GROUP/GENDER DIFFERENTIATION STUDY

Manual to run simulations

STARTING WITH THE MODEL

This model is based on the Leviathan model (Deffuant, Carletti and Huet, 2013: http://jasss.soc.surrey.ac.uk/16/1/5.html). It simplifies this model according to the work of (Huet and Deffuant, 2017: https://link.springer.com/chapter/10.1007/978-3-319-47253-9_13) , and considers, on the contrary to this preliminary studies, that there are two groups of agents in the population. This "2-groups" work, mainly aiming to understand the gender differentiation process, is presented further in a paper submitted to Nature Human Behaviour.

RUNNING THE MODEL

Since you have downloaded and unzipped the zip file DifferentiationStudy, you can read this manual. To launch the model, you need java installed on your computer. This is very often already the case. If not, install java or ask for help to the proper person. If you have some doubt, you can try the execution procedure as follows.

Then, you can execute the model by clicking on:

- Launch.bat if you work with Windows
- Launch.sh if you work with Mac OS or Linux

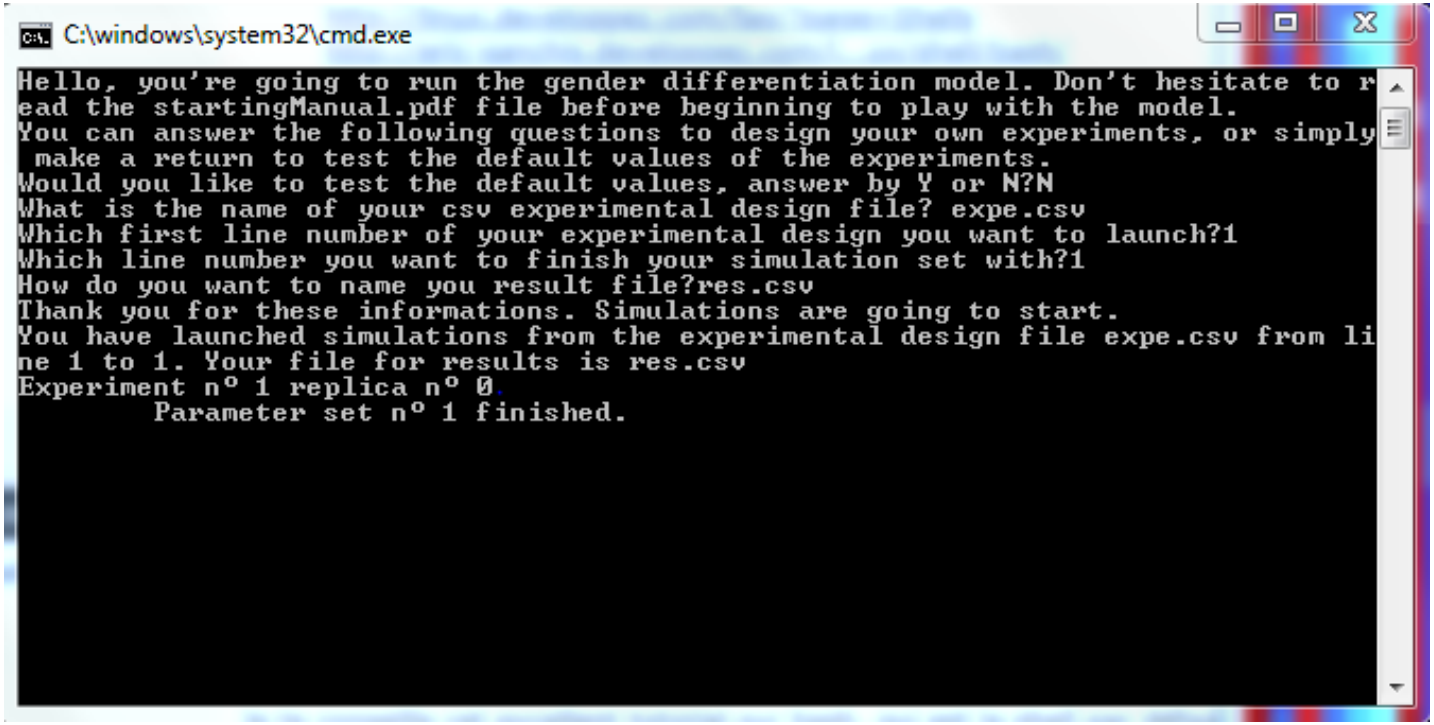
A console appears, similar to the one shown in the following. It guides you in your choice for running the model.

