename LopnrG lopnr

keep lopnr chassi_id

rename lopnr id

tostring chassi_id, replace

destring chassi_id, replace

** dropping duplicates(id)*

duplicates tag chassi_id, generate(dups)

drop if (dups > 0)

drop dups

** generate a variable to later find out which veh (1 or 2) has the highest chassi_id, need this for
sort id chassi_id*

generate delta = id - id[_n-1]

generate d1 = 1 if (delta == 0)

generate order = 1 if (d1 != 1)

replace order = (order[_n-1]+1) if (d1 == 1 & order != 1)

** change number "x" below on the line "order > x" to the number of vehicles you want to include
ps. this changes the dimension of matrix*

drop if (order > 5)

generate int order2=1

replace order2=(order2[_n-1]+1) if (id==id[_n-1])

drop delta d1 order

** reshaping to matrix form/wide form*

reshape wide chassi_id , i(id) j(order2)

generate int veh_nr=0

replace veh_nr=1 if

((chassi_id1!=.) & (chassi_id2==.) & (chassi_id3==.) & (chassi_id4==.) & (chassi_id5==.))

replace veh_nr=2 if

((chassi_id1!=.) & (chassi_id2!=.) & (chassi_id3==.) & (chassi_id4==.) & (chassi_id5==.))

replace veh_nr=3 if

((chassi_id1!=.) & (chassi_id2!=.) & (chassi_id3!=.) & (chassi_id4==.) & (chassi_id5==.))

replace veh_nr=4 if

((chassi_id1!=.) & (chassi_id2!=.) & (chassi_id3!=.) & (chassi_id4!=.) & (chassi_id5==.))

replace veh_nr=5 if

((chassi_id1!=.) & (chassi_id2!=.) & (chassi_id3!=.) & (chassi_id4!=.) & (chassi_id5!=.))

keep id veh_nr

save "\\Mfso01\p0374_barsia\$\DL_model\data\veh_mtx.dta", replace

** opening barnover18 database and identify parnets*

*odbc load, exec("select * from dbo.VTI2007_barnover18") dsn("P0374_VTI_KTH_Bilinnerhav")*

clear

keep lopnr lopnr_barn

rename lopnr id_parent

rename lopnr_barn id_child

sort id_child id_parent

generate delta = id_child - id_child[_n-1]

generate d1 = 1 if (delta == 0)

generate order = 1 if (d1 != 1)

```

replace order = (order[_n-1]+1) if (d1 == 1 & order != 1)
generate int order2=1
replace order2=(order2[_n-1]+1) if (id_child==id_child[_n-1])
drop delta d1 order
reshape wide id_parent , i(id_child) j(order2)
generate parents_nr=0
replace parents_nr=1 if ((id_parent1!=.)&(id_parent2==.))
replace parents_nr=2 if ((id_parent1!=.)&(id_parent2!=.))
gen int p1_veh=.
gen int p1_inc=.
gen str p1_SAMS=""
gen int p2_veh=.
gen int p2_inc=.
gen str p2_SAMS=""
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
*merge to vehicle file to get number of cars available from each parent
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent1
rename id_parent1 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename id id_parent1
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta", replace
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent2
drop if(id_parent2==.)
rename id_parent2 id

```



```

merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename id id_parent2
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta", replace
* put number of vehicles back in parents_matrix
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
merge m:m id_parent1 using "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta"
replace p1_veh=veh_nr if(_merge==3)
drop _merge veh_nr
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
merge m:m id_parent2 using "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta"
replace p2_veh=veh_nr if(_merge==3)
drop _merge veh_nr
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
* next fylling parents incomes
* individual data preparation
* reading from db and fixing the id's and deciding household head
*****
***** change here 2 *****
odbc load, exec("select * from dbo.VTI2007 ") dsn("P0374_VTI_KTH_Bilnehav") clear
*****
*****
duplicates drop Lopnr, force
keep Lopnr Aterpnr Fodar Kon Antalborn Bosams Sysstat Stud Arbsams Dispink Korkort
lopnrmak
rename Lopnr id
rename Aterpnr recycled_id
rename Fodar birth_year

```

```
rename Kon sex
rename Antalbarn children_nr
rename Bosams home_sams
rename Sysstat emp_status
rename Stud student
rename Arbsams work_sams
rename Dispink income
rename Korkort drv_lic
rename lopnrmak id_partner
* reformating
destring birth_year, replace
*destring home_sams, replace
destring emp_stat, replace
*destring work_sams, replace
destring drv_lic, replace
tostring children_nr, replace
destring children_nr, replace
*****
***** change here 3 ***
save "\\Mfso01\p0374_barsia$\DL_model\data\ind_07.dta", replace
*****
*****
*merging number of vehicles
merge 1:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename veh_nr own_veh
drop _merge
rename id own_id
rename id_partner id
merge m:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
```

```

rename veh_nr partner_veh
drop _merge
rename id id_partner
rename own_id id
*****
***** change here 3 ****
save "\\Mfso01\p0374_barsia$\DL_model\data\ind_07.dta", replace
*****
*****

keep id home_sams income
rename home_sams p_SAMS
rename income p_inc
save "\\Mfso01\p0374_barsia$\DL_model\data\p_data.dta", replace
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent1
rename id_parent1 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\p_data.dta"
rename id id_parent1
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta", replace
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent2
drop if(id_parent2==.)
rename id_parent2 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\p_data.dta"
rename id id_parent2
drop if(_merge!=3)

```

```

drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta", replace
* put income and SAMS-info back in parents_matrix
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
rename id_child id
merge m:m id_parent1 using "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta"
tostring p1_SAMS, replace
replace p1_SAMS=p_SAMS if(_merge==3)
replace p1_inc=p_inc if(_merge==3)
drop _merge p_SAMS p_inc
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
merge m:m id_parent2 using "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta"
tostring p2_SAMS, replace
replace p2_SAMS=p_SAMS if(_merge==3)
replace p2_inc=p_inc if(_merge==3)
drop _merge p_SAMS p_inc
*rename id_child id
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
*merge to all individuals
*****
***** change here 3 ****
use "\\Mfso01\p0374_barsia$\DL_model\data\ind_07.dta", clear
*****
*****
duplicates tag id, gen(dups)
drop if(dups!=0)
merge 1:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta"
drop _merge dups

