Charles University in Prague

Faculty of Social Sciences

Institute of Economic Studies

Bachelor Thesis

Generating firm size distribution with agent-based models

Author: Mgr. Peter Marko

Consultant: PhDr. Petr Švarc

Academic year: 2007/2008

I hereby declare that I have elaborated this thesis on my own and that I have used only the sources listed.

Prague, 20 May 2008 Peter Marko

Acknowledgment

I am deeply indebted to my supervisor PhDr. Petr Švarc, who directed me to the very motivating topic of agent-based computational simulations and gave me numerous essential and stimulating suggestions needed to process my efforts.

I would like to express my gratitude to my parents as well, who, apart from the everlasting patient support they have been giving me, helped me financially to successfully complete also my second enjoyable studies.

Title: Generating firm size distribution with agent-based models

Author: Mgr. Peter Marko

Consultant: PhDr. Petr Švarc

Abstract: The thesis concerns the agent-based computational economics topic on the example of company sizes generation. After a brief explanation of the methodology, an already existing Robert Axtell's model is described. Based on it, two own approaches considering employees' effort levels monitoring are subsequently suggested and fully implemented in JAVA. Resulting data are finally examined for power law fitting, since a broad spectrum of empirical studies observed this distribution in real world companies' data.

Keywords: agent-based computational modelling, computational economics, power law, company sizes distribution

Název práce: Generování distribuce velikostí firem pomocí agent-based modelů

Autor: Mgr. Peter Marko

Konsultant: PhDr. Petr Švarc

Abstrakt: Bakalářská práce se věnuje ekonomickým aplikacím agent-based (agentních) modelů na příkladu algoritmického generování velikostí firem. Po krátkém přiblížení metodologie je představen již existující model Roberta Axtella, který je následně základem pro dva nově navrhnuté přístupy uvažující monitoring snah zaměstnanců firem. Nový model je kompletně implementován v jazyce JAVA. V závěru práce jsou výsledná data testována na přítomnost power law distribuce, která byla odpozorována v mnoha empirických studiích velikostí firem reálného světa.

Klíčová slova: agent-based výpočetní modely, výpočetní ekonomie, power law, distribuce velikostí firem

Contents

Chapter 1

Introduction

There is a surprising empirical evidence, that many natural and social patterns follow power law distributions, examples being earthquake magnitudes, sizes of cities, word occurrences, etc. ([18]) Fact that even company sizes in various countries have been observed to behave according to power law ([5]) forms the fundamental motivation for this thesis.

Instead of inventing rigid mathematical models, a relatively new method of studying various phenomena called agent-based computational modelling is used. By focusing on microeconomic interactions simulation, it does not have to describe the entire complex environment. One way or another, that is usually not even feasible without assuming certain simplifications. Agent-based simulations do not have to do so and reality conforming models can be more easily created.

Aim of the work is to present this economic modelling methodology on the example of company sizes generation, which will subsequently be subject to a power law analysis. After a brief introductory explanation of power law (chapter 2.1) and agent-based modelling (2.2) an already existing model being extended herein is presented (2.3). Two extensions adding a feature of individuals' effort levels monitoring are afterwards described and fully implemented (3). Finally, obtained results are statistically analysed (4).

Chapter 2

Used framework

2.1 Power law and company size distributions

Since the birth of modern statistics, scientists have introduced plenty of more or less known probability distribution types. While some of them are useful mainly in theory (e.g. chi-squared, F-distribution), others are to various extents being observed in the real world environment as well (e.g. uniform, exponential, normal, lognormal). This thesis concerns a (at least for an average student) less known probability distribution type – *Power law*, also known as *Scaling Law*, *Pareto distribution* or *Zipf's law*. We will not go into details differing the terms and will simply consider them identical.

A quantity x is said to follow a power law, when "the probability of measuring a particular value of the quantity varies inversely as a power of that value"([18], 1), in other words where the probability density/mass function of the quantity x can be written as follows:

$$p(x) = Cx^{-\alpha} \tag{1}$$

where α is an exponent characterizing the power law. For (1) to be a density/mass function, $\alpha > 1$. The value of C is then given by α[1]. Besides being right skewed, the power law distributions do have a notable feature of linearity in log-log scales[2], where α represents the negative slope.

[1]For the area beyond the $p(x)$ curve to be equal to 1; see normalization in [18], page 9 and in [11], page 2.

[2]Could be easily seen after taking a logarithm of both sides of the equation (1) : $\log p(x) = \log C - \alpha \log x$

Furthermore, there is a fascinating empirical evidence of observing such relations in surprisingly many real phenomena, e.g. in word frequencies, intensities of wars, populations of the cities, aspects of the Internet traffic, diameters of earthquakes[3], etc. ([5], 2;[18], 5) Knowledge of such a fact can become a very practical information, for instance for predictions[4]. Significant for the thesis is that even the companies' sizes distribution seems to follow a power law ([5],[18],[19],[21]). In this case, it is practical to employ a different version of (1)[5]:

$$p(x, x_0, \alpha) = \left(\frac{x}{x_0}\right)^{-\alpha} \tag{2}$$

where x_0 (minimal company size, usually 1) is a constant and $x > x_0$ is a random variable of company size.

Studies vary in estimated value of α. On the data for 1997 from the U.S. Census Bureau, Axtell ([5]) estimates α as 2.059 if size defined by number of employees and 1.994 if by receipts. In ([4], 41) he uses another reported values of 2.23 for US and 2.11 for UK of Simon and Ijiri and confirms that it is not significant which of the two definitions is applied. Ramsden ([21]) successfully estimates the power value parameter of a similar *simple canonical law* for 20 countries and tries to explain the different results by a "temperature of the economy"[6]. Kaizoji ([14]) adds that there is no universality in the size distributions of firms. Nevertheless, power law distributions in company sizes are really generally observed, with different α for different countries, usually around 2[7].

For illustration purposes, the figure 2.1 displays a histogram of company sizes frequencies across the US economy of 2005. Data were taken from the US Census Bureau ([24]) in a form of counts of the companies in few categories. Despite the need of relying on a much softer set of data and proper methods, an estimation of α by OLS (similar to the one

[3]For instance, an empirical evidence for Zipf's law says that the occurrence of words used in a natural language (especially English) follows $p(n) = n^{-\alpha}$, where n is the rank of a word in the language with 1 being the most used word.

[4]Zipf's law in the word frequencies could be used in an optimal data structures design in computer science ([20], 31).

[5]As C is given by the normalization, it has been omitted.

[6]Pérez, Brown and Tun ([19]) receive similar results for a group of less developed countries.

[7]Clauset, Shalizi and Newman showed, that due to not using rigorous methods, sometimes "the power-law hypothesis is found to be incompatible with the observed data" ([11]). The results must therefore be taken with caution.

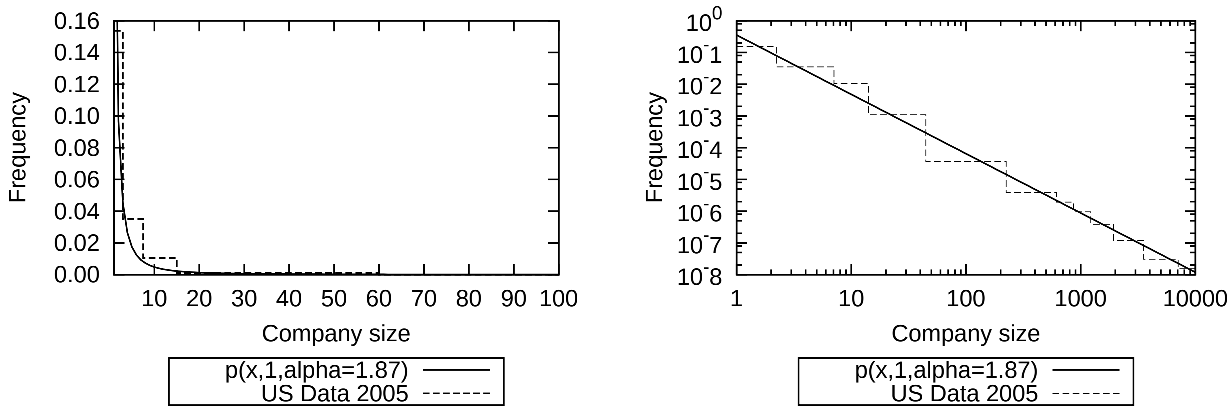

Figure 2.1: US Companies sizes for 2005 in normal and log-log scales following the power-law. Source of data: US Census Bureau([24])

Axtell did in [5]) was applied. If the histogram is redrawn in logarithmical scales (on the righthand side of the figure), some linearity could be roughly really observed (and OLS gives $\alpha = 1.87$). Such conclusions might however be spurious[8] and should be thus based on statistical methods with more accurate data only[9].

Observing a behaviour is clearly vital, yet the search for its causes is even more important. In some probability distribution types, the occurrence is unambiguously explained by the distribution's character (e.g. uniform), or using the theoretical concepts (e.g. central limit theorem for normal distributions). In power law distributions, finding the causes is however not so clear. This paper looks for the roots of the observed power law in company sizes by a relatively new technique of agent-based computational modelling. For sake of completeness, let us remind that "there exists a body of stochastic process models in which random draws from a symmetric distribution of growth rates yield distributions of firm

[8] One should really not take similar results seriously. Data used is actually only few numbers, while needed would be precise sizes of representative companies samples. In addition, the method of ordinary least squares does not take into account the normalization condition and estimates are "subject to systematic and potentially large errors"([11], 22).

[9] In the following chapters a more reliable method built on the maximum likelihood estimation by Clauset, Shalizi and Newman ([11]) will be used (Appendix A). Besides α, it estimates the lower bound for power law behaviour x_0 and can statistically say if the data indeed fit a power law distribution (conclusions based on displaying a graph in log-log scale only could easily be deceptive; the statistical test the authors introduced uses the Kolmogorov-Smirnov goodness-of-fit test). Finally, thanks to a computer generation of sizes a trustworthy data set will be examined.

sizes that are right skewed, following a Pareto distribution"([4]). This approach to explain the power law is based on so-called Gibrat's law saying that once growth rates and sizes are independent, company sizes distribution is right skewed. Growiec et al. present an econophysical model of proportional growth of firms, where firms embody products which grow proportionally to their count and sizes and so do the companies ([13]). Distribution of firm sizes follows power law in upper tails[10]. Another explanations might be also found (see for instance [18], page 12). Despite all those being relevant, the following chapters will be looking only for an agent-based way to examine observed behaviour.

2.2 Agent-based computational modelling

Economy is an unthinkably complicated and dynamic system. Models describing its rules, (ir)regularities, patterns have to deal with it. To explain even a partial process, complex, well-defined but cumbersome frameworks are often needed, though lighter solutions might be formed when seen the problem from the "bottom up" perspective. Plenty of global regularities arise from local interactions. Focusing on the "local" level of a general behaviour with the support of the current computational power it is possible to let a complex system emerge from a plain set of given rules in a simulation. This thesis's main methodology allowing such "bottom up" construction is called Agent-based computational economics (or modelling) (ACE).

The key term in ACE is agent. Agents could be practically anything - individuals (such as consumers), institutions (companies, states), physical entities or even strategies - generally any interactive unit. Usually, agents are meant to be independent, boundedly rational, heterogenous, autonomous and interacting with little or no central authority ([17], 4) - notions not often present in traditional methods. Characteristics may vary[11], however they define the environment and rules in which the agents "live" and form a complex organism representing the complicated model studied.

A common approach for simulating agents' life is to create an initial set of agents and then to process in steps. In each step a (randomly or deterministically) selected subset of agents is revived. Based on the defined rules agents then look around, evaluate their

[10] Almost the same group of authors similarly confirms that growth rates display heavy tails ([8]).

[11] Adaptation, backward learning, social networks for instance may be relevant too.

environment (i.e. fellow agents) and perform the defined tasks. Continuously, or after a certain number of periods their characteristics are gathered and examined. The initial parameters are commonly altered and the organism rerun. Analyzing the gained information, the entire system is studied. An agent-based model thus works simply as an affordable "economic laboratory" ([4]), where any situation could be simulated[12].

The ACE laboratories are constructed cheaply using programming languages. Despite the possibility to implement it procedurally, an elegant way is to make use of the object-oriented languages (such as Java, C++ or C#). Apart from other advantages, object oriented programming allows programmers to formulate code with a better resemblance of reality and suits the needs of ACE[13]. Based on the requirements, different techniques, such as genetic programming[14] may be applied, although similarly to the agent model as such, the implementation is purely up to the scientist and can vary from a model to another.

This chapter is not to be taken exhaustively. An agent-based model in its core is simply just a programme to simulate a certain economic system and can thus be understood in many ways. Its main strength is that it may overcome the cumbersomeness of robust but complicated mathematical models, while still allowing to study very complex and dynamic systems with no or little "heroic assumptions". No assumption of rationality, continuity, aggregation, homogenity of agents, equilibria reaching ([4], 90) is required, which results in producing more realistic theories. On the other hand, ACE models do not produce ultimate explanations, since these are hidden within agents configurations ([4], 89). It ought not to be considered a supplement, but rather a complement of other techniques ([22], 30).

A broad spectrum of different examples, where ACE is helpful could be presented. A famous example is the artificial stock market created at the Santa Fe Institute ([16]). Cournot oligopoly might be studied ([2]), business process could be modelled ([6]), etc. Agent based approach is in fact far more general, and economy is just one of the possible

[12]Described skeleton does not have to hold everywhere. The "life" or the data retrieval can even be understood entirely differently and still be modeling the system. The purpose of this introduction is just to outline the main properties.

