HAL
open science

# A Model for a complex economic system

Cornelia Metzig

## To cite this version:

Cornelia Metzig. A Model for a complex economic system. Economics and Finance. Université de Grenoble, 2013. English. NNT : 2013GRENS038 . tel-01557518

**HAL Id: tel-01557518**

**https://theses.hal.science/tel-01557518v1**

Submitted on 6 Jul 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE GRENOBLE

# THÈSE

Pour obtenir le grade de

## DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **CIA - Ingénierie de la Cognition, de l'interaction, de l'Apprentissage et de la création**

Arrêté ministériel :

Présentée par


## Cornelia Metzig


Thèse dirigée par **Mirta B. Gordon**
et codirigée par **Jean Pierre Nadal**

préparée au sein **Laboratoire d'Informatique de Grenoble (LIG)**
et de **EDISCE**

# A Model for a Complex Economic System

Thèse soutenue publiquement le ,
devant le jury composé de :

**M. Giulio Bottazzi**
Professor, Scuola Superiore Sant'Anna, Pisa, Rapporteur
**M. Pietro Terna**
Professor, Università di Torino, Rapporteur
**M. Eric Gaussier**
Professeur, Université Joseph Fourier, Examinateur
**M. Pablo Jensen**
Directeur de Recherche CNRS, ENS Lyon, Examinateur
**Mme Mirta B. Gordon**
Directeur de Recherche Emerite, CNRS, Directeur de thèse
**M. Jean-Pierre Nadal**
Directeur de Recherche CNRS, ENS Paris / EHESS, Co-Directeur de thèse

# Acknowledgements

I would like to thank all the people who helped me to do my best, my supervisors Mirta Gordon and Jean-Pierre Nadal, Giulio Bottazzi, Pietro Terna, Eric Gaussier, Pablo Jensen, the members of the group AMA of the laboratory LIG, Ioannis Partalas, Rohit Babbar, Sophie Neudert, all my discussion partners, and all my friends and family.

# Contents

**Acknowledgments** **3**

**Introduction** **7**
0.1 Modelling of Social Systems . . . 7
0.2 Introduction to this work . . . 8
0.3 Contributions of this thesis . . . 10

**1 The model – Theoretical Analysis** **11**
1.1 Introduction . . . 11
1.2 Context of related work . . . 12
1.3 The model . . . 13
1.3.1 The job market . . . 15
1.3.2 The goods market . . . 16
1.4 Preliminaries for analyzing the dynamics . . . 18
1.4.1 Brief introduction of existing models and power laws . . . 18
1.4.2 The Langevin equation . . . 19
1.4.3 Derivation via the master equation . . . 19
1.4.4 Derivation via the Fokker-Planck equation . . . 21
1.4.5 The relationship between multiplicative noise and entropy 22
1.5 Theoretical Analysis . . . 24
1.5.1 Discussion on additive and multiplicative noise . . . 24
1.5.2 The growth rate probability distribution . . . 26
1.6 Numerical implementation . . . 27
1.6.1 The sequence of events . . . 27
1.6.2 Rounding method . . . 27
1.6.3 Additive term $\xi$ in scenarios (i) and (ii) . . . 28
1.7 Numerical Results . . . 28
1.7.1 Size distribution . . . 31
1.7.2 Growth rate probability distribution . . . 36
1.8 Alternative implementation: growth of independent subunits . . 37
1.9 Results with different values of $\beta$ . . . 39
1.10 Comparison to other models . . . 41
1.11 Discussion . . . 44
1.12 Conclusion . . . 45

