

Institute of Economic Studies, Faculty of Social Sciences
Charles University in Prague

Firms formation and growth in the model with heterogeneous agents and monitoring

Peter Marko
Petr Švarc

IES Working Paper: 31/2008



**Institute of Economic Studies,
Faculty of Social Sciences,
Charles University in Prague**

[UK FSV – IES]

**Opletalova 26
CZ-110 00, Prague
E-mail : ies@fsv.cuni.cz
<http://ies.fsv.cuni.cz>**

**Institut ekonomických studií
Fakulta sociálních věd
Univerzita Karlova v Praze**

**Opletalova 26
110 00 Praha 1**

**E-mail : ies@fsv.cuni.cz
<http://ies.fsv.cuni.cz>**

Disclaimer: The IES Working Papers is an online paper series for works by the faculty and students of the Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague, Czech Republic. The papers are peer reviewed, but they are *not* edited or formatted by the editors. The views expressed in documents served by this site do not reflect the views of the IES or any other Charles University Department. They are the sole property of the respective authors. Additional info at: ies@fsv.cuni.cz

Copyright Notice: Although all documents published by the IES are provided without charge, they are licensed for personal, academic or educational use. All rights are reserved by the authors.

Citations: All references to documents served by this site must be appropriately cited.

Bibliographic information:

Marko, P., Švarc, P. (2008). “ Firms formation and growth in the model with heterogeneous agents and monitoring ” IES Working Paper 31/2008. IES FSV. Charles University.

This paper can be downloaded at: <http://ies.fsv.cuni.cz>

Firms formation and growth in the model with heterogeneous agents and monitoring

Peter Marko*

Petr Švarc#

*IES, Charles University Prague,
E-mail: peter.marko@gmail.com

#Institute of Information Theory and Automation,
Academy of Sciences of the Czech Republic,
E-mail: svarc@fsv.cuni.cz

November 2008

Abstract:

In this article we extend the agent-based model of firms' formation and growth proposed in [4]. In [4] the firms' creation, expansion or contraction results from the interaction of heterogeneous utility maximizers. While the original model was able to replicate the power law distribution in the firms' sizes agents in the model set their utility maximizing effort levels completely freely and undetected. This led to the emergence of free riding and influenced the overall dynamics of the model. Therefore we decided to extend the original model by introducing the monitoring which is seen in the economic literature, besides for example the proper incentive scheme ([18]), as a possible way how to make employees work harder. Our motivation is to compare the extended model with both to the original case without monitoring and empirical data about firms' sizes distribution.

Keywords: monitoring, firms' size, power law, agent-based model, simulation, heterogeneous agents

JEL: L11, C15, C16.

Acknowledgements:

This work is supported by the institutional grant MSMT 0021620841. All remaining errors are ours.

1. Introduction

Since pioneering work of Gibrat ([15]) economists have been continuously interested in the analysis of how firms form and grow and what is the long run distribution of firms' sizes. Recent empirical findings showed that like many natural and social patterns (e.g. earthquake magnitudes, sizes of cities, word occurrences, etc.) and also other economic variables (e.g. income and wealth distribution, stock market returns and trading volumes) the firms' sizes in various countries behave according to power law ([25]). The existence of power law in economic data has an important practical implications both for the risk control and management ([31]) and policy makers ([14]). It thus seems interesting from the economist's point of view to understand how such patterns arise and persist in the economy.

Many previously proposed models that addressed this topic took the form of simple statistical models based on the modifications of the so-called Gibrat law¹ (e.g. [29], [20], [23]). Despite the fact that these models were relatively successful in generating power-laws they provided little insights of how the power-laws emerge from the interaction of economic agents (e.g. entrepreneurs, employees, managers etc.).

To answer this question we follow a different approach. In this article we extend the agent-based model of firms' formation and growth proposed in [4]. In [4] the firms creation, expansion or contraction results from the interaction of heterogeneous utility maximizers. While the original model was able to replicate the power law distribution in the firms' sizes but agents in the model set their utility maximizing effort levels completely freely and undetected. This led to the emergence of free riding and influenced the overall dynamics of the model. Therefore, we would like to extend the original model by introducing the monitoring which is seen in the economic literature, besides for example the proper incentive scheme ([18]), as a possible way how to make employees work harder. Our motivation is to compare the extended model with both the original case without monitoring and empirical data about firms sizes distribution.

Workers can be monitored in many different ways: by co-workers ([8]), by the owner of the firm ([1]) or hired supervisor ([6]). Here we introduce a very simple ad-hoc monitoring mechanism which uses central authority personified by the boss of the firm who monitors and eventually punishes workers. The aim of this work is to introduce and implement a relatively simple initial model not very distant from reality, for which the outcomes might be easily analyzed and thus the effect of monitoring on the process of firms' formation and growth examined. We also treat it as a first attempt in this direction and use it as a basis for the future research.