[13]Concretely, the notion of objects and classes in this technique almost perfectly fits the notion of agents in ACE. A class defines agents by variables and methods resembling the parameters and rules of the agents, one object instance separately represents one agent, etc.

[14]In some agent-based models agents (when for instance representing strategies) are allowed to alter their selves by selections, mutations or recombination (e.g. in [2]).

usages, so many other more or less practical simulations might be run[15].

Despite the existence of numerous fascinating examples, this thesis concerns exclusively the topic of the companies' sizes generation. Agent-based models might offer different perspectives to the problematics studied once purely mathematically (as in already mentioned [13]). A spatial one-dimensional approach was introduced by Kuscsik and Horváth ([15]). Companies are randomly placed in the market area and have their own radii representing their sizes. The radii grow proportionally to their value, however so-called negative feedback takes place when they intersect, causing the firms to shrink. Power law distribution with $\alpha = 2.02$ is under some circumstances yielded.

In the following chapters, another ACE power law model is described ([4]) and its modification then proposed and studied for power law occurrences. Since this extended model had to be completely implemented, it can (together with the original model) be seen as a herein missing extensive example of agent based modelling.

2.3 Axtell's model

The underlying model subject to an extension proposed by the thesis in the chapter 3 is the one presented by Robert Axtell in [4]. It is a microeconomic approach able to yield empirically observed power law in company sizes distribution[16]. Despite the fact that it makes use of the production function, increasing returns on micro level, utility maximization or individual preferences, it "does not stand on equal footing with any of the conventional theories of the firm" – a firm is physically understood as a group of individuals and does not maximise its profit, there are no transaction costs or specialisation. Moreover, it is an agent-based model facilitating bounded rationality, individual heterogenity, local interactions and most importantly, non-equilibrium dynamics, a property widely present in the examined reality, but rarely in the theory. Since this very model forms the thesis's

[15]Imagine a public transport studied off the streets, analysis of logistic design patterns, distributed computing, workforce or portfolio management, etc. Even a quaint example similar to the thesis's topic of the size of wars generation exists ([9]).

[16]It also yields observed Laplace distribution of company growth rates, whose standard deviation follows logarithmically a power law too. Axtell argues that there does not exist an equilibrium microeconomic model explaining the causes; statistical solutions are however present (Gibrat's law).

basis, allow now its detailed description[17].

An agent is meant to be an individual. Let A be the set of all agents. Agents know few other agents. Let v be the count of known agents $\forall i \in A$. Each agent out of A either forms a standalone company or joins a company of one of its kins. The decision is based on the highest utility given by the possible companies' output shares. The *output* of any company C_j is defined by the production function in the form of

$$O(E_j) = aE_j + bE_j^\beta$$

where E_j is the sum of the *efforts* e_i of the member agents put in it, $E_j = \sum_{i \in C_j} e_i$. Parameters a, b and β are general for further analysis, with $b > 0 \land \beta > 1$ for increasing returns to production. All the agents belonging to a company C_j get equal shares of $\frac{O(E)}{|C_j|}$, which are then converted to their utilities[18]. *Utility function* of an agent $i \in C_j \subseteq A$ has a form of Coub-Douglas preferences

$$U^i\left(e_i; \theta_i; E_{j,\sim i}; N\right) = \left(\frac{O\left(e_i + E_{j,\sim i}\right)}{N}\right)^{\theta_i} \left(1 - e_i\right)^{(1-\theta_i)}$$

with $N = |C_j|$ being the count of agents in the company $C_j \ni i$ and $\theta_i \in [0;1]$ meaning the agent's *preference* for wage (first bracket's term) over leisure (second term). Agents may have different preferences. Agents do not know the efforts or preferencies of their colleagues. All they know is their count and their total remainder effort $E_{j,\sim i}$ (which can in fact be derived from e_i and $O(E_j)$ of C_j).

Axtell shows mathematically that the model is dynamically unstable and consequently the presented ACE model proves it too. The simulation works as follows: Initially, 1000 agents forming 1000 different singleton companies are created. Then a loop of single periods is executed. In each period, some agents are woken up (each one with uniform probability). According to its θ_i, each selected agent then computes possible utilities of staying in the company, forming a new company or joining some of (to him) known companies and chooses the best option of his future company and effort level. Only after all the selected agents have decided, changes are made – outputs, shares and new remainder effort levels are calculated and the next period starts. Program continuously gathers data for various statistics to be calculated after the termination.

[17] Axtell's notation with small amendments has been preserved.

[18] Companis sets' volumes and members change over time, term C_j is used just to indicate that there might be (and usually are) plenty of companies.

Agents' bounded rationality and autonomy is hidden in not seeing all the companies, not knowing other agents' decisions and calculating with only aggregate remainder effort levels of the previous period. Parameter θ_i values are the source of heterogenity, since they are uniformly distributed and do not change. Axtell demonstrates that local decisions and movements among the firms yield power law distributions of company sizes and it does not even matter how the size is defined – power law exponent α in (2) has a value of 2.28 when by number of member agents and 1.88 if by total output ([4], 40).

The power law is not the only remarkable result. Axtell presents a typical life cycle of a company in the simulation. Thanks to increasing returns to production on micro level, joining other companies in the beginnings pays out. As a firm grows larger, problem of free riding takes place. Utility gained from the equal share of agents becomes less sensitive to one's effort level changes. Growth declines, occasionally even collapses and the firm shrinks. Free riding limits a rampant expansion and causes near constant returns to production at the aggregate level ([4], 43).

The author subsequently investigates the model parameterization. Values of b, β, v, agents A count are separately altered. Distribution of θ_i and the utility functions are modified, new aspects of loyalty, hiring standards and alternative compensations schemes are introduced. Yielded properties, including the power law of company sizes are robust to the modifications, slope α changes to various extents. Only the change to entirely random decisions of agents impeded the distribution to arise.

All in all, an extensive work was done by Robert Axtell. What was left for future investigations is that in the existing model "shrinking goes completely undetected and unpunished" ([4], 81). There is a potential of improving the output, since "if economic organization meters poorly, productivity will be poor" ([1]). An extension to the Axtell's model, where monitoring of agents by companies is considered, is suggested in the next chapter.

Chapter 3

The extended model

"Clues to each input's productivity can be secured by observing behaviour of individual inputs." - A.A. Alchian, H. Demsetz ([1])

Despite already having introduced power law distributions and agent-based economic modelling, the main goal of the thesis is to present a complete and to certain extent unique ACE model. The model suggested herein emanates from the Axtell's model while it adds new features of monitoring of agents absent in the original. Even two different approaches of limiting the free riding problem are proposed and implemented – *demandingness* and *least effort out*. Both make use of a *boss* notion, which means that the boss – head of each company able to dismiss detected shirking agents is unambiguously defined.

In the following section, detailed description of the two approaches is given (3.1). Since the interpretation of results in ACE modelling lies within the simulation logic itself, a more technical section 3.2 deals with the simulation implementation issues. The results obtained are subsequently compared to the ones of Axtell's (chapter 4.1) and subject to simple modifications (4.2). A fully operating application capable of running under different input conditions in one of three possible modes (axtell, demandingness and least effort out) used for result analysis shown later is as a part of this work attached together with all used data on DVD (appendix B).

3.1 Boss monitoring method principles

In order to explain both methods, the notion of *boss* needs to be explained. Among all the agents in a company C_j, a boss is the one, who is in the company for the longest time. In the beginnings, it surely is the founder, nevertheless as time passes and woken agents join other companies, it does not have to be him, therefore the oldest member is always said to be the boss.

In both presented boss-monitoring approaches, the boss has the means to dismiss shirking member agents. If he is in his mother company for sufficiently long time, he can watch the other agents (who are members for sufficient number of periods too) working and thus estimate their individual effort levels put in the company according to their past performances. To implement such ability, recent periods' average effort levels of the members are visible to him. Based on their values, he can then decide of dismissals. A new parameter m defining the number of periods required to be in the company to monitor the average effort levels[1], but also to be monitored is added. Not only that the boss has to be in for m periods, but also the observed members[2]. Said explicitly, for each period t, company C_j and agent $i \in A$, let *observed average effort level* be $\bar{e}_i^{(t)}(j) = \frac{1}{m}\sum_{k=t-m}^{t-1} e_i^{(k)}$, where $e_i^{(k)}$ is the effort level of the agent i in the period k. Let $\bar{e}_i^{(t)}(j)$ be undefined if not all the m past periods were spent in the current company C_j. Value of the average is only visible to the boss and only if he has been a member for the last m periods.

The definitions of boss and observed average effort levels are equal in both variations. A distinctive aspect is the way in which bosses dismiss member agents:

- In the simplier ***demandingness*** approach, each agent, including bosses has another parameter defined - the *demandingness level* $d_i \in [0; 1]$ assigned randomly at the start of each simulation. In the existing implementation values can be drawn from either uniform or truncated normal distribution $N(0.5, 0.5^2)$. Each time a boss agent n of a company C_j is woken up in a period t, he looks at all the visible member agents' average effort levels of the past m periods and dismisses those, who have been recently giving less than his actual effort multiplied by his demandingness –

[1] The boss member does not have to be a boss for m periods, all he has to be is a boss in actual period and to be a member for sufficient time. Such behaviour seems to be in step with reality.

[2] Short-term members are left unnoticeable for couple of periods.

i.e. where $\bar{e}_i^{(t)}(j) < e_n^{(t)} \cdot d_n$.

The method may be understood as if boss agents had some personal and time shifting lower limits on the others' efforts, which nobody would be allowed to evade.

- In the ***least effort out*** approach, dismissals decisions will be based on possible increase of the utility level of a boss. When a boss agent n of a company C_j is woken up in a period t, he calculates his expected utility level $\hat{U}^n$ without monitored average effort level of members, however with the current amount of his effort level. He proceeds from the least diligent agents to the hardest workers. Each time the expected utility without agent's $i \in C_j$ effort level increases, he decides to dismiss the agent. The first comparison (decision about the member with the least average effort level) is with the current boss's utility level and the following ones iteratively with the expected utilities without previously already dismissed agents. Said precisely, he decides to dismiss a member agent $i_{(k)}$ if

$$\hat{U}^n(e_n^{(t)}, \theta_n, E_{j,\sim n} - \sum_{z=1}^{k} \bar{e}_{i_{(z)}}^{(t)}, N_j^{(t)} - k) > \hat{U}^n(e_n^{(t)}, \theta_n, E_{j,\sim n} - \sum_{z=1}^{k-1} \bar{e}_{i_{(z)}}^{(t)}, N_j^{(t)} - k + 1)$$

where $\hat{U}^n(e_n^{(t)}, \theta_n, E_{j,\sim n}, N_j^{(t)}) = U^n(e_n^{(t)}, \theta_n, E_{j,\sim n}, N_j^{(t)})$, $i_{(k)}$ is the member with the k^{th} least average effort level and $N_j^{(t)}$ the current size. The boss stops dismissing as soon as the utility without some agent does not increase. Since examined members are ordered by average effort levels, he would not find any other free rider. Note that effort levels of bosses are not optimized to the changes in remainder effort levels since the current values are applied.

The least effort out dismissal method is just a calculation, whether a boss would be better of without paying wages to the least active members even without having their average efforts, all other conditions stayed unchanged. If there are some free riders detected this way, there is no reason to keep them.

A woken boss looks at the averages only if he has been in the company for m years and he would never fire himself; therefore he ignores himself in the calculations. For sake of correctness of parallel decisions, boss calculations are done before agents (and bosses) decisions of next period's place of employment. Since not all the dismissed agents have to decide in the period of their dismissal, unemployed agents arise, nevertheless that is changed as soon as such agents are woken up in following periods.

If number of neighbours were low, it would however be very likely that they would join the same companies when woken up. In order to prevent this happening, a concept of *banned lists* is suggested – a company simply bans such agents to join again. As companies' dismissal politics depend on their bosses, who can be replaced, a ban lists is cleared whenever the company gets a new boss.

It is questionable if such extensions are conformable with reality. Especially the existence of demandingness of bosses could be attacked, since evaluating the past effort levels takes place in certain companies, e.g. in a way of monitoring the past performances and outcomes of employees according to their charged hours. Perhaps in smaller companies, there may exist an aspect such as demands of the founders on newcomers.

The methods presented herein could be distantly seen as simpliest implementations of free riders metering remarked by Alchian and Demsetz ([1]), moving Axtell's model "in a useful and realistic direction" ([10]). Adding a monitoring might affect firms' sizes and persistences, especially of the bigger ones, and is studied later (chapter 4). Yet, it does not take into account any (certainly existing) monitoring costs, or the fact that the one to monitor – the boss, who has the power of dismissal, can not encourage or threaten agents in order to make them perform faithfully. More sophisticated models might be considered, such as monitoring by "reciprocators" – employees willing to monitor shirkers for some residual claim, when benefits outweigh costs by Bowles et al. ([7]), etc. One should however be aware of the fact, that "all firms suffer, to a greater or lesser extent, from imperfect monitoring, and therefore the creation of economic models in which perfect monitoring obtains in equilibrium is a kind of quixotic undertaking, for which the only possible outcome can be disagreement with empirical data" ([4], 82).

The aim of this work is not to perfectly reflect monitoring concepts, but to introduce and implement a relatively simple initial agent-based model not very distant from reality, for which the outcomes might be easily analysed. And on which possible future improvements considering the aspects closer to metering free riders reality might stand on.