```

```
*****  
***** change here 3 *****
```

```
save "\\Mfso01\p0374_barsia$\DL_model\data\ind_07.dta", replace
```

```
*****  
*****
```

```
***** 2008 *****
```

```
* vehicle data preparation
```

```
*****  
*****
```

```
*****  
***** change here 1 *****
```

```
odbc load, exec("select * from dbo.vy_bildata2008") dsn("P0374_VTI_KTH_Bilinnehav") clear
```

```
*****  
*****
```

```
rename lopnr LopnrN
```

```
merge m:1 LopnrN using "\\Mfso01\p0374_barsia$\ownership\Data\Basic data\key.dta"
```

```
rename LopnrG lopnr
```

```
keep lopnr chassi_id
```

```
rename lopnr id
```

```
tostring chassi_id, replace
```

```
destring chassi_id, replace
```

```
* dropping duplicates(id)
```

```
duplicates tag chassi_id, generate(dups)
```

```
drop if (dups > 0)
```

```
drop dups
```

```
* generate a variable to later find out which veh (1 or 2) has the highest chassi_id, need this for
```

```
sort id chassi_id
```

```
generate delta = id - id[_n-1]
```

```

generate d1 = 1 if (delta == 0)
generate order = 1 if (d1 !=1)
replace order = (order[_n-1]+1) if (d1 == 1 & order != 1)
* change number "x" below on the line "order > x" to the number of vehicles you want to include
ps. this changes the dimension of matrix
drop if (order > 5)
generate int order2=1
replace order2=(order2[_n-1]+1) if (id==id[_n-1])
drop delta d1 order
* reshaping to matrix form/wide form
reshape wide chassi_id , i(id) j(order2)
generate int veh_nr=0
replace veh_nr=1 if
((chassi_id1!=.)&(chassi_id2==.)&(chassi_id3==.)&(chassi_id4==.)&(chassi_id5==.))
replace veh_nr=2 if
((chassi_id1!=.)&(chassi_id2!=.)&(chassi_id3==.)&(chassi_id4==.)&(chassi_id5==.))
replace veh_nr=3 if
((chassi_id1!=.)&(chassi_id2!=.)&(chassi_id3!=.)&(chassi_id4==.)&(chassi_id5==.))
replace veh_nr=4 if
((chassi_id1!=.)&(chassi_id2!=.)&(chassi_id3!=.)&(chassi_id4!=.)&(chassi_id5==.))
replace veh_nr=5 if
((chassi_id1!=.)&(chassi_id2!=.)&(chassi_id3!=.)&(chassi_id4!=.)&(chassi_id5!=.))
keep id veh_nr
save "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta", replace
* opening barnover18 database and identify parnets
odbc load, exec("select * from dbo.VTI2008_barnover18") dsn("P0374_VTI_KTH_Bilinnehav")
clear
keep lopnr lopnrbarn
rename lopnr id_parent
rename lopnrbarn id_child
sort id_child id_parent

```

```

generate delta = id_child - id_child[_n-1]
generate d1 = 1 if (delta == 0)
generate order = 1 if (d1 != 1)
replace order = (order[_n-1]+1) if (d1 == 1 & order != 1)
generate int order2=1
replace order2=(order2[_n-1]+1) if (id_child==id_child[_n-1])
drop delta d1 order
reshape wide id_parent , i(id_child) j(order2)
generate parents_nr=0
replace parents_nr=1 if ((id_parent1!=.) & (id_parent2==.))
replace parents_nr=2 if ((id_parent1!=.) & (id_parent2!=.))
gen int p1_veh=.
gen int p1_inc=.
gen str p1_SAMS=""
gen int p2_veh=.
gen int p2_inc=.
gen str p2_SAMS=""
save "\\Mfso01\p0374_barsia$DL_model\data\parents_mtx.dta", replace
*merge to vehicle file to get number of cars available from each parent
use "\\Mfso01\p0374_barsia$DL_model\data\parents_mtx.dta", clear
keep id_parent1
rename id_parent1 id
merge m:m id using "\\Mfso01\p0374_barsia$DL_model\data\veh_mtx.dta"
rename id id_parent1
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$DL_model\data\p1.dta", replace
use "\\Mfso01\p0374_barsia$DL_model\data\parents_mtx.dta", clear

```

```

keep id_parent2
drop if(id_parent2==.)
rename id_parent2 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename id id_parent2
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta", replace
* put number of vehicles back in parents_matrix
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
merge m:m id_parent1 using "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta"
replace p1_veh=veh_nr if(_merge==3)
drop _merge veh_nr
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
merge m:m id_parent2 using "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta"
replace p2_veh=veh_nr if(_merge==3)
drop _merge veh_nr
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
* next fylling parents incomes
* individual data preparation
* reading from db and fixing the id's and deciding household head

*****
***** change here 2 *****

odbc load, exec("select * from dbo.VTI_2008MS ") dsn("P0374_VTI_KTH_Bilinnehav") clear
*****
*****

duplicates drop lopnr, force
keep lopnr fodar kon antalbarn Bosams Sysstat Stud Arbsams Dispink Korkort lopnrmak

```



```
rename lopnr id
rename fodar birth_year
rename kon sex
rename antalbarn children_nr
rename Bosams home_sams
rename Sysstat emp_status
rename Stud student
rename Arbsams work_sams
rename Dispink income
rename Korkkort drv_lic
rename lopnrmak id_partner
* reformating
destring birth_year, replace
*destring home_sams, replace
destring emp_stat, replace
*destring work_sams, replace
*destring drv_lic, replace
tostring children_nr, replace
destring children_nr, replace
*****
***** change here 3 *****
save "\\Mfso01\p0374_barsia$\DL_model\data\ind_08.dta", replace
*****
*****
*merging number of vehicles
merge 1:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename veh_nr own_veh
drop _merge
rename id own_id
```

```

rename id_partner id
merge m:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename veh_nr partner_veh
drop _merge
rename id id_partner
rename own_id id
*****
***** change here 3 ****
save "\\Mfso01\p0374_barsia$\DL_model\data\ind_08.dta", replace
*****
*****

keep id home_sams income
rename home_sams p_SAMS
rename income p_inc
save "\\Mfso01\p0374_barsia$\DL_model\data\p_data.dta", replace
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent1
rename id_parent1 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\p_data.dta"
rename id id_parent1
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta", replace
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent2
drop if(id_parent2==.)
rename id_parent2 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\p_data.dta"