There exists a rich economic literature that considers different aspects of monitoring that are omitted here. For example for the insightful discussion of the motivation for the monitoring in firms we would like to refer interested reader to [1], [6] or [8]. We also simplify our analysis a lot by not considering barriers and costs of monitoring which are discussed for example in [17], [30] and [32] and leave such

¹Gibrat law states that if the growth rates of firms in a fixed population (i.e. with no entry and exit) are uncorrelated and independent of size, the resulting distribution of firms sizes is lognormal ([15])

features for future work.

The structure of the article is following. After a brief introductory explanation of power law (chapter 2.1.) and agent-based modeling (2.2.) an already existing model being extended herein is presented (2.3.). To extend the model two different monitoring mechanisms of the individuals' effort levels are described and fully implemented (3.). Finally, obtained results are statistically analyzed (4.).section 5. concludes.

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 article 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" ([25], 1), in other words where the probability density/mass function of the quantity x can be written as follows:

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

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^2 . Besides being right skewed, the power law distributions do have a notable feature of linearity in log-log scales³, where α represents the negative slope.

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 earthquakes⁴, etc. ([5], 2; [25], 5) Knowledge of such a fact can become a very practical information, for instance for predictions⁵. Significant for our work is that even the companies' sizes distribution seems to follow a power law ([5],[25],[26],[28]). In this case, it is

²For the area beyond the $p(x)$ curve to be equal to 1; see normalization in [25], page 9 and in [12], page 2.

³Could be easily seen after taking a logarithm of both sides of the equation (1) : $\log p(x) = \log C - \alpha \log x$

⁴For 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.

⁵Zipf's law in the word frequencies could be used in an optimal data structures design in computer science ([27], 31).

practical to employ a different version of (1)⁶:

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

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 ([28]) 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”⁷. Kaizoji ([19]) 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 2⁸.

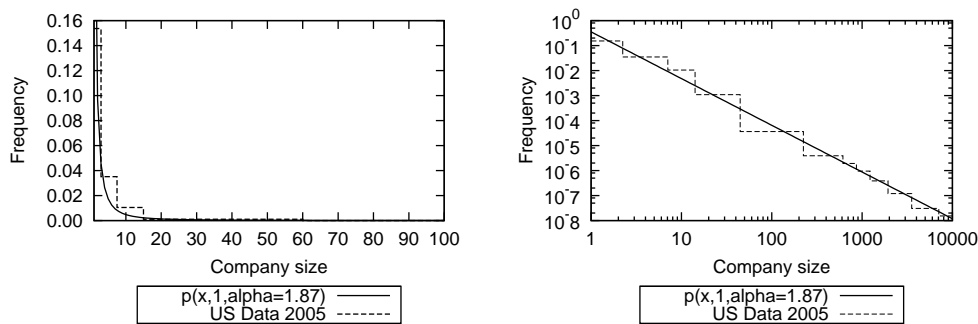


Figure 1: US Companies sizes for 2005 in normal and log-log scales following the powerlaw. Source of data: US Census Bureau([35])

For illustration purposes, the figure 1 displays a histogram of company sizes frequencies across the US economy of 2005. Data were taken from the US Census Bureau ([35]) 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 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 spurious⁹ and should be thus based on statistical methods with more

⁶As C is given by the normalization, it has been omitted.

⁷Pérez, Brown and Tun ([26]) receive similar results for a group of less developed countries.

⁸Clauset, 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” ([12]). The results must therefore be taken with caution.

⁹One 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”([12], 22).

accurate data only¹⁰.

Observing a behavior 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 modeling. 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 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 ([16]). Distribution of firm sizes follows power law in upper tails¹¹. Another explanations might be also found (see for instance [25], page 12). Despite all those being relevant, the following sections will be looking only for an agent-based way to examine observed behavior.

2.2. Agent-based computational modeling

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 behavior 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 articles's main methodology allowing such "bottom up" construction is called Agent-based computational economics (or modeling) (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, heterogeneous, autonomous and interacting with little or no central authority ([24], 4) - notions not often present in traditional methods. Characteristics may vary¹², however they define the environment and rules in which the agents "live" and form a complex organism representing the complicated model studied.

¹⁰In the following sections a more reliable method built on the maximum likelihood estimation by Clauset, Shalizi and Newman ([12]) will be used (Appendix A). Besides α , it estimates the lower bound for power law behavior 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.

¹¹Almost the same group of authors similarly confirms that growth rates display heavy tails ([9]).

¹²Adaptation, backward learning, social networks for instance may be relevant too.

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 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 simulated¹³.

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 ACE¹⁴. Based on the requirements, different techniques, such as genetic programming¹⁵ 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 section is not to be taken exhaustively. An agent-based model in its core is simply just a program to simulate a certain economic system and can thus be understood in many ways. Its main strength is that it may overcome the cumbersome 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, homogeneity 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 ([33], 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 ([22]). Cournot oligopoly might be studied ([2]), business process could be modeled ([7]), etc. Agent based approach is in fact far more general, and economy is just one of the possible usages, so many other more or less practical simulations might be run¹⁶.