3.2 Implementation notes

Similarly as in the approaches introduced, Axtell's paper influenced the very application of this work too. Nevertheless, it has been fully written from scratch and contains several amendments and modifications. Its underlying requirements, challenged briefly herein from programming point of view, were as follows:

- Despite being inspired, the application had to reflect new needs of the approaches of demandingness and least effort out.

- For giving a tool to compare all the methodologies with the original, it had to be capable of being executed in three different modes, where so-called "axtell" mode had to copy the originally presented program flow as much as possible. This way, simulations' outcomes under any admissible set of input parameters might be trusted in terms of comparing ability. Furthermore, results would not have to rely on Axtell's data gained by OLS and more reliable method (for power law estimations) of MLE of α and goodness-of-fit tests might be applied for all the models[3].

- As many as possible input parameters to be modifiable, so the outcomes might be studied to various changes by anybody. Specifically, number of agents, periods, neighbours v and average effort levels periods m as well as waking up probability, preferences (θ_i) and demandingness (d_i) distribution types, a and b output function parameters were chosen to be arbitrarily adjustable by users.

- Leaving $\beta = 2$ eliminated one dimension of possible modifications, however allowed the application to compute agents' optimal effort levels "precisely"[4].

- Application's output statistics to be stored in a clear format suitable for further analysis in any reasonable tool. When asked, it would print a human readable step-by-step process of agents' and bosses' thoughts and decisions[5].

[3]See appendix A

[4]Not that it would consider exact values such as $\sqrt{2}$ instead of a rounded float variable around 1.4142, but that optimal effort level formula $e_i^*(\theta_i, E_{\sim i}) = \max\left[0, \frac{-a-2b(E_{\sim i}-\theta_i)+\sqrt{a^2+4ab\theta_i^2(1+E_{\sim i})+4b^2\theta_i^2(1+E_{\sim i})^2}}{2b(1+\theta_i)}\right]$ would be applied. In the original Axtell's model, optimal effort levels were detected by a "line search over feasible range of efforts"([4], 26).

[5]See appendix D.

All the above-mentioned requirements were successfully fulfilled in the application implemented and attached. A brief user's manual of how to define and run it as well as what statistics are stored is a subject of the appendix C. The following paragraphs give few notes about the simulation programming principles. Object-oriented language used is Java in version 5. Three main classes are considered – classes for `Agent`, `Company` and `Simulation` object instances[6] which represent agents, companies in a simulation and the simulation itself.

`Agent` instance contains information about the agent's personal θ, current effort value and current company reference. It can compute Coub-Douglas utility and the optimal effort level in any given company if the remainder effort and size is given. Using the internal neighbourhood[7] agents lists, decision about the next periods' mother companies are then calculated. Additional aspects for the extensions include the variables for pass effort level values and count of periods in the current company as well as a method called after dismissals to handle unemployment correctly.

Similarly, instances of `Company` hold references to the actual member `Agent` instances and know who the bosses are. Using the members' variables, a company can compute the total output and effort and can give information about the current size and remainder for any agent. Main extended logic is hidden within two methods returning actual lists of free riders according to the approaches' definitions and will be reviewed shortly. Finally, each company has a list of banned agents, which is filled with agent references each time somebody is dismissed and cleared whenever a boss leaves the company.

Agent-based model run flow is determined within `Simulation` class. For illustration purposes, algorithm 1 depicts its main method's skeleton. Logic is similar to Axtell's, with free riders dismissal being processed prior to any new decision takes place. If *mode* is set to "axtell", $freeRiders$ set of agents returned is empty and the simulation almost precisely emulates the original. Otherwise, the appropriate free riders detection method of the boss current company is called. Note that when deciding, new decisions are not finalized immediately, but only after all the woken agents have done so too. Such conduct is necessary for correct parallel execution to avoid affecting the still not decided agents unpredictably.

[6]There is also a class for input parameters representation, but that is irrelevant herein.

[7]Neighbourhood does not have to necessarily be a symmetric relation.

Algorithm 1 Simulation logic in pseudocode

```
read input parameters
period ⇐ 1
initialize agents {randomizes their θ, d, neighbours as well}
found agents' singleton companies
prepare statistical files
while period < totalPeriodsCount do
    period ⇐ period + 1
    wokenAgents ⇐ wake randomly some agents
    for all wokenAgents do
        if wokenAgent was in his current company a boss in the last m periods then
            freeRiders ⇐ detect all free riders according to the current simulation mode
            dismiss and ban all the agents in freeRiders
        end if
    end for
    for all wokenAgents do
        currentUtil ⇐ compute expected optimal utility in the current company
        neighbourUtil, neighbour ⇐ compute highest expected optimal utility among the neighbourhood companies
        newCompanyUtil ⇐ compute expected optimal utility for a newly founded company
        decide to stay/join/found according to the highest among currentUtil, neighbourUtil, newCompanyUtil values
    end for
    for all Agents do
        finalize decision{stores current company and effort levels variables, handles unemployed passive agents correctly}
    end for
    write period's period entries into aggregate companies and agents statistical files
end while
generate final company sizes data
```

Algorithm 2 Free riders detection in *demandingness* mode of a `Company` instance

```
bossEffort ⇐ get current effort of this company's boss
bossDemandingness ⇐ get demandingness level of this company's boss
acceptableThreshold ⇐ bossEffort ∗ bossDemandingness
freeRiders ⇐ empty Agents list
for all memberAgents in this Company except the boss and those not members in the m last periods do
    agentAvEffort ⇐ get average effort of the agent
    if agentAvEffort < acceptableThreshold then
        add agent to the freeRiders list
    end if
end for
return freeRiders list
```

Algorithms 2 and 3 explain how the free riders are detected. In the demandingness mode, the boss of a company straightforwardly looks at its members´ effort averages and selects those agents, whose values are less than the *acceptableThreshold*. Shirkers under the least effort out are found differently. Starting from the lowest averages, utility of the boss without certain members, but with the current (and possibly not optimal) boss´s effort is calculated. Searching stops as soon as dismissing an agent would not make the boss better off. In both methods, bosses and agents not in the firm for sufficient time are ignored. Agents are finally banned and in the main `Simulation` method immediately dismissed. Those free riders, who are among the woken agents, decide of their future in the same period shortly after, whereas the others later.

For detailed process description, see the actual source code and Javadoc documentation[8]. Program is executable in a verbose mode, where all the agents´ thoughts and decisions are visible. Appendix D contains such example complementing simulation operation description situated herein. All in all, having described the underlying logic, let us move to the results received.

Algorithm 3 Free riders detection in *least effort out* mode of a `Company` instance

utilityPossible ⇐ get current utility of this company's boss
bossCurrEffort ⇐ get current effort level of this company's boss
bossTheta ⇐ get theta value of this company's boss
sumEffort ⇐ get company's total effort level
size ⇐ get company size
freeRiders ⇐ empty Agents list
for all *memberAgents* upwardly sorted by average effort levels, excluding the boss and not members in the m last periods **do**
  agAvEffort ⇐ get average effort of the *agent*
  utilityWOAnother ⇐ compute boss utility with *bossCurrEffort*, $sumEffort - agAvEffort$, $size - 1$, *bossTheta*
  if $utilityWOAnother < utilityPossible$ **then**
    utilityPossible ⇐ *utilityWOAnother*
    sumEffort ⇐ $sumEffort - agentAvEffort$
    size ⇐ $size - 1$
    add *agent* to the *freeRiders* list
  else
    break and leave the forall cycle
  end if
end for
return *freeRiders* list

[8]Source code being almost 2500 lines long was being carefully commented and should be comprehensible.

Chapter 4

Results

After the introductory and explanatory parts, results may finally be presented. At first, Axtell's set of parameters is applied to all the three modes (section 4.1). The new approaches outputs are correctly, although indirectly, compared to the original through "axtell" mode, since maximum likelihood estimation of α instead of ordinary least squares was applied. Secondly, few parameters are subject to a modification analysis (4.2).

All the data used are stored on the attached DVD. For it is an output of a program, it contains no measurement error or noise, what may be slightly unusual. For exponent estimations, application was always run exactly one hundred times with the same input parameters and the power law fit functions were then applied on the aggregated final company sizes[1]. On the other hand, companies and agents statistics (companies lifetimes, agent average effort levels, etc.) shown are derived from single application runs only.

4.1 Initial parameters' results

Default parameter settings are summarized in the table 4.1. All the values are shared, demandingness level distribution type takes place only in demandingness mode.

Having run the application with this set of parameters in each mode one hundred times and aggregating the final company sizes, exponent α was estimated using MLE as 2.92 (Axtell), 3.23 (demandingness) and 3.28 (least effort out). Lower bound x_0 was

[1] Single run outputs vary because of the differences in randomly assigned θ values and the differences in waking order. Aggregating unique final company sizes means aggregating independent values and suppressing deviant behaviors too.

parameter	value
Number of agents, $\lvert A \rvert$	1000
Number of periods	2000
Output function parameters a and b	1.0, 1.0
Probability of waking up	0.2
Boss monitoring periods m	2
Number of neighbours v	2
Distribution of θ preferences	uniform
(Distribution of demandingness levels d)	normal

Table 4.1: Initial parameters settings

estimated to be 7, 6 and 5 respectively[2]. Figure 4.1 depicts the data gained together with the best power law fits graphically. It shows cumulative distribution functions, therefore the right-hand ends are distorted less than in mass functions (such as in the figure 2.1)[3].

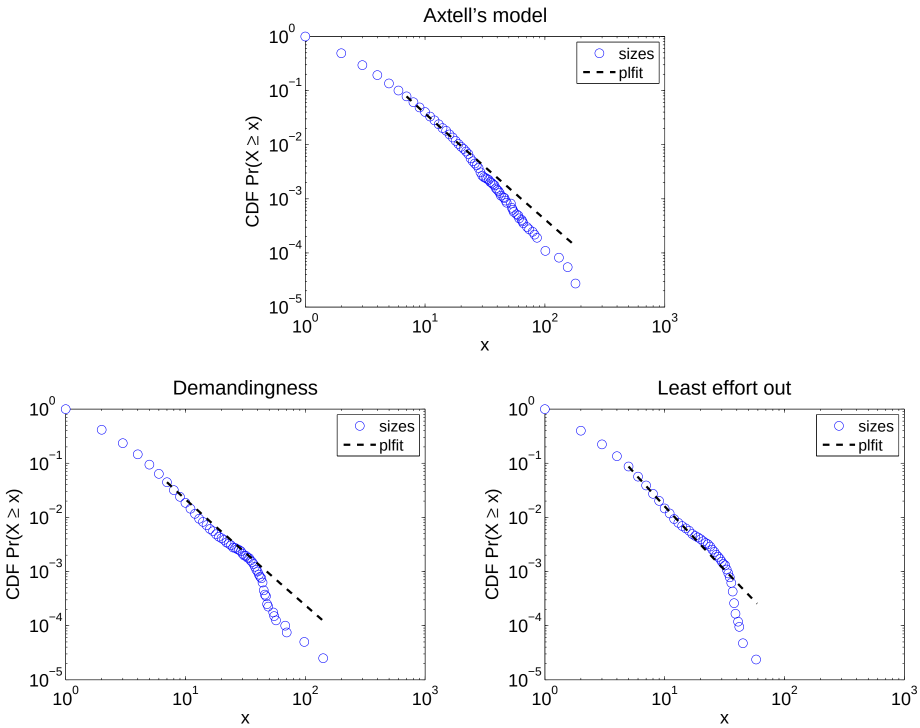

Figure 4.1: Power law MLE α and x_0 estimates for Axtell (2.92, 7), demandingness (3.23, 6) and leasteffort out modes (3.28, 5) in default environment

Note that the first exponent (2.92) is different to the one published by Robert Axtell

[2] Example of such estimations (concretely of 3.23) using the power law functions for `Matlab` and `R` (available in [12]) is illustrated in the Appendix A.

[3] Cumulative power law distribution is similar to line in log-log scales too - see [18].

(2.28), what may have been caused by applying MLE instead of OLS, different aggregation of the runs' outputs, random events, as well as the applications' variances themselves. Yet, the Axtell mode results are comparable to the new approaches through using the same application in different modes. For both of them, exponents are estimated to be higher, even above 3 - 3.23 and 3.28. Since the new methodologies generally adopt dismissing and banning the shirkers, this is not so surprising. Higher α means more small and less bigger companies operating on a market, causing the line slope in log log scales to rise absolutely, which may be the case. Around $10^{1.8} \doteq 63$, a drastic fall of the companies sizes CDF in both models takes place and it is even less likely to observe a larger company.

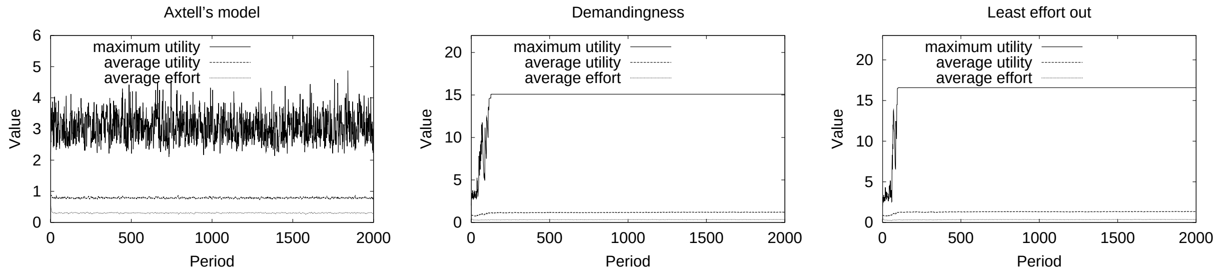

Figure 4.2: Single run statistics about agents in default environment

Free riders dismissals and bans to rejoin seem to create barriers not possible to be overstepped. Larger companies are extinct and work as "exclusive clubs". Such behavior is strengthened by a further data analysis. Figures 4.2 and 4.3 show single run statistics of agents and companies[4]. Effect of boss monitoring introduction looks alike for both new approaches, since all the graphs in the second and third columns are almost identical. Companies count always stabilizes at around 400 and the average size at around 2.5, even though ways there are different to the Axtell's (figure 4.3). Company average lifetimes are higher for the new approaches, what is clearly caused by unlimited maximum lifetime (probably of the biggest company; notice the constantly linear increases in the second row of 4.3), contrast to the oldest company changes in the original version. In addition, maximum company sizes do not fluctuate that rapidly[5] and stabilize around 40 (it fluctuates around 40 in the first mode too). Average values do not differ even for agents statistics (see 4.2), whereas maximal agent utilities do. The value is from a certain point constant,

[4]Tenth out of one hundred runs was selected in all three cases.

