Motivation

- EU-SILC and at risk of social exclusion (‘arose‘)
- Qualitatively high well-being indicators at national or NUTS1
- Lower NUTS-Levels usually yield poor estimates
- Methodology, which is easy to apply and yields better estimates on sub-national levels?

Methods

- Many techniques already exist for estimating indicators on sub-national levels.
- Already existing techniques, for example
  - Small area estimation
  - Use administrative data to impute variable of interest (see povmap)
- Modells need assumptions and administrative data is not always available
- Need a more harmonious approach R-Package $\rightarrow$ `surveysd`

STATISTIK AUSTRIA
Die Informationsmanager

- R-package for variance estimation on regional levels
- Variance estimation via bootstrap techniques
- Uses multiple (consecutive) waves of a survey
  - Similar approach as proposed by VIJAY
  - Average bootstrap replicates over waves
- Easy to use, even for R-Beginners

Main functionality

STATISTIK AUSTRIA
Die Informationsmanager

- Draw bootstrap replicates ‘draw.bootstrap()‘
- Calibrate bootstrap replicates ‘recalib()‘
- Estimate standard errors ‘calc.stError()‘

Draw bootstrap replicates

```
draw.bootstrap(dat,REP=1000,hid="DB030",weights="RB050",
               year="RB010",strata="DB040",cluster=NULL,
               totals=NULL,single.PSU=c("merge","mean"),
               boot.names=NULL,country=NULL,split=FALSE,pid=NULL)
```

- Rectangular data set with household identifier
- Column with sampling weight, year
- Define arbitrary sampling design with `strata` and `cluster`
- Automatic detection and dealing with single PSUs
- Bootstrap replicates are drawn for each year.
  - Applies rescaled bootstrap for stratified multistage sampling
- Replicates are taken forward to mimic rotational panel design.
  - Split households can be considered for this step, `split=TRUE`

Draw bootstrap replicates

STATISTIK AUSTRIA
Die Informationsmanager

```
draw.bootstrap(dat,REP=1000,hid="DB030",weights="RB050",
               year="RB010",strata="DB040",cluster=NULL,
               totals=NULL,single.PSU=c("merge","mean"),
               boot.names=NULL,country=NULL,split=FALSE,pid=NULL)
```

- Rectangular data set with household identifier
- Column with sampling weight, year
- Define arbitrary sampling design with `strata` and `cluster`
- Automatic detection and dealing with single PSUs
- Bootstrap replicates are drawn for each year.
  - Applies rescaled bootstrap for stratified multistage sampling
- Replicates are taken forward to mimic rotational panel design.
  - Split households can be considered for this step, `split=TRUE`

Calibrate Bootsrap Replicates

```
recalib(dat,hid="DB030",weights="RB050",
        b.rep=paste0("w",1:1000),year="RB010",
        country=NULL,conP.var=c("RB090"),
        conH.var=c("DB040","DB100"),...)
```

- Use output of `draw.bootstrap()` or
- Rectangular data set with household identifier and bootstrap replicates.
- Define households and/or personal variables to be calibrated onto
- Calibration with `ipu2()` from Package simPop

Estimate standard errors

```
calc.stError(dat,weights="RB050",b.weights=paste0("w",1:1000),
             year="RB010",var="HX080",fun="weightedRatio",
             cross_var=NULL,year.diff=NULL,year.mean=3,bias=FALSE,
             add.arg=NULL,size.limit=20,cv.limit=10,p=NULL)
```

- Use output of `recalib()` or rectangular data with bootstrap weights.
- Function fun is applied on Variable var for, using each bootstrap weight.
- Predefined functions available, also able to handle custom functions or functions from other packages
  - Must return double or integer and second argument is weight

Estimate standard errors

```
calc.stError(dat,weights="RB050",b.weights=paste0("w",1:1000),
             year="RB010",var="HX080",fun="weightedRatio",
             cross_var=NULL,year.diff=NULL,year.mean=3,bias=FALSE,
             add.arg=NULL,size.limit=20,cv.limit=10,p=NULL)
```

- Use output of `recalib()` or rectangular data with bootstrap weights.
- Function fun is applied on Variable var for, using each bootstrap weight.
- Predefined functions available, but custom functions or functions from other packages can be supplied
  - Must return double or integer and second argument is weight

Estimate standard errors

- Define subgroups of sample using `cross_var` (optional)
- Estimate standard errors for changes between years with `year.diff` (optional)
- Results of point estimates are averaged over `year.mean` years (optional)
  - Apply filter with equal filter weights over time series
- Estimate quantiles using parameter `p`.