Despite the existence of numerous fascinating examples, this article concerns exclusively the topic of the companies' sizes generation. Agent-based models might offer different perspectives to the problematics studied once purely mathematically

¹³Described 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.

¹⁴Concretely, 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.

¹⁵In 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]).

¹⁶Imagine 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 article's topic of the size of wars generation exists ([10]).

(as in already mentioned [16]). A spatial one-dimensional approach was introduced by Kuscsik and Horváth ([21]). 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 sections, 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 modeling.

2.3. Axtell's model

The underlying model subject to an extension proposed by our article in the section 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 distribution¹⁷. 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 maximize its profit, there are no transaction costs or specialization. Moreover, it is an agent-based model facilitating bounded rationality, individual heterogeneity, 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 article's basis, allow now its detailed description¹⁸.

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 $b > 0 \wedge \beta > 1$ 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 utilities¹⁹. *Utility function* of an agent $i \in C_j \subseteq A$ has a form of Coub-Douglas preferences

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

¹⁷It 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).

¹⁸Axtell's notation with small amendments has been preserved.

¹⁹Companies 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.

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 preferences 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.

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 heterogeneity, 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 “shirking 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 section.

3. Model with monitoring

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

Despite already having introduced power law distributions and agent-based economic modeling, the main goal of the article 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 modeling 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 can be obtained by the request from the author.

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 levels²⁰, but also to be monitored is added. Not only that the boss has to be in for m periods, but also the observed members²¹. 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.

²⁰The 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 behavior seems to be in step with reality.

²¹Short-term members are left unnoticeable for couple of 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 simpler **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 - 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 off 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 neighbors 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 simplest implementations of free riders metering remarked by Alchian and Demsetz ([1]), moving Axtell’s model ”in a useful and realistic direction” ([11]). 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. ([8]), 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 analyzed. 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 models²².

- 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, neighbors 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”²³.
- 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.

All the above-mentioned requirements were successfully fulfilled in the application implemented. 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 instances²⁴ 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 neighborhood²⁵ 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

²²See appendix A

²³Not 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).

²⁴There is also a class for input parameters representation, but that is irrelevant herein.

²⁵Neighborhood does not have to necessarily be a symmetric relation.

Algorithm 1 Simulation logic in pseudocode

```
read input parameters
period  $\leftarrow$  1
initialize agents {randomizes their  $\theta$ ,  $d$ , neighbours as well}
found agents' singleton companies
prepare statistical files
while period < totalPeriodsCount do
  period  $\leftarrow$  period + 1
  wokenAgents  $\leftarrow$  wake randomly some agents
  for all wokenAgents do
    if wokenAgent was in his current company a boss in the last  $m$  periods then
      freeRiders  $\leftarrow$  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  $\leftarrow$  compute expected optimal utility in the current company
    neighbourUtil, neighbour  $\leftarrow$  compute highest expected optimal utility among the neighborhood companies

    newCompanyUtil  $\leftarrow$  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
```

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.

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

```
bossEffort  $\leftarrow$  get current effort of this company's boss
bossDemandingness  $\leftarrow$  get demandingness level of this company's boss
acceptableThreshold  $\leftarrow$  bossEffort * bossDemandingness
freeRiders  $\leftarrow$  empty Agents list
for all memberAgents in this Company except the boss and those not members in the  $m$  last periods do
  agentAvEffort  $\leftarrow$  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.

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