[5]Finding another run with from some point constant maximum size is in fact very easy.

meaning that the highest utility agent's firm (very likely the maximum sized one) found its optimal employment structure. It does not let in anyone else and works as an "exclusive club". Banned agents are left to work in the remainder space.

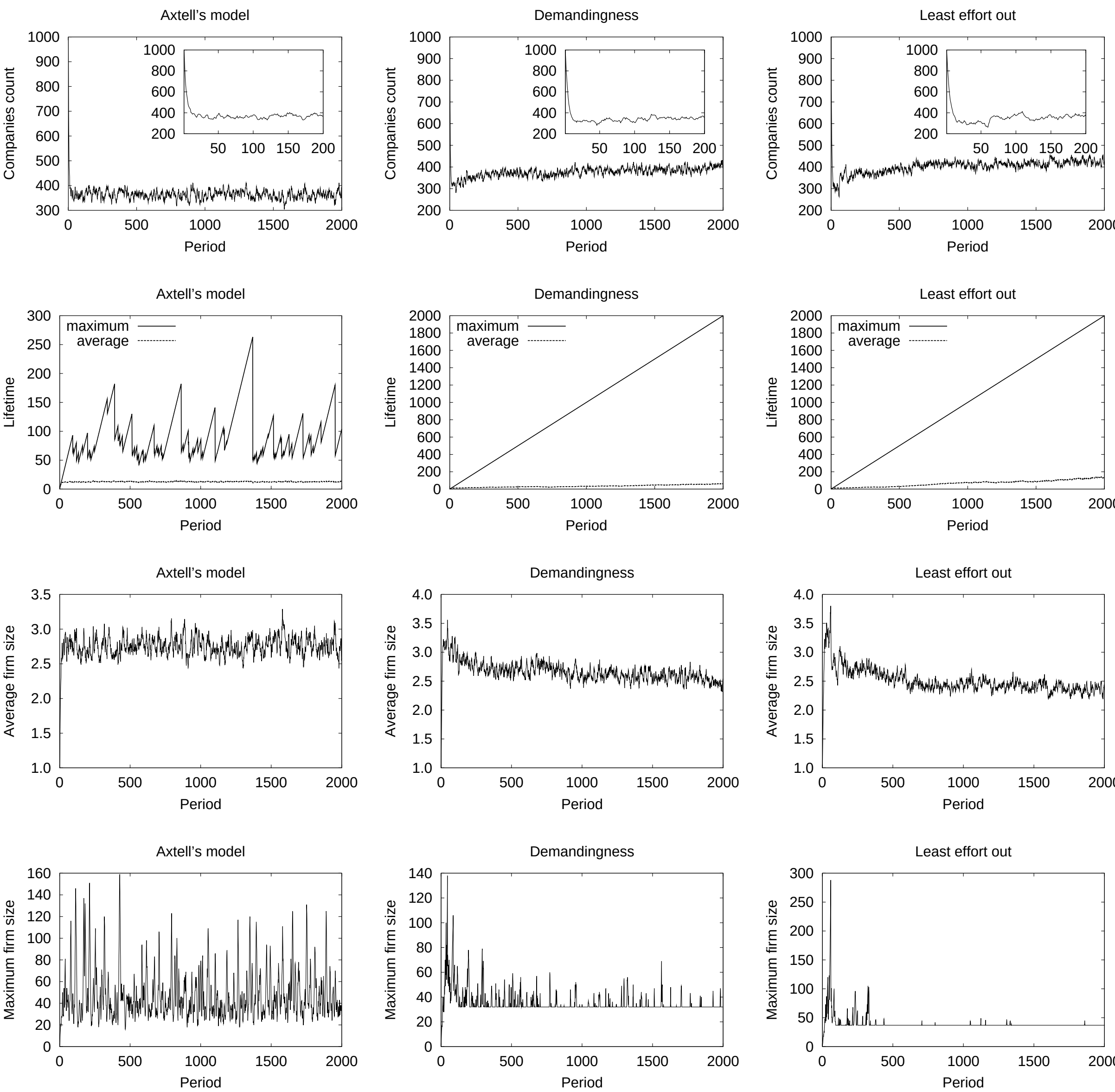

Figure 4.3: Single run statistics about companies in default environment

Although estimating higher exponents in the effort levels monitoring approaches may sound logical, limiting the total dynamics is not a desirable feature. One way to relieve seen strong constraints could be in changing the monitoring effort level methodology, e.g. considering a completely different approach or at least allowing bosses to forget about banned agents after a certain number of periods. Another way is to change the underlying parameters. Having only 2 neighbours in the environment of dismissals and banning may be too harsh, since for some agents, being dismissed twice is relatively easy. If such pairs of neighbourhood companies stabilize in time – meaning that their bosses would not want to leave them, it is impossible for the dismissed agents to join any company. Allowing to have more neighbours may then help and so may an alternation of another parameters.

4.2 Parameters adjustments analysis

Modifying various parameters may or may not have a significant impact on the model dynamics and therefore the output. This section deals specifically with the values of neighbours count v, average effort levels calculation periods count m and d demandingness levels and θ values distribution types[6] adjustments.

Exponents of all the variations were once again calculated using MLE applied on one hundred independent runs aggregates. Lower bound is always selected to be the value, from which the estimated power law distribution fits generated data the best in terms of Kolmogorov-Smirnov distance. P-values estimates of Kolmogorov-Smirnov goodness-of-fit tests are in the outline tables presented as well[7]. For comparison, initial parameter results from the previous section are in the greyed rows.

	Axtell				**Demandingness**				**Least effort out**			
m	α	x_0	p-value	OLS α	α	x_0	p-value	OLS α	α	x_0	p-value	OLS α
3	2.92	7	0	2.42	3.23	6	0.049	2.54	3.28	5	0	2.70
4	3.39	15	0.172	2.59	3.13	9	0.457	2.40	3.06	6	0.045	2.53
5	3.17	12	0.177	2.49	2.81	5	0	2.92	2.47	6	0.022	2.52
6	3.31	14	0.36	2.44	3.5	18	0.396	3.04	2.52	9	0.042	2.41

Table 4.2: Monitored periods count adjustments

Table 4.2 summarizes the results gained when changing the m parameter. Note that the exponents are not always higher for the Axtell's mode, and that they decrease in general for the new approaches with higher m. From the first column, it is clear, that the

[6]In the existing version of the software, truncated normal and uniform distributions on $[0; 1]$ are selectable.

[7]See Appendix A. For the specific calculations and the data values and aggregates consult the DVD containing around 4GB of results.

values vary among estimations all other conditions stayed unchanged even when aggregating one hundred different runs. Parameter m does not affect the logic of the simulation in the Axtell's mode, therefore the first column represents four equivalent simulations, although all leading to different estimations. One can consider it an error, nevertheless the estimations of x_0 vary much as well. Data sets are sensitive enough in terms of that even a small shift in few values can cause a better resemblance to power law from higher lower bounds, which causes the increases of α[8]. That is the root of comparing difficulties when x_0 differ a lot. Despite not guarantying normalization, unbiasedness and not meeting the theoretical requirements ([11]), OLS comes herein handy, therefore the ordinary least squares estimations were calculated too[9]. Roughly sad, the estimates are not very sensitive to the changes of m for the first mode, slighlty more for the other two. In addition, the first column values are closer to Axtell's 2.28 now, while the "better" MLE α differ significantly. That only underlines the need for careful applications, especially of the OLS estimator often used incontinently.

	Axtell				**Demandingness**				**Least effort out**			
d	α	x_0	p-value	OLS α	α	x_0	p-value	OLS α	α	x_0	p-value	OLS α
normal	2.92	7	0	2.42	3.23	6	0.049	2.54	3.28	5	0	2.70
uniform	-	-	-	-	3.17	6	0.192	2.50	-	-	-	-

Table 4.3: Demandingness levels distribution adjustments

	Axtell				**Demandingness**				**Least effort out**			
θ	α	x_0	p-value	OLS α	α	x_0	p-value	OLS α	α	x_0	p-value	OLS α
uniform	2.92	7	0	2.42	3.23	6	0.049	2.54	3.28	5	0	2.70
normal	3.5	7	0	3.18	3.5	6	0	2.74	3.5	5	0	3.39

Table 4.4: Preferences distribution adjustments

As seen from the table 4.3, a change in demandingness levels distribution type from normal to truncated uniform does not have any influence to the outputs. On the other hand, modifying the distribution of preferences expressed by θ values to truncated normal instead of uniform increases the exponents and not only when applying OLS, but even

[8]For instance, on the pictures 4.2 or 4.5 even a small shift of x_0 means higher α.

[9]OLS was applied on all the logarithmic sizes and frequencies and therefore does always assume lower bound to be $x_0 = 1$.

MLE (table 4.4; x_0 estimates stayed unchanged). A lower number of "extreme" and a higher count of "average" agents (with θ around 0.5) has a positive absolute impact on the slope.

	Axtell				**Demandingness**				**Least effort out**			
v	α	x_0	p-value	OLS α	α	x_0	p-value	OLS α	α	x_0	p-value	OLS α
2	2.92	7	0	2.42	3.23	6	0.049	2.54	3.28	5	0	2.70
3	2.49	7	0	2.10	2.59	7	0	2.21	2.35	4	0	2.34
4	2.55	14	0	1.98	2.48	9	0	2.02	2.54	10	0	2.18
5	3.24	67	0.14	1.76	2.66	14	0	2.00	3.5	31	0	2.13
6	3.06	56	0.05	1.75	3.5	67	0.43	1.91	3.5	35	0	2.03
7	3.32	80	0.35	1.63	3.5	43	0.18	1.90	3.5	36	0.12	1.91

Table 4.5: Neighbours count adjustments

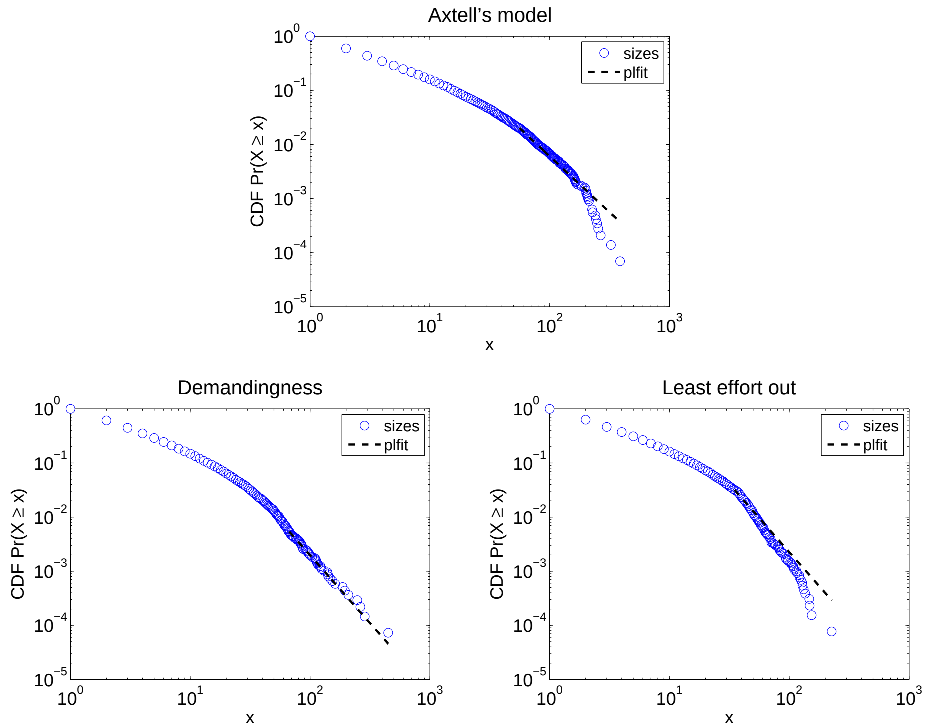

Figure 4.4: Power law MLE α and x_0 estimates for Axtell (3.06, 56), demandingness (3.5, 67) and leasteffort out modes (3.5, 35) in six neighbours environment

The most promising parameter to be analysed is the neighbours count summarized in the table 4.5. Where x_0 estimate stayed small (the first three rows of the table), α decreased with increased neighbours count bringing in more joining possibilities for agents. It is once again tricky to compare it with the remaining estimates, where lower bounds raised rapidly causing exponents to jump up. Having minimum company size for power law fit well above 50 within 1000 agents' environments undermines any estimation result, since it speaks only about a fragment of data set, even though it might be the best fit. Nonetheless, α OLS approximations there also imply substantial decreases of exponents with neighbour counts increases as it was already predicted in the previous section. Figure 4.4 demonstrates the reason of OLS and MLE with lower bound selection differences. While latter considers only the best right-hand part of the graph, former uses all the values for linear fit causing slope to fall down.

