Master of Science Thesis

# A Study of the Distribution of Firm Sizes

Applying Methods of Physics on Team Dynamics

Marcus Cordi

Supervisor at CFM, Paris: Jean-Philippe Bouchaud
Supervisor at ENS, Paris: Francesco Zamponi
Supervisor at KTH, Stockholm: Patrik Henelius

Department of Theoretical Physics,
School of Engineering Sciences
Royal Institute of Technology, SE-106 91 Stockholm, Sweden

Stockholm, Sweden 2014

Typeset in LaTeX

Examensarbete inom ämnet teoretisk fysik för avläggande av civilingenjörsexamen inom utbildningsprogrammet Teknisk fysik.

Graduate thesis on the subject Theoretical Physics for the degree of Master of Science in Engineering from the School of Engineering Sciences.

TRITA-FYS 2015:01
ISSN 0280-316X
ISRN KTH/FYS/–15:01–SE


Printed in Sweden by Universitetsservice US AB, Stockholm December 2014

# Abstract

The Agent-Based Model (ABM) proposed Robert L. Axtell in his paper 'Team Dynamics and the Empirical Structure of Firms' (2013) has been successfully reproduced, and various aspects of it have been studied. From this model more simplistic models have been derived, and in particular the power-law behaviour of the distribution of firm sizes generated by these models has been studied. The derived models have been amenable to analytical treatment, and certain results pertaining to the properties of these models have been obtained.



**Key words**: power-laws, firm sizes, econophysics, agent-based models.

# Preface

This thesis is the result of a degree project at the Department of Theoretical Physics at the Royal Institute of Technology (KTH). The work was conducted at Capital Fund Management (CFM), Paris and at the Department of Theoretical Physics, École Normale Supérieure (ENS), Paris during the summer and autumn of 2014.

## Overview

This thesis is divided into five chapters and four appendices.

In Chapter 1 the subject of this thesis is briefly introduced and put into its general scientific context. The purpose of this thesis is also discussed.

In Chapter 2 some necessary preliminaries are presented.

In Chapter 3 various models of varying complexity are presented and used for simulations. The results, including some comments, are also presented.

In Chapter 4 some analytical results and observations are presented, aided by references to the results of the simulations in the previous chapter.

Finally, in Chapter 5 the results are summarised, with some additional comments. A brief outlook for future research is also presented.

The first appendix contains the pseudo-code for the models in Chapter 3, detailing how they were implemented. The second appendix presents some background information on the Binder cumulant, and the third appendix presents a brief mathematical proof, used in Chapter 4.

# Acknowledgements

First and foremost, I would like to thank Professor Jean-Philippe Bouchaud for giving me the opportunity to come to Paris and work on this thesis. He has introduced me to this topic and helped me immensely by proposing different approaches and discussing the results.

Dr. Francesco Zamponi and Dr. Stanislao Gualdi have also helped me substantially with implementing the code for the simulations and solving all the small and large problems encountered on the way to the final results.

I would also like to thank Adrien Bilal for providing me with some initial input and for sharing some preliminary notes, and Professor Patrik Henelius for supporting me during the thesis work.

Finally, I would like to thank my family who have always supported me.

# Contents

Abstract . . . . . . . . . . iii

**Preface** **v**

**Acknowledgements** **vii**

**Contents** **ix**

**1 Introduction** **1**

**2 Preliminaries** **3**
2.1 Power-Laws . . . . . . . . . . 3
2.1.1 Definitions . . . . . . . . . . 3
2.1.2 Power-Law Distributions in Empirical data . . . . . . . . . . 5
2.2 Zipf Distribution of U.S. Firm Sizes (Axtell 2001) . . . . . . . . . . 5
2.3 Team Dynamics and the Empirical Structure of U.S. Firms (Axtell 2013) . . . . . . . . . . 6
2.4 Barabási-Albert's model - preferential attachment . . . . . . . . . . 8
2.5 Master Equation . . . . . . . . . . 9

**3 Simulations** **11**
3.1 Axtell Model . . . . . . . . . . 11
3.2 Axtell Mean Field Model . . . . . . . . . . 20
3.2.1 Base Case Axtell Mean Field Model . . . . . . . . . . 20
3.2.2 Reduced Axtell Mean Field Model . . . . . . . . . . 22
3.3 Axtell Mean Field Model without Utility . . . . . . . . . . 27

**4 Analytical Results** **37**
4.1 Calculation of the Optimal Utility . . . . . . . . . . 37
4.2 Rationale for the Preferential Attachment Mechanism . . . . . . . . . . 39
4.3 Calculation of $\langle E \rangle_n$ . . . . . . . . . . 40
4.4 Derivation of the Master Equation . . . . . . . . . . 42
4.5 Alternative Derivation of the Master Equation . . . . . . . . . . 47
4.6 Criterion for Condensation . . . . . . . . . . 48

**5 Conclusion** **53**
5.1 Summary of Results . . . . . . . . . . . . . . . . . . . . . . . . . . 53
5.2 Additional Comments . . . . . . . . . . . . . . . . . . . . . . . . . 54
5.3 Outlook . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 55

**A Pseudo-code** **57**

**B Binder Cumulant** **65**

**C Distribution of the Minimum of Exponential Variables** **69**

**Bibliography** **72**

# Chapter 1

# Introduction

In complex systems, there are generally many interacting units (agents). The interactions between these units lead to phenomena which would not be expected by just observing the behaviour of one individual unit. Systems like these have been extensively studied in physics. A typical example is spin glasses [8].

There are other domains, than physics, in which complex systems, or behaviour reminiscent of complex systems, appear. Examples of this are ecological systems in biology and socio-economic systems. This thesis will focus on the latter.

Agent-Based Models (ABMs) provide an important tool for analysing the aggregate, collective behaviour of various agents in economic systems. The behaviour displayed in these models may be analogous to the behaviour of systems encountered and modelled in, for example, statistical mechanics, and a similar terminology may be applied (phase-space, phase-transitions, criticality, etc.).

Commonly used macro-economic models provide an acceptable description of the phenomena in 'normal' situations, but fail to explain the complex behaviour observed particularly in times of economic instability. In this regard, the science of complexity and ABMs provide a promising framework to take into account some potentially important aspects, such as the heterogeneity of economic agents, their strong interactions that can lead to dramatic collective phenomena, out-of-equilibrium dynamics and network effects.

The material in this essay is substantially built on the work of Professor Robert L. Axtell[1], from his initial empirical study in 2001 [3] on the distribution of firm sizes to the development of his ABM in 2013 [2], which generates a non-trivial distribution of firms sizes. It is also partially based on previously gained expertise in the context of the CRISIS project[2] in which the teams of two of the supervisors of this thesis have studied stylized ABMs of the macro-economy through numerical simulations and methods from statistical mechanics and complex systems.

[1] George Mason University, Krasnow Institute for Advanced Study, Department of Computational Social Science and Center for Social Complexity, and Santa Fe Institute

[2] www.crisis-economics.eu

The purpose of this thesis is to investigate the application of (statistical) models used in physics to problems concerning socio-economic data, primarily with ABMs. The mechanisms behind the ABM presented by Axtell will be explored more in depth, and with this model as foundation various more simplistic models, which highlight different facets of the original model, will be proposed and studied. In particular, the distribution of firm sizes will be studied, aided by computer simulations and, when it is deemed possible, a more general mathematical characterisation. In this respect statistical mechanics and complex systems offer a useful set of tools and concepts.

# Chapter 2

# Preliminaries

In this chapter some necessary background information will be presented.

## 2.1 Power-Laws

A major theme in this thesis is the study of power-laws. A brief introduction to the essential mathematical tools used in their characterisation is therefore necessary.

### 2.1.1 Definitions

If the quantity $x$ is drawn from a power-law distribution, it has the following probability distribution

$$p(x) \propto x^{-(1+\alpha)} \tag{2.1}$$

where $\alpha$ is a constant parameter of the distribution known as the scaling parameter[1]. One typically finds in many socio-economic distributions, characterised by large disparities (e.g., frequency of use of words, net worth of Americans, and number of books sold in the U.S.), a scaling parameter in the range of $1 < \alpha < 2$, although exceptions do occur [9, 16][2].

It is seldom one sees empirical phenomena obeying power-laws for all values of $x$, instead the power-law might only apply for values greater than some minimum $x_{\min}$. The tails are usually of significant interest.

Power-law distributions can represent either continuous real numbers or a discrete set of values, usually positive integers. The primary focus of this thesis will be on discrete distributions since it is the distribution of firm sizes, in terms of number of employees, which is studied.

[1] $-(1+\alpha)$ is used here as the exponent, instead of the usual choice of $-\alpha$, in order to maintain a notation which is consistent with the one employed by Axtell.

[2] $1 < \alpha < 2$ implies a well-defined mean, but not a well-defined finite variance, i.e, 'black swan' behaviour.

The continuous power-law distribution may be represented by a probability density $p(x)$ such that

$$p(x)\,\mathrm{d}x = \Pr(x \leq X < x + \mathrm{d}x) = Cx^{-(1+\alpha)}\,\mathrm{d}x, \tag{2.2}$$

where $X$ is the observed value and $C$ is a normalisation constant. It is obvious that this density diverges as $x \to 0$, and it is therefore necessary to have a lower bound $x_{\text{min}} > 0$. If $\alpha > 0$, the normalisation constant may then be calculated as

$$p(x) = \frac{\alpha}{x_{\text{min}}} \left( \frac{x}{x_{\text{min}}} \right)^{-(1+\alpha)}. \tag{2.3}$$

Similarly for the discrete case, with a probability distribution which looks like

$$p(x) = \Pr(X = x) = Cx^{-(1+\alpha)}, \tag{2.4}$$

one finds that

$$p(x) = \frac{x^{-(1+\alpha)}}{\zeta(1+\alpha, x_{\text{min}})}, \tag{2.5}$$

where

$$\zeta(s, q) = \sum_{n=0}^{\infty} \frac{1}{(q+n)^s} \tag{2.6}$$

is the Hurwitz zeta function.

It is often useful to consider the complementary cumulative distribution function of a power-law distributed variable, which is defined for both the continuous and discrete case as $P(x) = \Pr(X \geq x)$. In the continuous case, $P(x)$ is found to be

$$P(x) = \int_x^{\infty} p(x')\,\mathrm{d}x' = \left( \frac{x}{x_{\text{min}}} \right)^{-\alpha}, \tag{2.7}$$

and in the discrete case

$$P(x) = \frac{\zeta(1+\alpha, x)}{\zeta(1+\alpha, x_{\text{min}})}. \tag{2.8}$$

Since the formulas for the continuous distributions tend to be easier to handle than those for the discrete distributions, one often approximates a discrete distribution with a continuous one. There are, however, several different ways of doing this; one relatively dependable way of doing this is to assume that the values of $x$ were generated from a continuous power law and then rounded to the nearest integer. The continuous approximation will be assumed throughout this thesis, and it will generally be assumed that it is justified.

In this thesis the complementary cumulative distribution function will be used to represent distributions visually. This is partially because the complementary cumulative distribution function is more robust against fluctuations due to finite sample sizes, particularly in the tail of the distribution, than the probability density function.

### 2.1.2 Power-Law Distributions in Empirical data

A frequently employed graphical method of identifying a power-law relation (and also of obtaining an estimate of the scaling parameter $\alpha$ and the lower-bound of the scaling region $x_{\min}$) in empirical data is the log-log plot. Taking the logarithm of both sides of Equation (2.2) yields

$$\ln p(x) = \ln C - (1+\alpha)\ln x, \tag{2.9}$$

i.e., it should follow a straight line on a log-log plot. Having plotted the data of interest, one may then by, for example, visual inspection determine from where the data follows a straight line, and thus determine $x_{\min}$. The scaling parameter $\alpha$ may be estimated by examining the absolute slope of the straight line.

There are however, understandably, several drawbacks to this method. One, which is very important, is that other distributions, for example, the log-normal, exponential or stretched exponential, may generate results which also give approximately straight lines. Establishing a straight line on a log-log plot should thus be considered a necessary but not sufficient condition for a power-law.

This method will, however, be used frequently in this thesis as an initiatory method of investigating the existence of a power-law relation, given the exploratory character of this thesis. The investigation will primarily be conducted by visually inspecting the log-log plots of the complementary cumulative distribution function, which avoids the introduction of an implicit bias in the representation of the data.

Methods which are more rigorous and accurate in obtaining the parameters of a power-law distribution, and of determining whether or not a given data set really does follow a power-law, are explored in a paper written by A. Clauset et al. [9].

## 2.2 Zipf Distribution of U.S. Firm Sizes (Axtell 2001)

In a report published in 2001 by Professor Robert L. Axtell [3], he claims that analyses of firm sizes have historically been based on data of limited samples of small firms, which typically can be described by log-normal distributions. In this report, however, he presents an empirical power-law distribution observed in the distribution of firm sizes of the entire population of tax-paying firms in the United States in 1997.

Using the following notation, the tail cumulative distribution function for a discrete Pareto-distributed random variable[3] S is

$$\Pr[S \geq s_i] = \left(\frac{s_0}{s_i}\right)^{\alpha}, s_i \geq s_0, \alpha > 0 \tag{2.10}$$

where $s_0$ is the minimum size of the random variable ($s_0 = 1$ for firms in this case), and in the special case of $\alpha = 1$ it is known as the Zipf distribution. When

[3] It is here assumed that the discrete distribution may simply be replaced by a continuous one.

investigating the distribution of U.S. firm sizes, including firms consisting of only one employee, he finds by Ordinary Least Squares (OLS) that $\alpha \approx 1.059$ with $R^2 = 0.992$[4].

An important conclusion made in Axtell's report is that since the power-law distribution well describes the data of firms with one to $10^6$ employees, it suggests that a universal mechanism behind the growth of firms operates on firms of all sizes, and that the individual employee is the fundamental unit of analysis. This is of great importance for the model described by Axtell in his 2013 article (presumably) aimed at replicating these empirical results.

## 2.3 Team Dynamics and the Empirical Structure of U.S. Firms (Axtell 2013)

In a working paper from 2013 [2], Professor Robert L. Axtell presents an agent-based model (ABM) which closely reproduces empirical results on the population of U.S. firms. The model manages to capture several important aspects of the dynamics of firms; they grow, they get smaller, new firms are started by entrepreneurs breaking off from existing ones, and finally they perish. The basic idea of the simulation of the model will be presented here; a more detailed pseudo-code will presented in Appendix A.

The economy simulated consists of a fix population of $N$ agents, and initially each agent is self-employed (singleton firms). An agent $i$ works with a certain effort $e_i \in [0, \omega]$ while working in a firm of $n$ agents, either singleton or non-singleton. An important consequence is thus that it is possible for an agent to be part of a firm and work without any effort ($e_i = 0$); a so called 'free rider'. It is assumed that all the agents have the same maximum effort level $\omega = 1$.

The other agents who are in the same firm as agent $i$ collectively work with an effort $E_{-i}$, which means that the total effort of the firm may be written as $E = e_i + E_{-i}$. The agents in the firm of agent $i$ produce an output as a function of their collective effort $E$. The output function is defined as

$$O(e_i, E_{-i}) = O(E) = aE + bE^{\beta}, \tag{2.11}$$

where $a \sim U\left[0, \frac{1}{2}\right]$, $b \sim U\left[\frac{3}{4}, \frac{5}{4}\right]$ and $\beta \sim U\left[\frac{3}{2}, 2\right]$ are the output parameters in the 'base case' configuration of the computational model. Each time a new firm is created, new output parameters are drawn[5].

What is of great importance here is that if $b > 0$ and $\beta > 1$, then there are increasing returns to production, which means that agents working together can produce more than they can as individuals. This means that agents have incentives

[4] It is not appropriate to use $R^2$ in this instance, see [9] for an explanation why.

[5] Axtell is not explicitly clear that this is how he uses these output parameters (although it seems to be the most reasonable); another alternative would be to assume that each agent has an intrinsic set of output parameter $a, b$ and $\beta$, and if an agent starts a new firm, these parameters make up the parameters of the output function of the firm as long as it exists.

to team-up and create firms with other agents. If $a > 0$ there are also constant returns. In other words, the total output function of a firm, in the 'base case' configuration, consists of both constant and increasing returns.

The output produced by a firm is shared equally among the agents belonging to it[6]. This is reflected in the utility function of an agent. The utility function of agent $i$ thus has a Cobb-Douglas form for income and leisure

$$u_i(e_i, E_{-i}) = \left(\frac{O(e_i, E_{-i})}{n}\right)^{\theta_i} (\omega - e_i)^{1-\theta_i}, \tag{2.12}$$

where $\theta_i \sim U[0,1]$ is a fix parameter assigned uniquely to each agent. $\theta_i$ is a parameter which determines what preference an agent has for income (work) or leisure, with a $\theta_i$ closer to 1 indicating a higher preference for income, and a $\theta_i$ closer to 0 indicating a preference for leisure, as is evident in Equation (2.12). Each agent continually makes decisions to remain in their current firm, join another firm or start a new firm, always with the aim of maximising its utility function by choosing an optimal effort $e_i^* \in [0, \omega]$. In other words, agents move between teams or start new teams when it is in their self-interest

Time is discrete, $t \in \mathbb{N}$, and for each point in time each agent is activated, in the sense of making a choice to join another firm, start a new firm or remain in its current firm, with a 4% activation probability. If the agent is not activated it remains in its current state (i.e., in its current firm). In other words, in one time step all agents have been subject to the activation process, and 4% of them have been activated and made a choice of whether to remain in their current state or change it.

Each agent has a fix social network consisting of $\nu_i$ other agents assigned randomly as $U \sim [2, 6]$ in an Erdős-Renyi graph. If an agent is activated it calculates its utility maximising choice, which may be to join one of the firms of its $\nu_i$ 'neighbours', start a new firm in which it is the only agent or remain in its current firm. The agent then chooses the option which yields the greatest utility, and all the relevant information which has subsequently changed is updated.

When Axtell ran the simulation, with the 'base case' configuration, as outlined above, and $N = 1.2 \cdot 10^8$ as the number of agents, long enough, i.e., until an approximately stationary macro state was reached[7], he claims to have obtained a distribution of firm sizes with statistics comparable to empirical data on U.S. firms. Specifically, he claims that the U.S. data are well fit by a power-law with the exponent $\alpha \approx 1.06$ in the complementary cumulative distribution function [3], and that the data generated by the simulation is in agreement[8]. This is central to this thesis.

[6]It is implicitly assumed here that at the end of each period all the output of a firm is sold for unit price, and each agent in the firm receives an equal share.

[7]Presumably, the stationary state may be defined as when the distribution of firms remains approximately constant.

[8]A line with the slope $-2.06$ is drawn in a plot of the probability density function.

## 2.4 Barabási-Albert's model - preferential attachment

In a paper published in 1999, Albert-László Barabási and Réka Albert presented a model which reproduces the stationary power law (scale-free) distributions observed empirically in growing networks [4]. Examples of such networks are the links of web pages, collaboration of actors in films, the electrical power-grid of the western U.S. and the citations between scientific papers.

