SCorPioN: Social Comparisons and Perception in Homophilic Networks

ODD Protocol of the Underlying Model

Daniel Mayerhoffer[1,2] **and Jan Schulz**[3]

[1]*Institute for Political Science, Chair for Political Theory, University of Bamberg, Feldkirchenstraße 21, 96052 Bamberg, Germany.*
[2]*DFG Project Political Epistemology, grant-no. 92/06070104*
[3]*Economics Department, Chair for International Economics, University of Bamberg, Feldkirchenstraße 21, 96052 Bamberg, Germany.*

*Correspondence should be addressed to *daniel.mayerhoffer@uni-bamberg.de*

Contents

Information on Related Projects

This ODD protocol[1] describes three versions of the SCorPioN model:

- The Baseline Model focuses on general inequality perceptions in homophilic networks:

  ⇒ Schulz, J., Mayerhoffer, D.M. and Gebhard, A. (2022). A network-based explanation of inequality perceptions. Social Networks, 70. $https://doi.org/10.1016/j.socnet.2022.02.007$

- The impact Expenditure Cascades extension investigates social consumption in those networks.

  ⇒ Working Paper available: Schulz, J. and Mayerhoffer, D.M. (2021). A Network Approach to Consumption. BERG working paper series, 173. $www.uni-bamberg.de/.../BERG_173$.

- The Perception and Privilege extension introduces two groups and a wage gap which the individuals estimate given homophilic ties in income, not privilege itself:

  ⇒ Mayerhoffer, D.M., and Schulz, J. (2022). Marginalisation and Misperception: Perceiving Gender and Racial Wage Gaps in Ego Networks. In: International Conference on Complex Networks and Their Applications.

For the sake of clarity and in line with the ODD standards, the protocol only presents the model and its design concepts. For detailed reasoning behind the modelling choices, see the corresponding papers.

[1] The presentation follows Müller, B. et al. (2013: Describing human decisions in agent-based models - ODD+D, an extension of the ODD protocol. Environmental Modelling & Software, 48, 37-48) in their extension of Grimm, V. et ak. (2010: The ODD protocol: a review and first update. Ecological modelling, 221(23), 2760-2768).

1 Overview

1.1 Purpose

The purpose of the model is to understand the impact of homophilic network formation on individual and aggregate perception and action (e.g., consumption). Furthermore, it presents a novel algorithm for creating Random Geometric Graphs that allows the specification of a minimum degree.

The model is designed as a tool for researchers who intend to study homophily in graph formation and localised perception. It is applicable not only to tie formation based on income but also other factors; the factors governing homophilic linkage and the ones that agents perceive or act upon may differ, too.

1.2 Entities, State Variables and Scales

The model consists of agents/nodes representing individuals or households and (undirected) links representing mutual connection or observation. The two components form a network space.

Agents possess an exogenously set income, income rank/decile, and storage of their links. In the consumption model, they also have a consumption; in the wage gap assessment model, they store their privilege group, too.

Links possess three state variables that result from a comparison of their two ends: income difference and whether they connect agents from different income deciles or different privilege groups.

1.3 Process Overview and Scheduling

The simulation composes of three distinct phases that take place after each other, shown in Figure 1. In the first phase, external income lists are created/loaded, and each agent is initialised with an income and respective decile from these lists; for perception and privilege, the privilege group membership is stored, too. In the second phase, the agents form links, and the resulting network is analysed using standard metrics. The network remains static in the third phase, and agents perceive the Gini/wage gap or consume based on it.