```

```

rename id id_parent2
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta", replace
* put income and SAMS-info back in parents_matrix
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
rename id_child id
merge m:m id_parent1 using "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta"
tostring p1_SAMS, replace
replace p1_SAMS=p_SAMS if(_merge==3)
replace p1_inc=p_inc if(_merge==3)
drop _merge p_SAMS p_inc
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
merge m:m id_parent2 using "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta"
tostring p2_SAMS, replace
replace p2_SAMS=p_SAMS if(_merge==3)
replace p2_inc=p_inc if(_merge==3)
drop _merge p_SAMS p_inc
*rename id_child id
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
*merge to all individuals
*****
***** change here 3 ****
use "\\Mfso01\p0374_barsia$\DL_model\data\ind_08.dta", clear
*****
*****
duplicates tag id, gen(dups)
drop if(dups!=0)
merge 1:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta"

```

drop _merge dups

***** *change here 3* *****

save "\\Mfso01\p0374_barsia\$\DL_model\data\ind_08.dta", replace

***** 2009 *****

** vehicle data preparation*

***** *change here 1* *****

*odbc load, exec("select * from dbo.vy_bildata2009") dsn("P0374_VTI_KTH_Bilinnehav") clear*

keep lopnrG chassi_id

tostring chassi_id, replace

destring chassi_id, replace

** dropping duplicates(id)*

duplicates tag chassi_id, generate(dups)

drop if (dups > 0)

drop dups

** generate a variable to later find out which veh (1 or 2) has the highest chassi_id, need this for*

rename lopnrG id

sort id chassi_id

generate delta = id - id[_n-1]

generate d1 = 1 if (delta == 0)

generate order = 1 if (d1 != 1)

```

replace order = (order[_n-1]+1) if (d1 == 1 & order != 1)

* change number "x" below on the line "order > x" to the number of vehicles you want to include
ps. this changes the dimension of matrix

drop if (order > 5)

generate int order2=1

replace order2=(order2[_n-1]+1) if (id==id[_n-1])

drop delta d1 order

* reshaping to matrix form/wide form

reshape wide chassi_id , i(id) j(order2)

generate int veh_nr=0

replace veh_nr=1 if
((chassi_id1!=.) & (chassi_id2==.) & (chassi_id3==.) & (chassi_id4==.) & (chassi_id5==.))

replace veh_nr=2 if
((chassi_id1!=.) & (chassi_id2!=.) & (chassi_id3==.) & (chassi_id4==.) & (chassi_id5==.))

replace veh_nr=3 if
((chassi_id1!=.) & (chassi_id2!=.) & (chassi_id3!=.) & (chassi_id4==.) & (chassi_id5==.))

replace veh_nr=4 if
((chassi_id1!=.) & (chassi_id2!=.) & (chassi_id3!=.) & (chassi_id4!=.) & (chassi_id5==.))

replace veh_nr=5 if
((chassi_id1!=.) & (chassi_id2!=.) & (chassi_id3!=.) & (chassi_id4!=.) & (chassi_id5!=.))

keep id veh_nr

save "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta", replace

* opening barnover18 database and identify parnets

odbc load, exec("select * from dbo.VTI2009_barnover18") dsn("P0374_VTI_KTH_Bilinnhav")
clear

keep LopNr LopNrBarn

destring LopNr, replace

destring LopNrBarn, replace

rename LopNr id_parent

rename LopNrBarn id_child

```

```

sort id_child id_parent
generate delta = id_child - id_child[_n-1]
generate d1 = 1 if (delta == 0)
generate order = 1 if (d1 != 1)
replace order = (order[_n-1]+1) if (d1 == 1 & order != 1)
generate int order2=1
replace order2=(order2[_n-1]+1) if (id_child==id_child[_n-1])
drop delta d1 order
reshape wide id_parent , i(id_child) j(order2)
generate parents_nr=0
replace parents_nr=1 if ((id_parent1!=.)&(id_parent2==.))
replace parents_nr=2 if ((id_parent1!=.)&(id_parent2!=.))
gen int p1_veh=.
gen int p1_inc=.
gen str p1_SAMS=""
gen int p2_veh=.
gen int p2_inc=.
gen str p2_SAMS=""
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
*merge to vehicle file to get number of cars available from each parent
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent1
rename id_parent1 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename id id_parent1
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta", replace

```

```

use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent2
drop if(id_parent2=.)
rename id_parent2 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename id id_parent2
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta", replace
* put number of vehicles back in parents_matrix
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
merge m:m id_parent1 using "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta"
replace p1_veh=veh_nr if(_merge==3)
drop _merge veh_nr
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
merge m:m id_parent2 using "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta"
replace p2_veh=veh_nr if(_merge==3)
drop _merge veh_nr
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
* next fylling parents incomes
* individual data preparation
* reading from db and fixing the id's and deciding household head
*****
***** change here 2 *****

odbc load, exec("select * from dbo.VTI2009 ") dsn("P0374_VTI_KTH_Bilinnerhav") clear
*****
*****

duplicates drop lopnr, force

```

keep lopnr Fodar Kon AntalBarn BoSAMS sysstat stud ArbSAMS Dispink korkort LopnrMak

rename lopnr id

rename Fodar birth_year

rename Kon sex

rename AntalBarn children_nr

rename BoSAMS home_sams

rename sysstat emp_status

rename stud student

rename ArbSAMS work_sams

rename Dispink income

rename korkort drv_lic

rename LopnrMak id_partner

** reformating*

save "\\Mfso01\p0374_barsia\$\DL_model\data\v1_mtx.dta", replace

destring birth_year, replace

**destring home_sams, replace*

destring emp_stat, replace

**destring work_sams, replace*

tostring children_nr, replace

destring children_nr, replace

***** *change here 3* *****

save "\\Mfso01\p0374_barsia\$\DL_model\data\ind_09.dta", replace

**merging number of vehicles*

merge 1:1 id using "\\Mfso01\p0374_barsia\$\DL_model\data\veh_mtx.dta"

rename veh_nr own_veh

drop _merge

rename id own_id

rename id_partner id

destring id, replace

merge m:1 id using "\\Mfso01\p0374_barsia\$\DL_model\data\veh_mtx.dta"

rename veh_nr partner_veh

drop _merge

rename id id_partner

rename own_id id

***** *change here 3* *****

save "\\Mfso01\p0374_barsia\$\DL_model\data\ind_09.dta", replace

keep id home_sams income

rename home_sams p_SAMS

rename income p_inc

save "\\Mfso01\p0374_barsia\$\DL_model\data\p_data.dta", replace

use "\\Mfso01\p0374_barsia\$\DL_model\data\parents_mtx.dta", clear

keep id_parent1

rename id_parent1 id

merge m:m id using "\\Mfso01\p0374_barsia\$\DL_model\data\p_data.dta"