If $p(k)$ is defined as the probability that a vertex in a network is connected with $k$ edges to other vertices (repetitions are ignored), then many dynamic networks have the property that $p(k)$ decays as a power law, i.e.,

$$p(k) \sim k^{-(1+\alpha)}, \tag{2.13}$$

indicating that large networks self-organise into a scale-free state.

There are two main features which are found to generate a scale-free power-law distribution in this model; the network expands continuously as new vertices are added, and new vertices attach themselves preferentially to already well-connected vertices, i.e., 'preferential attachment'[9].

If the network of citation between scientific papers is studied more closely, then vertices may represent scientific papers published in refereed journals, and the edges links to the articles cited in a paper. The intuition behind the dynamics is then that as new papers are added, it is more likely that they will cite an article which already has many citations. It has been shown that the probability that a paper is cited $k$ times follows a power-law with the scaling parameter $\alpha \approx 2$ [14].

As a simple illustration of how this would work, a network may be introduced in which vertices, indexed by the order of their birth, are continually born and form $m$ edges to existing vertices [1]. Let $k_i(t)$ represent the degree (number of edges) of a vertex, born at time $i$, at time $t$.

If uniform attachment is assumed, it entails that each new vertex at each time $t$ spreads its $m$ new edges randomly over the $t$ existing nodes at time $t$. This leads to the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t} k_i(t) = \frac{m}{t}, \tag{2.14}$$

with the initial condition $k_i(i) = m$ for all $i$[10]. From the solution of this differential equation, it is possible to derive an approximation of the degree distribution, which is exponential and stationary,

$$p(k) \sim e^{1-\frac{k}{m}}. \tag{2.15}$$

It is also possible to assume that the system starts with a group of $m$ vertices all connected to one another. New vertices with $m$ edges are still continually added

[9] Mechanisms similar to preferential attachment are also known as 'cumulative advantage' or 'rich-get-richer effect'

[10] Repetitions are ignored

to the network, but in this case the probability $\Pi(k_i(t))$ that an existing vertex with $k_i(t)$ edges will get a connection with one edge of the new node is proportional to the degree of the existing vertex, i.e., the system exhibits linear preferential attachment. Specifically, this probability may be expressed as

$$\Pi(k_i(t)) = \frac{k_i(t)}{\sum_j k_j(t)}, \tag{2.16}$$

which leads to the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t} k_i(t) = \frac{k_i(t)}{2t} \tag{2.17}$$

with the initial condition $k_i(i) = m$. From the solution of this differential equation it is possible to derive an approximation of the degree distribution, which follows a stationary power-law,

$$p(k) \sim k^{-3}. \tag{2.18}$$

An important distinction between the model proposed by Barabási and Albert, and the models studied in this thesis, is that their model is dynamic, i.e., constantly growing, while the models in this thesis have a fix number of vertices (agents). It does however seem reasonable that the dynamics, with a mechanism similar to the preferential attachment described above, might be similar. The scenario would then be that a firm might be regarded as a selection of vertices, all connected with one another, and where a larger number of employees implies a higher probability of an agent (vertex), searching for a new firm to join, joining the firm. Although this mechanism is not technically the same mechanism described by Barabási and Albert, it will be referred to as 'preferential attachment'. Barabási and Albert note however that they only witness scaling when there is linear preferential attachment.

In the simulations they performed, approximately $\alpha = 1.9 \pm 0.1$ was obtained.

## 2.5 Master Equation

Based on the Barabási-Albert model, Dorogovstev, Mendes and Samukhin used a different approach, the master equation, to obtain rigorous asymptotic results for the mean degree of the vertices [1][11].

If a similar set-up is used, and $p(k)$ still denotes the probability that a vertex in a network is connected with $k$ edges to other vertices, or equivalently the fraction of vertices in the network with degree $k$, then the probability that a new edge attaches to a vertex of degree $k$ is given by

$$\Pi(k) = \frac{kp(k)}{\sum_j jp(j)} = \frac{kp(k)}{2m}, \tag{2.19}$$

since the mean degree of the network is $2m$.

[11] A more rigorous derivation is given in [10].

The mean number of vertices of degree $k$ that gain an edge when a new vertex with $m$ edges is added to the network is thus given by $\frac{kp(k)}{2}$. This means that if $n$ is the total number of vertices, then the number of nodes with degree $k$, given by $np(k)$, will decrease by this amount.

However, the number of vertices with degree $k$ also increases because of an influx from vertices of degree $(k-1)$ which have received a new edge. An exception is however vertices of degree $m$ which have an influx of exactly 1.

Let $p_n(k)$ be the value of $p(k)$ when there are $n$ vertices in the network. It then holds that

$$\begin{cases} (n+1)p_{n+1}(k) - np_n(k) = \frac{1}{2}(k-1)p_n(k-1) - \frac{1}{2}kp_n(k), & k > m \\ (n+1)p_{n+1}(k) - np_n(k) = 1 - \frac{1}{2}kp_n(k), & k = m, \end{cases} \tag{2.20}$$

where the stationary solution $p(k) = p_n(k) = p_{n+1}(k)$ satisfies the equation

$$p(k) = \begin{cases} \frac{1}{2}(k-1)p(k-1) - \frac{1}{2}kp(k), & k > m \\ 1 - \frac{1}{2}kp(k), & k = m. \end{cases} \tag{2.21}$$

It can then be shown that

$$p(k) = \frac{2m(m+1)}{(k+2)(k+1)k}, \tag{2.22}$$

which in the limit of large $k$ results in a power-law distribution given by

$$p(k) \sim k^{-3}. \tag{2.23}$$

This is in agreement with the result obtained by Barabási and Albert.

The master equation approach, outlined in this section, will serve as inspiration for the attempt at formulating a master equation for the Axtell Model.

# Chapter 3

# Simulations

In this chapter various models, and the results obtained from the simulations of these models, are presented and briefly commented on.

## 3.1 Axtell Model

One of the primary goals of this thesis was to replicate the results obtained by Professor Robert L. Axtell in his paper 'Team Dynamics and the Empirical Structure of U.S. Firms' [2]. This model is referred to as the 'Axtell Model'. Gradually, this model was altered and reduced into less complex ones in order to study various facets of the model, and to aid and facilitate an attempt at an analytical treatment in Chapter 4. The 'base case' configuration of the parameters of the Axtell Model, as outlined in Section 2.3, will be referred to as the 'Base Case Axtell Model'.

The activation probability of 4% used by Axtell, was discarded completely in the 'Axtell Model' (and all other models). The reasoning behind this is that the activation probability only changes the unit of time, i.e., the dynamical evolution is exactly the same but with a different time unit. Therefore, what is referred to as one time step is when a random agent has been chosen $N$ times, and done its choices.

The results obtained by Axtell and presented in his 2013 article were replicated quite successfully, it seems. In Figure 3.1 the complementary cumulative distribution function (referred to as $P(n)$) for the Base Case Axtell Model at $t = 1000$ with $N = 1.2 \cdot 10^8$ agents, which was the population size used by Axtell, has been drawn.

Since a stationary state, or at least a state with a non-trivial distribution, is reached rather quickly (approximately after 4-5 time steps) in the Axtell Model, it should be possible to accumulate data from various points in time, as in Fig. 3.2[1].

[1] In order to be sure to only gather data from the stationary state, only data from $t = 100$ and later was used.

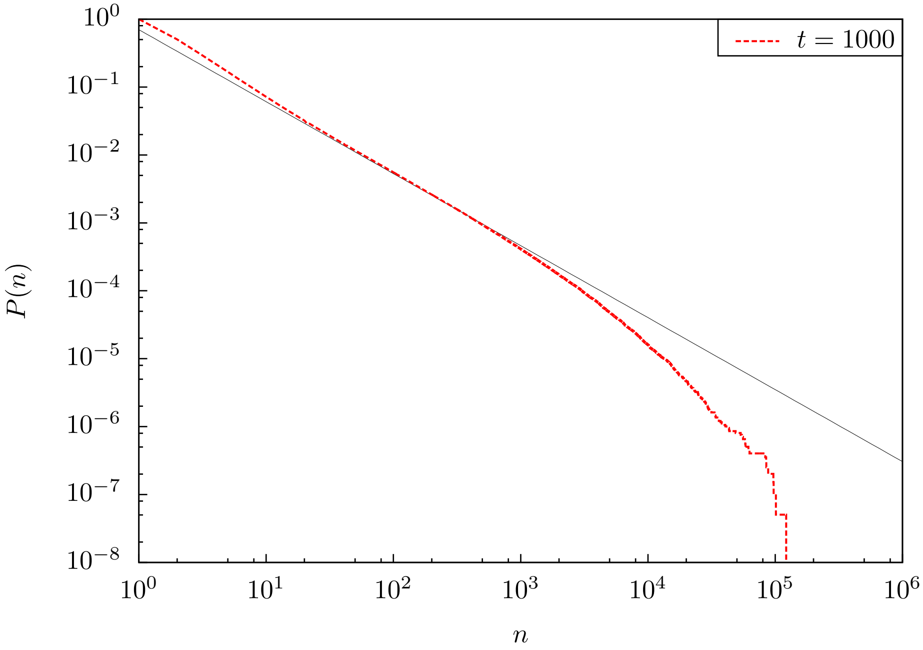


**Figure 3.1:** The complementary cumulative distribution function of the Base Case Axtell Model for $N = 1.2 \cdot 10^8$ when $t = 1000$. The continuous straight line has a slope corresponding to $\alpha = 1.06$ and has not been fitted.

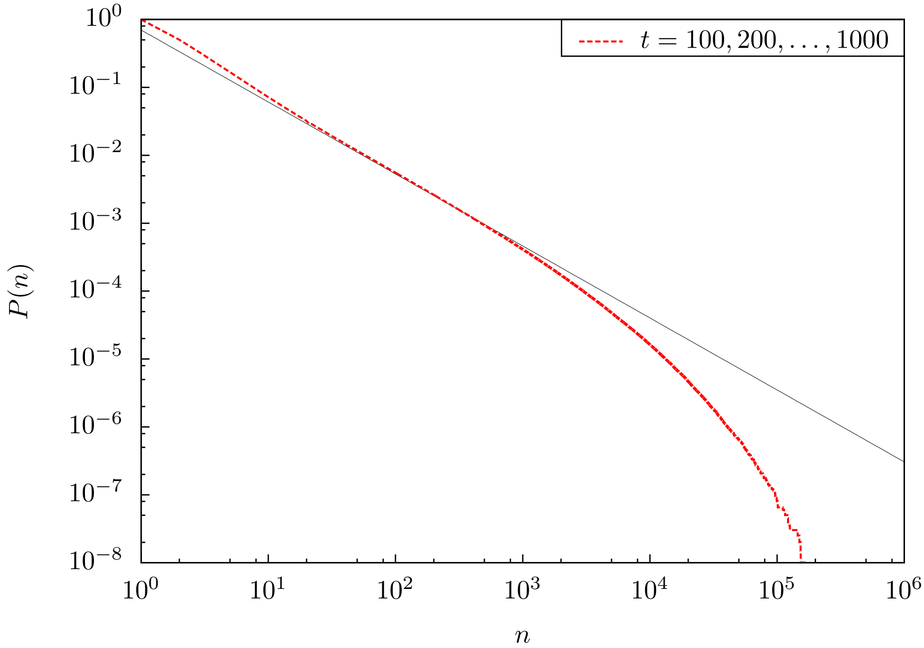


**Figure 3.2:** The complementary cumulative distribution function of the Base Case Axtell Model for $N = 1.2 \cdot 10^8$ when $t = 100, 200, \ldots, 1000$ (accumulating data). The continuous straight line has a slope corresponding to $\alpha = 1.06$ and has not been fitted.

This makes the distribution smoother, and hereinafter the plots will in general utilise this procedure.

It was also apparent that the presence of the largest firm affects the complementary cumulative distribution function; when examining the largest firm at $t = 100, 200, \ldots, 1000$, it seems to fluctuate in an interval without ever converging towards a final value. Therefore, it also seemed reasonable to accumulate data established at various points in time.

Although Axtell does not provide much quantitative material to examine whether this endeavour was successful or not, it seems as if this goal was achieved in the sense of reaching approximately the same distribution of firm sizes, as is evident in the plot of the probability density function (referred to as $p(n)$) in Figure 3.3[2].

[2] A line with the slope −2.06 is drawn in a plot of the probability density function. The slope of this line corresponds to the exponent observed in the distribution of U.S. firms, namely $\alpha \approx 1.06$. The probability density function in Figure 3.3 is quite similar to the one obtained by Axtell. The probability density might however be misleading, and in this thesis the complementary cumulative distribution function will primarily be used.

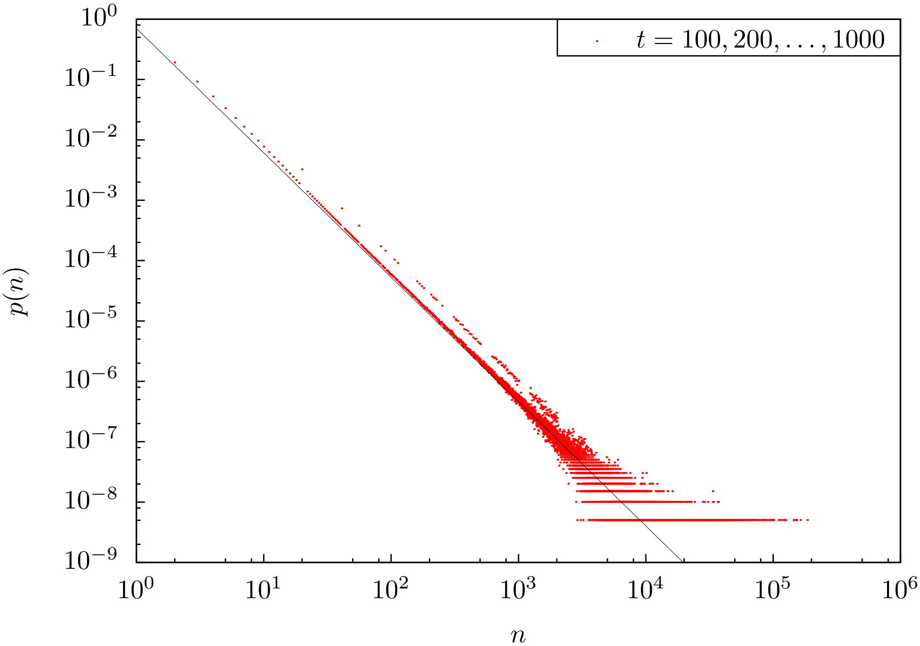


**Figure 3.3:** The probability density function of the Base Case Axtell Model for $N = 1.2 \cdot 10^8$ when $t = 100, 200, \ldots, 1000$ (accumulating data). The continuous straight line has a slope corresponding to $\alpha = 1.06$ and has not been fitted.

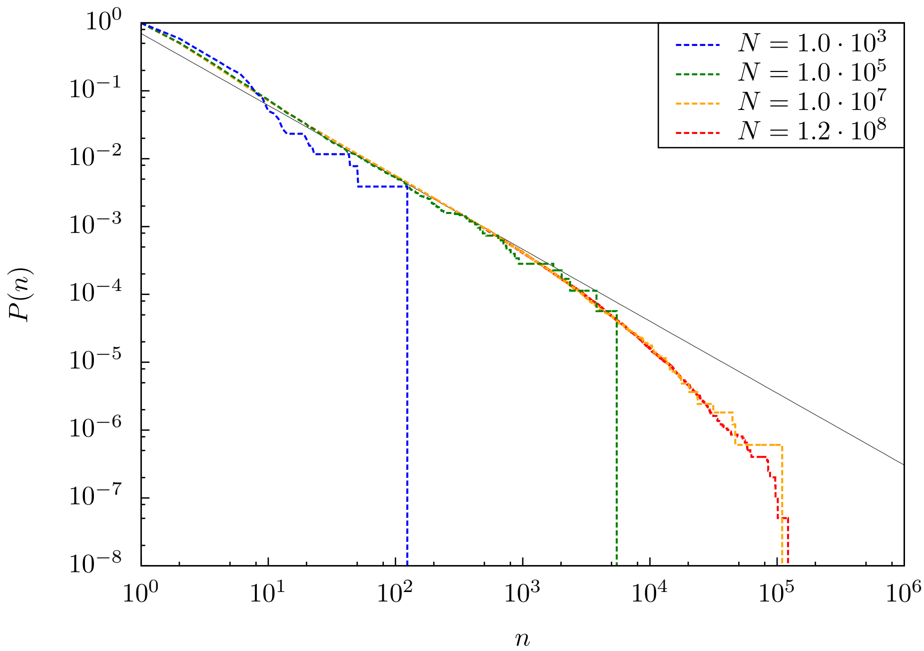


**Figure 3.4:** The complementary cumulative distribution function of the Base Case Axtell Model for $N = 10^3, 10^5, 10^7$ when $t = 1000$. The continuous straight line has a slope corresponding to $\alpha = 1.06$ and has not been fitted.

The implementation is described briefly in Section 2.3 and more extensively in Appendix A.

The same procedure for smaller population sizes was repeated in Figure 3.4, when $t = 1000$, and Figure 3.5, when $t = 100, 200, \ldots, 1000$ (accumulating data). It is of interest to note that the distributions for $N = 10^7$ and $N = 1.2 \cdot 10^8$ seem to be quite similar.

It is evident that the behaviour of the system seems to be relatively invariant with regards to the population size $N$. Specifically it seems as if the distributions for different population sizes approximately converge for smaller $n$ as more data is accumulated.

It was noted that if $N = 10^5$ was used as the population size, a reasonable compromise was reached between the time of the execution of the simulation and the validity of the results. If only $N = 10^3$ was used as the population size, significantly more data from not just the final state of the system, but also previous stationary states, had to be gathered to reach any valid results. For brevity's sake, $N = 10^5$ will henceforth primarily be the default choice for the population size, unless certain circumstances warrant the use of either smaller or larger population sizes.