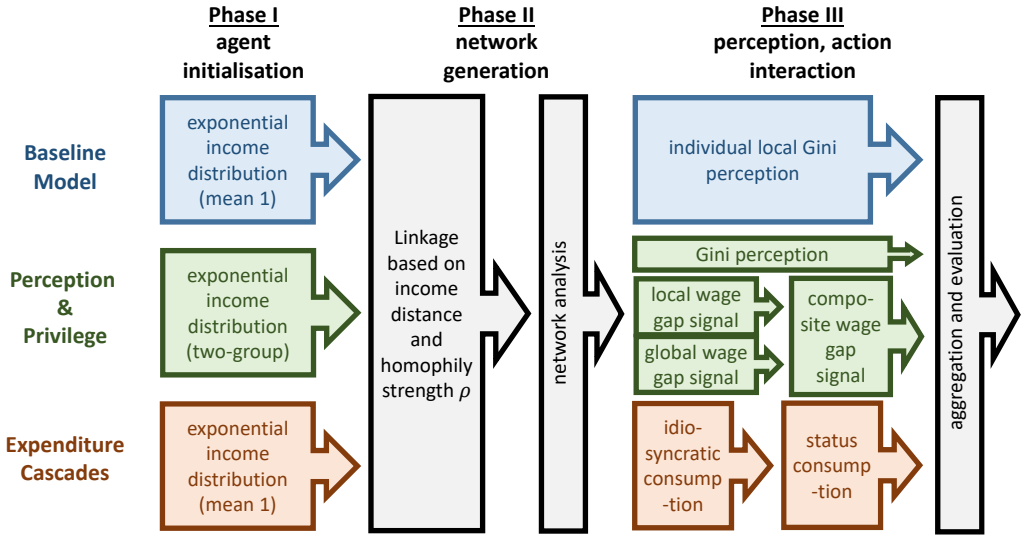

Figure 1: Process overview of all model variants

2 Design Concepts

2.1 Theoretical and Empirical Background

The model combines the notion of Random Geometric Graphs with a Preferential Attachment linkage. As a result, it applies a preferential attachment procedure to an input factor other than degree (namely income) and allows for a non-linear relation between values of the input and linkage behaviour. From this, complexity in individual localised perception emerges that mirrors empirical patterns if empirical income distributions are used as an input.

Links in the model are formed by agents' explicit decisions, not externally imposed on them. For inequality or wage gap perception, agents simply employ the unbiased approaches used to calculate the measures globally. However, they base their estimates on the biased information set that their links provide. The expenditure cascades model utilises the idea of status consumption.

2.2 Individual Decision Making

Decision making occurs during network generation and in the perception/action phase.

The decision to form links homophilically in income represents the empirically established notion of seeking similarity in social contacts, e.g., to avoid transaction costs of social interaction or use social capital efficiently. Thereby, one should understand income as a proxy because it correlates with factors like education, residential area, and occupation that agents explicitly select upon; moreover, empirical data also suggests explicit homophily in income that the above factors cannot explain alone.

The perception of general inequality and the wage gap assumes bounded rationality regarding information availability, i.e. the localised information sets governed by the agents' ego-networks. However, agents'cognitive capabilities are perfect, and they calculate the Gini and the wage gap of their individual sample correctly. These assumptions allow to focus on locally limited knowledge and invite extensions to study limited information processing.

In the expenditure cascades model, individual consumption consists of an idiosyncratic component and a status one. The weight of these components is identical for all agents. However, absolute numbers differ: Idiosyncratic consumption depends on income and status consumption implies an orientation on the highest observed consumption among link-neighbours. Such behaviour is in line with the theoretical and empirical literature and also micro-founded as a derivative of individual utility maximisation.

2.3 Learning

Agents do not learn or adapt their behavioural rules.

2.4 Individual Sensing

Agents perceive their own income/consumption and also their link-neighbours' incomes/consumption levels.

2.5 Individual Prediction

Agents do not make predictions.

2.6 Interaction

There is no direct interaction. Implicitly, an observation of link neighbours' incomes/consumption levels, of course, means that the corresponding agents interact or at least observe each other but such behaviour is represented by the links and not modelled in detail.

2.7 Collectives

There are no collectives that impact individual behaviour or constitute collective agency. In the perception and privilege model, there are two privilege groups, but agents merely perceive membership in these groups.

2.8 Heterogeneity

Behavioural/perception rules are identical across the population. However, the individual circumstances in which the agents apply these rules differ, leading to heterogeneous behavioural patterns.

2.9 Stochasticity

Chance impacts the formation of links, for one can understand this formation as a weighted draw of link-neighbours with weights being exponentially inversely given by income distance between the drawer and potential drawee. There are 100 Monte Carlo runs per model setting.

Drawing individual incomes from an input distribution also include chance, although effects are mitigated by averaging repeated draws.

2.10 Observation

After network creation, the model calculates network metrics and collects all individual perception and consumption data. Evaluation takes place on individual and aggregate levels.

