

# Compositional Impact of Migration (CMI)

*Francisco Rowe, Nikos Patias, Jorge Rodríguez-Vignoli*

This R package produces summary statistical indicators of the impact of migration on the socio-demographic composition of an area. Three measures can be used: ratios, percentages and the Duncan index of dissimilarity. The input data files are assumed to be in an origin-destination matrix format, with each cell representing a flow count between an origin and a destination area. Columns are expected to represent origins, and rows are expected to represent destinations. The first row and column are assumed to contain labels for each area. See Rodríguez-Vignoli and Rowe (2018) for technical details.

## Getting Started

These instructions will get you the CMI package running on your local machine and provide an example on how to interpret the various output indicators.

### Prerequisites

This package has no pre-requisites.

### Installing

To install **{CIM}** from CRAN, type:

```
#install.packages("CIM")
```

To install **{CIM}** from Github, type:

```
#devtools::install_github("fcorowe/cim")
```

Load **{CIM}**

```
library(CIM)
```

### Use

We present two examples.

#### Example 1: Sex ratio

First, we use the package to quantify the impact of internal migration on the sex ratio of the Greater Santiago in Chile drawing on 2008-2013 transition data from the 2013 CASEN survey. For simplicity, data for this example are aggregated into 3 broad areas.

Read input data

```
m <- male  
f <- female
```

Display male input data

```
m
```

```
## Greater.Santiago
## Greater Santiago 2542597
## Rest of the Metropolitan region 20364
## Rest of the country 66038
## Rest.of.the.Metropolitan.region
## Greater Santiago 6313
## Rest of the Metropolitan region 350989
## Rest of the country 2818
## Rest.of.the.country
## Greater Santiago 72591
## Rest of the Metropolitan region 7143
## Rest of the country 4381084
```

**NOTE:** The required input data must be in an origin-destination matrix, with origins as columns.

Compute and print the CIM outputs

```
CIM.ratio <- CIM(m, f, calculation = "ratio", numerator = 1, denominator = 2)
CIM.ratio
```

```
## $num_results
## FV CFV CIM CIM_PC
## Greater Santiago 87.42172 87.02433 0.3973829 0.4566343
## Rest of the Metropolitan region 94.05847 92.91429 1.1441815 1.2314376
## Rest of the country 90.16837 90.52613 -0.3577543 -0.3951945
## totalCol 89.36814 89.36814 0.0000000 0.0000000
## DIAG CIM_I CIM_O
## Greater Santiago 86.83030 0.591414697 -0.1940318
## Rest of the Metropolitan region 93.76409 0.294381660 0.8497999
## Rest of the country 90.17376 -0.005385734 -0.3523685
## totalCol 89.36814 0.000000000 0.0000000
## CIM_I_PC CIM_O_PC
## Greater Santiago 148.827403 -48.82740
## Rest of the Metropolitan region 25.728580 74.27142
## Rest of the country 1.505428 98.49457
## totalCol 0.000000 0.00000
```

## Interpretation

Interpreting the results from the table above:

Factual Value (FV) indicates the sex ratio at the end of the time interval (i.e. 2013).

Counterfactual Value (CFV) indicates the sex ratio at the start of the time interval (i.e. 2008). Alternatively, it can be interpreted as the counterfactual sex ratio; that is, what would have been the sex ratio if no migration had occurred.

Compositional Impact of Migration (CIM) is the difference between the FV and CFV and indicates the change in the area-specific sex ratio because of migration. The results indicate that internal migration contribute to increase the sex ratio of the Greater Santiago by 0.4.

CIM\_PC is the CIM divided by the CFV and indicates the percentage change of the CMI i.e. the percentage change in the sex ratio. The results indicate that internal migration contributed to increase the sex ratio of the Greater Santiago by 0.46% between 2008 and 2013.

DIAG corresponds to the diagonal of the origin-destination matrix and indicates the sex ratio of the no-migrant population. The results indicate that the sex ratio relating to those staying in the Greater Santiago was 86.83.

CIM\_I represents the change in the CMI due to migration inflows.

CIM\_O represents the change in the CMI due to migration outflows.

$CIM\_I\_PC = (CIM\_I / CIM) * 100$

$CIM\_O\_PC = (CIM\_O / CIM) * 100$

**NOTE:**  $CIM = CIM\_I + CIM\_O$

Taken together, the CIM\_I\_PC and CIM\_O\_PC tell us their respective contribution to changes in the CIM i.e. if changes in the CMI were due to migration inflows, migration outflows or both, and the extent of these influences. The results tell us that while migration inflows contributed to increase the sex ratio in the Greater Santiago by 148.83%, migration outflows operated to reduce it by 48.83%. Thus, in absence of migration outflows, migration would have increased the sex ratio by some additional 0.19.

## Example 2: Duncan index

Next, we measure the impact of internal migration on residential age segregation in the Greater London Metropolitan Area, England, drawing on one-year migration data by age bands (i.e. 1-14, 15-29, 30-34, 45-64 and 65+) at the local authority level, 2011 UK Censuses. Local authorities comprising outside the Greater London Metropolitan Area are collapsed into a single area, labelled “the Rest of the UK”. We use the same approach employed by Rodríguez-Vignoli and Rowe (2017) to measure the impact of internal migration on residential educational segregation in the Greater Santiago, Chile.