The launched test by default is based on :

- the experimental design file called expeLotDifferentiationModel.csv, available in the zip file DifferentiationStudy
- the beginning line number for the launched simulations is 1 and the final line is 273. These numbers correspond to the first line described an experiment and to the last line describing an experimental in the file expeLotDifferentiationModel.csv, but you can choose to run only a part of the file by indicating your number in your own design of simulations
- the file for saving the result is called resultsFile.csv by default.

Pay attention, if you have not chosen your own name for the resultsFile.csv but want to keep your previous results, you should rename the file containing your previous results because it is going to be deleted by the a new execution of the program if you have not changed it.

If you want to design your own experiment, you can answer N to the related question in the console. Thus you have to answer question related to the file names and the lines you want to test. Don't forget to precise the extension .csv in the name of the file.



THE EXPERIMENTAL DESIGN FILE

The basic file to test the model is named expeLotDifferentiationModel.csv. You can change it to design you own experiment keeping the same file name, or naming it differently (but you have to indicate the file to run when you launch the model - see "Running the model"). The file contains one title line indicating the information you should give to run the model and simulate the evolution of two groups of agents. Then each following line designs a parameter set for a simulation (which can be replicated). The parameter set is described by:

- visu: 1 means you want a visualization of the simulation running (see details further), 0 means you do not want to visualize but just to write down the results in the results file called basically resultsFile.csv (see details further for what is written in). 0 is the standard mode to choose if you want to make many replicated experiments.
- nbReplicats: number of times you want to replicate the parameter set
- nbStep: number of steps of a simulation
- sizePop: total size of the population of agents. The study is typically realized with 40 agents. You can use less or more but pay attention that the computation time increases with the size of the population, even more when you use the visualization. Thus, if you want results before being retired, it is better not to choose a size higher than 100 agents (:-).
- ka: parameter *k* of the model indicating how many people one agent talks about when he/she meets another agent
- freqSavings: frequency to write the results in the results file, in number of steps. Example : if you put 1000, the variables indicating the model state will be written in the file every 1000 steps
- sizes of the groups : the sizes of each of the two groups you want to simulate (please pay attention to the fact the sum of agents of groups should be equal to the total size of the population indicated previously)
- sigma of the groups : sigma of each of the two groups you simulate. Sigma is a float having the form 0.2. Interesting values to test can be chosen between 0.1 to 1.3 for example. It defines how strongly one uses the esteem he/she has for an agent to decide about the influence this agent has on him/her. Typically, the larger the sigma of an agent, the lower the difference of influence this agent gives to someone he/she hold in higher esteem compared to someone held in lower esteem. Details can be found in the papers related to the model (see section "Starting with the model")/
- delta of the groups : delta of each of the two groups you simulate. Delta is a float having the form 0.2. Interesting values to test can be chosen between 0.1 to 0.4 for example. Delta defines the level of noise in the communication between two agents. It is typically small compared to the length of the opinion space (equal to 2) and in the study we describe here maintained equal for the two groups.
- pIntergroupContat: this is the probability (i.e. a float having the form 0.15 for example) that someone discuss with an agent from his/her own group. Then, 0.5 corresponds to an equal probability to meet someone from his/her own group or from another group; 1 means agents meet only people from their own groups; 0 means agents meet only people from the other group.

An example of such file with 6 designed experiments, all with two groups, is given here (a common separates each field) :

```
visu,nbReplicats,nbStep,sizePop,ka,freqSavings,size G0,size G1,sigma G0,sigma
G1,deltaG0,deltaG1,pIntergroupContact
1,1,1500000,40,3,1000,20,20,0.2,0.3,0.2,0.2,0.5
0,2,500000,40,3,100000,5,35,1,1.1,0.2,0.2,0.5
0,2,500000,40,3,100000,5,35,2,2.1,0.2,0.2,0.5
0,2,500000,40,3,100000,5,35,3,3.1,0.2,0.2,0.5
0,2,1000000,40,3,100000,5,35,0.03,0.13,0.2,0.2,0.5
0,2,1000000,40,3,100000,5,35,0.05,0.15,0.2,0.2,0.5
```