**2 The model – Economic Specification** **47**

2.1 Introduction . . . . . 47

Publication 1 . . . . . 49

2.2 Theoretical description of the extended model . . . . . 69

2.2.1 Replicator dynamics . . . . . 69

2.2.2 Application to the firm growth model . . . . . 69

2.2.3 Combined dynamics . . . . . 71

2.2.4 The recentering of the effective margin . . . . . 72

2.2.5 Interest payments, Entry and Exit of Firms . . . . . 72

2.3 Effect of heretogeneous $\mu_i$ on the scaling exponent $\beta$ . . . . . 74

2.3.1 The prior distribution . . . . . 74

2.3.2 The conjugate prior distribution for job market constraints in the model . . . . . 75

2.4 The context of economic models . . . . . 80

2.4.1 Keynesian arguments . . . . . 80

2.4.2 Neoclassical arguments . . . . . 81

2.4.3 Framework of statistical market equilibrium [38] . . . . . 82

2.4.4 Evolutionary Economics . . . . . 82

2.4.5 Stock-flow consistent models . . . . . 82

2.4.6 Macroeconomic Agent-Based Modelling . . . . . 85

2.5 Concluding Remarks . . . . . 91

**3 Study of Bankruptcies in the model** **93**

3.1 Introduction . . . . . 93

Publication 2 . . . . . 94

3.2 Concluding Remarks and Outlook . . . . . 123

**4 Fat-tailed distributions in other Social Systems** **125**

4.1 Introduction . . . . . 125

Publication 3 . . . . . 127

**5 Conclusion** **139**

# Introduction

## 0.1 Modelling of Social Systems

The science of complex systems describes systems composed of interacting units, and infers macroscopic properties from these interactions. This concept is extremely successful in physics, where the statistical description of interacting particles provide a microscopic foundation to macroscopic phenomena in thermodynamics. It applies to all many-body descriptions within physics, i.e. gases, liquids and condensed matter, to fluids, as well as to molecular self-assembly [74]. It also applies to many fields beyond it, such as genetics and language formation [16], networks [7], biology [105], and social sciences [32]. In the latter, applications range from trading behaviour in financial markets [65], urbanism [13], segregation [83] to firm growth, which is the main focus of this thesis. Obtained macroscopic results are for instance equilibrium prices [38], price variations [6], city size distributions [68, 109], or growth rate distributions [97, 19]. When decribing social systems in the complexity perspective, one main difficulty is that several micro-specifications may exist, which explain equally well some macroscopic result [32]. An even more profound difficulty when describing a system of interacting humans is that they are typically heterogeneous.

> "(...) Any modelling of social agents inevitably involves a large and unwarranted simplification of the real problem. It is then clear that any investigation of models of social dynamics involves two levels of difficulty. The first is the very definition of sensible and realistic microscopic models; the second is the usual problem of inferring the macroscopic phenomenology out of the microscopic dynamics of such models. Obtaining useful results out of these models may seem a hopeless task. "[24]

However, phenomena in complex systems are often unsensitive to the details of the interactions which generate them, which allows that one can abstract from them [24]. A factor contributing to the success of models for social phenomena in the complexity perspective is that they can be validated on data more and more easily, thanks to growing availability of data and advances in computer science [41]. Besides this approach to formulate relatively abstract models, another strand has developed to describe complex systems. "The second approach is to

create more comprehensive and realistic models, usually in the form of computer simulations, which represent the interacting parts of a complex system, often down to minute details, and then to watch and measure the emergent behaviors that appear."[74] Both approaches are relevant for this work.

Also within the field of economics, the concept of modelling macrobehavior as the result of interactions at an underlying level is widespread. Tacitly, it is applied by any economic theory, in the sense that there is no doubt that macroeconomics results from human interactions. J. Epstein [32] formulates this for the example of neoclassical price formation: "Whether they realize it or not, when economists say "the market is at equilibrium", they are asserting that this type of dynamic "social neural net" has executed a computation – it has computed $P^*$, an equilibrium price vector. No individual has tried to compute this, but the society of agents has done so nonetheless." However, the transition between the two levels cannot always be explained, and what can be called "micro-macro" problem exists [90, 72]. One common way out of this is to assume that a representative agent at the micro-level is performing some optimization task, which was proposed as early as 1890 by A. Marshall [66]. Some consider this to be a microfoundation, albeit one where interactions among units at the microlevel do not provide an explanation of results at the macro-level. More explicitly, agent-based models for economics have been advocated by Orcutt[77]; first examples are so-called microsimulation models [11, 31]. Later, this strand of research has become known as 'agent-based computational economics' [101]. The focus lies on feedback between micro and macrobehaviour, and on sophisticated behaviour of agents that, in some models, are endowed with cognitive abilities. Some examples of such models are presented in this thesis, whose aim is among others the description of the feedback that causes business cycles. A limitation of this strand of research remains that the models are very complicated, and therefore difficult to describe and to analyze. Lately, agent-based models in economics have been advocated because their two most common alternatives, data-driven econometric-empirical statistical models and dynamical stochastic general equilibrium models were found to have flaws [36].