```

utilityPossible  $\leftarrow$  get current utility of this company's boss
bossCurrEffort  $\leftarrow$  get current effort level of this company's boss
bossTheta  $\leftarrow$  get theta value of this company's boss
sumEffort  $\leftarrow$  get company's total effort level
size  $\leftarrow$  get company size
freeRiders  $\leftarrow$  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  $\leftarrow$  get average effort of the agent
    utilityWOAnother  $\leftarrow$  compute boss utility with bossCurrEffort, sumEffort - agAvEffort, size - 1,
    bossTheta
    if utilityWOAnother < utilityPossible then
        utilityPossible  $\leftarrow$  utilityWOAnother
        sumEffort  $\leftarrow$  sumEffort - agentAvEffort
        size  $\leftarrow$  size - 1
        add agent to the freeRiders list
    else
        break and leave the forall cycle
    end if
end for
return freeRiders list

```

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 can be obtained by the request from the authors. 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 sizes²⁶. 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 1. All the values are shared, demandingness level distribution type takes place only in demandingness mode.

²⁶Single 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, $ A $	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 neighbors v	2
Distribution of θ preferences	uniform
(Distribution of demandingness levels d)	normal

Table 1: Initial parameters settings

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 estimated to be 7, 6 and 5 respectively²⁷. Figure 2 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 1)²⁸.

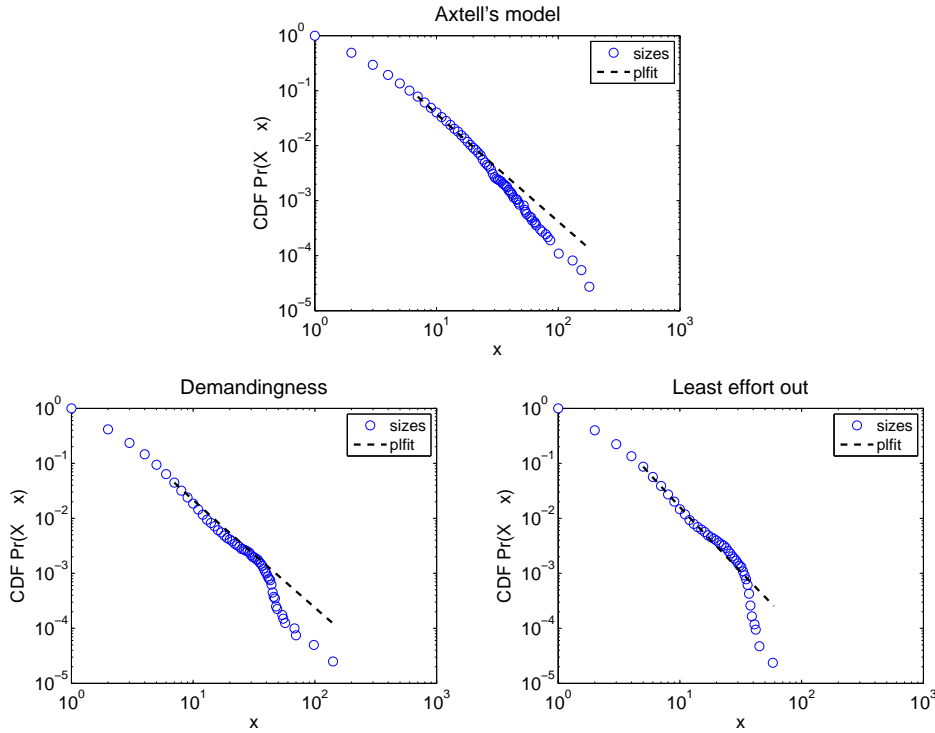


Figure 2: Power law MLE α and x_0 estimates for Axtell (2.92, 7), demandingness (3.23, 6) and leasteffort out modes (3.28, 5) in default environment

Note that the first exponent (2.92) is different to the one published by Robert Axtell (2.28), what may have been caused by applying MLE instead of OLS, dif-

²⁷Example of such estimations (concretely of 3.23) using the power law functions for Matlab and R (available in [13]) is illustrated in the Appendix A.

²⁸Cumulative power law distribution is similar to line in log-log scales too - see [25].

ferent 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.

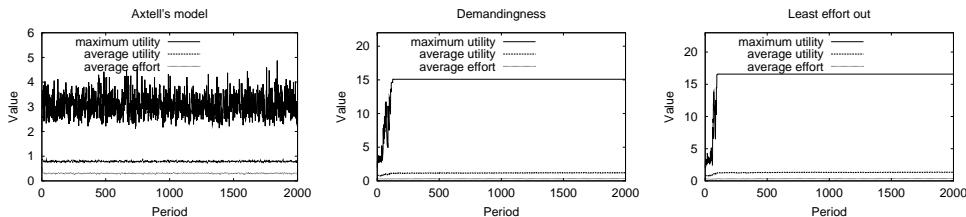


Figure 3: Single run statistics about agents in default environment

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 3 and 4 show single run statistics of agents and companies²⁹. 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). 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), contrast to the oldest company changes in the original version. In addition, maximum company sizes do not fluctuate that rapidly³⁰ and stabilize around 40 (it fluctuates around 40 in the first mode too). Average values do not differ even for agents statistics (see 3), whereas maximal agent utilities do. The value is from a certain point constant, 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.

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 neighbors 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 neighborhood companies stabilize in time

²⁹Tenth out of one hundred runs was selected in all three cases.