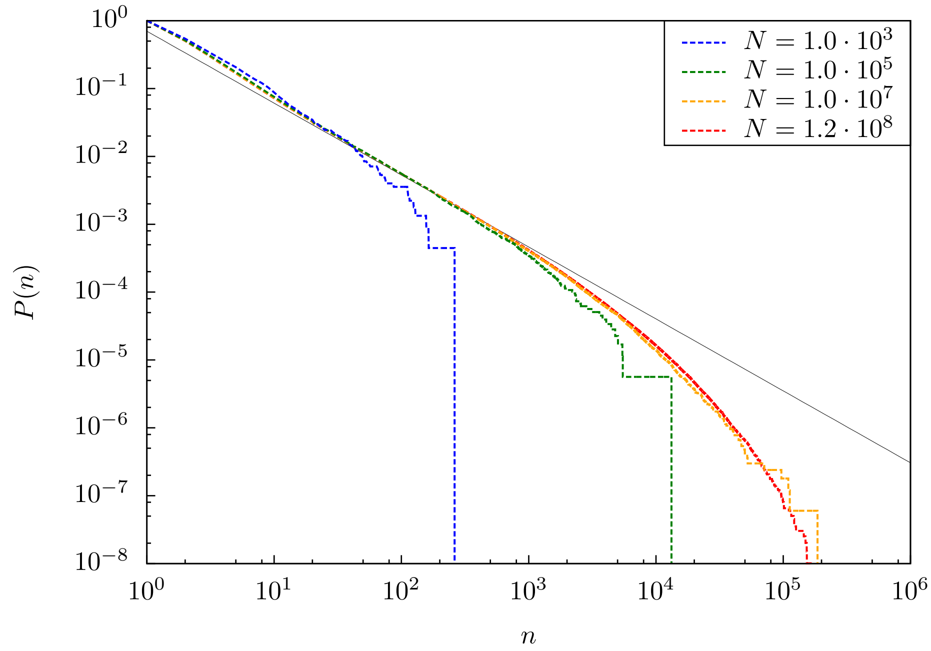


**Figure 3.5:** The complementary cumulative distribution function of the Base Case Axtell Model for $N = 10^3, 10^5, 10^7$ when $t = 100, 200, \ldots, 1000$ (accumulated data). The continuous straight line has a slope corresponding to $\alpha = 1.06$ and has not been fitted.

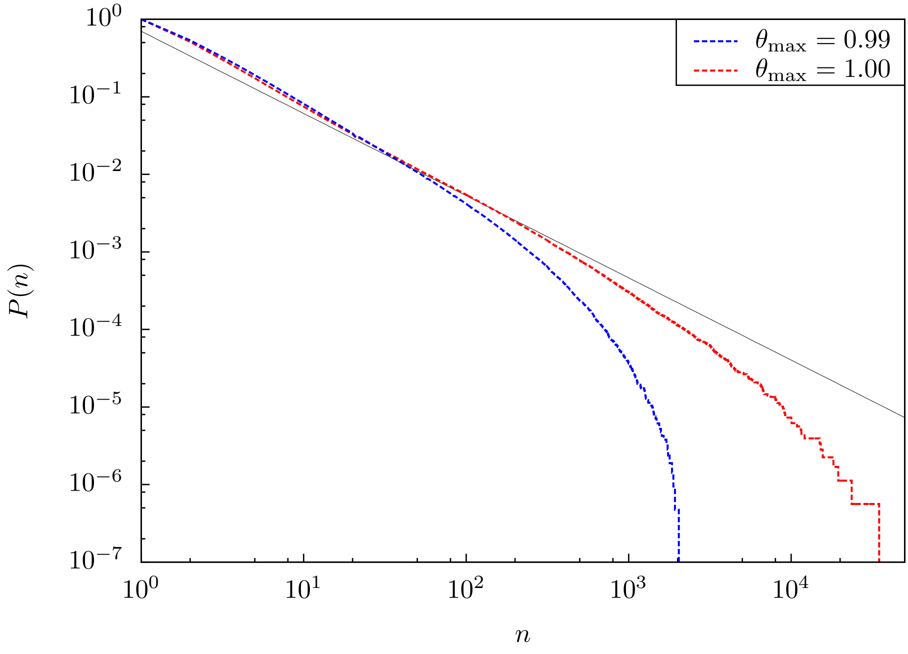


**Figure 3.6:** The complementary cumulative distribution function of the Base Case Axtell Model ($\theta_{max} = 1.00$) and the Base Case Axtell Model with $\theta_i$ restricted ($\theta_{max} = 0.99$) for $N = 10^5$ when $t = 100, 200, \ldots, 10\,000$ (accumulated data). The continuous straight line has a slope corresponding to $\alpha = 1.06$ and has not been fitted.

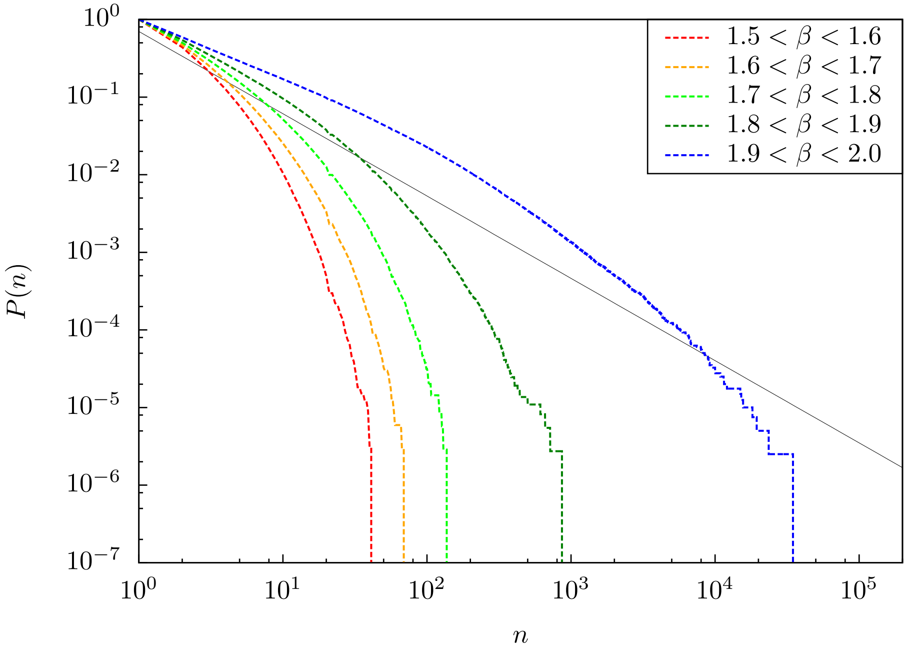


**Figure 3.7:** The complementary cumulative distribution function of the Base Case Axtell Model for $N = 10^5$ when $t = 100, 200, \ldots, 10\,000$ (accumulated data) where firms with the output parameter $\beta$ in a certain interval have been extracted and plotted in their own graphs. The continuous straight line has a slope corresponding to $\alpha = 1.06$ and has not been fitted.

Figure 3.6 quite clearly suggests that only slightly restricting $\theta_{max} = 1$ to $\theta_{max} = 0.99$ instead has a considerable effect on the distribution, especially in terms of inhibiting the appearance of significantly large firms. An analytical explanation as to why agents with a $\theta_i$ close to 1 give rise to larger firms, is given in Section 4.2.

In a similar fashion, it is shown in Figure 3.7 the importance of the output parameter $\beta$ for the appearance of large firms. Specifically, it seems as if the larger firms all have $\beta \approx 2$. A possible analytical explanation for this is given in Section 4.2.

Figure 3.8 shows how the average effort of a firm $\langle E \rangle_n$ depends on the number of employees in the firm, $n$. The firms producing zero effort were discarded. This is because it was noted that there would appear firms with one single employee performing more or less all the work, with the rest of the employees essentially being free-riders. Eventually the employee doing all the work would leave the firm, and would thus leave an 'empty' firm which would soon after disintegrate. Repeatedly

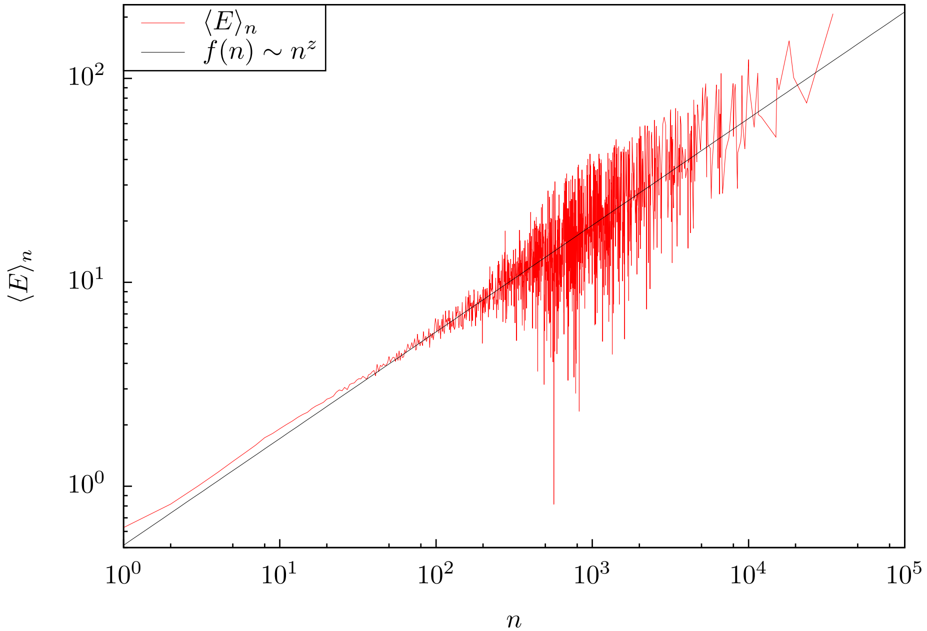


**Figure 3.8:** The average collective effort $\langle E \rangle_n$ of a firm with $n$ employees. The data has been gathered from a simulation of the Base Case Axtell Model with a population size of $N = 10^5$ when $t = 100, 200, \ldots, 10\,000$ (accumulated data). Fitting by least squares from $n = 100$ yields $z \approx 0.52$.

it happened that data was gathered (a 'snapshot') precisely after the moment the hard-working employee had left the firm, thereby creating an unrepresentative firm in terms of the collective effort.

Figure 3.8 quite strongly suggests it approximately holds that

$$\langle E \rangle_n \sim n^z, \tag{3.1}$$

where $z = \frac{1}{2}$. This further validates the implementation of the Axtell Model since Axtell finds that the average firm output scales linearly with size, which implies constant returns to scale [2, p. 24], i.e.,

$$O(E) \sim n. \tag{3.2}$$

Assuming Equation 2.11 may be written approximately as

$$O(E) \sim E^\beta, \tag{3.3}$$

one finds with $\beta \approx 2$ and $z \approx \frac{1}{2}$ that

$$O(E) \ \sim (n^z)^\beta \sim n, \tag{3.4}$$

which is consistent with Axtell's findings. An analytical explanation for Equation (3.1) is given in Section 4.3.

The reason why $\langle E \rangle_n$ is not smooth for larger $n$ is because there is insufficient data to establish a relevant average for larger $n$; in other words, it is rare to find a firm with exactly the same number of employees as another when $n$ is large.

## 3.2 Axtell Mean Field Model

### 3.2.1 Base Case Axtell Mean Field Model

In order to enable an analytical treatment of the Axtell model, certain features of the original model were gradually made more simplistic or removed, and new, less complex, models were developed.

The first feature modified was the network agents use to search for new firms to join. All of the other features of the original model were kept intact though, including the output parameters. This model is referred to as the 'Base Case Axtell Mean Field Model'.

In the Base Case Axtell Mean Field Model the agents are not connected with each other through an Erdős-Renyi network as in the 'Axtell Model', instead an agent randomly selects one of the other $(N-1)$ agents and examines if it is beneficial to join the firm which this agent belongs to.

In order to examine the validity of the results of the Base Case Axtell Mean Field Model, the parameter $\Pi$, which is defined as the probability that the agent will choose another random agent and examine if it is beneficial to join this randomly chosen agent's firm, and if so, join the firm, instead of examining its neighbours'

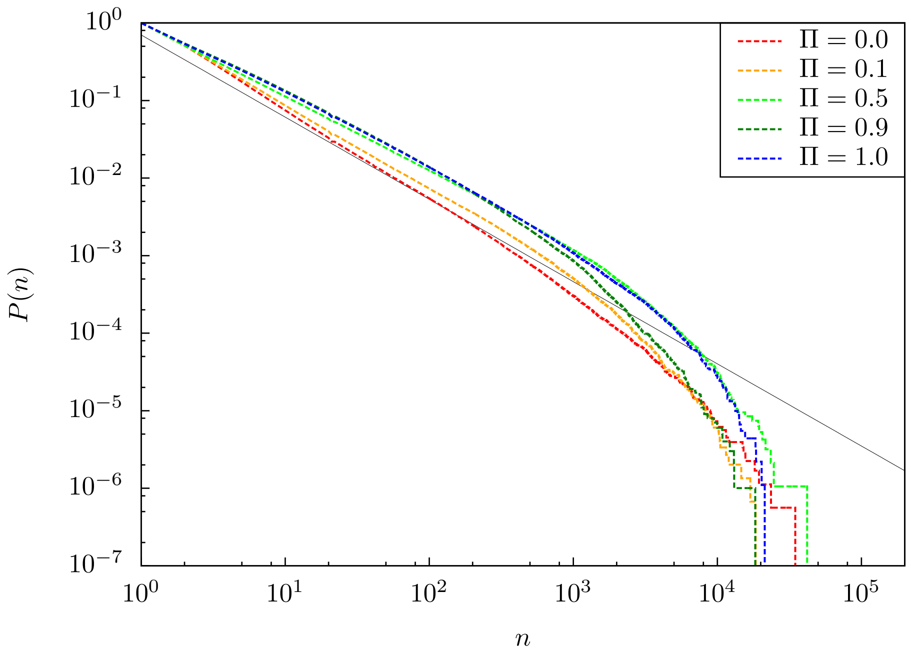


**Figure 3.9:** The complementary cumulative distribution function of the Base Case Axtell Mean Field Model for $N = 10^5$ when $t = 100, 200, \ldots, 10\,000$ (accumulated data) with $\Pi = 0.0, 0.1, 0.5, 0.9, 1.0$. The continuous straight line has a slope corresponding to $\alpha = 1.06$ and has not been fitted.

firms, is introduced (see Appendix A for more details). In other words, $\Pi$ is used to interpolate between the models, where $\Pi = 0$ corresponds to the Axtell Model, and $\Pi = 1$ corresponds to a complete mean field approximation.

The rationale for the term 'mean field' is that in statistical mechanics the mean field approximation, for example in the Ising model, involves replacing the interaction with the other magnetic atoms, with a calculated average magnetic field [13]. This approximation is performed as a consequence of the fact that all atoms interact with each other, and this is what the agents in the Axtell Mean Field Model do. In this sense it is justified to call the Axtell Mean Field Model, completely or partially without the Erdős-Renyi graph, a mean field approximation[3].

In Figure 3.9 the complementary cumulative distribution functions for the original Base Case Axtell Model ($\Pi = 0$), and the Base Case Axtell Mean Field Model ($\Pi = 0.1, 0.5, 0.9, 1.0$) have been drawn. It is clear that introducing $\Pi$ somewhat alters the distribution, specifically the distribution seems to 'bend' more (which

[3]It might be argued that the model using only the Erdős-Renyi graph also is a mean field approximation, although, in a different sense.

is indicative of an exponential distribution). The size of the largest firms in the distribution tends to be reduced as $\Pi$ is introduced. Although, an anomaly here is that $\Pi = 0.5$ seems to be optimal for generating the largest firm (without having a trivial distribution with all, or almost all, agents of the population in one single firm), apparently providing agents with an optimal balance between having an incentive to explore new firms to join, or continuing to exploit the current firm they are in, or firms their neighbours are part of.

It would seem reasonable, though, that generally smaller firms were obtained since agents can no longer get 'stuck' in a certain part of the network, having no better choice than to remain in their current firm. Instead they have, one could say, almost total mobility, yet a stationary state is still achieved. From here on out, when referring to the Axtell Mean Field Model, it will be assumed that $\Pi = 1.0$

### 3.2.2 Reduced Axtell Mean Field Model

In the 'Reduced Axtell Mean Field Model' further simplifications were made to the Base Case Axtell Mean Field Model[4]. Specifically, the output parameters, as defined in Equation 2.11, were reduced to $a = 0$ and $b = 1$, and only $\beta$ remained variable. In other words,

$$O(E) = E^{\beta}, \tag{3.5}$$

where $\beta \sim U\left[\frac{3}{2}, 2\right]$.

A comparison between the distributions of the original Axtell Model, Axtell Mean Field Model and Reduced Axtell Mean Field Model is made in Figure 3.10. It is obvious that the distribution of firms is altered as the modifications are introduced, gradually reducing the power-law regime and making the distribution more similar to an exponential distribution.

In Figure 3.10 the slope for the various model have also been estimated in the interval where there seems to be a power-law regime. It is evident that the simplifications introduced in the more simplistic models gradually decrease the estimated $\alpha$ of the distributions[5].

The exponent in the expression for the average effort of a firm $\langle E \rangle_n$, in equation 3.1, is estimated to still be in the vicinity of $z = \frac{1}{2}$; more precisely the approximation yields $z \approx 0.57$. It is clear that the curve for $\langle E \rangle_n$ is more curved upwards for the Reduced Axtell Mean Field Model, compared to the original Base Case Axtell Model.

In Figure 3.12 it is investigated if restricting $\theta_{\max}$ in the Reduced Axtell Mean Field Model has an impact similar to what was witnessed in the Axtell Model, and similarly for the output parameter $\beta$ in Figure 3.13. The results seem to indicate that the characteristic effect of these parameter choices are kept intact.

[4] The Reduced Axtell Mean Field Model is the model which is primarily dealt with analytically in Chapter 4.

[5] This procedure is of course quite primitive, but it is deemed satisfactory for an initial investigation.

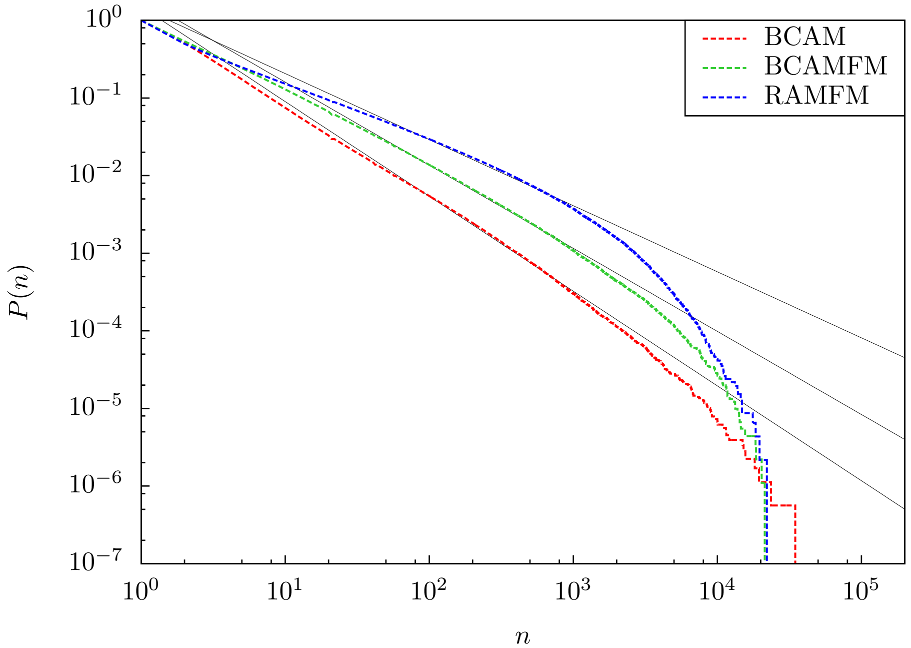


**Figure 3.10:** The complementary cumulative distribution function of the Base Case Axtell Model (BCAM), Base Case Axtell Mean Field Model (BCAMFM) and Reduced Axtell Mean Field Model (RAMFM) for $N = 10^5$ when $t = 100, 200, \ldots, 10\,000$ (accumulated data). The continuous straight lines have been fitted (OLS) between $n = 10$ and $n = 1000$ and have a slope corresponding to $\alpha = 1.22$ (for the BCAM), $\alpha = 1.07$ (for the BCAMFM) and $\alpha = 0.99$ (for the RAMFM).