rename id id_parent1

drop if(_merge!=3)

drop _merge

```

save "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta", replace
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent2
drop if(id_parent2==.)
rename id_parent2 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\p_data.dta"
rename id id_parent2
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta", replace
* put income and SAMS-info back in parents_matrix
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
rename id_child id
merge m:m id_parent1 using "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta"
tostring p1_SAMS, replace
replace p1_SAMS=p_SAMS if(_merge==3)
replace p1_inc=p_inc if(_merge==3)
drop _merge p_SAMS p_inc
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
merge m:m id_parent2 using "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta"
tostring p2_SAMS, replace
replace p2_SAMS=p_SAMS if(_merge==3)
replace p2_inc=p_inc if(_merge==3)
drop _merge p_SAMS p_inc
*rename id_child id
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
*merge to all individuals
*****
***** change here 3 *****

```

use "\\Mfso01\p0374_barsia\$\DL_model\data\ind_09.dta", clear

duplicates tag id, gen(dups)

drop if(dups!=0)

merge 1:1 id using "\\Mfso01\p0374_barsia\$\DL_model\data\parents_mtx.dta"

drop _merge dups

***** *change here 3* ***

save "\\Mfso01\p0374_barsia\$\DL_model\data\ind_09.dta", replace

***** 2010 *****

** vehicle data preparation*

***** *change here 1* ***

*odbc load, exec("select * from dbo.vy_bildata2010") dsn("P0374_VTI_KTH_Bilinnehav") clear*

rename lopnrG id

keep id chassi_id

tostring chassi_id, replace

destring chassi_id, replace

** dropping duplicates(id)*

duplicates tag chassi_id, generate(dups)

drop if (dups > 0)

drop dups

** generate a variable to later find out which veh (1 or 2) has the highest chassi_id, need this for sort id chassi_id*

generate delta = id - id[_n-1]

generate d1 = 1 if (delta == 0)

generate order = 1 if (d1 != 1)

replace order = (order[_n-1]+1) if (d1 == 1 & order != 1)

** change number "x" below on the line "order > x" to the number of vehicles you want to include ps. this changes the dimension of matrix*

drop if (order > 5)

generate int order2=1

replace order2=(order2[_n-1]+1) if (id==id[_n-1])

drop delta d1 order

** reshaping to matrix form/wide form*

reshape wide chassi_id , i(id) j(order2)

generate int veh_nr=0

replace veh_nr=1 if

((chassi_id1!=.) & (chassi_id2==.) & (chassi_id3==.) & (chassi_id4==.) & (chassi_id5==.))

replace veh_nr=2 if

((chassi_id1!=.) & (chassi_id2!=.) & (chassi_id3==.) & (chassi_id4==.) & (chassi_id5==.))

replace veh_nr=3 if

((chassi_id1!=.) & (chassi_id2!=.) & (chassi_id3!=.) & (chassi_id4==.) & (chassi_id5==.))

replace veh_nr=4 if

((chassi_id1!=.) & (chassi_id2!=.) & (chassi_id3!=.) & (chassi_id4!=.) & (chassi_id5==.))

replace veh_nr=5 if

((chassi_id1!=.) & (chassi_id2!=.) & (chassi_id3!=.) & (chassi_id4!=.) & (chassi_id5!=.))

keep id veh_nr

save "\\Mfso01\p0374_barsia\$DL_model\data\veh_mtx.dta", replace

** opening barnover18 database and identify parnets*

*odbc load, exec("select * from dbo.VTI2010_barnover18") dsn("P0374_VTI_KTH_Bilinnehav")*

clear

```

keep LopNr LopNrBarn
destring LopNr, replace
destring LopNrBarn, replace
rename LopNr id_parent
rename LopNrBarn id_child
sort id_child id_parent
generate delta = id_child - id_child[_n-1]
generate d1 = 1 if (delta == 0)
generate order = 1 if (d1 != 1)
replace order = (order[_n-1]+1) if (d1 == 1 & order != 1)
generate int order2=1
replace order2=(order2[_n-1]+1) if (id_child==id_child[_n-1])
drop delta d1 order
reshape wide id_parent , i(id_child) j(order2)
generate parents_nr=0
replace parents_nr=1 if ((id_parent1!=.)&(id_parent2==.))
replace parents_nr=2 if ((id_parent1!=.)&(id_parent2!=.))
gen int p1_veh=.
gen int p1_inc=.
gen str p1_SAMS=""
gen int p2_veh=.
gen int p2_inc=.
gen str p2_SAMS=""
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
*merge to vehicle file to get number of cars available from each parent
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent1
rename id_parent1 id

```

```

merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename id id_parent1
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta", replace
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent2
drop if(id_parent2==.)
rename id_parent2 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename id id_parent2
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta", replace
* put number of vehicles back in parents_matrix
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
merge m:m id_parent1 using "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta"
replace p1_veh=veh_nr if(_merge==3)
drop _merge veh_nr
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
merge m:m id_parent2 using "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta"
replace p2_veh=veh_nr if(_merge==3)
drop _merge veh_nr
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
* next fylling parents incomes
* individual data preparation
* reading from db and fixing the id's and deciding household head
*****
***** change here 2 *****

```

*odbc load, exec("select * from dbo.VTI2010 ") dsn("P0374_VTI_KTH_Bilnehav") clear*

rename LopnrMak Mak

duplicates drop lopnr, force

keep lopnr Fodar Kon AntalBarn BoSAMS sysstat stud ArbSAMS Dispink korkort Mak

rename lopnr id

rename Fodar birth_year

rename Kon sex

rename AntalBarn children_nr

rename BoSAMS home_sams

rename sysstat emp_status

rename stud student

rename ArbSAMS work_sams

rename Dispink income

rename korkort drv_lic

rename Mak id_partner

** reformating*

destring birth_year, replace

**destring home_sams, replace*

destring emp_stat, replace

**destring work_sams, replace*

tostring children_nr, replace

destring children_nr, replace

***** change here 3 ****

```

save "\\Mfso01\p0374_barsia$\DL_model\data\ind_10.dta", replace
*****
*****

*merging number of vehicles
merge 1:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename veh_nr own_veh

drop _merge

rename id own_id

rename id_partner id

destring id, replace
merge m:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename veh_nr partner_veh

drop _merge

rename id id_partner

rename own_id id

*****
***** change here 3 ****

save "\\Mfso01\p0374_barsia$\DL_model\data\ind_10.dta", replace
*****
*****

keep id home_sams income

rename home_sams p_SAMS

rename income p_inc

save "\\Mfso01\p0374_barsia$\DL_model\data\p_data.dta", replace
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear

keep id_parent1

rename id_parent1 id

