# Package 'migR'

May 15, 2019

**Type** Package

**Title** Migration and Homeowernship Dataprocessing (Oswald 2019)

**Version** 1.0

**Date** 2019-05-14

**Depends** R (>= 3.5.0)

**Description** This R package serves two purposes: First, to process raw data and summarise it so that it can be use as input to the structural model presented in Oswald (2019). Second, the package produces a set of stylized facts from the same raw data which are used to motivate the entire exercise. The package can be used as is, in the sense that I ship the final data products. Users interested in replicating the data acquisition and cleaning steps, need to follow further instructions to obtain SIPP raw data. This step is documented within the package.

**License** MIT + file LICENSE

**URL** https://github.com/floswald/migR

**BugReports** URL: https://github.com/floswald/migR/issues

**Imports** splines,
zoo,
copula,
grid,
quantmod,
stringr,
Hmisc,
xts,
survey,
texreg,
DBI,
data.table,
reshape2,
rjson,
xtable,
devtools,
ggplot2,
plyr,
MASS,
erer,
tikzDevice,
MultinomialCI

**Collate** 'migR-package.r'
'plots.r'
'CPS.r'
'plotjulia.R'
'RegionalPrices.r'
'SippAnalyse.r'
'SippPrepare.r'

**RoxygenNote** 6.1.0

# R topics documented:

## both_prices_output *Produce Table 6 in Main Text*

### Description

Collects R2 from all regressions in `reg_vs_state_y` and `reg_vs_state_p`

### Usage

```
both_prices_output()
```

## Clean.CPS *Clean CPS data*

### Description

Clean CPS data

### Usage

```
Clean.CPS(dta = "~/datasets/CPS/outdata/selected.dta")
```

### Details

cleans CPS downloaded from http://www.nber.org/data/current-population-survey-data.html I use the 2013 March supplement, documentation here http://www.nber.org/cps/cpsmar13.pdf

## Clean.Sipp *Clean Sipp Data*

### Description

take output from `Extract.wrap` and clean data. apply labels, account for missing vars. merge topical and core data. output two datasets, differing in time resolution (monthly or 4-monthly).

### Usage

```
Clean.Sipp(inpath = "~/Dropbox/research/mobility/data/SIPP",
  outpath = "~/git/migration/mig-pkg/data", TM.idx = list(p96 = c(3, 6,
  9, 12), p01 = c(3, 6, 9), p04 = c(3, 6), p08 = c(4, 7, 10)),
  agg.by = "age", verbose = TRUE)
```

### Arguments

| | |
|---|---|
| inpath | to output from Extract.wrap. These are called subsetxxxx.RData. |
| outpath | to save resulting dataset to disk. Object is called merged. |
| TM.idx | list with one index vector of Topic Module (TM) waves to use per panel. Name list elements like "p96" [panel 96] |
| agg.by | list of variable names by which to aggregate. those should be time variables present in the dataset like qtr, year, age etc |

### Details

Data is cleaned for inconsistencies across SIPP panels 1996-2008, merged with house price indices by state, and dollar denoted variables are deflated to 2012 as a base year using the US cpi. All dollar values are denoted in 1000s of US dollars. The SIPP can be cast at different time resolutions, i.e. you can look at monthly data quarterly data, annual, etc. you chose the level of aggregation by setting the argument agg.by

### Value

NULL. Saves 2 data.tables to dropbox.

---

## combine_BEA_fhfa *Get Macro Price Series*

---

### Description

Get Macro Price Series

### Usage

```
combine_BEA_fhfa()
```

### Details

This function writes out a dataset which contains the series q and p used in the paper for each Division. This function uses output from get_BEA_persincome and getFHFA_realPrices.

---

## correlograms *Regional Price Correlograms*

---

### Description

Regional Price Correlograms

### Usage

```
correlograms(path = "~/Dropbox/research/mobility/output/data/FHFA")
```

### Arguments

| | |
|---|---|
| path | file path to output figure |

### Details

Uses regional time series on q and p to illustrate the structure of the joint process (q,p) across regions.

### Value

Produces figure 2 in main text.

---

`CPS.distance` *Main Reason to Move: CPS data*

---

### Description

Main Reason to Move: CPS data

### Usage

```
CPS.distance(path = "~/Dropbox/mobility/output/data/cps")
```

### Arguments

| | |
|---|---|
| `path` | to output file |

### Details

Uses CPS data cleaned in `Clean.CPS` to produce a frequency table showing the main reasons to move.

### Value

produces table 2 in the main text.

---

`download.FHFA` *get FHFA state level HPI*

---

### Description

download FHFA state-level house price index 1990-2013, quarterly data.