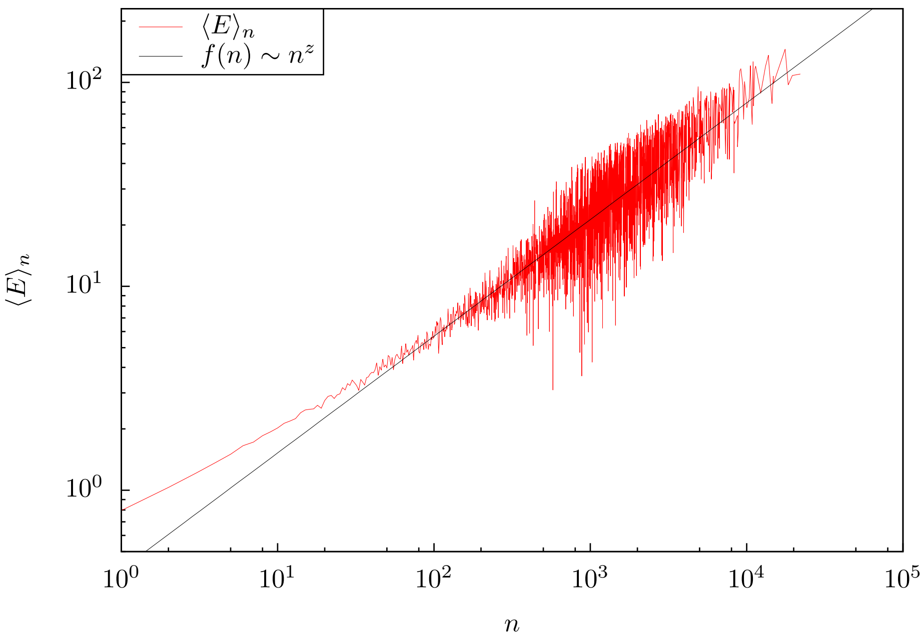


**Figure 3.11:** The average collective effort $\langle E\rangle_n$ of a firm with $n$ employees. The data has been gathered from a simulation of the Reduced Axtell Mean Field Model with a population size of $N = 10^5$ when $t = 100, 200, \ldots, 10\,000$ (accumulated data). Fitting by least squares from $n = 100$ yields $z \approx 0.57$.

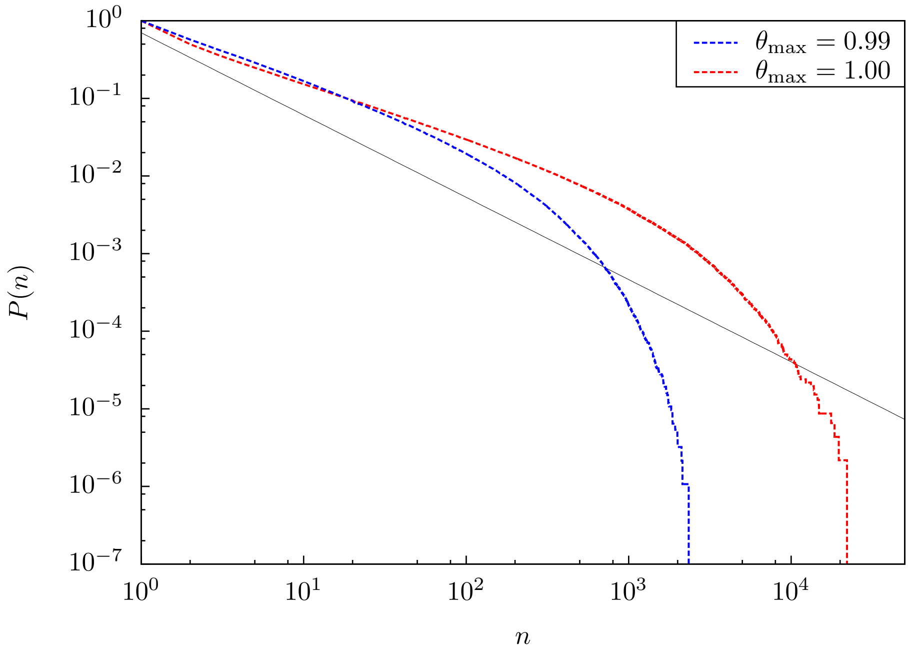


**Figure 3.12:** The complementary cumulative distribution function of the Reduced Axtell Mean Field Model ($\theta_{max} = 1.00$) and the Base Case Axtell Model with $\theta$ restricted ($\theta_{max} = 0.99$) for $N = 10^5$ when $t = 100, 200, \ldots, 10\,000$ (accumulated data). The continuous straight line has a slope corresponding to $\alpha = 1.06$ and has not been fitted.

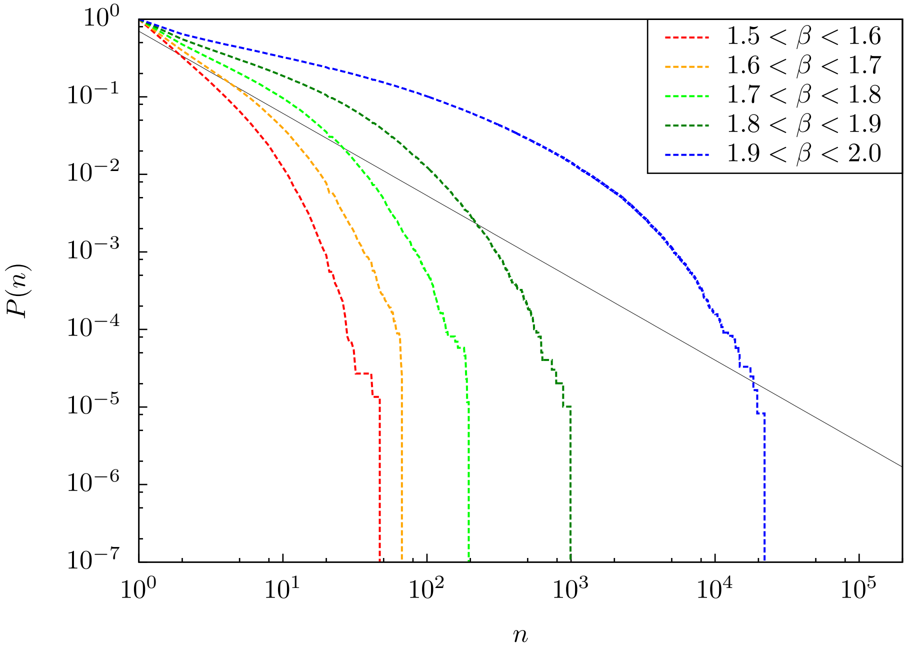


**Figure 3.13:** The complementary cumulative distribution function of the Reduced Axtell Mean Field Model for $N = 10^5$ when $t = 100, 200, \ldots, 10\,000$ (accumulated data) where firms with the output parameter $\beta$ in a certain interval have been extracted and plotted separately. The continuous straight line has a slope corresponding to $\alpha = 1.06$ and has not been fitted.

In total, it seems as if, albeit the distribution is changed, that some of the key features of the original model are kept in the Reduced Axtell Mean Field Model. This is a vital assumption for the analytical exploration which is carried out in Chapter 4.

## 3.3 Axtell Mean Field Model without Utility

In the analytical treatment in Section 4.2 it is shown that under certain circumstances agents exhibit something which may be characterised as preferential attachment, i.e., a tendency to prefer to join larger firms. This is a result of particular properties of the utility function, given by Equation (2.12). However, it is also evident that the agents sometimes have an incentive to remain in their current firm, join smaller firms (negative preferential attachment), or even start their own firm. Taking these aspects into consideration, it should be possible to reduce the Axtell Model even further by removing the utility function of the agents and instead introducing two new parameters, namely $p$ and $\phi$.

The set-up of the 'Axtell Mean Field Model without Utility' is considerably simpler than the previously introduced models; $\theta_i$ is not used any more, nor do the firms have any of the output parameters $a$, $b$, or $\beta$. When agents search for a new firm to possibly join, another random agent is randomly chosen, similar to the Axtell Mean Field Model, and instead of making a choice based on some utility-maximising criteria, the agent simply chooses to join the firm of the randomly chosen agent with probability $p$ if the firm is larger, and with probability $(1-p)$ if the firm is smaller. If the firm of the randomly chosen agent is equal in size, then the agent will join with probability $\frac{1}{2}$, or, consequently, stay in its current firm with probability $\frac{1}{2}$. Also, in order to mimic the tendency of agents to sometimes leave their current firm and have a start-up, the agent will, instead of examining the firm of another randomly chosen agent, have a start-up with probability $\phi$ (see Appendix A for more details).

This represents a simple and crude, 'bare-bone' version (or interpretation) of the Axtell Model. The parameters $\phi$ and $p$ make it, presumably, possible to capture the most basic mechanics of the model, where $p > \frac{1}{2}$ expresses the (weak) preferential attachment witnessed in the model, which also seems to be essential for modelling various empirical power-law distributions.

In Figure 3.14 the distribution function of the Axtell Mean Field Model without Utility is compared with some of the previous models. Given a specific $\phi$, it was possible to approximately find a $p$ (referred to as $p_c$) where the distribution approximately seems to follow a power-law. Above $p_c$, condensation characterises the distribution, that is, a significant part of the total number of agents are in one single firm. Below $p_c$, the distribution seems to be approximately exponential, with a quite sharp cut-off in the distribution for larger firms, as illustrated in Figure 3.15 for $\phi = 0.01$ and Figure 3.16 for $\phi = 0.1$.

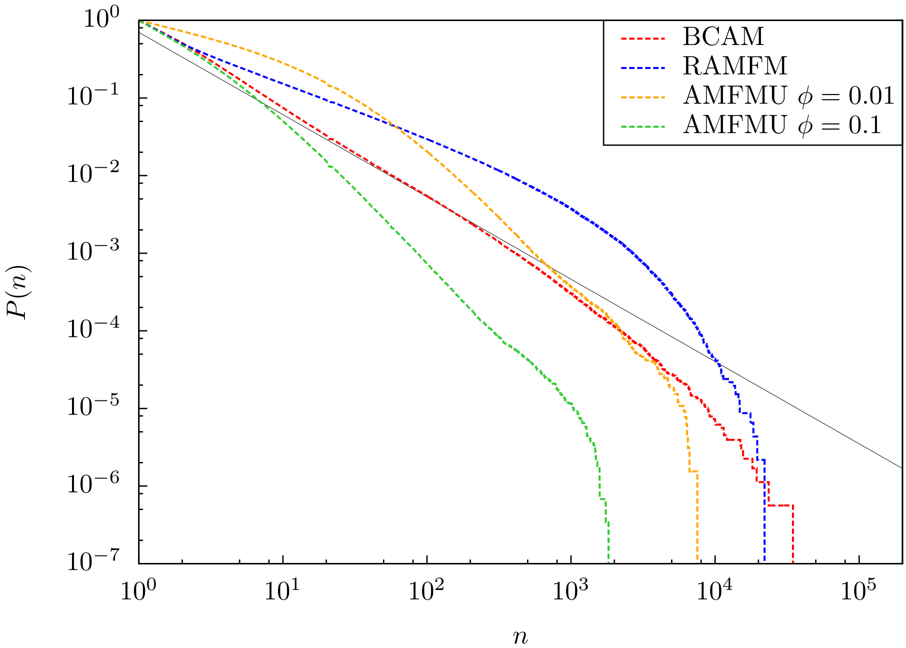


**Figure 3.14:** The complementary cumulative distribution function of the Base Case Axtell Model (BCAM) and Reduced Axtell Mean Field Model (RAMFM) for $N = 10^5$ when $t = 100, 200, \ldots, 10\,000$ (accumulated data), and Axtell Mean Field Model without Utility with $\phi = 0.01$, $p = 0.505078125$ (AMFMU $\phi = 0.01$) and with $\phi = 0.1$, $p = 0.555655$ (AMFMU $\phi = 0.1$) for $N = 10^5$ when $t = 1000, 1100, \ldots, 10\,000$ (accumulated data). The continuous straight line has a slope corresponding to $\alpha = 1.06$ and has not been fitted.

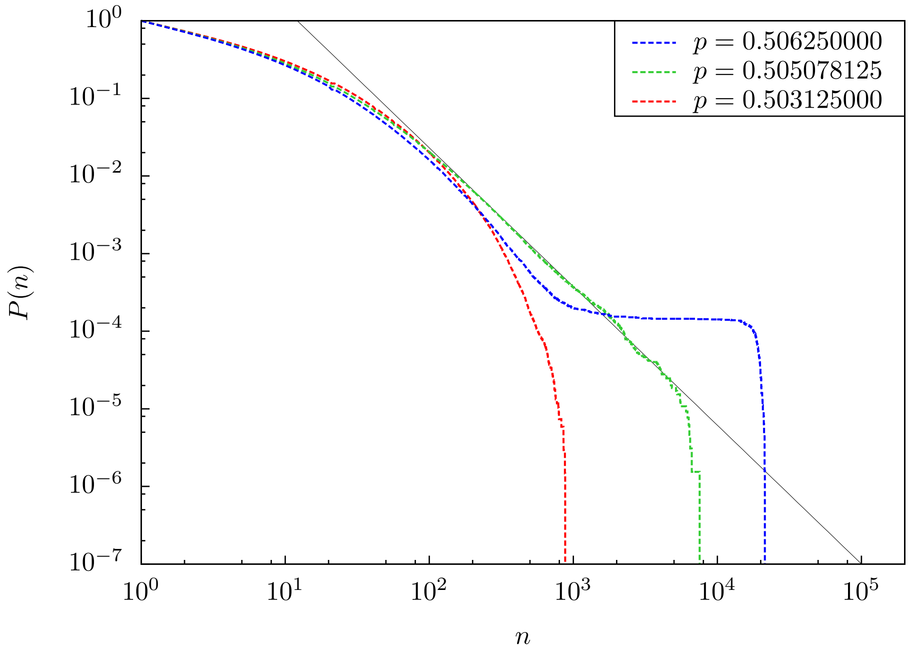


**Figure 3.15:** The complementary cumulative distribution function of the Axtell Mean Field Model without Utility for $N = 10^5$ with $\phi = 0.01$ and $p = 0.503125000, 0.505078125, 0.506250000$ when $t = 1000, 1100, \ldots, 10\,000$ (accumulated data). The continuous straight line has a slope corresponding to $\alpha = 1.79$ and has been fitted (OLS) between $n = 100$ and $n = 5000$.

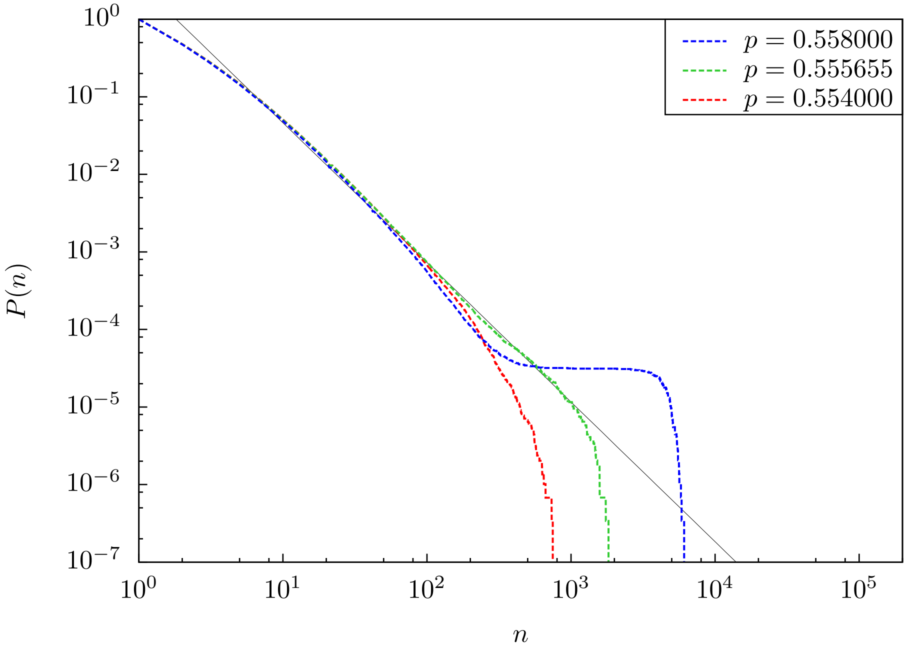


**Figure 3.16:** The complementary cumulative distribution function of the Axtell Mean Field Model without Utility for $N = 10^5$ with $\phi = 0.1$ and $p = 0.554000, 0.555655, 0.558000$ when $t = 1000, 1100, \ldots, 10\,000$ (accumulated data). The continuous straight line has a slope corresponding to $\alpha = 1.80$ and been fitted (OLS) between $n = 10$ and $n = 1000$.

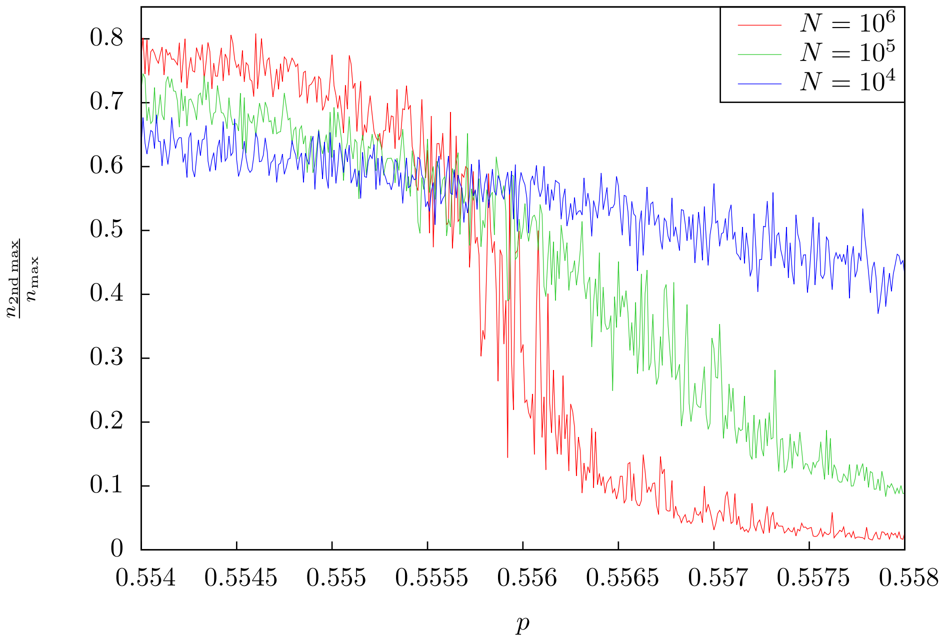


**Figure 3.17:** The average $\frac{n_{2\text{nd max}}}{n_{\max}}$ as a function of $p$ for $N = 10^4, N = 10^5, N = 10^6$ when $t = 1000, 1100, \ldots, 10\,000$ with $\phi = 0.1$.