³⁰Finding another run with from some point constant maximum size is in fact very easy.

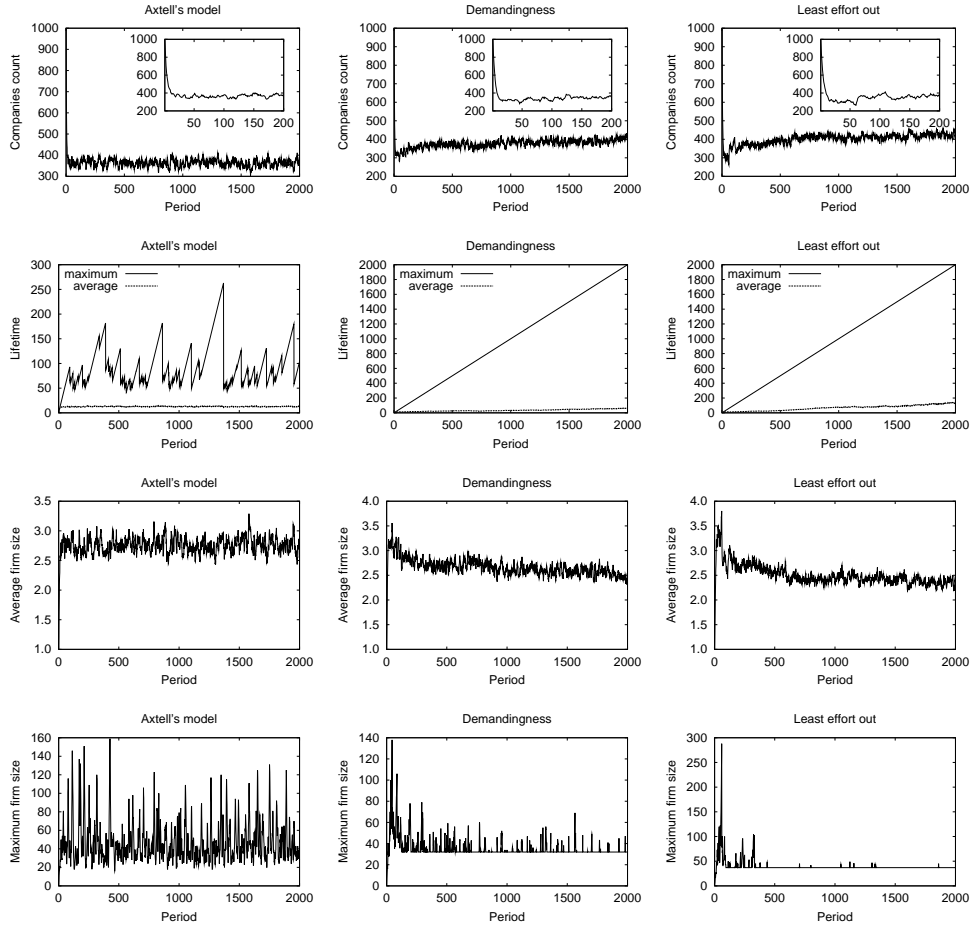


Figure 4: Single run statistics about companies in default environment

– 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 neighbors 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 neighbors count v , average effort levels calculation periods count m and d demandingness levels and θ values distribution types³¹ 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 well³². For comparison,

³¹In the existing version of the software, truncated normal and uniform distributions on $[0; 1]$ are selectable.

³²See Appendix A.

initial parameter results from the previous section are in the greyed rows.

m	Axtell				Demandingness				Least effort out			
	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 2: Monitored periods count adjustments

Table 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 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 ³³. That is the root of comparing difficulties when x_0 differ a lot. Despite not guaranteeing normalization, unbiasedness and not meeting the theoretical requirements ([12]), OLS comes herein handy, therefore the ordinary least squares estimations were calculated too³⁴. Roughly sad, the estimates are not very sensitive to the changes of m for the first mode, slightly 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.

d	Axtell				Demandingness				Least effort out			
	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 3: Demandingness levels distribution adjustments

θ	Axtell				Demandingness				Least effort out			
	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: Preferences distribution adjustments

As seen from the table 3, a change in demandingness levels distribution type from normal to truncated uniform does not have any influence to the outputs. On

³³For instance, on the pictures 3 or 6 even a small shift of x_0 means higher α .

³⁴OLS was applied on all the logarithmic sizes and frequencies and therefore does always assume lower bound to be $x_0 = 1$.

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 MLE (table 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.

v	Axtell				Demandingness				Least effort out			
	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 5: Neighbors count adjustments

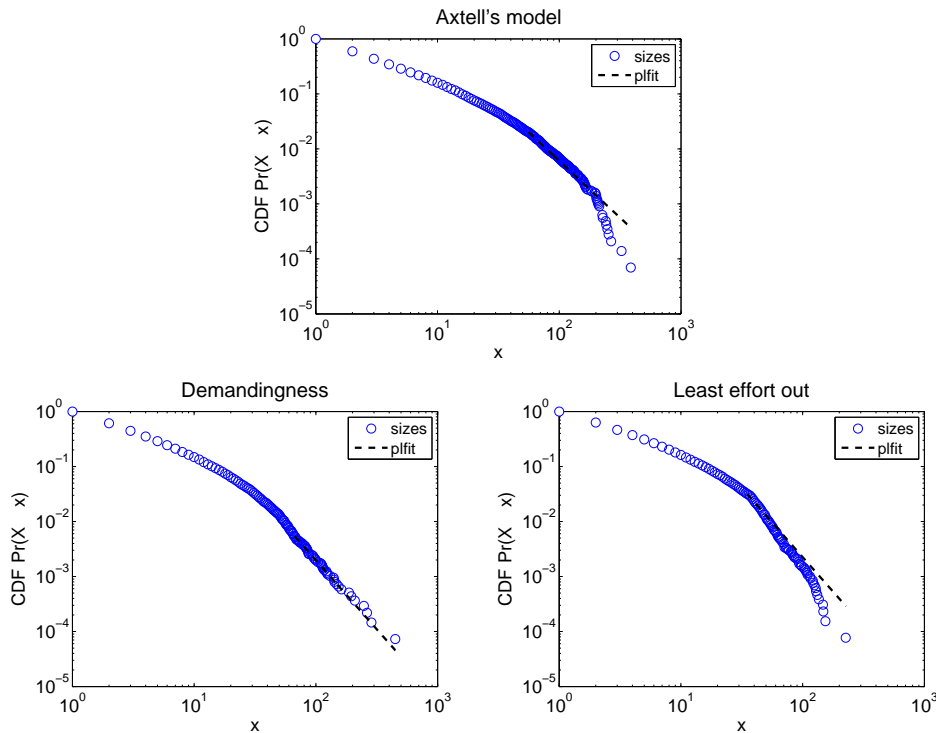


Figure 5: Power law MLE α and x_0 estimates for Axtell (3.06, 56), demandingness (3.5, 67) and leasteffort out modes (3.5, 35) in six neighbours environment

The most promising parameter to be analyzed is the neighbors count summarized in the table 5. Where x_0 estimate stayed small (the first three rows of the table), α decreased with increased neighbors 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 neighbor counts increases as it was already predicted in the previous section. Figure 5 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.

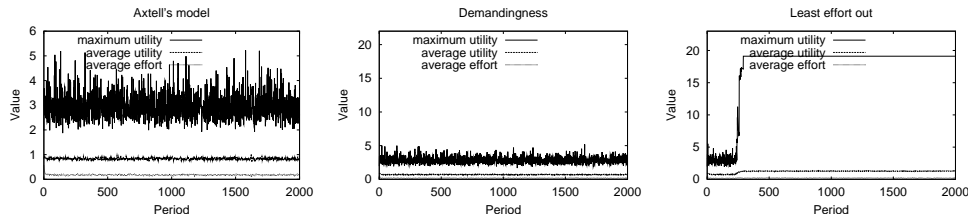


Figure 6: Single run statistics about agents in six neighbors environment