Compute and print the CIM outputs

```
CIM.duncan <- CIM(pop65over, pop1_14, pop15_29, pop30_44, pop45_64, calculation = "duncan", numerator =  
CIM.duncan$duncan_index
```

```
## [1] 0.01624249
```

The CIM for the Duncan index of dissimilarity indicates that internal migration has contributed to increase age segregation of the population aged 65 and over in the Greater London Metropolitan Area by 2.81% between 2010 and 2011 i.e. from 16.2% in 2010 to 19% in 2011.

To visualise where the population aged 65 and over in the Greater London Metropolitan Area is concentrating, we can map differences in the spatial distribution of this population across local authority districts.

First install and load the needed packages

```
#install.packages(c("rgdal", "dplyr", "tmap"))  
library("rgdal")
```

```
## Loading required package: sp
```

```
## rgdal: version: 1.3-6, (SVN revision 773)
```

```
## Geospatial Data Abstraction Library extensions to R successfully loaded
```

```
## Loaded GDAL runtime: GDAL 2.1.3, released 2017/20/01
```

```
## Path to GDAL shared files: /Library/Frameworks/R.framework/Versions/3.5/Resources/library/rgdal/gdal
```

```
## GDAL binary built with GEOS: FALSE
```

```
## Loaded PROJ.4 runtime: Rel. 4.9.3, 15 August 2016, [PJ_VERSION: 493]
```

```
## Path to PROJ.4 shared files: /Library/Frameworks/R.framework/Versions/3.5/Resources/library/rgdal/proj
```

```
## Linking to sp version: 1.3-1
```

```
library("dplyr")
```

```
##
```

```
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##   filter, lag
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

```
library("tmap")
```

NOTE: Download a shapefile containing the Greater London Local Authority Districts from the shapefile folder at: <https://github.com/fcorowe/cim>

NOTE: The Local Authority Districts for the City of London and Westminster in our shapefile are combined to make our shapefile consistent with our migration data.

Read the shapefile.

```
setwd("/Users/Franciscorowe/Dropbox/Francisco/research/in_progress/r package/shapefile")
Greater_London <- readOGR(dsn = ".", layer = "Greater_London_districts", stringsAsFactors = FALSE)
```

```
## OGR data source with driver: ESRI Shapefile
## Source: "/Users/Franciscorowe/Dropbox/Francisco/research/in_progress/r package/shapefile", layer: "G
## with 32 features
## It has 3 fields
```

Plot the shapefile

```
plot(Greater_London)
```



Obtain the differences in the spatial distribution of the population aged 65 and over across local authority districts using the CIM.Duncan function:

```
CIM.duncan <- CIM(pop65over, pop1_14, pop15_29, pop30_44, pop45_64, calculation = "duncan", numerator =
Dun_65over <- CIM.duncan$duncan_results
```

Visualise the results

```
head(Dun_65over)
```

```
##           ASFVShare_cg ASCFVShare_cg ASFVShare_ref
## Barking and Dagenham  0.001361764   0.002104807   0.002709112
## Barnet                0.004891941   0.005458619   0.006253946
## Bexley                0.002353451   0.002966390   0.002625476
```

```
## Brent          0.002338995  0.003478135  0.005567036
## Bromley        0.004154680  0.005204192  0.004251649
## Camden         0.002382364  0.002870979  0.005701721
##               ASFCVShare_ref ASShareFV_diff ASShareCFV_diff
## Barking and Dagenham 0.002839302  1.347349e-03  0.0007344958
## Barnet           0.006453164  1.362005e-03  0.0009945449
## Bexley           0.002777017  2.720244e-04  0.0001893730
## Brent            0.006133709  3.228041e-03  0.0026555736
## Bromley          0.004306069  9.696933e-05  0.0008981229
## Camden           0.005854386  3.319358e-03  0.0029834063
```

Append these data to the shapefile using the local authority names as joiner