### Usage

```
download.FHFA(from = "http://www.fhfa.gov/webfiles/25831/3q13hpists_expandeddata.txt",
 to = "~/git/migration/mig-pkg/data")
```

### Arguments

| | |
|---|---|
| `from` | url |
| `to` | location to save data |

## Details

source of data is http://www.fhfa.gov/Default.aspx?Page=87

## Value

TRUE

---

`Export.IncomeProcess` *Estimate and Export Individual Income Process*

---

## Description

Estimate and Export Individual Income Process

## Usage

```
Export.IncomeProcess(dat, writedisk, nocollege = FALSE,
  path = "~/Dropbox/research/mobility/output/model/fit")
```

## Arguments

| | |
|---|---|
| `dat` | A data.table of SIPP micro data |
| `writedisk` | boolean whether to save to disk |
| `nocollege` | boolean whether subsetting to no college degree or not. |
| `path` | string for graph output location |

## Details

Uses SIPP micro data to estimate an income process, which is used in the structural model to predict and simulate individual income. This is the implementation of equation (21) in the main text, further illustrated in online appendix C.1. This function is called from `Export.Julia`, so please refer for input arguments to that function. The procedure subsets sipp income data to leave out year 2007, which I found to be full of inconsistencies.

## Value

Implements equation (21) in main text, writes table C.1 in online appendix to disk and produces figure C.1 titled *Labor Income profiles for different q levels* also in online appendix.

## Export.Julia *Export Data to Julia Package*

### Description

Export Data to Julia Package

### Usage

```
Export.Julia(writedisk = TRUE, noCollege = TRUE)
```

### Arguments

| | |
|---|---|
| writedisk | boolean TRUE whether to save to disk |
| noCollege | boolean TRUE whether to subset data to individuals without college degree. |

### Details

Takes all output data from this R package (mainly: moments) and stores them on disk such that they can be used to run the structural julia model in `migration/mig/`.

## Export.VAR *Export Aggregate/Regional VAR Processes*

### Description

Export Aggregate/Regional VAR Processes

### Usage

```
Export.VAR(plotpath = "~/Dropbox/research/mobility/output/data/sipp",
  writedisk)
```

### Arguments

| | |
|---|---|
| plotpath | file path to output location |
| writedisk | boolean whether to save to disk |

### Value

produces

1. Table 5 in main text *Estimates for Aggregate VAR process*
2. Table B.4 in online appendix *Aggregate to Regional price mappings*
3. Figure 3 in main text *VAR fit to regional price data (p)*
4. Figure B.3 *VAR fit to regional productivity data (q)*
5. Figures B.1 and B.2 showing the raw data series, titled *Regional (p) and National (P) house price index* and *Regional (q) and National (Q) Labor Productivity index*
6. returns all parameters of the estimated models

## Extract.wrap *Extractor wrapper*

### Description

Selects variables from SIPP database and does some initial subsetting. This is an interface to ExtractorSippDB

### Usage

```
Extract.wrap(verbose = TRUE, which = paste0(c(1996, 2001, 2004, 2008)),
  dblocation = "~/datasets/SIPP/R",
  dropbox = "C:/Users/florian_o/Dropbox/mobility/data/SIPP")
```

### Arguments

| | |
|---|---|
| verbose | |
| which | names of datasets to extract. Names are: "1996", "2001", "2004", "2008", "Mig_2008" |
| dblocation | path to location of SIPP database, obtained as illustrated in https://github.com/floswald/asdfree/blob/master/SIPP/down1996.R. |
| dropbox | path to folder where to save this |

### Value

NULL saves subset data.tables into dropbox

## ExtractorSippDB *Extract data.tables from SIPP database*

### Description

Select variables and build data.tables from the SIPP database, downloaded and built with anthony damico's usgsd tools. selects ALL waves from coredata, but only selected waves from topical modules.

### Usage

```
ExtractorSippDB(dbfile, ck, which.core, which.tm, which.wgt, tk,
  subset = "", outfile, verbose, test = FALSE)
```

### Arguments

| | |
|---|---|
| dbfile | location of database |
| ck | string of variable names from core data to keep |
| which.core | numeric vector of which core waves to keep |
| which.tm | numeric vector of which topical modules to keep |
| which.wgt | character vector of name of weight tables |
| tk | list of character vectors of variable names from topical data to keep, one vector for each topical module |
| subset | SQL string for selecting from database |
| outfile | filename of where to save results |
| test | if TRUE extract only a short test dataset |