```



```

merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\p_data.dta"
rename id id_parent1
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta", replace
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent2
drop if(id_parent2==.)
rename id_parent2 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\p_data.dta"
rename id id_parent2
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta", replace
* put income and SAMS-info back in parents_matrix
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
rename id_child id
merge m:m id_parent1 using "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta"
tostring p1_SAMS, replace
replace p1_SAMS=p_SAMS if(_merge==3)
replace p1_inc=p_inc if(_merge==3)
drop _merge p_SAMS p_inc
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
merge m:m id_parent2 using "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta"
tostring p2_SAMS, replace
replace p2_SAMS=p_SAMS if(_merge==3)
replace p2_inc=p_inc if(_merge==3)
drop _merge p_SAMS p_inc

```

**rename id_child id*

save "\\Mfso01\p0374_barsia\$\DL_model\data\parents_mtx.dta", replace

**merge to all individuals*

***** *change here 3* *****

use "\\Mfso01\p0374_barsia\$\DL_model\data\ind_10.dta", clear

duplicates tag id, gen(dups)

drop if(dups!=0)

merge 1:1 id using "\\Mfso01\p0374_barsia\$\DL_model\data\parents_mtx.dta"

drop _merge dups

***** *change here 3* *****

save "\\Mfso01\p0374_barsia\$\DL_model\data\ind_10.dta", replace

***** *2011* *****

** vehicle data preparation*

***** *change here 1* *****

*odbc load, exec("select * from dbo.vy_bildata2011") dsn("P0374_VTI_KTH_Bilinnehav") clear*

rename lopnrG id

keep id chassi_id

```

tostring chassi_id, replace
destring chassi_id, replace
* dropping duplicates(id)
duplicates tag chassi_id, generate(dups)
drop if (dups > 0)
drop dups
* generate a variable to later find out which veh (1 or 2) has the highest chassi_id, need this for
sort id chassi_id
generate delta = id - id[_n-1]
generate d1 = 1 if (delta == 0)
generate order = 1 if (d1 != 1)
replace order = (order[_n-1]+1) if (d1 == 1 & order != 1)
* change number "x" below on the line "order > x" to the number of vehicles you want to include
ps. this changes the dimension of matrix
drop if (order > 5)
generate int order2=1
replace order2=(order2[_n-1]+1) if (id==id[_n-1])
drop delta d1 order
* reshaping to matrix form/wide form
reshape wide chassi_id , i(id) j(order2)
generate int veh_nr=0
replace veh_nr=1 if
((chassi_id1!=.) & (chassi_id2==.) & (chassi_id3==.) & (chassi_id4==.) & (chassi_id5==.))
replace veh_nr=2 if
((chassi_id1!=.) & (chassi_id2!=.) & (chassi_id3==.) & (chassi_id4==.) & (chassi_id5==.))
replace veh_nr=3 if
((chassi_id1!=.) & (chassi_id2!=.) & (chassi_id3!=.) & (chassi_id4==.) & (chassi_id5==.))
replace veh_nr=4 if
((chassi_id1!=.) & (chassi_id2!=.) & (chassi_id3!=.) & (chassi_id4!=.) & (chassi_id5==.))

```

```

replace veh_nr=5 if
((chassi_id1!=.)&(chassi_id2!=.)&(chassi_id3!=.)&(chassi_id4!=.)&(chassi_id5!=.))
keep id veh_nr
save "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta", replace
* opening barnover18 database and identify parnets
odbc load, exec("select * from dbo.VTI2011_barnover18") dsn("P0374_VTI_KTH_Bilinnhav")
clear
keep LopNr LopNrBarn
destring LopNr, replace
destring LopNrBarn, replace
rename LopNr id_parent
rename LopNrBarn id_child
sort id_child id_parent
generate delta = id_child - id_child[_n-1]
generate d1 = 1 if (delta == 0)
generate order = 1 if (d1 !=1)
replace order = (order[_n-1]+1) if (d1 == 1 & order != 1)
generate int order2=1
replace order2=(order2[_n-1]+1) if (id_child==id_child[_n-1])
drop delta d1 order
reshape wide id_parent , i(id_child) j(order2)
generate parents_nr=0
replace parents_nr=1 if ((id_parent1!=.)&(id_parent2==.))
replace parents_nr=2 if ((id_parent1!=.)&(id_parent2!=.))
gen int p1_veh=.
gen int p1_inc=.
gen str p1_SAMS=""
gen int p2_veh=.
gen int p2_inc=.

```

```

gen str p2_SAMS=""
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
*merge to vehicle file to get number of cars available from each parent
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent1
rename id_parent1 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename id id_parent1
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta", replace
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent2
drop if(id_parent2==.)
rename id_parent2 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename id id_parent2
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta", replace
* put number of vehicles back in parents_matrix
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
merge m:m id_parent1 using "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta"
replace p1_veh=veh_nr if(_merge==3)
drop _merge veh_nr
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
merge m:m id_parent2 using "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta"
replace p2_veh=veh_nr if(_merge==3)

