

Max Planck Institute for Demographic Research

MP116 Microsimulation with Socsim

Carl Mason
Emilio Zagheni
UC Berkeley, Demography

June 15 & 16, 2010

1 Project 4: Sandwich generation

In this project we will use Socsim to investigate the effect of the combination of declining fertility and increasing longevity on generational overlap. An attempt is made to use actual realistic rates – at least until the present time – so the results we find this time, might be relevant.

The goals of this adventure are:

- To gain experience with a “realistic” simulation.
- To learn about the likelihood of having small children and crumbling parents at the same time.
- To generate new and interesting questions about generational overlap that might be addressed today, tomorrow and beyond.

1.1 Setup

In order to do the complete exercise, it will be necessary to have a running recent version of Socsim and a recent version of R, a software environment for statistical computing and graphics. R is free and wonderful. Visit <http://www.r-project.org/> if you do not already have R.

See <http://lab.demog.berkeley.edu/socsim> if you still need to get Socsim running. The tar file referenced below does contain a sample of Socsim output so if you would like to learn about Socsim without actually running it, you can just do the R part of the exercise.

All of the Socsim files for this exercise are available in a directory called **Sandwich** which you can download as a compressed tar file from <http://lab.demog.berkeley.edu/socsim/Sandwich.tgz>.

The **Sandwich** directory contains a **usa.sup** file which references several rate files stored in the **HMDdata**, **HFDDdata** and **UNdata** directories. The **Rcode** directory contains a file called **sandwich.r** which contains enough R code to get you started.

The **Sandwich** directory layout is shown in Figure 1. Note that the **HFDDdata** and **HMDdata** directories are symbolic links to directories from the **MarriageSqueeze** example. In order for the **Sandwich** example to work, you must also download and unpack the **MarriageSqueeze.tgz** file such that **Sandwich** and **MarriageSqueeze** have the same parent directory. No doubt you have already done this. In operating systems which do not understand symbolic links, it may be necessary to make copies of the **HFDDdata** and **HMDdata** directories.

```
|-- GretchenRates
|-- HFDDdata -> ../MarriageSqueeze/HFDDdata/
|-- HMDdata -> ../MarriageSqueeze/HMDdata/
|-- Rcode
|-- SimResults
'-- UNdata
    |-- Austria
    |-- CzechRepublic
    |-- Netherlands
    |-- Republic
    |-- Russia
    |-- Slovakia
    |-- Sweden
    |-- Switzerland
    '-- USA
```

Figure 1: Sandwich directory structure

1.2 Socsim instructions

Have a look at the `usa.sup` file and notice the following things:

This is a long simulation Beginning in the misty past (1800) and leading well into the misty future (2145). It is necessary to run the simulation this long because we are interested in looking at 3 and at times even 4 generations over time. The length of the simulation, combined with our natural impatience requires that we start with a fairly small population (1218). This produces a final population of around 50,000 mostly dead. This size population is easy to work with in R, but we may get more stochastic variation than we like as a result.

As presently configured, the first segment runs for only 100 years. This could be enough to produce a stable age structure, but as an extension, one might increase this duration to increase the size of the final population and reduce stochastic variation.

Mortality and fertility rates The rates for the initial segments come from Gretchen Donehower’s excellent dissertation. The rates for that segment are a compromise between the 1900 and 1910 rates and are “guaranteed” to produce, after 100 years, a population that is close in terms of the age and sex distribution of the 1900 US population – close enough for science anyway. See the README file in **GretchenRates** for more details and other overlooked problems. Segments covering years prior to 2000 also come from Gretchen’s work. Segments for 2000-2008 use rates from the Human Mortality Database and the Human Fertility Database, just as in the Marriage Squeeze example. After 2008, we shift to rates that reflect the UN “Medium” projection for both mortality and fertility. The “high” and “low” projections of fertility are also available in Socsim-ready form. Rates for all of the HFD countries are also readily usable. Running the simulation for other countries is obvious extension, however, to do real science, one would have to tune those rates as Gretchen did for the US.

Marriage rates For the simulation up to 2000 we use the marriage rates developed by Gretchen Donehower. For the segments that extend into the future, we continue with the marriage rates from 1900-2000 segment. However, since fertility rates after 2000 are not marital status specific, we should be cautious about drawing conclusions that rely on marriage behavior after 2000.

A lot of good questions can be addressed in terms of mothers and maternal (great)grandmothers which will be unaffected by marriage.

1.3 Run the simulation

To run the simulation cd into the **Sandwich** directory (that you produced by unpacking the Sandwich.tgz file).

then type:

```
@:> pathtoSocsim usa.sup 13531
```

where the “13531” is an arbitrary number that serves as random number seed.

/path/to/socsim means the full filename of the executable that you compiled earlier.

If all goes well a bunch of now familiar stuff will scroll onto the screen indicating how things are going with the simulation. Output will be stored in the **SimResults** directory.

1.4 Analyzing simulation output

In the sub directory **Rcode** you’ll find a file called **sandwich.r** That file contains R instructions for reading the simulation output (the *.opop* file) and for generating some graphs relating to the the questions of generational overlap. The R code in **sandwich.r** file is a bit more involved than in previous examples. It is also intended to be more open ended. As you work through it, look for new questions that can be addressed. Remember, there is no handout for tomorrow, so if you want to have something to do between 1 and 3PM (besides watching the World Cup) you’ll need to come up with an idea yourself.