### Details

Notice that Damico's repo has since evolved. You can replicate what I did by using the code in a fork I created from his code, located at https://github.com/floswald/asdfree. In particular, you need to run the code in https://github.com/floswald/asdfree/blob/master/SIPP/down1996.R

To build data, don't use this function but the easier to use Extract.wrap

### References

https://github.com/floswald/asdfree,https://github.com/ajdamico/usgsd, http://www.asdfree.com/

---

## get.istate *auxiliary function to get movers origin and destination state in a data.table (AT THE BEGINNING OF CURRENT PERIOD, SAY)*

---

### Description

auxiliary function to get movers origin and destination state in a data.table (AT THE BEGINNING OF CURRENT PERIOD, SAY)

### Usage

```
get.istate(states, imove)
```

### Examples

```
ttab = data.table(pid = rep(c(1,2),each=5),state=c(3,3,4,4,4,6,7,7,8,9),istate=c(FALSE,FALSE,TRUE,FALSE,FALS
ttab[,c("from","to") := get.istate(states=state,imove=istate),with=FALSE]
```

## getFHFA_realPrices *Extend SIPP/FHFA data back to 1967*

### Description

Extend SIPP/FHFA data back to 1967

### Usage

```
getFHFA_realPrices()
```

### Details

This takes the mean house value by division in base year 2012 and uses several price indices to project this value backwards in time.

1. FHFA Division Index goes back until 1975
2. Use CPI to cover 1967-1975

## getHomeValues *get home values and adjust by inflation*

### Description

get home values and adjust by inflation

### Usage

```
getHomeValues(freq = "yearly")
```

## getMovers *Get Movers from full data*

### Description

Get Movers from full data

### Usage

```
getMovers(d)
```

## get_BEA_persincome *Get Personal Income Data from BEA*

### Description

Get Personal Income Data from BEA

### Usage

```
get_BEA_persincome()
```

### Details

this function depends on R package EconData available at https://github.com/floswald/EconData

## merge.idx *Merge SIPP cores and topical modules auxiliary function*

### Description

merges the most recent wealth module onto the corresponding core data. i.e. if breaks=c(3,6,9), the TM was asked in waves 3,6 and 9. therefore merge TM_3 onto cores 1-3, merge TM_6 onto 4-6, etc

### Usage

```
## S3 method for class 'idx'
merge(core, topic, breaks = c(3, 6, 9, 12),
  topic.names = NULL)
```

### Arguments

| | |
|---|---|
| core | list of core datasets |
| topic | list of topical datasets |
| breaks | numeric vector of waves where a TM was asked. |
| topic.names | NULL by default assumes names of topic are like "TM_2". if not, supply names here. |

### Examples

```
co <- lapply(1:12, function(x) data.table(ssuid=1:4,covar=rnorm(4),key="ssuid"))
br <- c(2,5,9,12)
tm <- lapply(1:5, function(x) data.table(ssuid=1:4,tmvar=10*c(1,br)[x] + sample(1:4,size=4),key="ssuid"))
names(tm) <- paste0("TM_",c(1,br)) # don't merge first TM
merge.idx(core=co,topic=tm,breaks=br)
```

migR *migR*

## Description

R companion package for *The Effect of Homeownership on the Option Value of Regional Migration* forthcoming in Quantitative Econmics, www.qeconomics.org.

## Author(s)

Florian Oswald <Florian.oswald@gmail.com>

multiplot *Multiple plot function ggplot objects can be passed in ..., or to plotlist (as a list of ggplot objects)*

## Description

Multiple plot function ggplot objects can be passed in ..., or to plotlist (as a list of ggplot objects)

## Usage

```
multiplot(..., plotlist = NULL, cols = 1, layout = NULL)
```

## Arguments

| | |
|---|---|
| plotlist | list of ggplots |
| cols | Number of columns in layout |
| layout | matrix specifying the layout. If present, 'cols' is ignored. If the layout is something like matrix(c(1,2,3,3), nrow=2, byrow=TRUE), then plot 1 will go in the upper left, 2 will go in the upper right, and 3 will go all the way across the bottom. |

## Author(s)

http://www.cookbook-r.com/Graphs/Multiple_graphs_on_one_page_(ggplot2)/

## PlotSippMigrationRates
*Plot Sipp Migration Rates by Age and Ownership Status*

### Description

Plot Sipp Migration Rates by Age and Ownership Status

### Usage

```
PlotSippMigrationRates(nocollege = FALSE)
```

### Arguments

| | |
|---|---|
| nocollege | Boolean whether to subset to no college population. |

### Details

Generates a scatter plot of mobility vs age and a plot showing ownership by age.

### Value

Figure 1 in the main text.