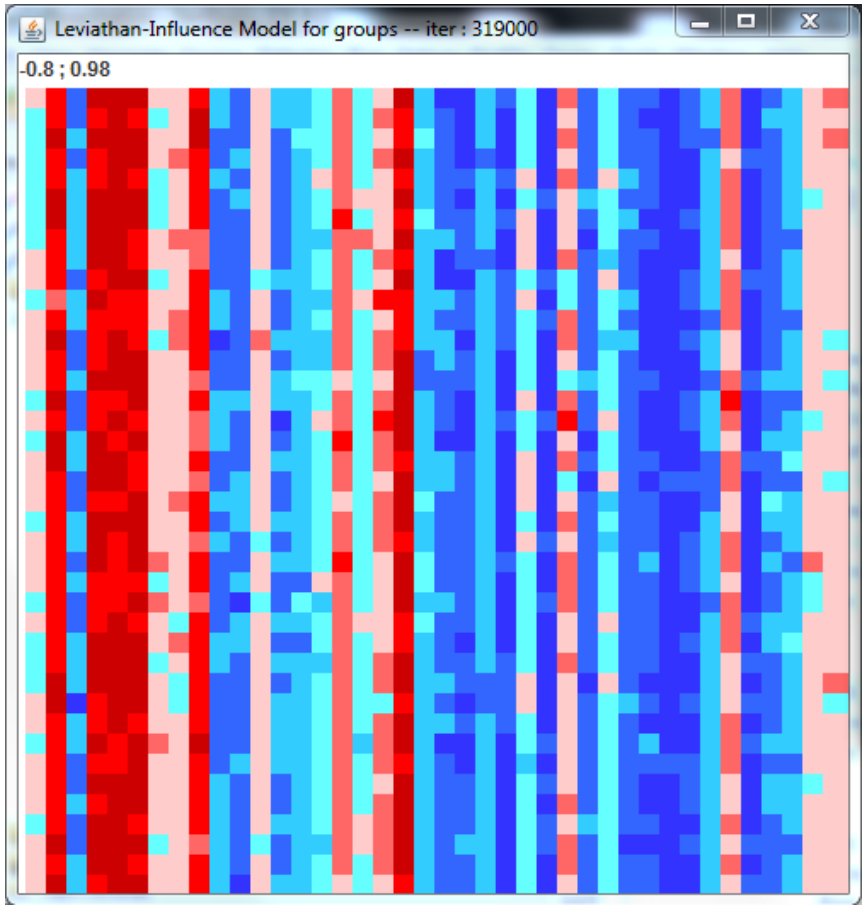
VISUALISATION OF THE SIMULATION

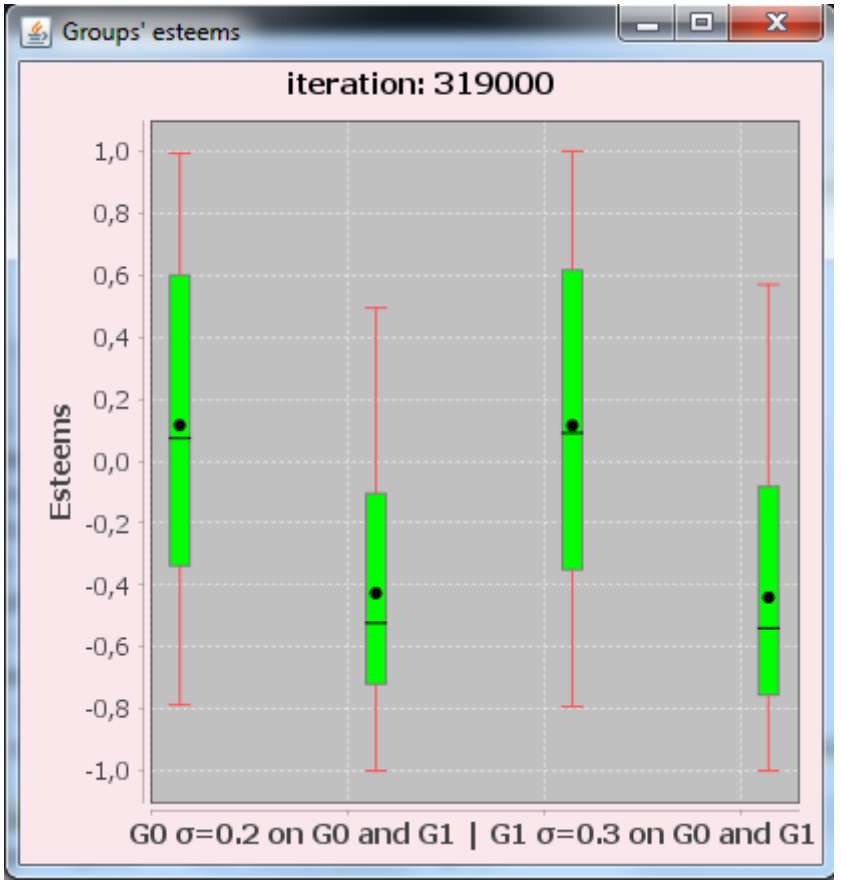
Two different visualizations appear when you run the model with the parameter "visu" set at 1 (see the following figures). The parameter used to run the illustrated visualization which follows is the first line of the experimental design file given in example in the section "The experimental design file". During a simulation, the visualization evolves to show all the states of the population at each *freqSavings* iterations (*freqSavings* is a parameter you have indicated in the experimental design files - see the previous paragraph for details).