```

```
drop _merge veh_nr
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
* next fylling parents incomes
* individual data preparation
* reading from db and fixing the id's and deciding household head
*****
***** change here 2 ****

odbc load, exec("select * from dbo.VTI2011 ") dsn("P0374_VTI_KTH_Bilinnerhav") clear
*****
*****

duplicates drop lopnr, force
keep lopnr Fodar Kon AntalBarn BoSAMS sysstat stud ArbSAMS Dispink korkort LopnrMak
rename lopnr id
rename Fodar birth_year
rename Kon sex
rename AntalBarn children_nr
rename BoSAMS home_sams
rename sysstat emp_status
rename stud student
rename ArbSAMS work_sams
rename Dispink income
rename korkort drv_lic
rename LopnrMak id_partner
*rename Barnover adult_child

* reformating
destring birth_year, replace
*destring home_sams, replace
```

```
destring emp_stat, replace
*destring work_sams, replace
tostring children_nr, replace
destring children_nr, replace
*****
***** change here 3 ***
save "\\Mfso01\p0374_barsia$\DL_model\data\ind_11.dta", replace
*****
*****
*merging number of vehicles
merge 1:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename veh_nr own_veh
drop _merge
rename id own_id
rename id_partner id

destring id, replace
merge m:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\veh_mtx.dta"
rename veh_nr partner_veh
drop _merge
rename id id_partner
rename own_id id
*****
***** change here 3 ***
save "\\Mfso01\p0374_barsia$\DL_model\data\ind_11.dta", replace
*****
*****

keep id home_sams income
rename home_sams p_SAMS
```

```

rename income p_inc
save "\\Mfso01\p0374_barsia$\DL_model\data\p_data.dta", replace
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent1
rename id_parent1 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\p_data.dta"
rename id id_parent1
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta", replace
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear
keep id_parent2
drop if(id_parent2==.)
rename id_parent2 id
merge m:m id using "\\Mfso01\p0374_barsia$\DL_model\data\p_data.dta"
rename id id_parent2
drop if(_merge!=3)
drop _merge
save "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta", replace
* put income and SAMS-info back in parents_matrix
use "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", clear

merge m:m id_parent1 using "\\Mfso01\p0374_barsia$\DL_model\data\p1.dta"
tostring p1_SAMS, replace
replace p1_SAMS=p_SAMS if(_merge==3)
replace p1_inc=p_inc if(_merge==3)
drop _merge p_SAMS p_inc
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace

```



```
merge m:m id_parent2 using "\\Mfso01\p0374_barsia$\DL_model\data\p2.dta"
tostring p2_SAMS, replace
replace p2_SAMS=p_SAMS if(_merge==3)
replace p2_inc=p_inc if(_merge==3)
drop _merge p_SAMS p_inc
rename id_child id
save "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta", replace
*merge to all individuals
*****
***** change here 3 ****
```

```
use "\\Mfso01\p0374_barsia$\DL_model\data\ind_11.dta", clear
*****
*****
duplicates tag id, gen(dups)
drop if(dups!=0)
```

```
merge 1:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\parents_mtx.dta"
drop _merge dups
*****
***** change here 3 ****
save "\\Mfso01\p0374_barsia$\DL_model\data\ind_11.dta", replace
```

```
*****extracting id from all years for sampling stage
use "\\Mfso01\p0374_barsia$\DL_model\data\ind_03.dta", clear
gen int year=2003
```

```
save "\\Mfso01\p0374_barsia$\DL_model\data\ind_03.dta", replace  
keep id
```

```
save "\\Mfso01\p0374_barsia$\DL_model\data\id_03.dta", replace
```

```
use "\\Mfso01\p0374_barsia$\DL_model\data\ind_04.dta", clear  
gen int year=2004
```

```
save "\\Mfso01\p0374_barsia$\DL_model\data\ind_04.dta", replace  
keep id
```

```
save "\\Mfso01\p0374_barsia$\DL_model\data\id_04.dta", replace
```

```
use "\\Mfso01\p0374_barsia$\DL_model\data\ind_05.dta", clear  
gen int year=2005
```

```
save "\\Mfso01\p0374_barsia$\DL_model\data\ind_05.dta", replace  
keep id
```

```
save "\\Mfso01\p0374_barsia$\DL_model\data\id_05.dta", replace
```

```
use "\\Mfso01\p0374_barsia$\DL_model\data\ind_06.dta", clear  
gen int year=2006
```

```
save "\\Mfso01\p0374_barsia$\DL_model\data\ind_06.dta", replace  
keep id
```

```
save "\\Mfso01\p0374_barsia$\DL_model\data\id_06.dta", replace
```

```
use "\\Mfso01\p0374_barsia$\DL_model\data\ind_07.dta", clear  
gen int year=2007
```

```
save "\\Mfso01\p0374_barsia$\DL_model\data\ind_07.dta", replace  
keep id
```

```
save "\\Mfso01\p0374_barsia$\DL_model\data\id_07.dta", replace
```

```
use "\\Mfso01\p0374_barsia$\DL_model\data\ind_08.dta", clear
gen int year=2008
save "\\Mfso01\p0374_barsia$\DL_model\data\ind_08.dta", replace
keep id
save "\\Mfso01\p0374_barsia$\DL_model\data\id_08.dta", replace
```

```
use "\\Mfso01\p0374_barsia$\DL_model\data\ind_09.dta", clear
gen int year=2009
save "\\Mfso01\p0374_barsia$\DL_model\data\ind_09.dta", replace
keep id
save "\\Mfso01\p0374_barsia$\DL_model\data\id_09.dta", replace
```

```
use "\\Mfso01\p0374_barsia$\DL_model\data\ind_10.dta", clear
gen int year=2010
save "\\Mfso01\p0374_barsia$\DL_model\data\ind_10.dta", replace
keep id
save "\\Mfso01\p0374_barsia$\DL_model\data\id_10.dta", replace
```

```
use "\\Mfso01\p0374_barsia$\DL_model\data\ind_11.dta", clear
gen int year=2011
save "\\Mfso01\p0374_barsia$\DL_model\data\ind_11.dta", replace
keep id
save "\\Mfso01\p0374_barsia$\DL_model\data\id_11.dta", replace
```

```
use "\\Mfso01\p0374_barsia$\DL_model\data\id_11.dta", clear
```

```
append using "\\Mfso01\p0374_barsia$\DL_model\data\id_10.dta"
append using "\\Mfso01\p0374_barsia$\DL_model\data\id_09.dta"
```

```
append using "\\Mfso01\p0374_barsia$\DL_model\data\id_08.dta"  
append using "\\Mfso01\p0374_barsia$\DL_model\data\id_07.dta"  
append using "\\Mfso01\p0374_barsia$\DL_model\data\id_06.dta"  
append using "\\Mfso01\p0374_barsia$\DL_model\data\id_05.dta"  
append using "\\Mfso01\p0374_barsia$\DL_model\data\id_04.dta"  
append using "\\Mfso01\p0374_barsia$\DL_model\data\id_03.dta"  
save "\\Mfso01\p0374_barsia$\DL_model\data\id_all.dta", replace  
duplicates drop id, force  
save "\\Mfso01\p0374_barsia$\DL_model\data\id_all.dta", replace
```