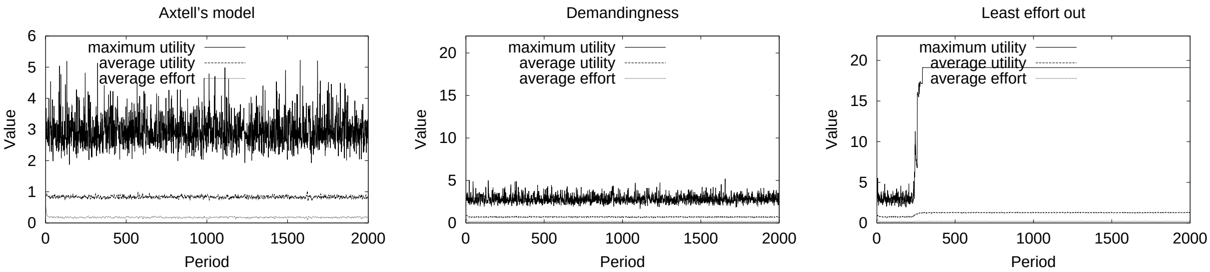

Figure 4.5: Single run statistics about agents in six neighbours environment

Relaxing the tight constraint of having only two neighbours creates a wider possibilities space. A dismissed and banned agent may still have a chance to join another company. It is questionable then, whether blocking all the new possibilities might be only a matter of time or if the possibilities' relief results also in a greater dynamics in bosses' career movements outweighing the blockages. The figures 4.5 and 4.6 depict the same set of agents and companies statistics for all the three methodologies as in the figures 4.2 and 4.3. Once again, the tenth simulation runs of six neighbours were selected. Interestingly, company average sizes amplified everywhere causing companies counts fell down proportionally. Maximum firm sizes show no stagnation for demandingness and least effort out approaches anymore. The biggest companies are opened to new employees. Would that prevail even in a longer run? Note that maximum lifetime varies now for demandingness too, while for least effort out it does not. If more graphs with number of neighbours greater than 2 were plotted, it would not be uncommon to see some mixture of such two behaviours. Dynamics in the beginnings occasionally turning into linear segments in the ends, meaning that an "exclusive" company has arisen[10], strengthened by the maximum agent utilities

[10]For the given least effort out example, "exclusive" company does not have to be the largest company

graph (see the last two charts of the figure 4.6). Unfortunately, all that one can say is that with higher neighbours count dynamics improve[11], however may be overcome observing more periods.

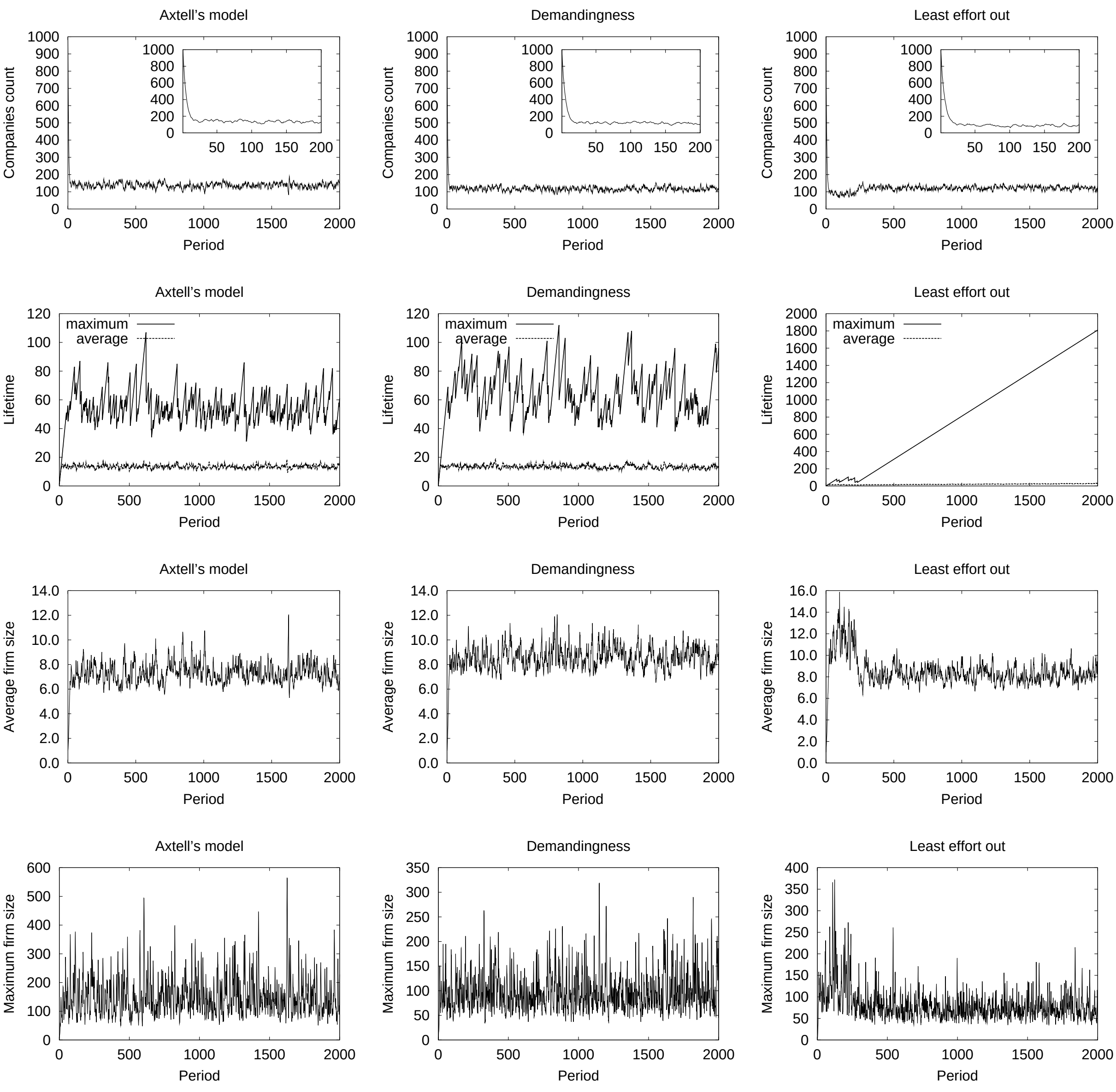

Figure 4.6: Single run statistics about companies in six neighbours environment

now, as maximum firm size fluctuates a lot.

[11]Even the data on the figure 4.4 do not fell down as drastically as they did on the 4.1.

4.3 Final remarks and future work

More than 3000 simulations were run in order to aggregately quantify exponents and lower bounds of power law distribution fitting along with 11 different sets of parameters. Comparing solely the MLE α values is problematic in the situations where x_0 is determined to be higher causing the upward shift of the corresponding exponents. If we accept OLS outputs as being reliable (while still not forgetting their insufficiency for power law fitting), then a decent reduction of α can be observed with higher number of neighbours relieving the tight constrains levied on agents in 2 neighbours mode. Adjusting the boss monitoring periods and demandingness levels distribution types does not affect the estimations.

Tables and 4.2, 4.3, 4.4 and 4.5 contain values, which have been constantly ignored up to this point – p-values of goodness-of-fit tests. P-value "quantifies the probability that generated data were drawn from the hypothesized power law distribution" ([11]). Having selected a statistical significance level as for instance 5%, this hypothesis can be rejected if corresponding p-value falls below it. In the presented results, it unfortunately happens quite often. Only few MLE results have high p-values – for instance in Axtell's approach with $m = 4, 5, 6$ or in demandingness with $m = 4, 6$, uniform d values and in few others cases. The highest p-value obtained is within demandingness with six neighbours - 0.43, meaning that the null hypothesis of fitting the power law with $\alpha = 3.5$ and $x_0 = 67$ cannot be rejected at 5% significance level. Second chart of the figure 4.4 shows a nice fit there. Yet, two problems are still persistent: not rejecting is not equal to accepting and fitting from 67 covers only a fragment of data.

One can see such results rather unsatisfactory. It may be so, however the work only confirmed Clauset's warning of relying solely on OLS ([11]). If this fitting methodology had been used herein exclusively, results would have unambiguously conformed the ones of Axtell. Applying the recommended MLE reveals the existing problems in power law fitting procedures.

In the agent-based models, economic roots lie within the simulations' logics and so it is in the model analysed. Heterogeneous boundedly rational agents maximizing their own utilities lead to company size distributions. Since the results depend on indeterministic factors such as random θ assignment and waking order, it is only the interaction rules that explain the results. In the model described, interactions do not regularly bring in power law firm sizes, when applying the proper estimation methodology.

Future offers few paths in order to further study company sizes distribution using agent-based modelling - either an improvement of proposed monitoring approaches or building a new one. Dynamics enhancing features, such as banned agents' memory continual clearing, bosses rotations or elections, more complex hierarchic structures and different methods of dismissals might be added. Interesting idea would be the mutual monitoring by "reciprocators" of Bowles et el. ([7]) and considering the monitoring costs.

Chapter 5

Conclusion

Agent-based computational modelling offers new ways of examining economic patterns from the bottom-up perspective, where microeconomic interactions lead to an aggregate dynamics. A scientist focuses in particular on correct definition of agents' behaviour and programming languages offer him unlimited spectrum of usable possibilities. Relatively recent economic notions of bounded rationality, individuals' heterogenity, etc. are often natural characteristics of an agent-based simulation, whereas including these in traditional economic models is usually a hard nut to crack.

In the thesis, an own agent-based computational model for company sizes distribution generation based on the ideas of Robert Axtell ([4]) was proposed and even fully implemented. Motivation for examining this economic and statistical pattern comes from numerous empirical studies, which have observed power law distributions in real firms' sizes. The idea behind the implemented model is to let agents form various sized companies in search for utility maximization. As companies grow, free-riders problem arises. The thesis adds a notion of aggregate effort levels detectability by bosses, who are given the right to dismiss the shirkers.

Simulation always generates a distribution, where it is much more likely to see a small than a big company, however fit to a power law is irregular. Applying correct statistical methodologies, only certain parameter settings led to power law, attenuated by the sensitivity of indeterministic factors. Even running the simulation one hundred times and then once again may result in different exponents when lower bounds are estimated too. Widely used OLS neglects these problems and since its theoretical requirements are not met, it should be avoided. Nevertheless, if one accepts it, then the results are similar to the ones of Axtell and others. Including the monitoring, dynamics tightens and power law estimated exponents increase. As predicted, relaxing default constraints of neighbour counts on the other hand supports the dynamics. Both OLS and MLE (to certain extent) prove that exponents decrease and a higher number of bigger companies is observable.

All in all, apart from few fitting problems, thesis' main goal of presenting agent-based modelling on the example of company sizes distribution generation was successfully satisfied. Attached source code of the implemented application permits its future extensions into more advanced computational simulations with free riders detection.

Appendix A

Power law exponent α estimation

For reliable analysis of data generated by the application, methods presented by Clauset, Shalizi and Newman in [11] were used. The authors proved that well-founded methods, such as least squares fitting, produce biased estimates and what is even worse, they can not be trusted in power law presence conclusions drawn.

Their method makes use of:

- Maximum likelihood estimation of α from a dataset given[1].
- Estimating the lower bound x_0 by selecting the value from which on the probability distribution of generated data and the best fitting power law MLE estimation is the closest possible measured by Kolmogorov-Smirnov statistic.
- Assessing if the data following the estimated power law distribution hypothesis can not be rejected by means of Kolmogorov-Smirnov goodness-of-fit test[2].

For the details, see their excellent paper [18]. The authors implemented `R` and `Matlab` functions for estimations following the methodology[3], which came very useful for this thesis as well. Example power law estimations using their functions on the aggregated final data of *demandingness* mode under the default set of parameters (chapter 4.1) is given in the listings A.1 and A.2.

[1]For the discrete case of power law distribution, approximation of $\hat{\alpha} = 1 + n \left[\sum_{i=i}^{n} \ln \frac{x_i}{x_0 - \frac{1}{2}} \right]$ is used.

[2]They present even a method of deciding whether the data were not drawn from competing distributions, e.g. lognormal or exponential.

[3]See Aaron Clauset webpage at `http://www.santafe.edu/~aaronc/powerlaws/`.

```
>> company_sizes_raw_combined = csvread('company_sizes_raw_combined.csv');
>> [alpha, xmin, D] = plfit(company_sizes_raw_combined);
>> [alpha, xmin, D]
   ans =
      3.2300    6.0000    0.0138
>> [p,gof]=plpva(company_sizes_raw_combined, xmin, 'silent');
>> plplot(company_sizes_raw_combined,xmin,alpha)
   ans =
    380.0015
    381.0010
>> p
   p =
      0.0490
>> gof
   gof =
      0.0138
```

Listing A.1: Matlab power law estimation procedure example

```
> data<-read.table("company_sizes_raw_combined.csv");
> plfit(data[,1]);
   $xmin
   [1] 6

   $alpha
   [1] 3.23

   $D
   [1] 0.01376551
```

Listing A.2: R power law estimation procedure example

Exponent of $\alpha = 3.23$, lower bound of $x_0 = 6$ were estimated. Kolmogorov-Smirnov distance to the fitted power law is 0.01376551 and p-value of the corresponding goodness-of-fit test is 0.0490 meaning that the power law hypothesis can be rejected even on the significance level of 5%.