There were difficulties determining $p_c$ when $\phi > 0.2$ because the dynamics of the system then become very sensitive when $p$ is close to $p_c$; specifically condensation would occur more abruptly (in the sense of a large firm 'quickly' appearing and dominating the distribution) when $p > p_c$ compared with a smaller $\phi$. The possible reason for this might be that a large $\phi$ has a too large impact on the dynamics and thus significantly reduce the probability of obtaining larger firms. Therefore, the analysis is restricted to two cases when $\phi$ is relatively small, namely $\phi = 0.01$ and $\phi = 0.1$. The transition from an exponential distribution to a condensed state may be regarded as a phase transition, more precisely, as is apparent in Figure 3.17 and Figure 3.18, a continuous (or second order) phase transition marked by a smooth transition in the order parameter, which $\frac{n_{2\text{nd max}}}{n_{\max}}$, i.e., the ratio of the second largest firm and the largest firm, might be regarded as.

It was noticed that it took significantly more time to reach a stationary state compared with the previous models, therefore only accumulated data gathered from $t = 1000$ and onwards could be used in the Axtell Mean Field Model without Utility.

When $\phi = 0$ and $p \geq \frac{1}{2}$, it seemed as if the distribution would (asymptotically) reach a condensed state where all agents were in one giant firm. Unsurprisingly, having $\phi > 0$ would hinder this kind of condensation. There remained, however, the

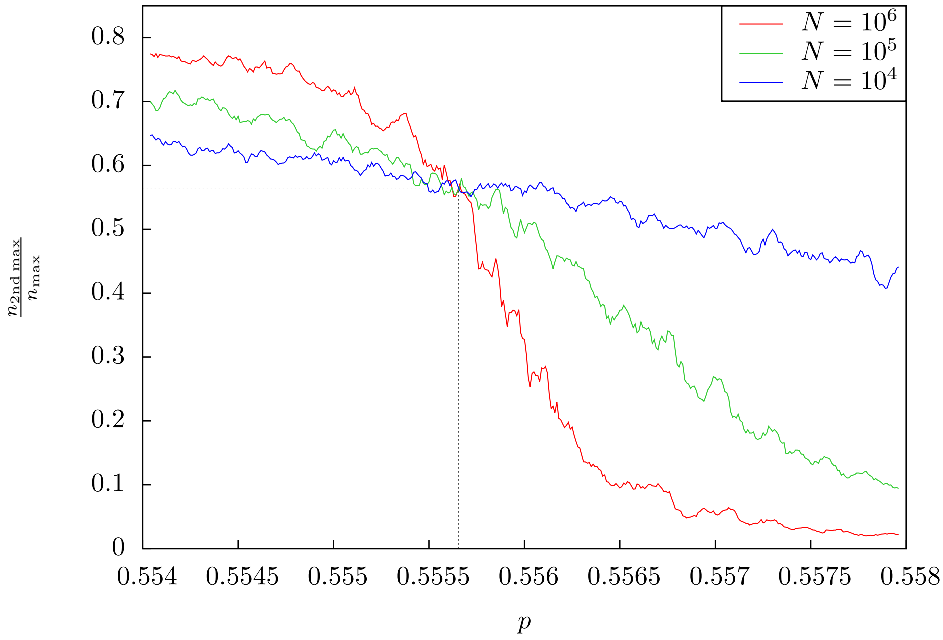


**Figure 3.18:** The average $\frac{n_{2\text{nd max}}}{n_{\max}}$, calculated with Equation (3.7) as a nine point moving average ($q = 4$), as a function of $p$ for $N = 10^4, N = 10^5, N = 10^6$ when $t = 1000, 1100, \ldots, 10\,000$ with $\phi = 0.1$. The dotted lines mark the point of the intersection of the three $\frac{n_{2\text{nd max}}}{n_{\max}}$ curves, where $p = p_c \approx 0.555655$ and $\frac{n_{2\text{nd max}}}{n_{\max}} \approx 0.5632$.

possibility of obtaining a distribution where one firm would significantly dominate it, if $p$ was large enough. Therefore, the condensation criterion had to be carefully considered.

Various methods for determining condensation were tried, e.g., examining if more than 1% or 50% of the total number of agents were in one firm, and if so, ascertain that the distribution was in a condensed state. These criteria yielded, however, imprecise results, and were not robust if the parameters were changed.

The criterion for condensation which was finally decided upon was when the ratio between the second largest firm and the largest firm,

$$\frac{n_{\text{2nd max}}}{n_{\text{max}}}, \tag{3.6}$$

becomes significantly smaller. This is still rather ambiguous though. In physics, as described in Appendix B, analogous situations are observed where scale invariance plays a crucial role in determining criticality. A similar approach should be viable here as well[6]. In other words, for a given $\phi$, $p_c$ should be independent of the size of the population of agents $N$.

In Figure 3.17 three different population sizes have been used, namely $N = 10^4$, $N = 10^5$ and $N = 10^6$. For a given $p$ ($0.554 < p < 0.558$), the average $\frac{n_{\text{2nd max}}}{n_{\text{max}}}$ when $t = 1000, 1100, \ldots, 10\,000$ was calculated. The plots are somewhat noisy, although calculating the average reduces the noise considerably. It is however obvious that the behaviour of the systems are as expected, with the possible exception of the curve for $N = 10^4$ not displaying a characteristic 'drop' in the vicinity of $p_c$[7], and there seems to be a unique point in which the three $\frac{n_{\text{2nd max}}}{n_{\text{max}}}$ curves intersect.

In order to establish the point where the intersection possibly might occur, further noise reducing techniques were performed in Figure 3.18. Specifically, an unweighted mean was calculated, from an equal number of data on either side of a central value, with the formula

$$X'_n = \frac{X_{n-q} + X_{n-q+1} + \ldots + X_n + \ldots + X_{n+q-1} + X_{n+q}}{2q+1}, \tag{3.7}$$

where $q = 4$, i.e., a nine point moving average, was deemed a suitable choice, in the sense of reducing the noise yet not distorting the underlying data significantly [7].

It is in Figure 3.18 evident that there is one point in which all three curves intersect. According to it, $p_c \approx 0.555655$, which is in accordance with the $p_c$ established for $N = 10^5$ with $\phi = 0.1$ in Figure 3.16. In other words, two different methods independently established approximately the same $p_c$.

In order to confirm the scale invariance, which is suggested by Figure 3.17 and Figure 3.18, simulations of the smaller and larger population sizes with $p$ presumably in the exponential regime, $p_c$ and $p$ in the condensed state, were performed.

[6] In Section 4.6 a more thorough explanation as to why this could be applicable here as well, and why the results would seem to indicate that this would be a fairly reasonable criterion, is given.

[7] This is possibly due to finite size effects being too strong when $N = 10^4$ only.

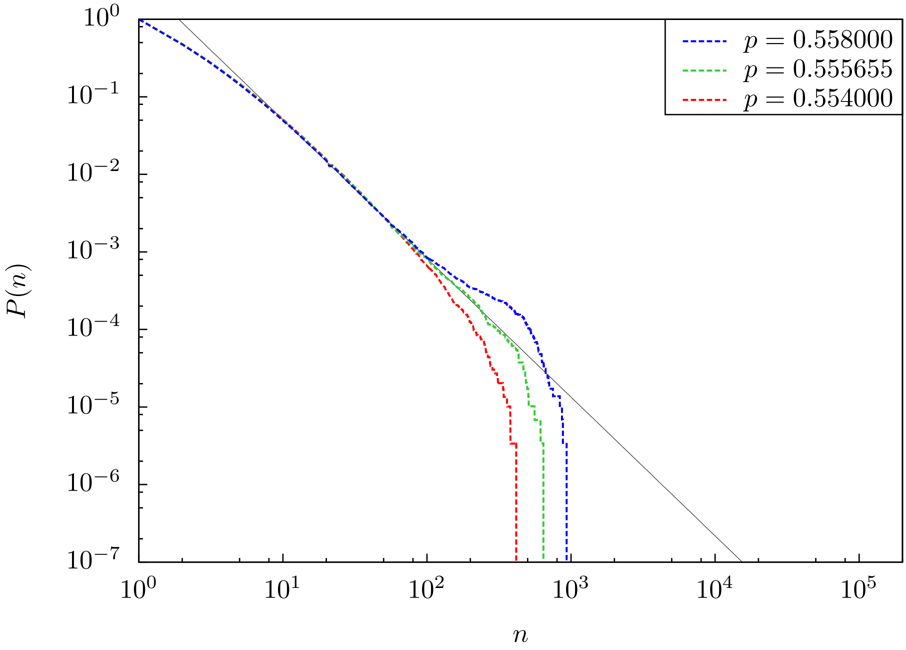


**Figure 3.19:** The complementary cumulative distribution function of the Axtell Mean Field Model without Utility for $N = 10^4$ with $\phi = 0.1$ and $p = 0.554000, 0.555655, 0.558000$ when $t = 1000, 1100, \ldots, 10\,000$ (accumulated data). The continuous straight line has a slope corresponding to $\alpha = 1.78$ and has been fitted (OLS) between $n = 10$ and $n = 200$.

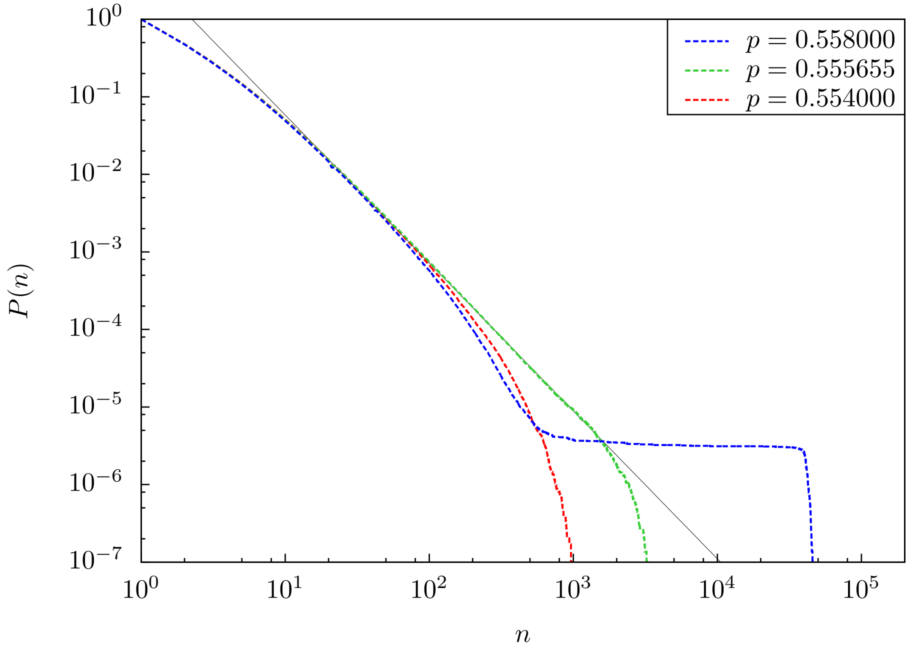


**Figure 3.20:** The complementary cumulative distribution function of the Axtell Mean Field Model without Utility for $N = 10^6$ with $\phi = 0.1$ and $p = 0.55400, 0.555655, 0.558000$ when $t = 1000, 1100, \ldots, 10\,000$ (accumulated data). The continuous straight line has a slope corresponding to $\alpha = 1.91$ and has been fitted (OLS) between $n = 10$ and $n = 1500$.

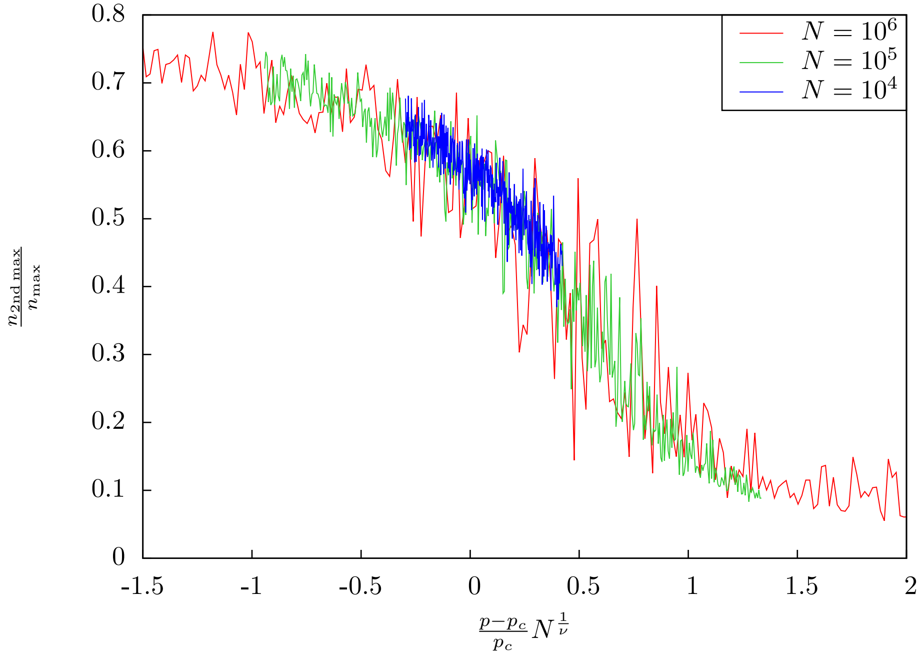


**Figure 3.21:** The average $\frac{n_{\text{2nd max}}}{n_{\max}}$ as a function of $\frac{p-p_c}{p_c}N^{\frac{1}{\nu}}$, where $p_c = 0.555655$ and $\nu = 2$, for $N = 10^4, N = 10^5, N = 10^6$ when $t = 1000, 1100, \ldots, 10\,000$.

The results in Figure 3.19 and Figure 3.20 seem to verify the scale invariance, i.e., there seems to be an (approximate) power-law regime for $\phi = 0.1$ when $p = p_c$ independently of the population size $N$.

The scale invariance is further validated by Figure 3.21, where $\frac{n_{\text{2nd max}}}{n_{\max}}$ has been plotted as a function of $\frac{p-p_c}{p_c}N^{\frac{1}{\nu}}$. If $\nu = 2$, the three curves approximately overlap each other in a certain interval, which indicates that the scaling function

$$\frac{n_{\text{2nd max}}}{n_{\max}} = f\left(\frac{p-p_c}{p_c}N^{\frac{1}{\nu}}\right) \tag{3.8}$$

appropriately describes the phenomenon.

It is of interest to note that the absolute value of the approximate exponent of the cumulative distribution obtained in Axtell Mean Field Model without Utility, ($\alpha \approx 1.8$ for the smaller systems or $\alpha \approx 1.9$ for the largest) is roughly equivalent to the one obtained analytically in Barabási-Albert's model ($\alpha \approx 2$), as explained in Section 2.4, but even closer to the one obtained in their simulations ($\alpha \approx 1.9$), indicating, possibly, a deeper underlying relationship.

The Axtell Mean Field Model without Utility, outlined in this section, will serve as the foundation for the master equation developed in Section 4.4.

# Chapter 4

# Analytical Results

The thesis work has involved analytical calculations done in parallel with the simulations. The simulations have played a vital role in confirming (and rejecting) analytical results. It appeared that some parts were amenable to an analytical treatment, while others were not. In this chapter the results are presented and briefly commented on. Included are also some speculations about the parts which did not yield any precise results.

## 4.1 Calculation of the Optimal Utility

In the Axtell Mean Field Model it is assumed that the agents are not connected with each other through an Erdős-Renyi graph. Furthermore, in the Reduced Axtell Mean Field Model it is also assumed, for the sake of gaining analytical tractability, that $a = 0$, $b = 1$. and $\frac{3}{2} < \beta < 2$. The output function of the firm, as defined in Equation (2.11), is thus reduced to

$$O\left(e_i, E_{-i}\right) = \left(e_i + E_{-i}\right)^{\beta}, \tag{4.1}$$

which leads to an utility function, as defined in Equation (2.12), reduced to

$$u(e_i, E_{-i}) = \left(\frac{(e_i + E_{-i})^{\beta}}{n}\right)^{\theta_i} (\omega - e_i)^{1-\theta_i}. \tag{4.2}$$

Given a certain $E_{-i}$ of the firm, the agent may choose an optimal effort $e \in [0, \omega]$ which maximises its utility, i.e., the maximum of the function

$$u(e_i) = \left(\frac{(e_i + E_{-i})^{\beta}}{n}\right)^{\theta_i} (\omega - e_i)^{1-\theta_i}. \tag{4.3}$$

The derivative may be written as

$$u'(e_i) = \frac{\left(e_i(\theta(1-\beta)-1) + E_{-i}(\theta-1) + \beta\theta\omega\right)\left(\frac{(e_i+E_{-i})^\beta}{n}\right)^{\theta_i}(\omega-e_i)^{1-\theta_i}}{e_i+E_{-i}}. \quad (4.4)$$

Since the utility function is concave and analytic, there is a well-defined unique maximum of the utility function[1]. The optimal effort $e_i^*$ of an agent, and thereby the effort the agent chooses to have, may be obtained by solving for $e_i^*$ in

$$u'(e_i^*) = 0, \quad (4.5)$$

which leads to

$$e_i^* = \frac{\gamma_i\omega - E_{-i}}{\gamma_i + 1}, \quad (4.6)$$

where

$$\gamma_i = \frac{\beta\theta_i}{1-\theta_i}. \quad (4.7)$$

It is important to note here that $e^*$ does not explicitly depend on $n$, the number of agents in the firm, but rather implicitly through the collective effort of the firm, $E$, which is found to depend on $n$ in Section 3.1 and Section 3.2.2, and which is also demonstrated analytically in Section 4.3.

A striking feature of Equation (4.6) is that if $E_{-i} > \gamma_i\omega$, then it would entail that the agent would produce negative effort, which is defined as impossible. Therefore, in that case the agent chooses to produce zero effort; it is in other words a free-rider. This behaviour of the agent is reflected in Equation (4.8).

Concludingly, in the Reduced Axtell Mean Field Model an agent $i$ with income preference $\theta_i$ in a firm with output parameter $\beta$ which is of size $n$ and where the other agents provided effort $E_{-i}$ in the previous period, the optimal effort produced by the agent is given by

$$e_i^* = \begin{cases} \frac{\gamma_i\omega - E_{-i}}{1+\gamma_i}, & E_{-i} \le \gamma_i\omega \\ 0, & E_{-i} > \gamma_i\omega. \end{cases} \quad (4.8)$$

Its optimal utility $u_i^*$ is then given by

$$u_i^* = \begin{cases} \frac{c_i}{n^\theta}(\omega + E_{-i})^{\eta_i}, & E_{-i} \le \gamma_i\omega \\ \frac{\omega^{1-\theta_i}}{n^{\theta_i}} E_{-i}^{\beta\theta_i}, & E_{-i} > \gamma_i\omega \end{cases} \quad (4.9)$$

where $\eta_i = 1 + \theta_i(\beta - 1)$ and $c_i = \frac{\gamma_i^{\beta\theta_i}}{(1+\gamma_i)^{\eta_i}}$.