3 Details

3.1 Implementation Details

The model and its input data (income lists) are available on GitHub:
$https://github.com/mayerhoffer/Inequality-Perception$

There is a NetLogo and a Julia implementation. While the NetLogo model includes all features, the Julia one does not contain the Perception and Privilege part, and network analysis is relatively limited. However, the Julia implementation has a much better performance and also has integrated analysis features; thus, the Julia version is recommended for all application cases that it features.

3.2 Initialisation

Since each model run features only a single step, initialisation mainly includes attribution of income and homophily strength. The model primarily relies on an exponential income distribution, which fits empirical data for industrialised countries (except for the income-riches 1-3%). To evaluate the relationship between actual and perceived inequality, lognormal distributions that allow for varying the Gini value are used, too. The used input distributions are normalised to mean 1 and the result of repeated draws and averaging the values for each income rank to get a more representative sample. The perception and privilege model features two input distributions for privileged and underprivileged agents. The values of the underprivileged distribution are reduced according to the wage gap g. The homophily strength ρ is identical for each agent within a run and varies between 0 (random network) and 15. Each agents draws five link neighbours in line with the empirical literature, citing this as the closest layer of interaction that the model wants to capture. However, the number of link draws per agent is implemented as available because of sensitivity analyses carried out. There are 100 Monte Carlo runs per parameter setting.

3.3 Input Data

The NetLogo version uses external income lists for convenience, while the Julia version calculates the input itself.

3.4 Submodels

Income Allocation

Each agent is assigned a value from the input income distribution. In the perception and privilege model, both groups are initialised separately. To account for the wage gap g, the actual underprivileged agents' incomes are straightforwardly calculated as follows: $Y_i(actual) = Y_i(input) * (1 - g)$

Network Formation

Each agent draws a number of other agents to link to. The model works with five link creations per agent but is validated for higher numbers, too. Thereby, agent j's weight in agent i's draw is denoted by w_{ij} and determined as follows: $w_{ij} = \frac{1}{\exp[\rho\ |Y_j - Y_i|]}$

Links are undirected and unweighted. If two agents both draw each other as link neighbours, there is still only a single link, and neither agent draws a new neighbour. Hence, the order in which agents draw link neighbours does not matter; there is a minimum but no maximum degree but a maximum number of overall links.

Perception, Action, Interaction

In the baseline model, agents count their richer and poorer link-neighbours to determine their own perceived income decile. Then, they take the mean of all possible pairwise income comparisons between agents in the set containing their link-neighbours and themselves. Afterwards, they divide the resulting value by twice the mean income in the set. This is the standard Gini calculation for their sample.

For privilege and perception, agents estimate the localised wage gap al_i as follows: $l_i = \frac{\bar{Y}_P^i - \bar{Y}_U^i}{\frac{1}{2}(\bar{Y}_P^i + \bar{Y}_U^i)}$ with $\bar{Y}_P^i$ being the mean income of their privileged neighbours (and themselves, if privileged) and $\bar{Y}_U^i$ being the mean income of their underprivileged neighbours (and themselves, if underprivileged). If a privileged agent has no underprivileged neighbours or vice versa, their wage gap estimate is set 0. The global signal is simply the true wage gap, and hence the composite perception depends on the weight put on that true global signal and the localised estimate: $p_i = (1 - w_i) \cdot l_i + w_i \cdot g$.

Consumption in the model consists of an indiosyncratic consumption (weighted by the parameter $w \in [0, 1]$) and status consumption (weighted by $1 - w$), and is thus a two-step process. First, agents carry out their idiosyncratic consumption as a fixed proportion $b \in [0, 1]$ of their income. Then, in reverse order of income ranks, they add the status consumption. This is a set fraction of the highest

consumption that they observe among link-neighbours. Hence, the consumption function reads as follows: $C(i) = w \cdot b\, Y_i + (1 - w) \cdot c\, (C(j|i) - w \cdot bY(i))$

The model analysis calculates average perceptions or consumptions for the same income rank over Monte Carlo runs. In all model versions, any external parameters are identical for all agents. Averages within and over runs are always unweighted, and especially the number of links (i.e., individual sample sizes) is not included.