Appendix B

DVD contents

This thesis has a supplementary DVD with the additional author's work attached. It contains the implemented application in the form of compiled classes and jar file prepared to be directly executed in conforming Java Virtual Machine (version 5 and higher)[1] as well as the application's source code with Javadoc documentation. Additionally, all the used simulation and US Census Bureau data are stored there too. Directory `thesis` contains an electronic version of this document together with few downloadable papers referenced.

Directory listing:

- `CompanySizesSimulation/` - application's main folder
- `CompanySizesSimulation/src/` - Java source files
- `CompanySizesSimulation/target/` - compiled code
- `CompanySizesSimulation/lib/` - used open source library of `javacsv`
- `CompanySizesSimulation/javadoc/` - complete Javadoc documentation
- `CompanySizesSimulation/build.xml` - `ANT` script for compilation
- `CompanySizesSimulation/parameters.csv` - example of the application parameters file
- `thesis/` - text of the thesis in `PDF` and `DVI` formats
- `thesis/bibliography` - some of the freely downloadable referenced papers
- `data/runs` - simulation outputs used in the thesis
- `data/US Census Bureau 2005 data` - company sizes data of US Census Bureau ([24])

[1] For user's manual, see the appendix C.

- `data/plfit functions` - used functions for power law α, x_0 estimations and goodness-of-fit tests ([12])

Subdirectories of `data/runs/` contain graphs and their `gnuplot` creating scripts. Power law estimations ([12]) function outputs are stored within these directories in the files called `plfit_r.txt` and `plfit_matlab.txt`.

Appendix C

Company Sizes Simulation user's manual

C.1 Compiling and running the application

Java 5 and Java CSV library[1] is required both to compile (J2SE 5) and run the application (at least JRE) placed on the attached DVD in `CompanySizesSimulation/` directory.

- To **compile**, Java classpath must be correctly set to contain the application source classes as well as the Java CSV library. For instance, when inside the parent directory, the command may look like:

  `javac.exe -cp "lib\javacsv.jar" src\net\hejja\muro\companysizes\* -d target` in Windows

  `javac -cp ".:lib/javacsv.jar" src/net/hejja/muro/companysizes/* -d target/` in UNIX

  Easier way is to use a prepared ant script (within the directory) by running `ant` or `ant jar` to create a JAR package, however ant has to be installed.

- Having already compiled the application, running is simple[2]. One must correctly set the classpath and specify `net.hejja.muro.companysizes.Simulation` as the class with the main static method:

  `java -classpath .;.\target;.\lib\javacsv.jar; net.hejja.muro.companysizes.Simulation` in Windows

  `java -classpath ./:./target::./lib/javacsv.jar net.hejja.muro.companysizes.Simulation` in UNIX[3]

[1] `http://www.csvreader.com/java_csv.php`
[2] Application on the DVD was already compiled and even a JAR file was created.
[3] Executing scripts are stored in `runACEsimulation.bat` and `runACEsimulation.sh`

Ant can be used to run the application by `ant run` command too. The easiest way however is to run JAR package by `java -jar companySizesSimulation.jar`. Remember that `javacsv.jar` must be placed in the `lib/` subdirectory.

C.2 Input parameters

Simulation's input parameters may be specified either by command line or from a parameters file. Command line parameters are superior:

```
[-pfile parametersfile] [-periods number_int] [-agents number_int] [-a number_notnegative]
[-demandingness] [-axtell] [-leasteffortout] [-bossperiods number_int] [-demandingnessuniform]
[-b number_notnegative] [-wakeup number_0to1] [-neighbours number_int] [-thetanormal] [-verbose]
[-progress] [-odir outputdirectory] [-aint aggregationinterval_int]
```

Specifying parameters is better using a parameter file. By default, `parameters.csv` is read, even though the command line parameter `-pfile` can override it. Individual agents and companies to generate statistics for might be specified too. Correct format and parameters meaning is explained in the example file `CompanySizesSimulation/parameters.csv`.

C.3 Application's output

When a simulation finishes, output statistics are placed within the set output directory. Following files are always generated:

- `agents_aggregate.csv` contains aggregate agents statistics (e.g. average effort, maximum utility) for each period
- `companies_aggregate.csv` contains aggregate companies statistics (e.g. maximum lifetime, average effort or average sizes) for each period
- `company_sizes_aggregated_final.csv` contains final existing company sizes and their count
- `company_sizes_aggregated_X.csv` analogically after each `aggregationinterval` periods
- `company_sizes_raw_final.csv` contains final company sizes (useful for `plfit` functions)
- `company_sizes_raw_X.csv` analogically after each `aggregationinterval` periods
- `agent_X.csv` file for each monitored agent contains the agent's career path
- `company_X.csv` files similarly contain specified company paths
- `parameters.txt` simulation's and agents' parameters