**** make samples*

**merging attributes to sampled individuals*

```
use "\\Mfso01\p0374_barsia$\DL_model\data\id_all.dta", clear  
sample 10  
save "\\Mfso01\p0374_barsia$\DL_model\data\sample_10.dta", replace
```

```
use "\\Mfso01\p0374_barsia$\DL_model\data\sample_10.dta", clear  
merge 1:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\ind_03.dta"  
drop if(_merge!=3)  
save "\\Mfso01\p0374_barsia$\DL_model\data\sample_10_03.dta", replace
```

```
use "\\Mfso01\p0374_barsia$\DL_model\data\sample_10.dta", clear  
merge 1:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\ind_04.dta"  
drop if(_merge!=3)  
save "\\Mfso01\p0374_barsia$\DL_model\data\sample_10_04.dta", replace
```

```
use "\\Mfso01\p0374_barsia$\DL_model\data\sample_10.dta", clear  
merge 1:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\ind_05.dta"
```

```
drop if(_merge!=3)
save "\\Mfso01\p0374_barsia$\DL_model\data\sample_10_05.dta", replace

use "\\Mfso01\p0374_barsia$\DL_model\data\sample_10.dta", clear
merge 1:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\ind_06.dta"
drop if(_merge!=3)
save "\\Mfso01\p0374_barsia$\DL_model\data\sample_10_06.dta", replace

use "\\Mfso01\p0374_barsia$\DL_model\data\sample_10.dta", clear
merge 1:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\ind_07.dta"
drop if(_merge!=3)
save "\\Mfso01\p0374_barsia$\DL_model\data\sample_10_07.dta", replace

use "\\Mfso01\p0374_barsia$\DL_model\data\sample_10.dta", clear
merge 1:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\ind_08.dta"
drop if(_merge!=3)
save "\\Mfso01\p0374_barsia$\DL_model\data\sample_10_08.dta", replace

use "\\Mfso01\p0374_barsia$\DL_model\data\sample_10.dta", clear
merge 1:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\ind_09.dta"
drop if(_merge!=3)
save "\\Mfso01\p0374_barsia$\DL_model\data\sample_10_09.dta", replace

use "\\Mfso01\p0374_barsia$\DL_model\data\sample_10.dta", clear
merge 1:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\ind_10.dta"
drop if(_merge!=3)
save "\\Mfso01\p0374_barsia$\DL_model\data\sample_10_10.dta", replace
```

```
use "\\Mfso01\p0374_barsia$\DL_model\data\sample_10.dta", clear
merge 1:1 id using "\\Mfso01\p0374_barsia$\DL_model\data\ind_11.dta"
drop if(_merge!=3)
save "\\Mfso01\p0374_barsia$\DL_model\data\sample_10_11.dta", replace
```

```
use "\\Mfso01\p0374_barsia$\DL_model\data\sample_10_11.dta", clear
append using "\\Mfso01\p0374_barsia$\DL_model\data\sample_10_10.dta"
append using "\\Mfso01\p0374_barsia$\DL_model\data\sample_10_09.dta"
destring student, replace
destring id_partner, replace
append using "\\Mfso01\p0374_barsia$\DL_model\data\sample_10_08.dta"
append using "\\Mfso01\p0374_barsia$\DL_model\data\sample_10_07.dta"
append using "\\Mfso01\p0374_barsia$\DL_model\data\sample_10_06.dta"
append using "\\Mfso01\p0374_barsia$\DL_model\data\sample_10_05.dta"
append using "\\Mfso01\p0374_barsia$\DL_model\data\sample_10_04.dta"
append using "\\Mfso01\p0374_barsia$\DL_model\data\sample_10_03.dta"
```

```
tostring sex, replace
```

```
destring sex, replace
```

```
tostring drv_lic, replace
```

```
destring drv_lic, replace
```

```
tostring parents_nr, replace
```

```
destring parents_nr, replace
```

** some observaions in 2009 to 2011 have null for drv_lic, checked, they shoudl replaced with zero*

```
replace drv_lic=0 if(drv_lic==.)
```

```
sort id year
```

```
order id year drv_lic
```

```
gen age=year-birth_year
```

drop if age==17

** fix a order number to enable identification of first year of observation*

sort id year

gen int ord=1 if(id[_n-1]!=id)

order id year ord age

**getting rid of extra observation year after drv-lic is acquired (older than 18 when they acquire drv-lic), since they would have been censored for years after they got drv-lic we delete those years from data set and only keep the years until they get drv-lic*

sort id year

gen int indicator2=77 if(drv_lic==1)

** some individuals seem to get drv_lic while after few years they seem to not have it. they will be kept only until they acquire drv_lic first time, the following years will be dropped*

replace indicator2=77 if((indicator2[_n-1]==77)&(id==id[_n-1]))

drop if((indicator2==77)&(indicator2[_n-1]==77))

order id ord indicator1 indicator2 year age

**get rid of few individuals to be safe, who acquire drv-lic while next year they dont have it,*

gen indicator3=66 if((id==id[_n-1])&(indicator2[_n-1]==77)&(indicator2==.))

order id ord indicator1 indicator2 indicator3 year age

gsort id -ord

replace indicator3=66 if((indicator3[_n-1]==66)&(id==id[_n-1]))

drop if indicator3==66

** getting rid of individuals which have missing years in data (with gaps in data)*

sort id year

gen test2=1 if((id==id[_n-1])&(year!=year[_n-1]+1))

replace test2=1 if((id==id[_n-1])&(test2[_n-1]==1))

gsort id -test2

replace test2=1 if((id==id[_n-1])&(test2[_n-1]==1))

drop if test2==1

drop test2

```

sort id year

drop indicator1 indicator2 indicator3 _merge recycled_id int_single_year_drv

gen int t_org=(birth_year+18)

* preparing variables

sort id year

replace sex=0 if sex==1

replace sex=1 if sex==2

gen int married=0

replace married=1 if((id_partner!=.)&(id_partner!=0))

gen int veh_nr=0

replace veh_nr=(own_veh+partner_veh)

replace veh_nr=0 if veh_nr==.

gen int acc_veh=0

replace acc_veh=(p1_veh+p2_veh)

replace acc_veh=0 if(acc_veh==.)

gen inc_parents=0

replace inc_parents=(p1_inc+p2_inc)

gen ln_inc =ln(income)

replace ln_in=0 if(ln_inc==.)

gen lninc_parents=ln(inc_parents)

replace lninc_parents=0 if lninc_parents

gen int first_age=age if(ord==1)

replace first_age=first_age[_n-1] if((id==id[_n-1]))

* fix employment 5 and 6, unemployed

replace emp_status=0 if(emp_status==5)

replace emp_status=0 if(emp_status==6)

*6 individuals have strange data, drop them