## 0.2 Introduction to this work

Motivated by the aim of better understanding these feedback mechanisms, this work started in the context of macroeconomic agent-based models, comprising firms, households, a bank and possibly a state, which model the production cycle. While building up a model similar to [21], it turned out that already a much simpler model yields very interesting results. In that sense, the model in this thesis is studied in two versions: once as an abstract mathematical model in chapter 1, then as a more complete model. In chapter 1, related simple stochastic growth models are presented in section 1.2. Next, the model is introduced. A central assumption of such a simple setting is that firms have a homogeneous expected profit margin, i.e. firms differ only in their size. These firms are in competition, either for workforce, or for the purchasing power of households

(section 1.3). In order to describe the model in the context of stochastic processes, some preliminaries are recalled in section 1.4. In the following, (section 1.5) the model is analyzed theoretically. Results from this stochastic process are a fat-tailed size distribution, a size-dependent growth rate variance, and a tent-shaped growth rate distribution. In chapter 2, a more comprehensive model is analyzed, which has additional features: margin heterogeneity, interest payments to a bank, the possibility of failure, as well as an ageing process. The introduction of these features is presented in a paper, followed by further analysis in the two successive sections. Where possible, the results are compared to empirical studies of firm data. In addition, the extensions make it possible to analyze the profit rate distribution, as well as typical life cycles of firms. The dynamics with margin heterogeneity can be described theoretically by replicator equations, which are derived in section 2.2. This heterogeneity modifies the scaling of the growth rates' standard deviation with respect to the simple model in chapter 1, which is shown theoretically and numerically in section 2.3. With these additional features, the model can be compared to different economic models, which are presented in section 2.4. Relevant fields of economics are notably evolutionary economics [85], stock-flow consistent models[47, 55, 40], and some recent macroeconomic agent-based models like [21, 45, 87, 29, 26], although the presented model remains simpler. Chapter 3 presents a working paper where the age distribution of firms and bankruptcy statistics are presented and analyzed. It is the continuation of the analysis of the model, in essence as it is introduced in chapter 2. These results allow for an extended comparison to empirical studies. Results are for example the effect of financial constraints on growth [35, 20], the empirical age distribution and bankruptcy probabilities of firms [57, 27], and the theoretical and empirical results on the sizes of bankrupt firms [79, 3]. These comparisons point out some limitations of the model, which are discussed in section 3.2. Chapter 4 is thematically less related to the preceding chapters. It presents results of a collaboration, where the existence of fat-tailed distributions in social systems has been exploited for the purpose of hierarchical classification in machine learning. Chapter 5 concludes and discusses strengths and weaknesses of the presented work.

## 0.3 Contributions of this thesis

1. Description of the results of a model for firm growth in the context of solutions of the Langevin equation
2. Identification of the problem of compatibility between a power law distribution for firm size, and the standard deviation of a firm's growth rate scaling as $\sigma \propto n^{-\beta}$ with firm size $n$.
3. Combination of a stochastic firm growth model with heterogeneity that leads to relative growth of firms, described by replicator equations.
4. The description of bankruptcies theoretically and numerically, and comparison to empirical studies
5. Explanation of the presence of power laws in large web-databases and its exploitation for machine learning purposes

# Chapter 1

# The model – Theoretical Analysis

## 1.1 Introduction

In this chapter, a simple model for firm growth is presented. It is motivated by some important stylized facts from relevant literature, regarding firm size distribution, growth rate distribution and growth rate variance scaling. Since the firm size distribution has been widely reported to follow a power law, some existing theory on the formation of power laws is presented, which are widely used in the field of complex systems. Further empirical evidence is that the growth rate distribution deviates from a Gaussian distribution and is tent-shaped, and that the growth rate variance is dependent on a firm's size. These stylized facts are reproduced by a simple agent-based model which is presented in this chapter, albeit not with the correct scaling exponent.

This chapter is organized as follows. In section 1.2, some related work is discussed. 1.3, the model is introduced. In section 1.4, some preliminaries for its analysis are recalled: the formation of normal distributions and power laws in the framework of the Langevin equation, as well as in the framework of entropy maximization. In section 1.5, the model is discussed theoretically in comparison to existing theory. In section 1.7, numerical results of the introduced model are provided and discussed. Some consequences of binning data are also discussed here. Section 1.8 presents an alternative implementation of the model, which yields a slightly different growth rate distribution. This is illustrated by numerical results. Section 1.9 presents a comparison with other scaling exponents than the one resulting from this model, illustrated by simulations. In section 1.10, the model is compared to two similar models from the literature. In section 1.11, a theoretical aspect of the model is discussed. Section 1.12 concludes and points out possible extensions.