Appendix D

Example of agents activity

The developed application can be run in a verbose mode, where all the activity is traceable for debugging purposes. In other words, agent calculations and decisions, dismissals, etc. are visible. Together with the explanations given in the chapter 3.1 it might help describing how a simulation works. Following pages contain an example simulation in the *least effort out* boss monitoring method suggested in the thesis. The simulation was run just with six agents for eight periods for illustrational purposes, average effort levels are calculated based on two periods only.

```
<<<<<< Company sizes Agent-Based model v0.1 started <<<<<<
Original model by Robert Axtell
Version with the effort monitoring and application by Peter Marko
Created as a part of a bachelor thesis at the Institute of Economic Studies,
Faculty of Social Sciences, Charles University in Prague
May 2008
http://muro.hejja.net/
>> SIMULATION EXECUTED ON <Mon May 12 23:52:56 CEST 2008> WITH PARAMETERS:
Number of agents: 6
Number of periods: 8
Output function parameter a: 1.0
Output function parameter b: 1.0
Probability of waking up: 0.2
Distribution of preferences: UNIFORM
Number of neighbours: 2
Simulation mode: LEASTEFFORTOUT
Count of (last) periods used for boss effort levels monitoring: 2
Verbose mode: true
Output directory: output
Aggregation interval: 500
Statistics for agents: 1 2 3 4 5
Statistics for companies: 1 2 3 4 5
<<<<<<<<<<<<<<<<<< Simulation commenced <<<<<<<<<<<<<<<<<<
AGENTS PROPERTIES
   AGENT 1 theta = 0.23571414  neighbours = 2 3
   AGENT 2 theta = 0.7050098   neighbours = 6 1
   AGENT 3 theta = 0.61339694  neighbours = 4 6
   AGENT 4 theta = 0.08556491  neighbours = 2 6
```

```
   AGENT 5 theta = 0.20210814  neighbours = 2 1
   AGENT 6 theta = 0.97940814  neighbours = 4 3
==================== PERIOD 1 ====================
   --> Agent 1 decided to stay in the company 1 and put effort of 0.27243248
   --> Agent 2 decided to stay in the company 2 and put effort of 0.77441907
   --> Agent 3 decided to stay in the company 3 and put effort of 0.6908728
   --> Agent 4 decided to stay in the company 4 and put effort of 0.09211727
   --> Agent 5 decided to stay in the company 5 and put effort of 0.23129217
   --> Agent 6 decided to stay in the company 6 and put effort of 0.98614544
FINAL AGENTS LIST AFTER THE PERIOD 1
 AGENT 1 - last active in period 1 (company 1 - boss - joined in period 1, in the company for last 1 periods)
      effort = 0.27243248 utility = 0.6109113 share = 0.34665194 neighbourhood companies = 2 3
      average total effort of the last 2 periods = N/A and in the current company = N/A
 AGENT 2 - last active in period 1 (company 2 - boss - joined in period 1, in the company for last 1 periods)
      effort = 0.77441907 utility = 0.80638903 share = 1.374144 neighbourhood companies = 6 1
      average total effort of the last 2 periods = N/A and in the current company = N/A
 AGENT 3 - last active in period 1 (company 3 - boss - joined in period 1, in the company for last 1 periods)
      effort = 0.6908728 utility = 0.69870734 share = 1.168178 neighbourhood companies = 4 6
      average total effort of the last 2 periods = N/A and in the current company = N/A
 AGENT 4 - last active in period 1 (company 4 - boss - joined in period 1, in the company for last 1 periods)
      effort = 0.09211727 utility = 0.7521066 share = 0.100602865 neighbourhood companies = 2 6
      average total effort of the last 2 periods = N/A and in the current company = N/A
 AGENT 5 - last active in period 1 (company 5 - boss - joined in period 1, in the company for last 1 periods)
      effort = 0.23129217 utility = 0.62893295 share = 0.28478825 neighbourhood companies = 2 1
      average total effort of the last 2 periods = N/A and in the current company = N/A
 AGENT 6 - last active in period 1 (company 6 - boss - joined in period 1, in the company for last 1 periods)
      effort = 0.98614544 utility = 1.768773 share = 1.9586282 neighbourhood companies = 4 3
      average total effort of the last 2 periods = N/A and in the current company = N/A
FINAL COMPANIES LIST AFTER THE PERIOD 1
 COMPANY 1 (Boss 1 - member for 1 period(s)) - size=1 output=0.34665194 totaleffort=0.27243248 members=1
 COMPANY 2 (Boss 2 - member for 1 period(s)) - size=1 output=1.374144 totaleffort=0.77441907 members=2
 COMPANY 3 (Boss 3 - member for 1 period(s)) - size=1 output=1.168178 totaleffort=0.6908728 members=3
 COMPANY 4 (Boss 4 - member for 1 period(s)) - size=1 output=0.100602865 totaleffort=0.09211727 members=4
 COMPANY 5 (Boss 5 - member for 1 period(s)) - size=1 output=0.28478825 totaleffort=0.23129217 members=5
 COMPANY 6 (Boss 6 - member for 1 period(s)) - size=1 output=1.9586282 totaleffort=0.98614544 members=6
==================== PERIOD 2 ====================
AGENTS WOKEN in this period: 1 6
FREE RIDERS PROCESSING OF THE PERIOD 2
 Agent 1 is the boss in the company 1, however can not decide who to dismiss (not a member for at least 2 periods).
 Agent 6 is the boss in the company 6, however can not decide who to dismiss (not a member for at least 2 periods).
DECISIONS OF THE PERIOD 2
 Decision of the agent 1 /boss of the company 1/
   A1 optimal effort level utility in the current company:0.6109113
   A1 highest optimal effort level utility in a neighbourhood company:0.91533303
   A1 optimal effort level utility in a newly founded company:0.6109113
   --> Agent 1 from the company 1 decided to move to the company 2 and put effort of 0.0
 Decision of the agent 6 /boss of the company 6/
   A6 optimal effort level utility in the current company:1.768773
   A6 highest optimal effort level utility in a neighbourhood company:2.0250225
   A6 optimal effort level utility in a newly founded company:1.768773
   --> Agent 6 from the company 6 decided to move to the company 3 and put effort of 0.97840714
 Boss of the company 1 (Agent 1) left, clearing the ban list.
 Boss of the company 6 (Agent 6) left, clearing the ban list.
 Closing down the company 1
 Closing down the company 6
FINAL AGENTS LIST AFTER THE PERIOD 2
 AGENT 1 - last active in period 2 (company 2 - joined in period 2, in the company for last 1 periods)
      effort = 0.0 utility = 0.91533303 share = 0.687072 neighbourhood companies = 2 3
      average total effort of the last 2 periods = 0.13621624 and in the current company = N/A
 AGENT 2 - last active in period 1 (company 2 - boss - joined in period 1, in the company for last 2 periods)
      effort = 0.77441907 utility = 0.49466997 share = 0.687072 neighbourhood companies = 3 2
      average total effort of the last 2 periods = 0.77441907 and in the current company = 0.77441907
 AGENT 3 - last active in period 1 (company 3 - boss - joined in period 1, in the company for last 2 periods)
      effort = 0.6908728 utility = 1.0382028 share = 2.2278876 neighbourhood companies = 4 3
```

```
     average total effort of the last 2 periods = 0.6908728 and in the current company = 0.6908728
 AGENT 4 - last active in period 1 (company 4 - boss - joined in period 1, in the company for last 2 periods)
     effort = 0.09211727 utility = 0.7521066 share = 0.100602865 neighbourhood companies = 2 3
     average total effort of the last 2 periods = 0.09211727 and in the current company = 0.09211727
 AGENT 5 - last active in period 1 (company 5 - boss - joined in period 1, in the company for last 2 periods)
     effort = 0.23129217 utility = 0.62893295 share = 0.28478825 neighbourhood companies = 2 2
     average total effort of the last 2 periods = 0.23129217 and in the current company = 0.23129217
 AGENT 6 - last active in period 2 (company 3 - joined in period 2, in the company for last 1 periods)
     effort = 0.97840714 utility = 2.0250225 share = 2.2278876 neighbourhood companies = 4 3
     average total effort of the last 2 periods = 0.9822763 and in the current company = N/A
FINAL COMPANIES LIST AFTER THE PERIOD 2
 COMPANY 2 (Boss 2 - member for 2 period(s)) - size=2 output=1.374144 totaleffort=0.77441907 members=2 1
 COMPANY 3 (Boss 3 - member for 2 period(s)) - size=2 output=4.4557753 totaleffort=1.6692799 members=3 6
 COMPANY 4 (Boss 4 - member for 2 period(s)) - size=1 output=0.100602865 totaleffort=0.09211727 members=4
 COMPANY 5 (Boss 5 - member for 2 period(s)) - size=1 output=0.28478825 totaleffort=0.23129217 members=5
==================== PERIOD 3 ====================
AGENTS WOKEN in this period: 3 6
FREE RIDERS PROCESSING OF THE PERIOD 3
 Agent 3 is the boss in the company 3 and is a member for at least 2 periods.
     Current boss utility is 1.0382028
     Known average (company) effort levels of the employees are:
     No agent to dismiss.
 Agent 6 in the company 3 can not decide who to dismiss (not a boss).
DECISIONS OF THE PERIOD 3
 Decision of the agent 3 /boss of the company 3/
   A3 optimal effort level utility in the current company:1.11268
   A3 highest optimal effort level utility in a neighbourhood company:0.5102351
   A3 optimal effort level utility in a newly founded company:0.69870734
   --> Agent 3 decided to stay in the company 3 and put effort of 0.4373918
 Decision of the agent 6 /company 3/
   A6 optimal effort level utility in the current company:2.0250225
   A6 highest optimal effort level utility in a neighbourhood company:1.0236082
   A6 optimal effort level utility in a newly founded company:1.768773
   --> Agent 6 decided to stay in the company 3 and put effort of 0.97840714
FINAL AGENTS LIST AFTER THE PERIOD 3
 AGENT 1 - last active in period 2 (company 2 - joined in period 2, in the company for last 2 periods)
     effort = 0.0 utility = 0.91533303 share = 0.687072 neighbourhood companies = 2 3
     average total effort of the last 2 periods = 0.0 and in the current company = 0.0
 AGENT 2 - last active in period 1 (company 2 - boss - joined in period 1, in the company for last 3 periods)
     effort = 0.77441907 utility = 0.49466997 share = 0.687072 neighbourhood companies = 3 2
     average total effort of the last 2 periods = 0.77441907 and in the current company = 0.77441907
 AGENT 3 - last active in period 3 (company 3 - boss - joined in period 1, in the company for last 3 periods)
     effort = 0.4373918 utility = 1.11268 share = 1.7101427 neighbourhood companies = 4 3
     average total effort of the last 2 periods = 0.5641323 and in the current company = 0.5641323
 AGENT 4 - last active in period 1 (company 4 - boss - joined in period 1, in the company for last 3 periods)
     effort = 0.09211727 utility = 0.7521066 share = 0.100602865 neighbourhood companies = 2 3
     average total effort of the last 2 periods = 0.09211727 and in the current company = 0.09211727
 AGENT 5 - last active in period 1 (company 5 - boss - joined in period 1, in the company for last 3 periods)
     effort = 0.23129217 utility = 0.62893295 share = 0.28478825 neighbourhood companies = 2 2
     average total effort of the last 2 periods = 0.23129217 and in the current company = 0.23129217
 AGENT 6 - last active in period 3 (company 3 - joined in period 2, in the company for last 2 periods)
     effort = 0.97840714 utility = 1.5629107 share = 1.7101427 neighbourhood companies = 4 3
     average total effort of the last 2 periods = 0.97840714 and in the current company = 0.97840714
FINAL COMPANIES LIST AFTER THE PERIOD 3
 COMPANY 2 (Boss 2 - member for 3 period(s)) - size=2 output=1.374144 totaleffort=0.77441907 members=2 1
 COMPANY 3 (Boss 3 - member for 3 period(s)) - size=2 output=3.4202855 totaleffort=1.4157989 members=3 6
 COMPANY 4 (Boss 4 - member for 3 period(s)) - size=1 output=0.100602865 totaleffort=0.09211727 members=4
 COMPANY 5 (Boss 5 - member for 3 period(s)) - size=1 output=0.28478825 totaleffort=0.23129217 members=5
==================== PERIOD 4 ====================
AGENTS WOKEN in this period: 2
FREE RIDERS PROCESSING OF THE PERIOD 4
 Agent 2 is the boss in the company 2 and is a member for at least 2 periods.
     Current boss utility is 0.49466997
     Known average (company) effort levels of the employees are: 0.0 (agent 1),
```

```
      Boss 2 would be better off without Agent 1 who put in effort of 0.0 on average in the company.
      Without this effort, the boss could reach the utility of 0.80638903 even with his last period's effort level of 0.77441907.
      -->Agent 1 was dismissed from the company 2, effort level set to 0, banned to join again until a boss change.
DECISIONS OF THE PERIOD 4
 Decision of the agent 2 /boss of the company 2/
   A2 optimal effort level utility in the current company:0.80638903
   A2 highest optimal effort level utility in a neighbourhood company:1.2635714
   A2 optimal effort level utility in a newly founded company:0.80638903
   --> Agent 2 from the company 2 decided to move to the company 3 and put effort of 0.5133328
 Boss of the company 2 (Agent 2) left, clearing the ban list.
 Closing down the company 2
FINAL AGENTS LIST AFTER THE PERIOD 4
 AGENT 1 - last active in period 2 (no company)
      effort = 0.0 utility = 0.0 share = 0.0 neighbourhood companies = 3 3
      average total effort of the last 2 periods = 0.0 and in the current company = 0.0
 AGENT 2 - last active in period 4 (company 3 - joined in period 4, in the company for last 1 periods)
      effort = 0.5133328 utility = 1.2635714 share = 1.8835603 neighbourhood companies = 3
      average total effort of the last 2 periods = 0.64387596 and in the current company = N/A
 AGENT 3 - last active in period 3 (company 3 - boss - joined in period 1, in the company for last 4 periods)
      effort = 0.4373918 utility = 1.1805938 share = 1.8835603 neighbourhood companies = 4 3
      average total effort of the last 2 periods = 0.4373918 and in the current company = 0.4373918
 AGENT 4 - last active in period 1 (company 4 - boss - joined in period 1, in the company for last 4 periods)
      effort = 0.09211727 utility = 0.7521066 share = 0.100602865 neighbourhood companies = 3 3
      average total effort of the last 2 periods = 0.09211727 and in the current company = 0.09211727
 AGENT 5 - last active in period 1 (company 5 - boss - joined in period 1, in the company for last 4 periods)
      effort = 0.23129217 utility = 0.62893295 share = 0.28478825 neighbourhood companies = 3
      average total effort of the last 2 periods = 0.23129217 and in the current company = 0.23129217
 AGENT 6 - last active in period 3 (company 3 - joined in period 2, in the company for last 3 periods)
      effort = 0.97840714 utility = 1.7179779 share = 1.8835603 neighbourhood companies = 4 3
      average total effort of the last 2 periods = 0.97840714 and in the current company = 0.97840714
FINAL COMPANIES LIST AFTER THE PERIOD 4
 COMPANY 3 (Boss 3 - member for 4 period(s)) - size=3 output=5.650681 totaleffort=1.9291317 members=3 6 2
 COMPANY 4 (Boss 4 - member for 4 period(s)) - size=1 output=0.100602865 totaleffort=0.09211727 members=4
 COMPANY 5 (Boss 5 - member for 4 period(s)) - size=1 output=0.28478825 totaleffort=0.23129217 members=5
==================== PERIOD 5 ====================
AGENTS WOKEN in this period: 4
FREE RIDERS PROCESSING OF THE PERIOD 5
 Agent 4 is the boss in the company 4 and is a member for at least 2 periods.
      Current boss utility is 0.7521066
      Known average (company) effort levels of the employees are:
      No agent to dismiss.
DECISIONS OF THE PERIOD 5
 Decision of the agent 4 /boss of the company 4/
   A4 optimal effort level utility in the current company:0.7521066
   A4 highest optimal effort level utility in a neighbourhood company:1.0300024
   A4 optimal effort level utility in a newly founded company:0.7521066
   --> Agent 4 from the company 4 decided to move to the company 3 and put effort of 0.0
 Boss of the company 4 (Agent 4) left, clearing the ban list.
 Closing down the company 4
FINAL AGENTS LIST AFTER THE PERIOD 5
 AGENT 1 - last active in period 2 (no company)
      effort = 0.0 utility = 0.0 share = 0.0 neighbourhood companies = 3 3
      average total effort of the last 2 periods = 0.0 and in the current company = 0.0
 AGENT 2 - last active in period 4 (company 3 - joined in period 4, in the company for last 2 periods)
      effort = 0.5133328 utility = 1.0316129 share = 1.4126703
      average total effort of the last 2 periods = 0.5133328 and in the current company = 0.5133328 neighbourhood companies = 3
 AGENT 3 - last active in period 3 (company 3 - boss - joined in period 1, in the company for last 5 periods)
      effort = 0.4373918 utility = 0.98960865 share = 1.4126703 neighbourhood companies = 3 3
      average total effort of the last 2 periods = 0.4373918 and in the current company = 0.4373918
 AGENT 4 - last active in period 5 (company 3 - joined in period 5, in the company for last 1 periods)
      effort = 0.0 utility = 1.0300024 share = 1.4126703 neighbourhood companies = 3 3
      average total effort of the last 2 periods = 0.046058636 and in the current company = N/A
 AGENT 5 - last active in period 1 (company 5 - boss - joined in period 1, in the company for last 5 periods)
      effort = 0.23129217 utility = 0.62893295 share = 0.28478825 neighbourhood companies = 3
```

```
      average total effort of the last 2 periods = 0.23129217 and in the current company = 0.23129217
 AGENT 6 - last active in period 3 (company 3 - joined in period 2, in the company for last 4 periods)
      effort = 0.97840714 utility = 1.296139 share = 1.4126703 neighbourhood companies = 3 3
      average total effort of the last 2 periods = 0.97840714 and in the current company = 0.97840714
FINAL COMPANIES LIST AFTER THE PERIOD 5
 COMPANY 3 (Boss 3 - member for 5 period(s)) - size=4 output=5.650681 totaleffort=1.9291317 members=3 6 2 4
 COMPANY 5 (Boss 5 - member for 5 period(s)) - size=1 output=0.28478825 totaleffort=0.23129217 members=5
==================== PERIOD 6 ====================
AGENTS WOKEN in this period: 4 5
FREE RIDERS PROCESSING OF THE PERIOD 6
 Agent 4 in the company 3 can not decide who to dismiss (not a boss).
 Agent 5 is the boss in the company 5 and is a member for at least 2 periods.
      Current boss utility is 0.62893295
      Known average (company) effort levels of the employees are:
      No agent to dismiss.
DECISIONS OF THE PERIOD 6
 Decision of the agent 4 /company 3/
   A4 optimal effort level utility in the current company:1.0300024
   A4 highest optimal effort level utility in a neighbourhood company:0.0
   A4 optimal effort level utility in a newly founded company:0.7521066
   --> Agent 4 decided to stay in the company 3 and put effort of 0.0
 Decision of the agent 5 /boss of the company 5/
   A5 optimal effort level utility in the current company:0.62893295
   A5 highest optimal effort level utility in a neighbourhood company:1.0250337
   A5 optimal effort level utility in a newly founded company:0.62893295
   --> Agent 5 from the company 5 decided to move to the company 3 and put effort of 0.0
 Boss of the company 5 (Agent 5) left, clearing the ban list.
 Closing down the company 5
FINAL AGENTS LIST AFTER THE PERIOD 6
 AGENT 1 - last active in period 2 (no company)
      effort = 0.0 utility = 0.0 share = 0.0 neighbourhood companies = 3 3
      average total effort of the last 2 periods = 0.0 and in the current company = 0.0
 AGENT 2 - last active in period 4 (company 3 - joined in period 4, in the company for last 3 periods)
      effort = 0.5133328 utility = 0.88144314 share = 1.1301363 neighbourhood companies = 3
      average total effort of the last 2 periods = 0.5133328 and in the current company = 0.5133328
 AGENT 3 - last active in period 3 (company 3 - boss - joined in period 1, in the company for last 6 periods)
      effort = 0.4373918 utility = 0.8630166 share = 1.1301363 neighbourhood companies = 3 3
      average total effort of the last 2 periods = 0.4373918 and in the current company = 0.4373918
 AGENT 4 - last active in period 6 (company 3 - joined in period 5, in the company for last 2 periods)
      effort = 0.0 utility = 1.0105228 share = 1.1301363 neighbourhood companies = 3 3
      average total effort of the last 2 periods = 0.0 and in the current company = 0.0
 AGENT 5 - last active in period 6 (company 3 - joined in period 6, in the company for last 1 periods)
      effort = 0.0 utility = 1.0250337 share = 1.1301363 neighbourhood companies = 3
      average total effort of the last 2 periods = 0.11564609 and in the current company = N/A
 AGENT 6 - last active in period 3 (company 3 - joined in period 2, in the company for last 5 periods)
      effort = 0.97840714 utility = 1.0416867 share = 1.1301363 neighbourhood companies = 3 3
      average total effort of the last 2 periods = 0.97840714 and in the current company = 0.97840714
FINAL COMPANIES LIST AFTER THE PERIOD 6
 COMPANY 3 (Boss 3 - member for 6 period(s)) - size=5 output=5.650681 totaleffort=1.9291317 members=3 6 2 4 5
==================== PERIOD 7 ====================
AGENTS WOKEN in this period: 1 2 4
FREE RIDERS PROCESSING OF THE PERIOD 7
 Agent 1 in the no company can not decide who to dismiss (not a boss).
 Agent 2 in the company 3 can not decide who to dismiss (not a boss).
 Agent 4 in the company 3 can not decide who to dismiss (not a boss).
DECISIONS OF THE PERIOD 7
 Decision of the agent 1 /no company/
   A1 optimal effort level utility in the current company:0.0
   A1 highest optimal effort level utility in a neighbourhood company:0.98596054
   A1 optimal effort level utility in a newly founded company:0.6109113
   --> Agent 1 from the no company decided to move to the company 3 and put effort of 0.0
 Decision of the agent 2 /company 3/
   A2 optimal effort level utility in the current company:0.88144314
   A2 highest optimal effort level utility in a neighbourhood company:0.0
```

```
  A2 optimal effort level utility in a newly founded company:0.80638903
  --> Agent 2 decided to stay in the company 3 and put effort of 0.5133328
 Decision of the agent 4 /company 3/
  A4 optimal effort level utility in the current company:1.0105228
  A4 highest optimal effort level utility in a neighbourhood company:0.0
  A4 optimal effort level utility in a newly founded company:0.7521066
  --> Agent 4 decided to stay in the company 3 and put effort of 0.0
FINAL AGENTS LIST AFTER THE PERIOD 7
 AGENT 1 - last active in period 7 (company 3 - joined in period 7, in the company for last 1 periods)
     effort = 0.0 utility = 0.98596054 share = 0.94178015 neighbourhood companies = 3 3
     average total effort of the last 2 periods = 0.0 and in the current company = N/A
 AGENT 2 - last active in period 7 (company 3 - joined in period 4, in the company for last 4 periods)
     effort = 0.5133328 utility = 0.7751231 share = 0.94178015 neighbourhood companies = 3 3
     average total effort of the last 2 periods = 0.5133328 and in the current company = 0.5133328
 AGENT 3 - last active in period 3 (company 3 - boss - joined in period 1, in the company for last 7 periods)
     effort = 0.4373918 utility = 0.771702 share = 0.94178015 neighbourhood companies = 3 3
     average total effort of the last 2 periods = 0.4373918 and in the current company = 0.4373918
 AGENT 4 - last active in period 7 (company 3 - joined in period 5, in the company for last 3 periods)
     effort = 0.0 utility = 0.9948807 share = 0.94178015 neighbourhood companies = 3 3
     average total effort of the last 2 periods = 0.0 and in the current company = 0.0
 AGENT 5 - last active in period 6 (company 3 - joined in period 6, in the company for last 2 periods)
     effort = 0.0 utility = 0.98795 share = 0.94178015 neighbourhood companies = 3 3
     average total effort of the last 2 periods = 0.0 and in the current company = 0.0
 AGENT 6 - last active in period 3 (company 3 - joined in period 2, in the company for last 6 periods)
     effort = 0.97840714 utility = 0.8713373 share = 0.94178015 neighbourhood companies = 3 3
     average total effort of the last 2 periods = 0.97840714 and in the current company = 0.97840714
FINAL COMPANIES LIST AFTER THE PERIOD 7
 COMPANY 3 (Boss 3 - member for 7 period(s)) - size=6 output=5.650681 totaleffort=1.9291317 members=3 6 2 4 5 1
==================== PERIOD 8 ====================
AGENTS WOKEN in this period: 3 4 5
FREE RIDERS PROCESSING OF THE PERIOD 8
 Agent 3 is the boss in the company 3 and is a member for at least 2 periods.
     Current boss utility is 0.771702
     Known average (company) effort levels of the employees are: 0.97840714 (agent 6), 0.5133328 (agent 2), 0.0 (agent 4), 0.0 (agent 5),
     Boss 3 would be better off without Agent 4 who put in effort of 0.0 on average in the company.
     Without this effort, the boss could reach the utility of 0.8630166 even with his last period's effort level of 0.4373918.
     Boss 3 would be better off without Agent 5 who put in effort of 0.0 on average in the company.
     Without this effort, the boss could reach the utility of 0.98960865 even with his last period's effort level of 0.4373918.
     -->Agent 4 was dismissed from the company 3, effort level set to 0, banned to join again until a boss change.
     -->Agent 5 was dismissed from the company 3, effort level set to 0, banned to join again until a boss change.
 Agent 4 in the no company can not decide who to dismiss (not a boss).
 Agent 5 in the no company can not decide who to dismiss (not a boss).
DECISIONS OF THE PERIOD 8
 Decision of the agent 3 /boss of the company 3/
  A3 optimal effort level utility in the current company:0.9992832
  A3 highest optimal effort level utility in a neighbourhood company:0.0
  A3 optimal effort level utility in a newly founded company:0.69870734
  --> Agent 3 decided to stay in the company 3 and put effort of 0.3091532
 Decision of the agent 4 /no company/
  A4 optimal effort level utility in the current company:0.0
  A4 highest optimal effort level utility in a neighbourhood company:0.0
  A4 optimal effort level utility in a newly founded company:0.7521066
  --> Agent 4 from the no company decided to found a new company and put effort of 0.09211727
 Decision of the agent 5 /no company/
  A5 optimal effort level utility in the current company:0.0
  A5 highest optimal effort level utility in a neighbourhood company:0.0
  A5 optimal effort level utility in a newly founded company:0.62893295
  --> Agent 5 from the no company decided to found a new company and put effort of 0.23129217
FINAL AGENTS LIST AFTER THE PERIOD 8
 AGENT 1 - last active in period 7 (company 3 - joined in period 7, in the company for last 2 periods)
     effort = 0.0 utility = 1.0561903 share = 1.2610272 neighbourhood companies = 3 3
     average total effort of the last 2 periods = 0.0 and in the current company = 0.0
 AGENT 2 - last active in period 7 (company 3 - joined in period 4, in the company for last 5 periods)
     effort = 0.5133328 utility = 0.95224404 share = 1.2610272 neighbourhood companies = 3 3
```

```
     average total effort of the last 2 periods = 0.5133328 and in the current company = 0.5133328
 AGENT 3 - last active in period 8 (company 3 - boss - joined in period 1, in the company for last 8 periods)
     effort = 0.3091532 utility = 0.9992832 share = 1.2610272 neighbourhood companies = 7 3
     average total effort of the last 2 periods = 0.37327248 and in the current company = 0.37327248
 AGENT 4 - last active in period 8 (company 7 - boss - joined in period 8, in the company for last 1 periods)
     effort = 0.09211727 utility = 0.7521066 share = 0.100602865 neighbourhood companies = 3 3
     average total effort of the last 2 periods = 0.046058636 and in the current company = N/A
 AGENT 5 - last active in period 8 (company 8 - boss - joined in period 8, in the company for last 1 periods)
     effort = 0.23129217 utility = 0.62893295 share = 0.28478825 neighbourhood companies = 3 3
     average total effort of the last 2 periods = 0.11564609 and in the current company = N/A
 AGENT 6 - last active in period 3 (company 3 - joined in period 2, in the company for last 7 periods)
     effort = 0.97840714 utility = 1.1597135 share = 1.2610272 neighbourhood companies = 7 3
     average total effort of the last 2 periods = 0.97840714 and in the current company = 0.97840714
FINAL COMPANIES LIST AFTER THE PERIOD 8
 COMPANY 3 (Boss 3 - member for 8 period(s)) - size=4 output=5.044109 totaleffort=1.8008931 members=3 6 2 1
 COMPANY 7 (Boss 4 - member for 1 period(s)) - size=1 output=0.100602865 totaleffort=0.09211727 members=4
 COMPANY 8 (Boss 5 - member for 1 period(s)) - size=1 output=0.28478825 totaleffort=0.23129217 members=5
<<<<<<<<<<<<<<<<< Simulation finalized <<<<<<<<<<<<<<<<<
```