Relaxing the tight constraint of having only two neighbors 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 6 and 7 depict the same set of agents and companies statistics for all the three methodologies as in the figures 3 and 4. Once again, the tenth simulation runs of six neighbors 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 behaviors. Dynamics in the beginnings occasionally turning into linear segments in the ends, meaning that an "exclusive" company has arisen³⁵, strengthened by the maximum agent utilities graph (see the last two charts of the figure 7). Unfortunately, all that one can say is that with higher neighbors count dynamics improve³⁶, however may be overcome observing more periods.

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

³⁵For the given least effort out example, "exclusive" company does not have to be the largest company now, as maximum firm size fluctuates a lot.

³⁶Even the data on the figure 5 do not fell down as drastically as they did on the 2.

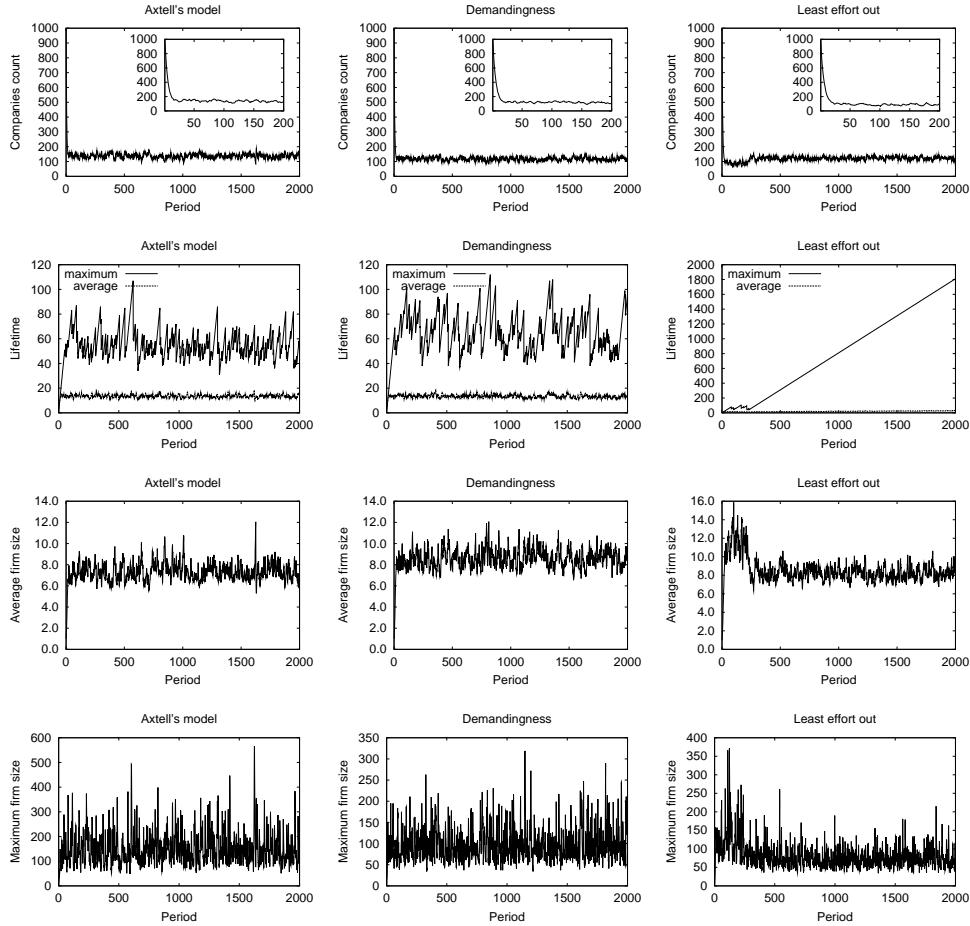


Figure 7: Single run statistics about companies in six neighbors environment

insufficiency for power law fitting), then a decent reduction of α can be observed with higher number of neighbors relieving the tight constrains levied on agents in 2 neighbors mode. Adjusting the boss monitoring periods and demandingness levels distribution types does not affect the estimations.

Tables and 2, 3, 4 and 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” ([12]). 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 neighbors - 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 5 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 ([12]). 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 analyzed. 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.

5. Conclusion

Agent-based computational modeling 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' behavior and programming languages offer him unlimited spectrum of usable possibilities. Relatively recent economic notions of bounded rationality, individuals' heterogeneity, 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 this article, 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 article 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 neighbor 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.

Future offers few paths in order to further study company sizes distribution using agent-based modeling - 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 al. ([8]) and considering the monitoring costs.

References

- [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.
- [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] Bohn H. *Monitoring Multiple Agents: The Role of Hierarchies*. Journal of Economic Behavior and Organization 8, pp. 279-305, North-Holland, 1987.
- [7] Bošanský B. *A virtual company simulation by means of autonomous agents*. Faculty of Mathematics and Physics, Charles University in Prague, 2007.
- [8] 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>.
- [9] 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 Association, pp. 574-584, May 2007, viewed 11th May 2008, <http://cps-www.bu.edu/hes/articles/bgprs07.pdf>.
- [10] 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>.
- [11] 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.

- [12] Clauset A., Shalizi C. R., Newman M. E. J. *Power-law distributions in empirical data*. 2007, <http://arxiv.org/pdf/0706.1062v1>.