The first one represents the esteem (or opinion) every agent has about agents (left figure). A square represents the esteem of an agent for another agent, or him/herself. A line represents the set of esteem one agent has for agent, included him/herself. A column represents the sets of esteem that every agent has for one particular agent. The reputation of one agent is the average of esteem of one column, excluded the esteem the agent has for him/herself. The esteem goes from -1 to +1. The dark blue color stands for -1 while the dark red color stands for +1. Starting from -1, the increase of esteem implies a lighter and lighter blue, and then a darker and darker red after the esteem has reached the value 0. Accordingly, the esteem 0 is represented by a blank square. During a simulation, the visualization evolves to show all the esteems of the model for every *freqSavings* you have indicated in the experimental design files (see the previous paragraph for details). You can observe on the following figure that the 20 agents located on the left (those having a lower sigma value) are held in higher esteem in red on average, than the agents located on the right part of the figure (those having a larger sigma value) who are held in a negative esteem on average. This difference can be observed in a clearer way on the right figure.

The second visualization, the right figure, represents the esteems at the group level. These are boxplots with the green box representing data from the first to the third quartile, the red lines the minimum and the maximum of the data excluding the outliers, the small line in the green box representing the median while the black dot represents the mean. Small dots located close to the extrema values correspond to outliers (outliers are very extreme singular values - there are none in the presented example). Overall the right figure show, from the left to the right:

- Esteems of the group G0 (called S in the paper) having the smaller sigma (here 0.2) for: its members in the first boxplot, and the group G1's members (called L in the paper) having the larger sigma (here 0.3) in the second boxplot;
- Esteems of the group G1 (called L in the paper) having the larger sigma (here 0.3) for: the group G0's members (called S in the paper) having the smaller sigma (here 0.2) in the third boxplot, and its own members in the fourth boxplot.





THE RESULTS WRITTEN IN THE resultsFile.csv

The file describing the results in a csv file containing the following fields (note for French people: if you don't want any trouble reading the file, this is better to change every "," in the file in ";", and every dot in ",") :

- numSimu : simulation identifier (ie correspond to the line number in the experimental design file)
- NumReplicat : replica identifier
- TaillePop : size of the population

- NbIter : total length of the simulation in steps (step = iteration)
- k: parameter *k* of the model indicating how many people one agent talks about when he/she meets another agent
- freqSauvegarde: frequency to write the results in the results file, in number of steps
- nb of groups: number of groups, here always 2
- pIntergroupcontact: this is the probability (i.e. a float having the form 0.15 for example) that someone discuss with an agent from his/her own group
- size of G0 : size of the group with the lower sigma
- size of G1: size of the group with the larger sigma
- sigma of G0 : sigma of the group G0 (called S in the paper)
- sigma of G1: sigma of the group G1 (called L in the paper)
- delta of G0: delta of the group G0
- delta of G1: delta of the group G1
- iteration: step that is described by the following written results
- avgAllOpinions: average esteem over all agents

Then, for each group G0 (= S group in the paper) and G1 (= L group in the paper), this is written the following indicators:

- G?reputAllHaveOnAverage: average reputation* of G? members in the population
- G?reputAllHaveOnStd: standard-deviation of the average reputation* of G? members in the population
- G?reputAllHaveOnMin: minimum of the reputations* of G? members in the population
- G?reputAllHaveOn1stQuartile: first quartile of the reputations* of G? members in the population
- G?reputAllHaveOnMediane: median of the reputations* of G? members in the population
- G?reputAllHaveOn3rdQuartile: third quartile of the reputations* of G? members in the population
- G?reputAllHaveOnMax: maximum of the reputations* of G? members in the population
- G?selfOpinionAverage: average self-esteem of G? members
- G?SelfOpinionStd: standard-deviation of the average self-esteem of G? members
- G?SelfOpinionMin: minimum self-esteem of G? members
- G?SelfOpinion1stQuartile: first quartile of the self-esteems of G? members
- G?SelfOpinionMediane: median of the self-esteems of G? members
- G?SelfOpinion3rdQuartile: third quartile of the self-esteems of G? members
- G?SelfOpinionMax: maximum of the self-esteems of G? members

- G?OpOnTheirGroupAverage: average esteem of G? members for their own group members
- G?OpOnTheirGroupStd: standard-deviation of the average esteem of G? members for their own group members

- G?OpOnTheirGroupMin: minimum of the esteems of G? members for their own group members
- G?OpOnTheirGroup1stQuartile: first quartile of the esteems of G? members for their own group members
- G?OpOnTheirGroupMediane: median of the esteems of G? members for their own group members
- G?OpOnTheirGroup3rdQuartile: third quartile of the esteems of G? members for their own group members
- G?OpOnTheirGroupMax: maximum of the esteems of G? members for their own group members

- G?OpOnGroupG?-Average: average esteem of G? members for the members of the other group
- G?OpOnGroupG?-Std: standard-deviation of the average esteem of G? members for the members of the other group
- G?OpOnGroupG?-Min: minimum of the esteems of G? members for the members of the other group
- G?OpOnGroupG?-1stQuartile: first quartile of the esteems of G? members for the members of the other group
- G?OpOnGroupG?-Mediane: median quartile of the esteems of G? members for the members of the other group
- G?OpOnGroupG?-3rdQuartile: third quartile of the esteems of G? members for the members of the other group
- G?OpOnGroupG?-Max: maximum of the esteems of G? members for the members of the other group

- G?OpOnGroupG?-Average: average esteem of G? members for the members of the other group
- G?OpOnGroupG?-Std: standard-deviation of the average esteem of G? members for the members of the other group
- G?OpOnGroupG?-Min: minimum of the esteems of G? members for the members of the other group
- G?OpOnGroupG?-1stQuartile: first quartile of the esteems of G? members for the members of the other group
- G?OpOnGroupG?-Mediane: median quartile of the esteems of G? members for the members of the other group
- G?OpOnGroupG?-3rdQuartile: third quartile of the esteems of G? members for the members of the other group
- G?OpOnGroupG?-Max: maximum of the esteems of G? members for the members of the other group

Some indicators are the same, but written with a different title for sake of simplicity in analysis, for example to simplify identify "the esteem for my own group".

There are then indicators about probability of agents with the larger sigma (G1 or L agents) to occupy different rank/status in the population:

- partLSigmaTop5: probability of L agents (ie G1 members agents) to occupy one of the 5 top ranks in terms of reputation*
- partLSigmaBottom35: probability of L agents (ie G1 members agents) to occupy one of the 5 bottom ranks in terms of reputation for a population of 40 agents. If the population has a different size, this is the probability to occupy one of the bottom ranks with the maximum rank being the 35th rank.
- partLSigmaTop10: probability of L agents (ie G1 members agents) to occupy one of the 10 top ranks in terms of reputation

- partLSigmaBottom30: probability of L agents (ie G1 members agents) to occupy one of the 10 bottom ranks in terms of reputation for a population of 40 agents. If the population has a different size, this is the probability to occupy one of the bottom ranks with the maximum rank being the 30th rank.
- partLSigmaTop15: probability of L agents (ie G1 members agents) to occupy one of the 15 top ranks in terms of reputation
- partLSigmaBottom25: probability of L agents (ie G1 members agents) to occupy one of the 15 bottom ranks in terms of reputation for a population of 40 agents. If the population has a different size, this is the probability to occupy one of the bottom ranks with the maximum rank being the 25th rank.
- partLSigmaTop20: probability of L agents (ie G1 members agents) to occupy one of the 20 top ranks in terms of reputation
- partLSigmaBottom20: probability of L agents (ie G1 members agents) to occupy one of the 20 bottom ranks in terms of reputation for a population of 40 agents. If the population has a different size, this is the probability to occupy one of the bottom ranks with the maximum rank being the 20th rank.

* the reputation of one agent is the average esteem other agents have for him/her