Listing D.1: Verbose output example in *least effort out* mode

Allow a brief discussion on what is happening in the output listing D.1: Initially, agents 4,1 and 5 have been randomly given the lowest three θ values, while agent 6 is the hardest worker. After the unexciting period 1, agents 1 and 6 woken up in the period 2 join their neighbours with the highest θ possible (Agent 2 and 3). They could not yet decide of dismissals since they had not been members for 2 previous periods. In the period 3, boss 3 does not have anybody to monitor, since he would not monitor himself and agent 6 is a member for one period only. Agents decide to stay. Waking up boss 2 in the period 4 means red light for agent 1 putting in average effort of 0 recently – agent is therefore dismissed and becomes unemployed until the period 7. Agent 2 becomes the only employee and it is better for him to join a neighbour, company 2 closes down. By the period 7, everybody is a member of company 3 – see how the neighbourhood companies adjusted to the agents' movements. In the period 8, the boss of this company – agent 3 is woken up and finds out free riders – agent 4 and 5. He would not dismiss agent 2, as this agent is valuable for the company (in terms of his own utility). Notice, that the agent 3 did the calculations even with his former effort level of 0.437 and would optimize it to 0.309 in the same period to reach higher utility in a while. Agents 4 and 5, who have been woken up too, can not join their only neighbourhood company 3, since they are banned there now and are forced to found new ones.

Bibliography

[1] Alchian A. A., Demsetz H.: *Production, Information Costs, and Economic Organization*. The American Economic Review, Vol. 62, No. 5., pp. 777-795, December 1972, viewed May 13th, `http://www.jstor.org/page/termsConfirm.jsp?redirectUri=/stable/pdfplus/1815199.pdf`.

[2] Alkemade F.: *Evolutionary Agent-Based Economics*. Technische Universiteit Eindhoven, 2004.

[3] Axelrod R., Tesfatsion L.: *On-Line Guide for Newcomers to Agent-Based Modeling in the Social Sciences*. Iowa State University, 2007, viewed 22 November 2007, `http://www.econ.iastate.edu/tesfatsi/abmread.htm`.

[4] Axtell R.: *The Emergence of Firms in a Population of Agents: Local Increasing Returns, Unstable Nash Equilibria, And Power Law Size Distributions*. Brookings Institution, 1999.

[5] Axtell R.: *Zipf Distribution of U.S. Firm Sizes*. Science 293, AAAS, 2007, viewed 22th November 2007, `www.sciencemag.org/cgi/reprint/293/5536/1818.pdf`.

[6] Bošanský B.: *A virtual company simulation by means of autonomous agents*. Faculty of Mathematics and Physics, Charles University in Prague, 2007.

[7] Bowles S., Carpenter J., Gintis H.*Mutual Monitoring in Teams: Theory and Evidence on the Importance of Residual Claimancy and Reciprocity* . Working paper, January 2001, viewed 10th April 2008, `http://www.umass.edu/preferen/gintis/mumonit.pdf`.

[8] Buldyrev S.V., Growiec J., Pammolli F., Riccaboni M., Stanley H.E.: *The growth of business firms: Facts and theory*. Journal of the European Economic Asso-

ciation, pp. 574-584, May 2007, viewed 11th May 2008, http://cps-www.bu.edu/hes/articles/bgprs07.pdf.

[9] Cederman L.E.: *Modeling the Size of Wars.* American Political Science Review, Vol. 97, pp. 135-150, 2003, http://www.econ.iastate.edu/tesfatsi/LarsErikCederman.ModelingSizeOfWars.pdf.

[10] Chang M.-H., Harrington J.E. jr.: *Agent-based Models of Organizations.* Handbook of Computational Economics, Vol. 2: Agent-Based Computational Economics, Elsevier/North-Holland, May 2006.

[11] Clauset A., Shalizi C. R., Newman M. E. J.: *Power-law distributions in empirical data.* 2007, http://arxiv.org/pdf/0706.1062v1.

[12] Clauset A.: *Power-law Distributions in Empirical Data.* Santa Fe Institute, 2008, viewed 5th May 2008, http://www.santafe.edu/~aaronc/powerlaws/.

[13] Growiec J., Pammolli F., Riccaboni M.,Stanley H.E.: *On the size distribution of business firms.* Economic Letters 98, pp. 297-212, 2008, viewed 11th May 2008, http://cps-www.bu.edu/hes/articles/gprs08.pdf.

[14] Kaizoji T., Iyetomi H., Yuichi Ikeda Y.: *Re-examination of the size distribution of firms.* Evolutionary and Institutional Economics Review 2-2, 2006, http://arxiv.org/pdf/physics/0512124.

[15] Kuscsik Z, Horváth D.: *Statistical properties of agent-based market area model.* Department of Theoretical Physics and Astrophysics, University of P.J. Šafárik, Košice, 2007, viewed 13th March 2008, http://arxiv.org/PS_cache/arxiv/pdf/0710/0710.0459v1.pdf.

[16] LeBaron B.: *Building the Santa Fe Artificial Stock Market.* Brandeis University, 2002, http://people.brandeis.edu/~blebaron/wps/sfisum.pdf.

[17] Macy M. W.: *From Factors to Actors: Computational Sociology and Agent-Based Modeling.* Cornell University, 2001, http://www.econ.iastate.edu/tesfatsi/Macy_Factors_2001.pdf.

[18] Newman M. E. J.: *Power laws, Pareto distributions and Zipf's law.* Contemporary Physics 46, 323-351, 2005, http://arxiv.org/pdf/cond-mat/0412004.

[19] Hernández-Pérez R., Angulo-Brown F. , Tun D.: *Company size distribution for developing countries*. Physica A 359 (2006) 607–618, 2006.

[20] Pokorný J., Žemlička M.: *Základy implementace souborů a databází*. Karolinum Praha, 2004.

[21] Ramsden J. J., Kiss-Haypál G.: *Company size distribution in different countries*. Physica A, 277:220-227, 2000.

[22] Tesfatsion L.: *Agent-Based Computational Economics: A Constructive Approach to Economic Theory*. Economics Department, Iowa State University, 2005.

[23] Tesfatsion L.: *Agent-Based Computational Economics, Growing Economies from the Bottom Up*. Material link collection, Iowa State University, 2007, viewed 22 November 2007, `http://www.econ.iastate.edu/tesfatsi/ace.htm`.

[24] *Statistics of U.S. Businesses: 2005*. U.S. Census Bureau, viewed 10th April 2008, `http://www.census.gov/epcd/susb/latest/us/US--.HTM`.