[1]The second derivative is always negative.

## 4.2 Rationale for the Preferential Attachment Mechanism

If it is assumed that $\omega = 1$, then it holds that

$$e_i^* \sim \begin{cases} \frac{\gamma_i - E_{-i}}{1+\gamma_i}, & E_{-i} \leq \gamma_i \\ 0, & E_{-i} > \gamma_i \end{cases} \tag{4.10}$$

and

$$u_i^* \sim \begin{cases} \frac{1}{n^{\theta_i}}(1 + E_{-i})^{1+\theta_i(\beta-1)}, & E_{-i} \leq \gamma_i \\ \frac{1}{n^{\theta_i}} E_{-i}^{\beta\theta_i}, & E_{-i} > \gamma_i. \end{cases} \tag{4.11}$$

For an agent with $\theta_i \approx 1$ it is observed that its $\gamma_i$, as defined in Equation (4.7), assumes a large value. This implies that, most likely, $E_{-i} \leq \gamma_i$, and therefore that the agent's effort is $e_i \approx 1$ (close to maximum effort); in other words a 'workaholic'.

If the collective effort is written as in Equation (3.1), then the utility of an agent may approximately be written as

$$u_i^*(n) \sim \begin{cases} n^{z(1+\theta_i(\beta-1))-\theta_i}, & Cn^z \leq \gamma_i \\ n^{\theta_i(z\beta-1)}, & Cn^z > \gamma_i, \end{cases} \tag{4.12}$$

where $C$ is a constant. When certain conditions are met, the exponent in the expression for $u_i^*(n)$ in Equation (4.12) is larger than 0. This is something akin to a preferential attachment mechanism, as described in Section 2.4, in the sense that there is a higher probability of a firm with a larger number of employees acquiring an agent, searching for a new firm to join, as a new employee.

If $Cn^z \leq \gamma_i$, then in order to have this preferential attachment mechanism the exponent should fulfil

$$z(1 + \theta_i(\beta - 1)) - \theta_i > 0, \tag{4.13}$$

which is equivalent to

$$\theta_i < \frac{z}{1 + z(1 - \beta)}. \tag{4.14}$$

If $z = \frac{1}{2}$ and $\beta = 2$, then Inequality (4.14) is equivalent to

$$\theta_i < 1, \tag{4.15}$$

i.e., all agents have a preferential attachment mechanism. On the other hand if $z = \frac{1}{2}$ and $\beta = \frac{3}{2}$, only agents with

$$\theta_i < \frac{2}{3} \tag{4.16}$$

fulfilled have something reminiscent of the preferential attachment mechanism. It is thus established that both $\beta_{\max} = 2$ and $\theta_{\max} = 1$ are critical for the appearance of large firms.

It is observed that when $Cn^z > \gamma_i$, a similar search for conditions which would lead to a preferential attachment mechanism, leads to the condition that agents, irrespective of what $\theta_i$ is, should be in a firm with

$$\beta > 2, \tag{4.17}$$

which is impossible with the standard choice of parameters. In fact, this implies a negative preferential attachment mechanism. This might act as a counterweight to the previously demonstrated preferential attachment mechanism, preventing the system from assuming a condensed state, defined as all agents, or at least a significant part of the total number of agents in one single firm[2].

This might also explain why it was observed in the simulations that if $\beta_{max} \gtrsim 2$, instead of the usual choice of $\frac{3}{2} < \beta < 2$, condensation would occur[3], i.e., the preferential attachment mechanism is too strong in the sense that it is beneficial for all agents, no matter what their $\theta_i$ is, to join certain firms.

## 4.3 Calculation of $\langle E \rangle_n$

In Section 3.1, the analysis of the data generated by the simulations seems to indicate (universally) that the average effort of a firm with $n$ employees roughly obeys the relation

$$\langle E \rangle_n \propto n^z \tag{4.18}$$

where $z \approx \frac{1}{2}$.

If Equation (4.18) is a valid approximation, then $\langle \delta E \rangle |_{n \to n+1}$ may be defined as the average change in effort conditional on a new agent joining a firm with $n$ employees. Then the following relation should hold

$$\langle E \rangle_{n+1} = \langle E \rangle_n + \langle \delta E \rangle |_{n \to n+1} \tag{4.19}$$

where

$$\langle \delta E \rangle \ |_{n \to n+1} = \frac{\partial \langle E \rangle_n}{\partial n}, \tag{4.20}$$

and if it is assumed $n$ is large

$$\frac{\partial \langle E \rangle_n}{\partial n} \propto z n^{z-1}, \tag{4.21}$$

which leads to

$$\langle \delta E \rangle |_{n \to n+1} \propto z n^{z-1}. \tag{4.22}$$

The distribution of $\theta_i$ is $\theta_i \sim U[0,1]$, with the probability density function $\rho_\theta(\theta_i) = 1$, and $\gamma_i$ is defined in Equation (4.7) as $\gamma_i = \frac{\beta\theta_i}{1-\theta_i}$. The following then holds

$$\rho_\gamma(\gamma_i)\, \mathrm{d}\gamma_i = \rho_\theta(\theta_i)\, \mathrm{d}\theta_i, \tag{4.23}$$

[2]In Section 4.6 a more precise condensation criterion is established for the Axtell Mean Field Model without Utility.

[3]In this case all agents were in one giant firm

which means that

$$\rho_\gamma(\gamma_i) = \frac{\mathrm{d}\theta_i}{\mathrm{d}\gamma_i}. \tag{4.24}$$

Furthermore, one may obtain

$$\frac{\mathrm{d}\gamma_i}{\mathrm{d}\theta_i} = \frac{\beta}{1-\theta_i} + \frac{\beta\theta_i}{(1-\theta_i)^2} = \frac{\beta}{(1-\frac{\gamma_i}{\beta+\gamma_i})^2} = \frac{(\beta+\gamma_i)^2}{\beta}. \tag{4.25}$$

The result is thus

$$\rho_\gamma(\gamma_i) = \frac{\beta}{(\beta+\gamma_i)^2} \tag{4.26}$$

Under the most general conditions, any agent, no matter what their $\theta_i$ is, may join a firm of size $n$. It is then assumed that a firm always consists of a spectrum of agents (in terms of their $\theta_i$). Nonetheless, it is the agents with $\theta_i \approx 1$, and consequently high $\gamma_i$, which are assumed to be the most significant contributors to the total effort in a firm in the following calculations.

Using that Equation (4.10) may be regarded as a function of $\gamma_i$, i.e.,

$$e_i^*(\gamma_i) = \begin{cases} \frac{\gamma_i\omega - E_{-i}}{1+\gamma_i}, & \text{if } E_{-i} \leq \gamma_i\omega \\ 0, & \text{if } E_{-i} > \gamma_i\omega, \end{cases} \tag{4.27}$$

where $\gamma_i = \frac{\beta\theta_i}{1-\theta_i}$, $\omega = 1$ and $E_{-i} \approx \langle E\rangle_n \propto n^z$, it is possible to obtain

$$\langle\delta E\rangle\,|_{n\to n+1} = \int_0^\infty e^*(\gamma_i)\rho_\gamma(\gamma_i)\,\mathrm{d}\gamma_i = \int_{E_{-i}}^\infty \frac{(\gamma_i - E_{-i})\beta}{(1+\gamma_i)(\beta+\gamma_i)^2}\,\mathrm{d}\gamma_i$$

$$\approx \int_{E_{-i}}^\infty \frac{\gamma_i - E_{-i}}{\gamma_i^3}\,\mathrm{d}\gamma_i \propto \int_{n^z}^\infty \left(\frac{1}{\gamma_i^2} - \frac{n^z}{\gamma_i^3}\right)\mathrm{d}\gamma_i = \frac{1}{n^z} - \frac{n^z}{2(n^z)^2} = \frac{1}{2}n^{-z}. \tag{4.28}$$

It is also known, however, that $\langle\delta E\rangle\,|_{n\to n+1} \propto zn^{z-1}$, which thus leads to the self-consistent equation

$$\frac{1}{2}n^{-z} = zn^{z-1} \tag{4.29}$$

with the solution

$$z = \frac{1}{2}. \tag{4.30}$$

The same result is obtained if it is assumed $n$ is large and the system displays a simple kind of preferential attachment mechanism, where agents simply join a firm if it is larger than the one they are currently in, i.e.,

$$
\begin{aligned}
\langle \delta E \rangle \,|_{n \to n+1} = & \int_1^\infty \int_0^\infty e^*(\gamma_i)\rho_\gamma(\gamma_i) \frac{n' p(n')}{\langle n \rangle} H(n - n') \, \mathrm{d}\gamma_i \, \mathrm{d}n' \\
& = \underbrace{\int_1^n \frac{n' p(n')}{\langle n \rangle} \, \mathrm{d}n'}_{\approx 1} \int_0^\infty e^*(\gamma_i)\rho(\gamma_i) \, \mathrm{d}\gamma_i = \ldots = \frac{1}{2} n^{-z}, \quad (4.31)
\end{aligned}
$$

where $p(n)$ is the probability density function of firm sizes and $H(x)$ is the Heaviside function.

## 4.4 Derivation of the Master Equation

Inspired by the master equation approach used by Dorogovstev et al. (see Section 2.5), an attempt was made at formulating a master equation capturing some of the dynamics of the Axtell Model. Specifically, an attempt was made at modelling the Axtell Mean Field Model without Utility, with approximately the same basic structure and mechanisms. There are, however, various inconsistencies which have not been resolved yet within the framework of this thesis.

As a first step, a quite simple master equation, which describes how the probability density function of firm sizes $p(n)$ approximately develops, was formulated. A simple preferential attachment mechanism, where agents simply choose to join a firm if it is larger than the one they are currently in, and with probability $\frac{1}{2}$ if it is of equal size, is assumed in this model. This simplistic master equation is defined as

$$
\begin{aligned}
\frac{\mathrm{d}p_t(n)}{\mathrm{d}t} = p_{t+1}(n) - p_t(n) = W^t_{n+1 \to n} p_t(n+1) + W^t_{n-1 \to n} p_t(n-1) \\
- (W^t_{n \to n+1} + W^t_{n \to n-1}) p_t(n) - (1 - \delta_{n,1}) \frac{\phi^* n}{\mathcal{N}} + \delta_{n,1} \phi^* \sum_{n'=2}^{N} n', \quad (4.32)
\end{aligned}
$$

where the transition probabilities are given by

$$
W^t_{n \to n+1} = n \sum_{n'=1}^{n-1} \frac{n' p(n')}{\mathcal{N} \langle n \rangle} + \frac{n p(n)}{2 \langle n \rangle} \quad (4.33)
$$

and

$$
W^t_{n \to n-1} = n \sum_{n'=n+1}^{N} \frac{n' p(n')}{\mathcal{N} \langle n \rangle} + \frac{n p(n)}{2 \langle n \rangle}, \quad (4.34)
$$

the normalisation constant is given by

$$
\mathcal{N} = N(1 + \phi^*), \quad (4.35)
$$

$\langle n \rangle$ is the average firm size, $\phi^*$ is a parameter which describes the propensity of agents in firms to have start-ups (singleton firms), and the Dirac delta function is defined as

$$\delta_{m,n} = \begin{cases} 1, & m = n \\ 0, & m \neq n. \end{cases} \tag{4.36}$$

The initial condition, i.e., all agents start as singleton firms, may be written as

$$p_0(n) = \delta_{n,1}. \tag{4.37}$$

It is worth mentioning that $t$ may be interpreted differently in this context, compared to how it previously has been used, i.e., one time step in this context refers to how on average the distribution develops after one agent has made a choice to remain in its current firm, join another firm or have a start-up, while it previously has referred to when all agents, on average, in the population have made their choices.

The parameter $\phi^*$, as the parameter $\phi$ does in Section 3.3 for the Axtell Mean Field Model without Utility, prevents the system from reaching a condensed state by continually bringing agents to the $n = 1$ state, as described by the last two terms in Equation (4.32).

The choice of normalisation constant is somewhat arbitrary; the one currently used is the smallest possible. The numerical solutions indicate that it does not significantly affect the dynamics of the system, at least not in the stationary state.

The last term in Equation (4.33) models how an agent if confronted with the choice of joining a firm of equal size chooses to do so with probability $\frac{1}{2}$. This term is important because it enables the system to leave the initial state. Consequently, for the sake of symmetry, a similar term is also present in Equation (4.33).

The master equation, as defined in Equation (4.32) with the transition probabilities given by Equation (4.33) and Equation (4.34), lacks however a parameter equivalent to $p$, used in Section 3.3 for the Axtell Mean Field Model without Utility to model a 'weak' preferential attachment, i.e., an agent is willing to join a larger firm with probability $p$ and a smaller firm with probability $(1 - p)$.

In order to model this, the parameter $p^*$ is introduced, and only the transition probabilities have to be modified. They are now defined as

$$W^t_{n \to n+1} = p^* \left( n \sum_{n'=1}^{n-1} \frac{n' p(n')}{\mathcal{N} \langle n \rangle} + \frac{n p(n)}{2 \langle n \rangle} \right) + (1 - p^*) \left( n \sum_{n'=n+1}^{N} \frac{n' p(n')}{\mathcal{N} \langle n \rangle} + \frac{n p(n)}{2 \langle n \rangle} \right) \tag{4.38}$$

and

$$W^t_{n \to n-1} = p^* \left( n \sum_{n'=n+1}^{N} \frac{n' p(n')}{\mathcal{N} \langle n \rangle} + \frac{n p(n)}{2 \langle n \rangle} \right) + (1 - p^*) \left( n \sum_{n'=1}^{n-1} \frac{n' p(n')}{\mathcal{N} \langle n \rangle} + \frac{n p(n)}{2 \langle n \rangle} \right). \tag{4.39}$$

The previously defined transition probabilities are thus a special case of these transition probabilities when $p^* = 1$.

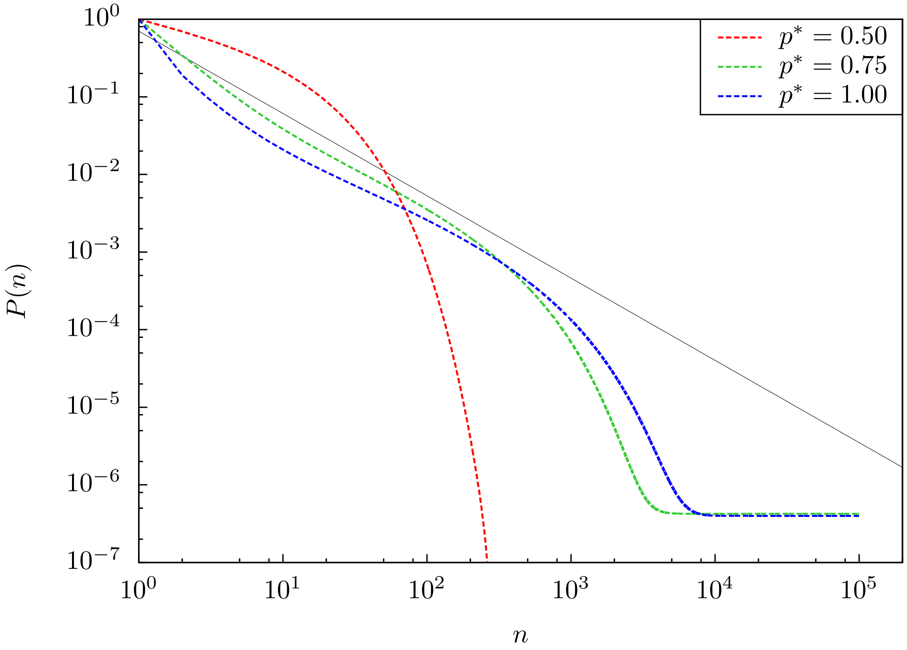


**Figure 4.1:** Numerical solution of the master equation represented in the form of the complementary cumulative distribution function for $N = 10^5$ with $\phi^* = 0.001$ and $p^* = 0.5, 0.75, 1.00$ when $t = 10\,000\,000$. The continuous straight line has a slope corresponding to $\alpha = 1.06$ and has not been fitted.

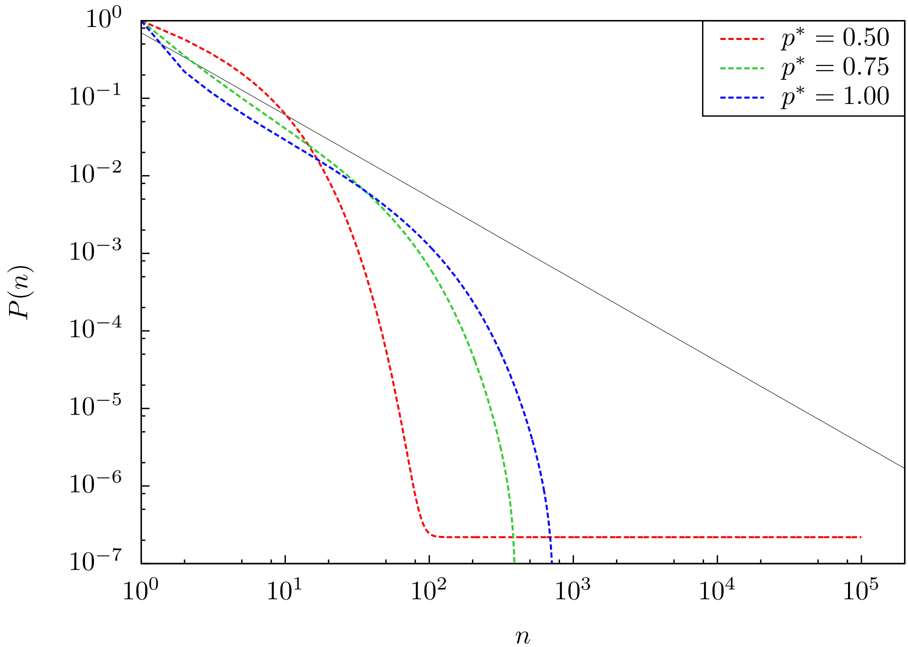


**Figure 4.2:** Numerical solution of the master equation represented in the form of the complementary cumulative distribution function for $N = 10^5$ with $\phi^* = 0.01$ and $p^* = 0.5, 0.75, 1.00$ when $t = 10\,000\,000$. The continuous straight line has a slope corresponding to $\alpha = 1.06$ and has not been fitted.