- [13] Clauset A. *Power-law Distributions in Empirical Data*. Santa Fe Institute, 2008, viewed 5th May 2008, <http://www.santafe.edu/~aaronc/powerlaws/>.
- [14] Gallegati M., Kirman A., Palestrini A. *Implications of Scalling Laws for Policy-Makers*. In Salzano M. and Colander D. *Complexity Hints for Economic Policy*, pp. 291-302, Springer, 2007.
- [15] Gibrat R. *Une loi des réparations économiques: l'effet proportionnel*. Bull. Statist. gén Fr. 19, 469, 1930.
- [16] 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>.
- [17] Guillermo A. Calvo, Wellisz S. *Supervision, Loss of Control, and the Optimum Size of the Firm*. The Journal of Political Economy, Vol. 86, No. 5, pp. 943-952, October, 1978.
- [18] Holmstrom B. *Moral Hayard in Teams*. The Bell Journal of Economics, Vol. 13, No. 2, pp. 324-340, Autumn, 1982.
- [19] 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>.
- [20] Kesten H. *Random Difference Equations and Renewal Theory for Products of Random Matrices*. Acta Mathematica, 131, 207-248, 1973.
- [21] 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.
- [22] LeBaron B. *Building the Santa Fe Artificial Stock Market*. Brandeis University, 2002, <http://people.brandeis.edu/~blebaron/wps/sfisum.pdf>.
- [23] Levy M., Solomon S. *Power-laws are Logarithmic Boltzmann Laws*. International Journal of Modern Physics C, 7, 595, 1996.
- [24] 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.
- [25] 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>.
- [26] Hernández-Pérez R., Angulo-Brown F. , Tun D. *Company size distribution for developing countries*. Physica A 359 (2006) 607–618, 2006.

- [27] Pokorný J., Žemlička M. *Základy implementace souborů a databází*. Karolinum Praha, 2004.
- [28] Ramsden J. J., Kiss-Haypál G. *Company size distribution in different countries*. Physica A, 277:220-227, 2000.
- [29] Simon H., Bonini C. *The Size Distribution of Business Firms*. American Economic Review, 48, 607-617, 1958.
- [30] Singh N. *Monitoring and Hierarchies: The Marginal Value of Information in a Principal-Agent Model*. The Journal of Political Economy, Vol. 93, No. 3, pp. 599-609, June, 1985.
- [31] Malevergne Y., Sornette D. *Extreme Financial Risks: From Dependence to Risk Management*. Springer, 2006.
- [32] Strausz R. *Delegation of Monitoring in a Principal-Agent Relationship*. The Review of Economic Studies, Vol. 64, No. 3, pp. 337-357, July, 1997.
- [33] Tesfatsion L. *Agent-Based Computational Economics: A Constructive Approach to Economic Theory*. Economics Department, Iowa State University, 2005.
- [34] 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>.
- [35] *Statistics of U.S. Businesses: 2005*. U.S. Census Bureau, viewed 10th April 2008, <http://www.census.gov/epcd/susb/latest/us/US--.HTM>.

A Power law exponent α estimation

For reliable analysis of data generated by the application, methods presented by Clauset, Shalizi and Newman in [12] 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 given³⁷.
- 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 test³⁸.

For the details, see their excellent paper [25]. The authors implemented R and Matlab functions for estimations following the methodology³⁹, which came very useful for this work 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 1 and 2.