Estimate standard errors

```
calc.stError(UDB_AT_calib,weights="RB050",
             year="RB010",b.weights=paste0("w",1:10),
             var="HX080",cross_var=list("DB040",c("RB090","DB040")))

## Calculated point estimates for variable(s)
##
##  HX080
##
## using function weightedRatio
##
## Results hold 448 point estimates for 9 years in 28 subgroups
##
## Estimted standard error exceeds 10 % of the the point estimate in 271 c
```

Estimate standard errors

```
# Apply function which is not in package 'surveysd'
# take the gini - index
library(laeken,quietly=TRUE)
# simulate income
set.seed(1234)
UDB_AT_calib[,income:=
               exp(rnorm(.N,mean=sample(7:10,1),sd=1)),
             by=list(DB100)]

# gini() returns list
# calc.stError needs function that returns double or integer
help_gini <- function(x,w){
  return(gini(x,w)$value)
}
```

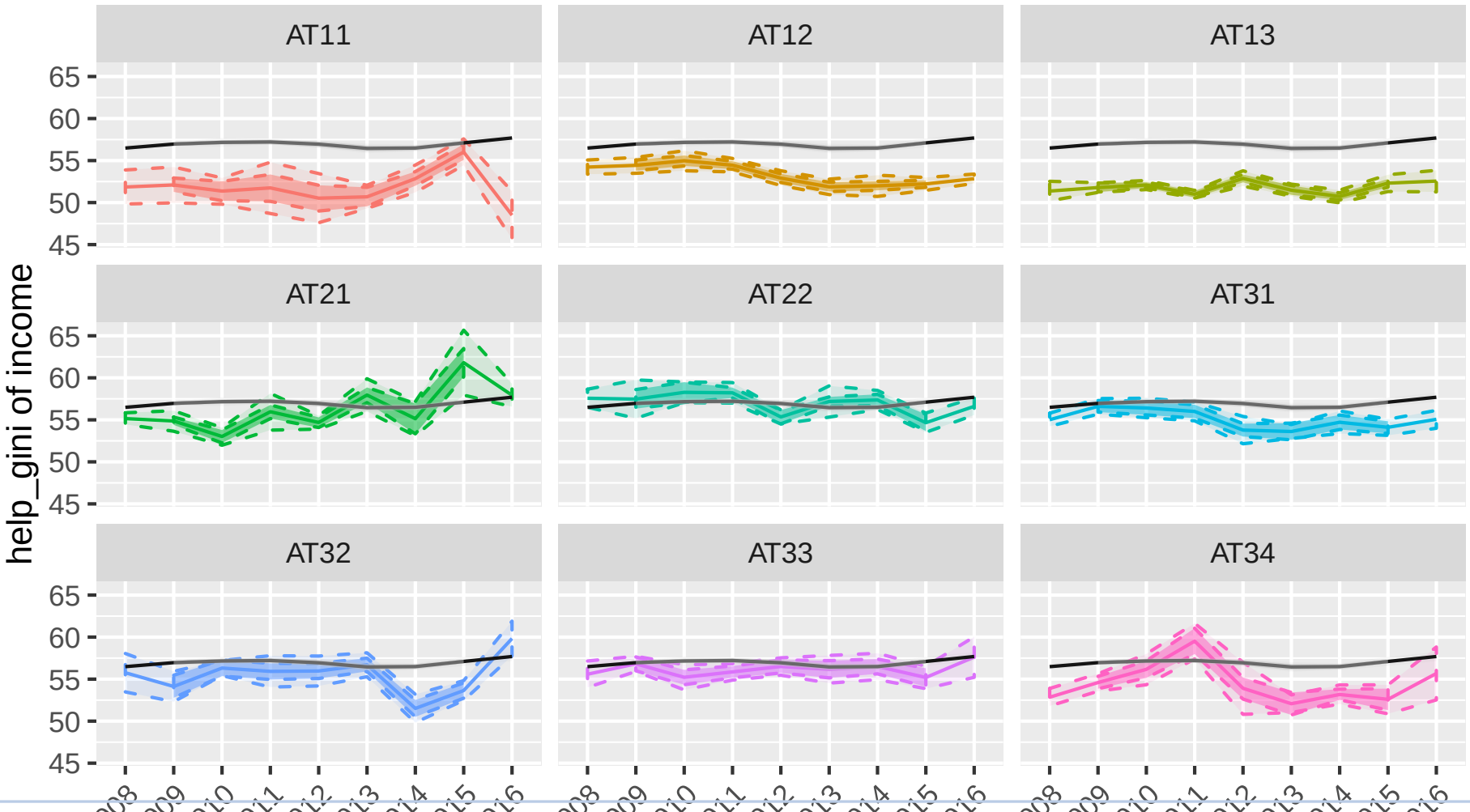
Estimate standard errors

```
calc.stError(UDB_AT_calib,fun="help_gini",
             weights="RB050",year="RB010",b.weights=paste0("w",1:10),
             var="income",cross_var=list("DB040",c("RB090","DB040")),
             year.diff=c("2014-2008"),p=c(.025,.975))

## Calculated point estimates for variable(s)
##
##  income
##
## using function help_gini from .GlobalEnv
##
## Results hold 504 point estimates for 9 years in 28 subgroups
##
## Estimted standard error exceeds 10 % of the the point estimate in 22 ca
```

Plot Method

```
plot(res_inc,type="grouping",
     groups="DB040",sd.type="ribbon")
```



Final Remarks

STATISTIK AUSTRIA
Die Informationsmanager

- Simple to use R-Package
- Supports a harmonious approach for estimating standard errors on surveysd with rotating panel design
  - Achieve more accuracy by averaging over multiple years
  - No need for administrative data or modelling assumptions
- Check it out on github: https://github.com/statistikat/surveysd