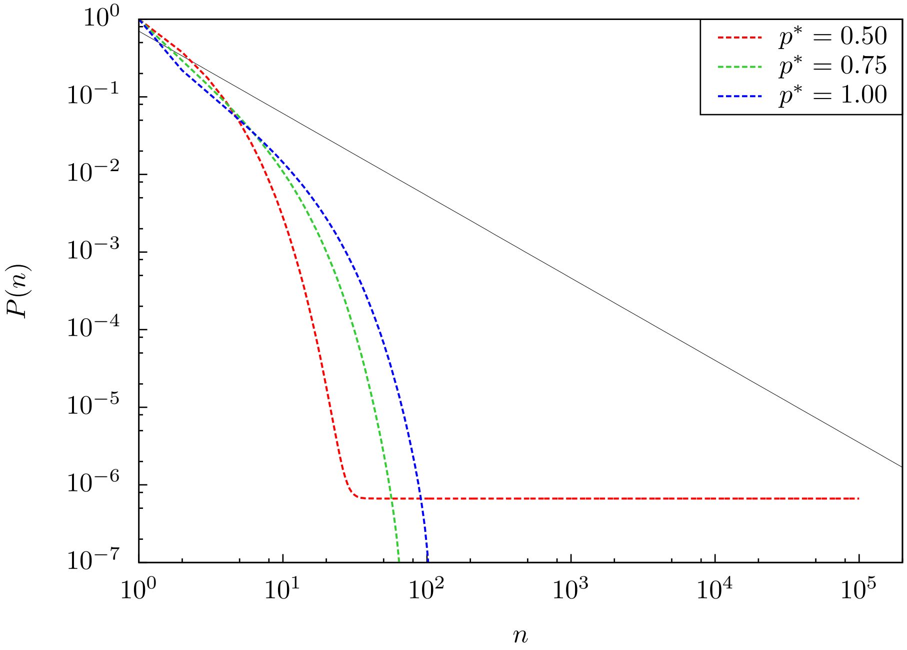


**Figure 4.3:** Numerical solution of the master equation represented in the form of the complementary cumulative distribution function for $N = 10^5$ with $\phi^* = 0.1$ and $p^* = 0.5, 0.75, 1.00$ when $t = 10\,000\,000$. The continuous straight line has a slope corresponding to $\alpha = 1.06$ and has not been fitted.

Similar to the procedure in Section 3.3 for the Axtell Mean Field Model without Utility, the analysis was restricted to certain values of $\phi^*$, namely $\phi^* = 0.001$, $\phi = 0.01$ and $\phi = 0.1$. In Figure 4.1, Figure 4.2 and Figure 4.3, the numerical results for these values of $\phi^*$, with various values of $p^*$, when $t = 10\,000\,000$ and the population size is $N = 10^5$ is displayed in the form of the complementary cumulative distribution function.

There are various problems with this implementation of the master equation. It does not, for example, take into account the precise changes which occur in the distribution when an agent joins another firm. Another problem is that once a stationary state has been reached, the boundary states $n = 1$ and $n = N$, and consequently adjacent states, have somewhat semi-absorbing qualities, i.e., the probability density function assume values here which are unrealistically high. Furthermore, compared to the Axtell Mean Field Model without Utility, it is not as obvious in this representation of the model when, or if, the system, for a given $\phi^*$, reaches a condensed state above a certain critical point of the parameter $p^*$, i.e., $p_c^*$.

Even though the results are quite different compared to the results obtained in Section 3.3, which is probably because the approximations are too extensive, they still indicate that the basic mechanisms encountered earlier are present, i.e., a larger $\phi^*$ results in a distribution more centred on smaller firms while a higher $p^*$ results in a more stretched distribution.

This model was first considered for a more detailed analysis, however, given the various problems with the master equation approach, it turned out that the Axtell Mean Field Model without Utility was more suitable for a more detailed analysis, e.g., the determination of $p_c$ given a certain $\phi$. Also, the relations between $\phi$ and $\phi^*$, and $p$ and $p^*$, remain unclear. Furthermore, a suitable, analytical treatment of the master equation leading to a derivation of the scaling parameter $\alpha$ was not possible, only numerical calculations could be carried out.

## 4.5 Alternative Derivation of the Master Equation

An alternative to working with $p(n)$, might be to consider $Q_n$, the probability of finding a randomly selected agent in a firm of size $n$, i.e.,

$$Q_n = \frac{np(n)N_f}{N}, \tag{4.40}$$

where $N_F$ is the number of firms. If a set-up similar to the previous one is used, then the calculation of $\frac{dQ_n}{dt}$ could, possibly, be

$$\begin{aligned}\frac{\mathrm{d}Q_n}{\mathrm{d}t} &= \phi^*\delta_{n,1}(1-Q_1) - nQ_n\left(\phi^*(1-\delta_{n,1}) + (1-\phi^*)\left(p^*\sum_{n'=1}^{n-1}Q_{n'}\right.\right. \\ &\left.\left. +p^*\sum_{n'=n+1}^{N}Q_{n'} + (1-p^*)\sum_{n'=1}^{n-1}Q'_n + (1-p^*)\sum_{n'=n+1}^{N}Q_{n'}\right)\right) + n\left(\phi^*Q_{n+1} + \right. \\ &(1-\phi^*)\left[p^*Q_{n+1}(1-R_{n+2}) + p^*Q_{n-1}R_{n-1} + (1-p^*)Q_{n+1}R_{n+1}\right. \\ &\left.\left. + (1-p^*)Q_{n-1}(1-R_n) + \frac{1}{2}Q_{n+1}^2 + \frac{1}{2}Q_{n-1}^2\right]\right) \\ &= -nQ_n\left[(1-\phi^*) + \phi^*(1-\delta_{n,1}\right] + \phi^*\delta_{n,1}(1-Q_1) + n\phi^*Q_{n+1} + \\ &n\left[Q_{n+1}[p^*(1-R_{n+2}) + (1-p^*)R_{n+1} + Q_{n-1}[p^*R_{n-1} + (1-p^*)(1-R_n)] + \frac{1}{2}\left(Q_{n+1}^2 + Q_{n-1}^2\right)\right],\end{aligned} \tag{4.41}$$

where $R_n = \sum_{n'=1}^{n-1} Q_{n'}$. For the special case $n = N$, Equation (4.41) is reduced to

$$\begin{aligned}\frac{\mathrm{d}Q_N}{\mathrm{d}t} &= -NQ_N\left[(1-\phi^*)(1-p^*)\sum_{n'=1}^{N-1}Q_{n'} + \phi^*\right] \\ &\quad + N\left(Q_{N-1}\left[p^*R_{N-1} + (1-p^*)(1-R_N) + \frac{1}{2}Q_{N-1}^2\right]\right) \\ &= -NQ_N\left[(1-\phi^*)(1-p^*)(1-Q_N) + \phi^*\right] \\ &\quad + N\left(Q_{N-1}\left[p^*R_{N-1} + (1-p^*)Q_N + \frac{1}{2}Q_{N-1}^2\right]\right).\end{aligned} \tag{4.42}$$

It is important to note that $(1-\phi^*)$ must not necessarily be used since it is possible to rescale to an arbitrary $\phi^{*'} = \frac{\phi^*}{1-\phi^*}$.

The use of $Q_n$, instead of $p(n)$, allows for a more intuitive and exact treatment of the changes which occur in the distribution as an agent joins a new firm or has a start-up. Although this model probably is more precise than the previous one, the problem with somewhat semi-absorbing boundary states still remain and must be considered more carefully.

## 4.6 Criterion for Condensation

In the Axtell Model certain parameter choices, e.g., $\frac{3}{2} < \beta < 2$, are crucial for avoiding condensation, roughly defined as all agents, or at least a significant (unrealistic) part of the total number of agents being in one single firm

In the Axtell Mean Field Model without Utility it was possible to ascertain by visual examination of the distribution function approximately where in the phase space there was an exponential regime, power law or condensed state. Further tests seemed to indicate these results were valid.

It is, however, not obvious how a criterion for condensation should be formulated. Various alternatives were tested until the final one, when the ratio between the second largest firm and the largest firm,

$$\frac{n_{\text{2nd max}}}{n_{\max}}, \tag{4.43}$$

becomes significantly smaller, was found. Since this is of course somewhat ambiguous, it was hypothesized that the situation here is analogous to the one encountered in statistical physics, as described in Appendix B, when experimentally obtaining the critical temperature in the Ising model via the Binder cumulant. This criterion was used in Section 3.3 to determine $p_c$ in the Axtell Mean Field Model without Utility, and the results seemed to indicate this hypothesis in fact was correct, and quite accurate results of $p_c$ were obtained. Analytical arguments also indicate that the hypothesis is correct.

If it is assumed in the exponential regime that the firm size distribution follows an exponential distribution with rate parameter $\lambda$ and independent and identically distributed (i.i.d) random variables, then it may be shown (see Appendix C) that the minimum of $k$ random variables is itself an exponentially distributed random variable with rate parameter $k\lambda$. If $X_i$ is the $i$th smallest of these $k$ variables, then it holds that

$$E[X_1] = \frac{1}{k\lambda}. \tag{4.44}$$

Since exponential random variables are continuous, there is zero probability that any two random variables will have the same value. This means that the $k-1$ other random variables all have values larger than $T_1$. Furthermore, if $X_i$ is exponentially distributed then it holds that it is 'memoryless', i.e.,

$$\Pr(X_i > t + s \mid X_i > s) \ = \Pr(X_i > t). \tag{4.45}$$

This means that knowledge of $X_1$ implies that the difference between $X_1$ and $X_2$ follows the same distribution as the minimum of $(k-1)$ i.i.d. exponentially distributed random variables. Therefore, one finds that

$$E[X_2 - X_1] = \frac{1}{(k-1)\lambda}, \tag{4.46}$$

and consequently, for $1 \leq i \leq k-1$,

$$E[X_{i+1} - X_i] = \frac{1}{(k-i)\lambda}. \tag{4.47}$$

Finally, this means that the expected value of the maximum of the $k$ random values is given by

$$
\begin{aligned}
E[X_k] &= E\left[X_1 + \sum_{i=1}^{k-1}(X_{i+1} - X_i)\right] \\
&= \sum_{i=0}^{k-1} \frac{1}{(k-i)\lambda} = \sum_{i=1}^{k} \frac{1}{i\lambda}.
\end{aligned}
\tag{4.48}
$$

If it is assumed $k$ is large, it holds that

$$
E[X_k] \approx \int_1^k \frac{1}{i\lambda}\,\mathrm{d}i = \frac{1}{\lambda}\ln(k)\,. \tag{4.49}
$$

The expected values of $X_k$ and $X_{k-1}$ are thus found to be

$$
E[X_k] = \frac{1}{\lambda}\ln(k) \tag{4.50}
$$

and

$$
E[X_{k-1}] = \frac{1}{\lambda}\ln(k) - \frac{1}{\lambda}. \tag{4.51}
$$

In Section 3.3, for the Axtell Mean Field Model without Utility, $k$ could represent the total number of firms in the exponential distribution regime, $X_k$ the largest firm $n_{\max}$, and $X_{k-1}$ the second largest firm $n_{\text{2nd max}}$. As the population size $N$ increases, $k$ should also increase, which thus leads to the fact that the ratio

$$
\frac{n_{\text{2nd max}}}{n_{\max}} = 1 - \frac{1}{\ln(k)} \tag{4.52}
$$

should therefore converge to 1 in the exponential distribution regime as $N$, and consequently $k$ also, increases, i.e.,

$$
\lim_{N\to\infty} \frac{n_{\text{2nd max}}}{n_{\max}} = 1. \tag{4.53}
$$

This behaviour of the ratio for larger system sizes is verified in Figure 3.17 and Figure 3.18.

The ratio $\frac{n_{\text{2nd max}}}{n_{\max}}$ may be regarded as analogous to the Binder cumulant $U_L$ used for determining the critical temperature in the Ising model; $p < p_c$, i.e., when the distribution of firms follows what approximately seems to be an exponential distribution, is equivalent to the ordered state in the Ising model when $T < T_c$, and $p > p_c$, i.e., when the system is in a condensed state, is equivalent to the disordered state in the Ising model when $T > T_c$.

In the thermodynamic limit, when $L \to \infty$, $T < T_c$ and the result in Equation (B.7) in Appendix B is obtained, corresponds to letting the number of agents

$N \to \infty$, when $p < p_c$ and obtaining the result in Equation (4.53). It also seems reasonable that when $p > p_c$, i.e., when $T > T_c$ in the Ising model, the ratio would mimic the behaviour of the Binder cumulant in the thermodynamic limit, i.e., converge towards zero when $N \to \infty$.

In other words, when $T < T_c$ in the Ising model and for some $L' > L'$, the Binder cumulant behaves like $U_{L'} > U_L$, $T < T_c$ would be replaced by $p < p_c$ and $U_L$ by $\frac{n_{\text{2nd max}}}{n_{\text{max}}}$, and similarly when $U_{L'} < U_L$ and $T > T_c$ would be replaced by $p > p_c$, There should therefore be a point of intersection for systems of different sizes for the Axtell Mean Field Model without Utility also.

Furthermore, the scaling form of the Binder cumulant in Equation B.10 should correspond to a scaling form of the ratio $\frac{n_{\text{2nd max}}}{n_{\text{max}}}$ in a system of size $N$, i.e.,

$$\frac{n_{\text{2nd max}}}{n_{\text{max}}} = f\left(\frac{p - p_c}{p_c} N^{\frac{1}{\nu}}\right). \tag{4.54}$$

This is verified in Figure 3.21.

In summary, the results indicate that this model, like Barabási-Albert's model, offers a mechanism accounting for the scale-invariant nature witnessed in real networks [4].

# Chapter 5

# Conclusion

In this final chapter the results are summarised, some additional comments are provided and a brief outlook for future research is presented.

## 5.1 Summary of Results

In this thesis the ABM presented by Professor Robert L. Axtell in [2] has been treated. The results obtained by Axtell have, most likely, been successfully reproduced.

For the agent's $\theta_i$ it has been shown that only slightly restricting $\theta_{\max} = 1$ to $\theta_{\max} = 0.99$ has a large impact on the distribution of firm sizes, resulting in a distribution considerably more centred on smaller sizes. Similarly, it is shown that the parameter $\beta$ is strongly related to the size of firms. It is also found that the average total effort of firm with $n$ employees roughly scales as $\langle E \rangle_n \sim n^z$, where $z \approx \frac{1}{2}$. This is also supported by Axtell's results.

Gradually the Axtell Model has been reduced to more simplistic models. As a first step, in the Axtell Mean Field Model, the fix Erdős-Renyi graph, which constitutes the 'neighbour network' used by the agents to search for new firms to join, was removed, and replaced by a fully connected network where agents randomly select another agent when they search for a new firm to join. In the Reduced Axtell Mean Field Model, further simplifications are made, namely the output function is reduced from depending on the output parameters $a$, $b$ and $\beta$, to only depending on $\beta$.

Although the distribution is significantly altered in the various simplifications of the original model, certain salient features are still present in the Reduced Axtell Mean Field Model, specifically the observed behaviour of the system with regards to the agent parameter $\theta_{\max}$, the output parameter $\beta$ and the tendency of scaling behaviour for the average total effort of firms. It is thus assumed that an analytical treatment of the Reduced Axtell Mean Field Model will result in generally valid results. Thereby, it is then possible to show analytically how the parameters $\theta_i$ and $\beta$ are important for the appearance of a mechanism enabling the formation of larger firms, referred to as 'preferential attachment', and to derive the scaling parameter $z \approx \frac{1}{2}$ for the total effort of firms $\langle E \rangle_n$.

In the Axtell Mean Field Model without Utility, the concept of agents making choices to remain in their current firm, join another firm or have a start-up in order to maximise their utility, is replaced by the parameters $p$ and $\phi$, where $\frac{1}{2} < p < 1$ represents the inclination of agents to strive to join larger firms, although it is not ever-present, and $\phi$ represents the continual process of agents leaving their current firm and having start-ups.

In this model, for a given $\phi$, it is possible to find, by fine-tuning the parameter $p$, a regime where the distribution of firms is close to an exponential distribution with a sharp cut-off ($p < p_c$), a regime where there seems to be a power-law distribution ($p = p_c$), and a condensed regime where one firm dominates the distribution. An order parameter to distinguish between these different states is suggested, namely the ratio between the second largest firm and the largest $\left(\frac{n_{\text{2nd max}}}{n_{\text{max}}}\right)$, which turns out to serve its role quite successfully. $p_c$ is solved successfully by employing a technique similar to the way the Binder cumulant is used in statistical physics to investigate criticality, i.e., to study the scale invariant behaviour displayed by systems of various sizes.

Finally, a master equation is proposed to model the behaviour of the Axtell Model. Although some results are obtained, they are of questionable validity. A more rigorous approach involving the master equation is, however, outlined as well.

## 5.2 Additional Comments

A major theme in this thesis has been the investigation of the apparent power-law behaviour observed in the distribution of firm sizes. The approximate values of the scaling parameter $\alpha$ found in the distribution of firm sizes generated by the various models, and, for reference, the one found empirically in the distribution of U.S. firms, are displayed in Table 5.1[1].

As the simplifications are introduced, it seems as if the distribution of firm sizes gradually loses its power-law appearance. Furthermore, it seems as if the exponent decreases as well, except for the Axtell Mean Field Model without Utility, which has a significantly higher exponent. If the utility function is not removed, most of the characteristics of the original model are still present.

The Axtell Mean Field Model without Utility model also requires substantial fine-tuning of the parameters to be at a critical point, but in some sense it might be argued that the other models require some sort of fine-tuning as well to be at a

[1]The estimations of $\alpha$ are quite rudimentary and correspond only to the sections where it has been estimated, by visual examination, a power-law behaviour seems present.

**Table 5.1:** Exponents found in various distributions

| Distribution | $\alpha$ |
|---|---|
| Empirical distribution of U.S. firms | 1.06 |
| Base Case Axtell Model | 1.22 |
| Base Case Axtell Mean Field Model | 1.07 |
| Reduced Axtell Mean Field Model | 0.99 |
| Axtell Mean Field Model without Utility | 1.9 |

critical point, specifically the output parameters and $\theta_i$. The observed power-laws may therefore not be regarded as generic and self-organising.

In a loose sense, it might be argued that heterogeneity plays a role in inducing power laws; it is introduced to the models, for example, via the variable parameters $\theta_i$, $a$, $b$ and $\beta$. This heterogeneity creates a natural variety in the distribution, and the more heterogeneous the model is, e.g., by adding variable parameters to polynomial that is the output function, the stronger the power-law behaviour seems to be. It is, however, fairly obvious that the preferential attachment mechanism, which is associated with some of these parameters, plays an important role in inducing power-laws.

In traditional macroeconomics one often uses a representative agent (RA) to simulate the aggregate behaviour of all agents, or one firm to represent all the firms [2]. This assumption is adequate if, for example, a Gaussian distribution of firm sizes is prevalent, i.e., the average truly represents something meaningful. However, if it is true that the distribution is more heterogeneous, e.g., if a power-law can be found, then this assumption breaks down, and it is thereby not meaningful to use that kind of approximation. A new kind of approach is thus required, and hopefully research along the lines of this thesis work will yield interesting results in the future.