```
>> 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 1: Matlab power law estimation procedure example

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

  $alpha
```

³⁷For 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.

³⁸They present even a method of deciding whether the data were not drawn from competing distributions, e.g. lognormal or exponential.

³⁹See Aaron Clauset webpage at <http://www.santafe.edu/~aaronc/powerlaws/>.

```
[1] 3.23
```

```
$D
```

```
[1] 0.01376551
```

Listing 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%.

IES Working Paper Series

2008

1. Irena Jindrichovska, Pavel Körner : *Determinants of corporate financing decisions: a survey evidence from Czech firms*
2. Petr Jakubík, Jaroslav Heřmánek : *Stress testing of the Czech banking sector*
3. Adam Geršl : *Performance and financing of the corporate sector: the role of foreign direct investment*
4. Jiří Witzany : *Valuation of Convexity Related Derivatives*
5. Tomáš Richter : *Použití (mikro)ekonomické metodologie při tvorbě a interpretaci soukromého práva*
6. František Turnovec : *Duality of Power in the European Parliament*
7. Natalie Svarciva, Petr Svarc : *Technology adoption and herding behavior in complex social networks*
8. Tomáš Havránek, Zuzana Iršová : *Intra-Industry Spillovers from Inward FDI: A Meta-Regression Analysis*
9. Libor Dušek, Juraj Kopecsni : *Policy Risk in Action: Pension Reforms and Social Security Wealth in Hungary, Czech Republic, and Slovakia*
10. Alexandr Kuchynka : *Volatility extraction using the Kalman filter*
11. Petr Kadeřábek, Aleš Slabý, Josef Vodička : *Stress Testing of Probability of Default of Individuals*
12. Karel Janda : *Which Government Interventions Are Good in Alleviating Credit Market Failures?*
13. Pavel Štika : *Možnosti analytického uchopení reciprocity v sociálních interakcích*
14. Michal Bauer, Julie Chytilová: *A Model of Human Capital, Time Discounting and Economic Growth*
15. Milan Rippel, Petr Teplý : *Operational Risk – Scenario Analysis*
16. Martin Gregor : *The Strategic Euro Laggards*
17. Radovan Chalupka, Petr Teplý : *Operational Risk Management and Implications for Bank's Economic Capital – a Case Study*
18. Vít Bubák : *Value-at-Risk on Central and Eastern European Stock Markets: An Empirical Investigation Using GARCH Models*
19. Petr Jakubík, Petr Teplý : *The Prediction of Corporate Bankruptcy and Czech Economy's Financial Stability through Logit Analysis*

20. Elisa Gaelotti : *Do domestic firms benefit from geographic proximity with FDI? Evidence from the privatization of the Czech glass industry*
21. Roman Horváth, Marek Rusnák : *How Important Are Foreign Shocks in Small Open Economy? The Case of Slovakia*
22. Ondřej Schneider : *Voting in the European Union - Central Europe's lost voice*
23. Fabricio Coricelli, Roman Horváth : *Price Setting and Market Structure: An Empirical Analysis of Micro Data*
24. Roman Horváth, Kamila Koprnická : *Inflation Differentials in EU New Member States: An Empirical Evidence*
25. Michal Franta, Branislav Saxa, Kateřina Šmídková : *Inflation Persistence: Is It Similar in the New EU Member States and the Euro Area Members?*
26. Jakub Seidler : *Implied Market Loss Given Default: structural-model approach*
27. Radovan Chalupka, Juraj Kopecsni : *Modelling Bank Loan LGD of Corporate and SME Segments: A Case Study*
28. Michal Bauer, Julie Chytilová, Jonathan Morduch: *Behavioral Foundations of Microcredit: Experimental and Survey Evidence From Rural India*
29. Jiří Hlaváček, Michal Hlaváček : *Mikroekonomické modely trhu s externalitami, zobecněný Coaseho teorém*
30. Václav Hausenblas, Petr Švarc : *Evoluční dynamika vězňova dilematu: Vliv topologie interakcí a imitace na vývoj kooperativního chování*
31. Peter Marko, Petr Švarc : *Firms formation and growth in the model with heterogeneous agents and monitoring*

All papers can be downloaded at: <http://ies.fsv.cuni.cz>