```
Duncan_65p <- merge(Greater_London, Dun_65over, by.x = "name", by.y = 0)
head(Duncan_65p@data)
```

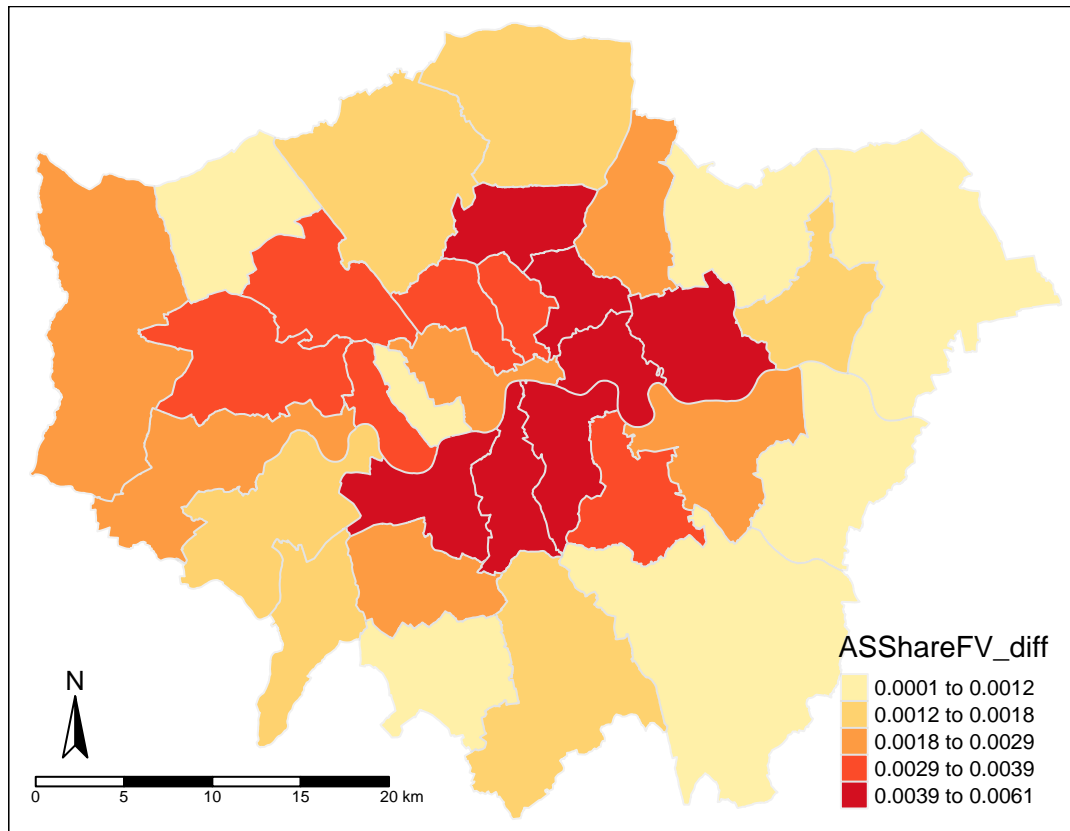
```
##               name label ons_label ASFVShare_cg ASFCVShare_cg
## 5              Bromley 02AF      00AF  0.004154680  0.005204192
## 27 Richmond upon Thames 02BD      00BD  0.002431514  0.003053126
## 17             Hillingdon 02AS      00AS  0.002726419  0.003501265
## 16             Havering  02AR      00AR  0.003021323  0.003298880
## 21 Kingston upon Thames 02AX      00AX  0.001679798  0.002237803
## 29              Sutton  02BF      00BF  0.002237803  0.002674377
##               ASFVShare_ref ASFCVShare_ref ASShareFV_diff ASShareCFV_diff
## 5      0.004251649  0.004306069  9.696933e-05  8.981229e-04
## 27      0.003643078  0.003640028  1.211564e-03  5.869023e-04
## 17      0.004594703  0.004611559  1.868285e-03  1.110294e-03
## 16      0.002584059  0.002668819  4.372638e-04  6.300606e-04
## 21      0.003456221  0.003301309  1.776423e-03  1.063506e-03
## 29      0.002454350  0.002632218  2.165477e-04  4.215847e-05
```

Set to a static map view and create a map using tmap

```
tmap_mode('plot')
```

```
## tmap mode set to plotting
```

```
tm_shape(Duncan_65p) +
  tm_polygons("ASShareFV_diff", style="quantile",border.alpha = 0.1, palette = "YlOrRd",
             title="ASShareFV_diff")+
  tm_compass(position = c("left", "bottom")) +
  tm_scale_bar(position = c("left", "bottom"))
```



Or, even better we can create an interactive map! by setting an interactive map view

```
#tmap_mode('view')
#tm_shape(Duncan_65p) +
#  tm_polygons("ASShareFV_diff", style="quantile",border.alpha = 0.1, palette = "YlOrRd",
#              title="ASShareFV_diff")+
#  tm_compass(position = c("left", "bottom")) +
#  tm_scale_bar(position = c("left", "bottom"))
```

## Citation

How to cite if you use the package:

Rowe, F., Patias, N., Rodríguez-Vignoli, J. (2018). CIM: Compositional Impact of Migration. R package version 1.0.0.

If you use the method:

Rodríguez-Vignoli, J. and Rowe, F., 2018. How is internal migration reshaping metropolitan populations in Latin America? A new method and new evidence. Population studies, 72(2), pp.253-273.

## License

This project is licensed under the MIT License - see the LICENSE.md file for details

## References

Rodríguez-Vignoli, J.R. and Rowe, F., 2017. The Changing Impacts of Internal Migration on Residential Socio-Economic Segregation in the Greater Santiago. 28th International Population Conference of the International Union for the Scientific Study of Population (IUSSP), Cape Town, South Africa.

Rodríguez-Vignoli, J. and Rowe, F., 2018. How is internal migration reshaping metropolitan populations in Latin America? A new method and new evidence. *Population studies*, 72(2), pp.253-273.