Even if the more simplistic models do not fully reproduce the results of the Axtell Model, they have a worth of their own in the sense of being able to give rise to non-trivial distributions. It has been challenging to know precisely how far one can approximate these models, with the aim of analytical tractability, and still obtain interesting and valid results.

## 5.3 Outlook

A natural next step in this area of research would be to investigate more closely the validity of the power-laws found in the distributions. Although there seem to be strong indications that distributions which follow power-laws have been generated by these new models, further tests have to be performed to validate these claims. This is however beyond the scope of this thesis work, which has had a rather exploratory character. It also seems as if more rigorous studies of the underlying empirical data on firms are necessary. Furthermore, the proposed master equation needs more work, especially with regards to calibration of the various parameters and the definitions of the boundary conditions.

Another possibility would be to perhaps create a model which incorporates growth of the population of agents, which would be similar to the preferential attachment models presented by Barabási and Albert.

# Appendix A

# Pseudo-code

In this appendix the pseudo-code for the Axtell Model, Axtell Mean Field Model and the Axtell Mean Field Model without Utility is presented. Certain details and the parts of the code needed to generate the data output are omitted. The pseudo-code should be easy to follow. It is, however, useful to clarify a few notational conventions:

- The declaration of a variable $a$ of, for example, integer type `int`, will be written in C syntax as `int` $a$.
- A function is written as, e.g., $e(x)$, where $x$ it the argument of the function.
- **for**$(t \leftarrow 0; t < T; t \leftarrow t+1)$ means that $t$ is set to 0 before the loop starts, $t$ is increased by 1 at the end of each iteration, and the loop continues if the condition $t < T$ is true.
- The notation $x \leftarrow k$ will denote the operation of assigning the value $k$ to the variable $x$.
- Vectors of objects will be used[1]. For example, `vector<int>` will denote an ordered set (array) of integers. Moreover:
  - A vector of size N will be declared as `vector<int>` $A(N)$, and by default the declaration `vector<int>` $A$ means that the set $A$ is initiated as empty.
  - $|A|$ will denote the size of the set $A$.
  - The notation $A[i]$ will denote the $i$-th element of the set (with $i = 0, \ldots, |A|-1$).
  - To denote an ordered iteration over an ordered set $A$, a loop **for**$(a \in A)$ will be used.
  - A simple function, e.g., $a.x$ is defined as returning the value of the parameter $x$ of the object $a$.

[1] http://www.cplusplus.com/reference/vector/vector/

- For the firms, the C++ class `list` will be used, with a terminology similar to the one used for the class `vector`. The use of the class `list` allows for a much shorter execution time.
- For logical operators, the C++ convention[2] is used, which should be quite transparent.

[2] `http://en.wikipedia.org/wiki/Operators_in_C_and_C++#Logical_operators`

**Algorithm 1** Axtell Model

```
Require: G_i Graph of the network of agents; T Total time; θ_max Maximum θ_i an
  agent may be assigned

  procedure Initialisation

    int N ← 0                                   ▷ Number of agents (obtained from the graph)
    vector<vector<int>> G_f                     ▷ G[i] represents the 'neighbours' of agent i

    procedure Read graph(G_i)                   ▷ Obtain N and G_f
    end procedure

    vector θ(N)                                 ▷ A vector of each agent's θ_i
    list<firms> F(N)                            ▷ The initial N firms
    vector e(N)                                 ▷ A vector of each agent's effort

    for (f ∈ F) do
        |f| ← 1                                 ▷ Each firm consists of one agent initially
        f.E ← 0                                 ▷ Total effort is 0
        f.a ← a_min + (a_max − a_min)random     ▷ Output parameters
        f.b ← b_min + (b_max − b_min)random
        f.β ← β_min + (β_max − β_min)random
    end for

    for (i ∈ θ) do
        θ[i] = θ_max random                     ▷ The agents' θ_is
    end for

    for (i ∈ e) do
        e[i] ← e*(θ[i], f[i].a, f[i].b, f[i].β, f[i].E)     ▷ Optimal effort
        f[i].E ← e[i]
    end for

  end procedure

  for (t ← 0; t < NT; t ← t + 1) do             ▷ Main loop

    int i ← N random                            ▷ A random agent is chosen

                ▷ Compute the utility if the agent does not change its situation
    if |f[i]| == 1 then
        E_−i ← 0
    else
        E_−i ← f[i].E − e[i]
    end if
    e_best ← e*(θ[i], f[i].a, f[i].b, f[i]β, E_−i)
    u_best ← u*(θ[i], f[i].a, f[i].b, f[i]β, E_−i, e_best, |f[i]|)
```

**Algorithm 1** Axtell Model (continued)

```
                                                        ▷ Consider changing firm
    for (k ← 0; k < |G_f[i]|; k ← k + 1) do
        int j ← G_f[i][k]                    ▷ Go through the neighbours of agent i
        if f[i]!=f[j] then             ▷ Check if agent i and j are in the same firm
                          ▷ Compute the utility if agent i joins the firm of agent j
            e_new ← e*(θ[i], f[j].a, f[j].b, f[j].β, f → E)
            u_new ← u*(θ[i], f[j].a, f[j].b, f[j].β, f[j].E, e_new, |f[j]|+1)
            if u_new > u_best then
                u_best ← u_new
                e_best ← e_new
            end if
        end if
    end for
                                                        ▷ Consider start-up
    firm f_new                                                ▷ New firm
    |f_new|← 1
    f_new.a ← a_min + (a_max − a_min)random      ▷ Draw new output parameters
    f_new.b ← b_min + (b_max − b_min)random
    f_new.β ← β_min + (β_max − β_min)random
    if |f[i]|> 1 then
        e_new ← e*(θ[i], f_new.a, f_new.b, f_new.β, 0)
        u_new ← u*(θ[i], f_new.a, f_new.b, f_new.β, 0, e_new, 1)
    end if
    if u_new > u_best then
        u_best ← u_new
        e_best ← e_new
    end if

    procedure UPDATE                     ▷ Update all relevant information
    end procedure

end for
```

**Algorithm 2** Axtell Mean Field Model

**Require:** $N$ Number of agents; $T$ Total time; $\theta_{\max}$ Maximum $\theta_i$ an agent may be assigned

**procedure** INITIALISATION

  `vector` $\theta(N)$ ▷ A vector of each agent's $\theta_i$
  `list<firms>` $F(N)$ ▷ The initial $N$ firms
  `vector` $e(N)$ ▷ A vector of each agent's effort

  **for** $(f \in F)$ **do**
    $|f| \leftarrow 1$ ▷ Each firm consists of one agent initially
    $f.E \leftarrow 0$ ▷ Total effort is 0
    $f.a \leftarrow a_{\min} + (a_{\max} - a_{\min})$`random` ▷ Output parameters
    $f.b \leftarrow b_{\min} + (b_{\max} - b_{\min})$`random`
    $f.\beta \leftarrow \beta_{\min} + (\beta_{\max} - \beta_{\min})$`random`
  **end for**

  **for** $(i \in \theta)$ **do**
    $\theta[i] = \theta_{\max}$`random` ▷ The agents' $\theta_i$s
  **end for**

  **for** $(i \in e)$ **do**
    $e[i] \leftarrow e^*(\theta[i], f[i].a, f[i].b, f[i].\beta, f[i].E)$ ▷ Optimal effort
    $f[i].E \leftarrow e[i]$
  **end for**

**end procedure**

**for** $(t \leftarrow 0; t < NT; t \leftarrow t + 1)$ **do** ▷ Main loop

  `int` $i \leftarrow N$`random` ▷ A random agent is chosen

  ▷ Compute the utility if the agent does not change its situation
  **if** $|f[i]| == 1$ **then**
    $E_{-i} \leftarrow 0$
  **else**
    $E_{-i} \leftarrow f[i].E - e[i]$
  **end if**
  $e\_best \leftarrow e^*(\theta[i], f[i].a, f[i].b, f[i]\beta, E_{-i})$
  $u\_best \leftarrow u^*(\theta[i], f[i].a, f[i].b, f[i]\beta, E_{-i}, e\_best, |f[i]|)$

**Algorithm 2** Axtell Mean Field Model (continued)

```
                                                          ▷ Consider changing firm
    int j ← N random                                      ▷ Another random agent is chosen

    if f[i]!=f[j] then                    ▷ Check if agent i and j are in the same firm
                              ▷ Compute the utility if agent i joins the firm of agent j
        e_new ← e*(θ[i], f[j].a, f[j].b, f[j].β, f → E)
        u_new ← u*(θ[i], f[j].a, f[j].b, f[j].β, f[j].E, e_new, |f[j]|+1)
        if u_new > u_best then
            u_best ← u_new
            e_best ← e_new
        end if
    end if

                                                          ▷ Consider start-up
    firm f_new                                            ▷ New firm
    |f_new|← 1
    f_new.a ← a_min + (a_max − a_min)random               ▷ Draw new output parameters
    f_new.b ← b_min + (b_max − b_min)random
    f_new.β ← β_min + (β_max − β_min)random
    if |f[i]|> 1 then
        e_new ← e*(θ[i], f_new.a, f_new.b, f_new.β, 0)
        u_new ← u*(θ[i], f_new.a, f_new.b, f_new.β, 0, e_new, 1)
    end if
    if u_new > u_best then
        u_best ← u_new
        e_best ← e_new
    end if

    procedure UPDATE                                      ▷ Update all relevant information
    end procedure

end for
```

**Algorithm 3** Axtell Mean Field Model without Utility

**Require:** $N$ Number of agents; $T$ Total time; $\phi$ Start-up parameter, $p$ Preferential attachment parameter

```
procedure INITIALISATION

    list<firms> F(N)                              ▷ The initial N firms

    for (f ∈ F) do
        |f| ← 1                                   ▷ Each firm consists of one agent initially
    end for

end procedure

for (t ← 0; t < NT; t ← t + 1) do                 ▷ Main loop

    int i ← N random                              ▷ A random agent is chosen

    if random < φ then
        firm f_new                                ▷ Agent i has a start-up
        |f_new| ← 1
    else
        int j ← N random                          ▷ Another random agent is chosen

        if f[i]!=f[j] then                        ▷ Check if agent i and j are in the same firm
            if |f[i]| == |f[j]| && random < 0.5 then
                f[i] ← f[j]                       ▷ Agent i joins the firm of agent j
            else
                if random < p then
                    if |f[i] < |f[j]| then        ▷ Join if the firm of j is larger
                        f[i] ← f[j]
                    end if
                else
                    if |f[i] > |f[j]| then        ▷ Join if the firm of j is smaller
                        f[i] ← f[j]
                    end if
                end if
            end if
        end if
    end if

    procedure UPDATE                              ▷ Update all relevant information
    end procedure

end for
```

# Appendix B

# Binder Cumulant

In Section 3.3 and Section 4.6 a technique borrowed from physics for determining a critical point is used, namely the Binder cumulant. This appendix will present relevant background information for the Binder cumulant.

The Ising model is a simple model in which phase transitions may be studied. In it, $N$ magnetic atoms are arranged on a regular lattice. Each atom can be in only one of two spin states, specifically up or down.

The Hamiltonian for an isotropic spin-$\frac{1}{2}$ Ising model on a square lattice may be defined as

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} S_i S_j, \tag{B.1}$$

where $S_i = \pm 1$ is the state of the atom at site $i$, $J > 0$ denotes the ferromagnetic interaction between spins at neighbouring lattice sites, and $\langle i, j \rangle$ denotes a sum over nearest-neighbour sites $i$ and $j$ [6]. At a certain temperature the expectation value of the magnetization may be defined as

$$m = \langle S_i \rangle . \tag{B.2}$$

$m$ may be regarded as the order parameter of the system. In the ordered phase ($T < T_c$), $m = \pm 1$, while in the disordered phase ($T > T_c$), $m = 0$ [13]. There is an exact solution of the critical temperature $T_c$ for the square lattice, namely

$$\frac{k_B T_c}{J} = \frac{2}{\ln(\sqrt{2} + 1)} \approx 2.26918, \tag{B.3}$$

where $k_B$ is the Boltzmann constant [15].

The Binder cumulant for a system of size $L$ is a fourth-order cumulant, which is defined as

$$U_L = 1 - \frac{\langle M^4 \rangle}{3 \langle M^2 \rangle^2}. \tag{B.4}$$

where the magnetization is $M = \sum_i S_i$ [12]. This expression is related to the kurtosis of a random variable $X$, commonly defined as

$$\kappa_2 = \frac{E[(X - E[X])^4]}{(E[(X - E[X])^2])^2}. \tag{B.5}$$

This is a measure of whether a distribution is peaked or flat relative to a normal distribution. Distributions with high kurtosis (leptokurtosis) tend to have a sharp peak near the mean, while distributions with low kurtosis (platykurtosis) tend to have more of a flat top near the mean rather than a sharp peak [11].

When $T > T_c$ and $L >> \xi$, where $\xi$ is the correlation length[1], it is possible to show that

$$U_L \propto L^{-d}, \tag{B.6}$$

where $d$ is the dimension of the system. Furthermore, when $T < T_c$ and $L >> \xi$, one can show that

$$\lim_{L \to \infty} U_L = \frac{2}{3} \tag{B.7}$$

In the thermodynamic limit, when the system size $L \to \infty$, one therefore obtains a discontinuous Binder cumulant, according to

$$U_\infty = \begin{cases} \frac{2}{3}, & T < T_c \\ U^*, & T = T_c \\ 0, & T > T_c, \end{cases} \tag{B.8}$$

where the critical Binder cumulant is $U^* \approx 0.61069$ for a square lattice, applying full periodic boundary conditions. This means that when $U = 0$ there is a peak in the distribution function of the magnetization centred on $m = 0$, approaching a Gaussian function, while if $U = \frac{2}{3}$ this peak has flattened and two new peaks have appeared at $m \pm 1$ [15].

[1]The correlation length is in this case defined as the average size of a cluster of atoms with the same spin, or the distance over which spins influence each other. At the critical temperature it goes to infinity.

On the other hand when $L << \xi$, $U_L$ varies only weakly with temperature and linear dimension. This means that if $U_L$ is plotted against T for some size $L$, and similarly for $U_{L'}$ for some size $L' > L$, one obtains

$$\begin{aligned} U_{L'} > U_L, &\quad T < T_c \\ U_{L'} < U_L, &\quad T > T_c, \end{aligned} \tag{B.9}$$

which means that there should be a point of intersection. Furthermore, the theory of finite size scaling indicates that the Binder cumulant assumes the scaling form of

$$U_L = f\left(\frac{T - T_c}{T} L^{\frac{1}{\nu}}\right), \tag{B.10}$$

where $f$ is a universal function and $\nu$ is the critical exponent of the correlation length. It is thus obvious that when $T = T_c$, all graphs of $U_L$, independent of $L$, should all pass through the same single point. This scale invariance provides therefore a convenient estimate of the critical temperature [5, 12].

In Section 3.3 and Section 4.6, the Binder cumulant serves as inspiration for the formulation of the condensation criterion, and for finding the critical value of the parameter $p$ ($p_c$).

# Appendix C

# Distribution of the Minimum of Exponential Variables

If $X_1, \ldots, X_k$ are independent exponentially distributed random variables with rate parameters $\lambda_1, \ldots, \lambda_k$, then the complementary cumulative distribution function of the minimum of the exponentially distributed random variables may be written as

$$\begin{aligned}\Pr(\min\{X_1, \ldots, X_k\} > x) &= \Pr(X_1 > x, \ldots, X_k > x) \\ = \prod_{i=1}^{k} \Pr(X_i > x) &= \prod_{i=1}^{k} e^{-x\lambda_i} = \exp\left(-x\sum_{i=1}^{k}\lambda_i\right).\end{aligned} \tag{C.1}$$

Since the relationship between the cumulative distribution function, $F(x)$, and the probability density function, $f(x)$, of a continuous random variable is

$$\frac{\mathrm{d}}{\mathrm{d}x}F(x) = f(x), \tag{C.2}$$

then the probability density function of the minimum of the exponentially distributed random variables can be written as

$$\begin{aligned}f_{\min\{X_1,\ldots,X_k\}}(x) &= \frac{\mathrm{d}}{\mathrm{d}x}\left(1 - \exp\left(-x\sum_{i=1}^{k}\lambda_i\right)\right) \\ &= \sum_{i=1}^{k}\lambda_i \exp\left(-x\sum_{i=1}^{k}\lambda_i\right),\end{aligned} \tag{C.3}$$

which means that the distribution of the minimum of $k$ exponentially distributed random variables with rate parameters $\lambda_1, \ldots, \lambda_k$, is itself exponentially distributed with rate parameter $\lambda_1 + \cdots + \lambda_k$ [11].

# Bibliography


[1] Acemoglu, D. *Growing Random Networks and Power Laws*, lecture notes distributed in the course 6.207/14.15 Networks at MIT, Fall 2009.

[2] Axtell, R. *Team Dynamics and the Empirical Structure of Firms*, CRISIS Publications and working paper at George Mason University (2013).

[3] Axtell, R. *Zipf Distribution of U.S. Firm Sizes*, Science **293**, 1818-1820 (2001).

[4] Barabási, A.L., and Albert, R. *Emergence of Scaling in Random Networks*, Science **286**, 509 (1999).

[5] Binder, K. *Finite Size Scaling Analysis of Ising Model Block Distribution Functions,* Zeitschrift für Physik B - Condensed Matter **43**, 119-140 (1981)

[6] Blundell, S.J. and Blundell, K.M. *Concepts in Thermal Physics, 2nd edition* (Oxford University Press, 2006).

[7] Brockwell, P.J. and Davis, R.A. *Introduction to Time Series and Forecasting, 2nd edition* (Springer, New York, 2010).

[8] Cava, R.J *et al.*, *Low Temperature Spin Freezing in $Dy_2 Ti_2 O_7$ Spin Ice*, Phys. Rev. B **69** 064414 (2004)

[9] Clauset, A. *et al.*, *Power-law distributions in empirical data*, SIAM Review **51**, 661-703 (2009).

[10] Dorogovstev, S.N. *et al.*, *Structure of Growing Networks: Exact Solution of the Barabási-Albert's Model*, Phys. Rev. Lett. **85**, 4633 (2000).

[11] Koski, T. *Lecture Notes: Probability and Random Processes at KTH*, lecture notes distributed in the course SF2940 Probability Theory, Fall 2013.

[12] Meyer, P. *Computational Studies of Pure and Dilute Spin Models*, PhM thesis, University of Derby, 2000.

[13] Plischke, M. and Bergersen, B. *Equilibrium Statistical Physics, 3rd edition* (World Scientific, Singapore, 2006).

[14] Redner, S. *How popular is your paper? An empirical study of the citation distribution*, European Physical Journal B **4**, 131 (1998).

[15] Selke, W. *Critical Binder cumulant for isotropic Ising models on square and triangular Lattices*, J. Stat. Mech. **P04008** (2007)

[16] Taleb, N.N. *The Black Swan* (Penguin Books, 2007)