## PlotSippTransitionMatrix
*Plot Sipp Transition matrix*

### Description

Plot Sipp Transition matrix

### Usage

```
PlotSippTransitionMatrix(ttable,
  path = "~/Dropbox/mobility/output/data/sipp")
```

### Examples

```
## Not run:
load('~/Dropbox/mobility/SIPP/Sipp_aggby_age.RData')
tt <- merged[from!=to,table(from,to)]
PlotSippTransitionMatrix(tt)

## End(Not run)
```

## plot_moment_fit *Plot Data vs Model Moments*

### Description

Plot Data vs Model Moments

### Usage

```
plot_moment_fit(path = "~/Dropbox/research/mobility/output/model/fit")
```

### Arguments

| | |
|---|---|
| path | location where julia objective function saves moments in file moms.json |

### Details

takes structural model-generated moments and compares them to data moments in a scatter plot.

### Value

produces

1. Figure 4 in main text *Graphical device to show model fit*
2. Figure D.1 in appendix *Auxiliary Models* and *Wealth*

## reg_vs_state_p *Compare State vs Division Level p Indices*

### Description

Compare State vs Division Level p Indices

### Usage

```
reg_vs_state_p()
```

### Details

produces table B.6 in online appendix

## reg_vs_state_y *Compare State vs Division Level q Indices*

### Description

Compare State vs Division Level q Indices

### Usage

```
reg_vs_state_y()
```

### Details

produces table B.7 in online appendix

## Sipp.moments *Calculate Moments from SIPP data*

### Description

Calculate Moments from SIPP data

### Usage

```
Sipp.moments(d, svy, ages = c(20, 50))
```

### Arguments

| | |
|---|---|
| d | a SIPP data.table |
| svy | a SIPP dataset in survey format. Can be obtained by calling SippSvyDesign |

### Details

Computes moments form SIPP data to be used in structural model SMM estimation routine

### Value

a datatable with moment names, values and standard deviations

## Sipp.own_in_j_rent_in_k *Fraction of non-resident landlords*

### Description

How many owners live in region k while renting out a flat in region d?

### Usage

```
Sipp.own_in_j_rent_in_k(path = "~/Dropbox/mobility/output/data/sipp")
```

### Arguments

| | |
|---|---|
| path | to save graphs |

## Sipp.SumStats *SIPP Summary Statistics*

### Description

SIPP Summary Statistics

### Usage

```
Sipp.SumStats(path = "~/Dropbox/research/mobility/output/data/sipp",
  nocollege = FALSE)
```

### Arguments

| | |
|---|---|
| path | to save output |
| nocollege | boolean whether to subset stats to no college degree. |

### Details

Produces summary statistics from SIPP micro dataset

### Value

writes table 3 in main text to disk (*Annual moving rate in percent of the population*).

## Sipp.wage_residual_copulas *Estimation of Movers' z Copula*

### Description

Estimation of Movers' z Copula

### Usage

```
Sipp.wage_residual_copulas(path = "~/Dropbox/research/mobility/output/data/sipp")
```

### Arguments

| | |
|---|---|
| path | to save graphs |

### Details

Implements the estimation of movers' z transition via a normal copula with SIPP micro data. This is the implementation described in section C.1 of the online appendix.

### Value

1. Table C.2 in online appendix reporting estimated Copula parameters.
2. Figure C.2 illustrating the marginal distributions of z in periods before and after move
3. Figure C.3 is produced in the julia package, not here.

## SippProbitMove *Probit: Determinants of Cross Division Moves*

### Description

Probit: Determinants of Cross Division Moves

### Usage

```
SippProbitMove(d, path = "~/Dropbox/mobility/output/data/sipp")
```

### Arguments

| | |
|---|---|
| d | a svy object. Can be obtained by calling SippSvyDesign. |
| path | to save output |

### Details

Uses SIPP micro data to estimate a probit model relating observables to a indicator of whether a move took place in period t.

### Value

produces table 4 in main text: *Determinants of cross census division moves in SIPP data*.

## SippSvyDesign *Create SIPP survey Design object*

### Description

Create SIPP survey Design object

### Usage

```
SippSvyDesign(merged = NULL)
```

### Details

Uses variable `HHweight` to build a `svydesign` object.

## Var.impulse *Illustrate translation of Aggregate to Regional Shocks*

### Description

Illustrate translation of Aggregate to Regional Shocks

### Usage

```
VAR.impulse(plotpath = "~/Dropbox/research/mobility/output/data/sipp")
```

### Arguments

| | |
|---|---|
| `plotpath` | location to output figure. |

### Details

This produces figure B.4 in the online appendix titled "10 percent shock to Y"

# Index