```



```
drop if
((id==16850314)|(id==11859317)|(id==10482083)|(id==10880389)|(id==12487182)|(id==65
40914))
```

** fix for population calss, need to first make municipality id from home_sams and then merge with population data*

** already in string format, otherwise change format, tostring home_sams, replace*

```
gen str muni_id=substr(home_sams, 1, 4)
```

```
destring muni_id, replace
```

```
merge m:1 muni_id using "\\Mfso01\p0374_barsia$\DL_model\data\final_data\bef.dta"
```

```
drop _merge
```

```
rename Totalt population
```

** population data is missing for Heby municipality, population 13450*

```
replace population=13450 if(muni_id==1917)
```

```
gen int pop_class=.
```

```
replace pop_class=1 if(population<16000)
```

```
replace pop_class=2 if((population>16000)&(population<44000))
```

```
replace pop_class=3 if((population>44000)&(population<200000))
```

```
replace pop_class=4 if(population>200000)
```

```
tabstat pop_class, by(pop_class) stat(n)
```

```
gen ln_population=ln(population)
```

**** new pop cateogires*

```
gen int pop_cat=.
```

```
replace pop_cat=10 if((population<=10000))
```

```
replace pop_cat=20 if((population>10000)&(population<=20000))
```

```
replace pop_cat=40 if((population>20000)&(population<=40000))
```

```
replace pop_cat=60 if((population>40000)&(population<=60000))
```

replace pop_cat=80 if ((population>60000)&(population<=80000))
replace pop_cat=100 if ((population>80000)&(population<=100000))
replace pop_cat=120 if ((population>100000)&(population<=120000))
replace pop_cat=140 if ((population>120000)&(population<=140000))
replace pop_cat=160 if ((population>140000)&(population<=160000))
replace pop_cat=161 if ((population>160000))

gen int pop_cat10=0
replace pop_cat10=1 if pop_cat==10
gen int pop_cat20=0
replace pop_cat20=1 if pop_cat==20
gen int pop_cat40=0
replace pop_cat40=1 if pop_cat==40
gen int pop_cat60=0
replace pop_cat60=1 if pop_cat==60
gen int pop_cat80=0
replace pop_cat80=1 if pop_cat==80
gen int pop_cat100=0
replace pop_cat100=1 if pop_cat==100
gen int pop_cat120=0
replace pop_cat120=1 if pop_cat==120
gen int pop_cat140=0
replace pop_cat140=1 if pop_cat==140
gen int pop_cat160=0
replace pop_cat160=1 if pop_cat==160
gen int pop_cat161=0
replace pop_cat161=1 if pop_cat==161

** solving age problem in register*

** discontinuity in age*

sort id year

gen test1=1 if((id==id[_n-1])&(age!=age[_n-1]+1))

replace test1=1 if((id==id[_n-1])&(test1[_n-1]==1))

gsort id -year

replace test1=1 if((id==id[_n-1])&(test1[_n-1]==1))

** now we delete them, because we can not be sure which of the lines with same age is correct, after all it is about 11 observations or so*

drop if test1==1

drop test1

** fix for sex*

sort id year

gen sex2=sex if(ord==1)

replace sex2=sex2[_n-1] if(id==id[_n-1])

replace sex=sex2

drop sex2

save "\\Mfso01\p0374_barsia\$\DL_model\data\final_data\sample10.dta", replace

** need to expand for truncated years and generate j and event,*

** first drop unwanted variables*

*drop id_partner own_veh partner_veh id_parent1 id_parent2 p1_veh p1_inc p1_SAMS p2_veh
p2_inc p2_SAMS*

sort id year

gen int t_enter=year if(ord==1)

replace t_enter=t_enter[_n-1] if(id==id[_n-1])

gen int trunc_nr=t_enter-t_org+1

```

expand trunc_nr if(ord==1)

* no fixing year and age for the truncated years
gen year2=year if ord>=2

gsort id -year

replace year2=year2[_n-1]-1 if((id==id[_n-1])&(ord==1))
replace year2=year if((trunc_nr==1)&(ord==1))
gen x1=1 if year2==.

* Those who are observed only once will not have year 2
replace year2=year if((id!=id[_n-1])&(year2==.))
replace year2=year2[_n-1]-1 if((id==id[_n-1])&(year2==.))

* fixing age
replace age =age[_n-1]-1 if((id==id[_n-1])&(x1==.))

sort id year

replace year=year2

drop year2 x1

replace age=year-birth_year

sort id year

bysort id: ge j=_n

lab var j "elapsed time in years"

bysort id: ge event=drv_lic==1 & _n==_N

lab var event "event=1, otherwise=0"

replace year=t_org+j-1

replace t_enter=year if((ord==1)&(ord[_n+1]==2))
replace t_enter=t_enter[_n-1] if t_enter[_n-1]!=. & id==id[_n-1]

gsort id -year

replace t_enter=t_enter[_n-1] if t_enter[_n-1]!=. & id==id[_n-1]

```

sort id year

** few observations (152) seems to have entered 2001 and 2002, need ot delete them*

drop if t_enter<2003

gen lnj=ln(j)

gen jq2=j^2

gen jq3=j^3

** categoris for income*

egen inc_cat= cut(income), at(20000, 50000, 100000, 150000, 200000, 250000)

replace inc_parents=0 if inc_parents==. / replacing thoes with no parents from "." to "0" */*

replace inc_parents=0 if inc_parents<0 / few parents have income debts, we set them to zero */*

egen inc_par_cat= cut(inc_parents), at(50000, 100000, 150000, 200000, 250000, 3000000, 350000, 400000, 450000, 500000, 550000, 600000, 700000)

compress

save "\\Mfso01\p0374_barsia\$\DL_model\data\final_data\samp10_expand.dta", replace

Modeller

*cloglog event lnj age sex emp_status student new_chld ln_incpa_cat new_accveh ln_pop,
vce(cluster id) noconstant*

estimates store clog_lnj

*cloglog event jq2 age sex emp_status student children_nr lninc_parents acc_veh ln_pop,
vce(cluster id) noconstant*

estimates store clog_jq2

```
cloglog event jq3 age sex emp_status student children_nr lninc_parents acc_veh ln_pop,  
vce(cluster id) noconstant
```

```
estimates store clog_jq3
```

```
estimates table clog_lnj clog_jq2 clog_jq3, star(.01 .05 .1)
```

```
estimates stats clog_lnj clog_jq2 clog_jq3
```