## 1.2 Context of related work

Economic growth processes have been the object of active research since the ground-laying work of R. Gibrat [46] who described growth as a multiplicative stochastic process. By assuming that growth rates are independent and identically distributed random variables, and by studying the time evolution of the system, he obtained a lognormal distribution of company sizes, i.e. a heavy-tailed distribution. The topic has since then attracted much interest [96]. Recent empirical studies suggest that the firm size distribution follows a Zipf (power law) distribution [18, 5, 76]. Various different models exist for the formation of such distributions with power-law tails. In the context of firm growth, the well-known model by H. Simon [89] explains a power law for firm size distribution based on a process introduced by Yule [73]. In his model, the number of firms is constantly growing in the system, and individual firms cannot shrink, i.e. they grow continuously. The exponent of the power law depends on the frequency of new firms, and approaches 1 if this frequency is low. However, the requirement that the system has to grow continuously for the formation of a power law limits its applicability [73, 44]. For systems of constant global size, which are the focus of this thesis, models exist that explain the formation of fat tails by multiplicative stochastic processes. Although for such systems with multiplicative noise, no result of the generality of the central limit theorem exists [71], it has been shown that the stationary distribution has a power law tail in the lowest order approximation, if in addition additive noise is present[98, 15, 93, 64].

More recently, Stanley et al. [95] uncovered two empirical features that an accurate theory on firm growth should explain. The first is that the growth rate frequencies exhibit exponential (Laplacian) decay, i.e. it gives rise to a tent shape in logarithmic scale. The second is that the variance of the growth rate scales with company size $n$ as $\sigma(n) \propto n^{-\beta}$. This means that the simple assumptions by Gibrat and others, who assume multiplicative noise to be independent of firm size, are at odds with the data. The empirically determined values of $\beta$ (typically $\approx 0.2$) depend on the studied system. A number of papers have provided further evidence of these two findings in various growth processes: firm growth [95] ($\beta = 0.15$), [75] ($\beta = 0.18$), [42] ($\beta = 0.28$), [86] ($\beta = 0.3$) (the latter authors also consider bird populations and mutual funds), a country's GDP growth [63] ($\beta = 0.15$), citations in scientific journals [78] ($\beta = 0.22$) and the growth rate of crime [2] ($\beta = 0.36$). The multitude of examples suggest that the process generating a tent-shaped growth rate distribution, a scaling exponent $\beta \neq 0$ for the growth rate standard deviation, and a fat tailed size distribution is simple and universal.

A number of sophisticated models giving rise to a tent shaped growth rate distribution have been proposed. They follow different approaches. Bottazzi and Secchi [19] predict a tent-shaped (Laplacian) growth rate distribution as being the result of a number of abstract shocks drawn from a Polya urn, without addressing the question of the standard deviation's scaling exponent. In order to obtain the scaling of the standard deviation, many models assume that firms have an internal structure, i.e. they are composed of subunits [107, 42, 99]. In

Schwarzkopf et al.[86], the probability that the firms' subunits reproduce themselves follows a power law. As a result, the aggregated growth rate distribution is tent-shaped with a power law decay: it is not a collective phenomenon but holds at the individual level. Another interesting model [78] assumes that the growth rate variance depends on the size of the elements (which are citations in their case), and numerically obtains a fat-tailed size distribution. The tent-shaped growth rate distribution is however not a result of the model but instead a hypothesis at the individual level.

Most models explaining the tent-shaped growth rate distribution and the variance scaling relation do not attempt to simultaneously explain the formation of the fat-tailed firm size distribution. Rather, existing models for power law tails via multiplicative noise assume the growth rate to be independent of the firms' size, which seems in conflict with a scaling exponent $\beta \neq 0$. However, there is empirical evidence for both a power law firm size distribution and a scaling exponent $\beta > 0$ for the standard deviation.

This issue is addressed in this work, both theoretically and numerically with a simple agent-based model comprising firms and employees. A distinction is made between collective phenomena at the firm level and at the level of the macroeconomy. In contrast to the models cited above, I investigate the hypothesis that the same process accounts for the tent-shaped growth rate distribution, its standard deviation's scaling exponent $\beta > 0$ and the fat-tailed size distribution of firms.

## 1.3 The model

In this section, a simple firm growth model is introduced, and subsequently analyzed in the following sections. Chapters 2 and 3 build on this simple model with additional features. In this first scenario, the agent-based model comprises firms and workers, $N_f$ and $N_w$, which are both constant numbers. Firms are characterized by an expected profit margin $\mu =$ (expected sales-expenses)/expenses, which allows them to determine the necessary amount of workers $n$ in order to produce a given quantity $q$ of commodity goods. Goods are sold at a fixed price $p$, workers earn a wage $w$. The expected margin of a firm $i$ is defined as

$$\mu_i = \frac{q_i\, p - n_i\, w}{n_i\, w}\,. \tag{1.1}$$

This relation allows a firm to calculate the neccessary number of workers in order to produce a certain amount of goods,

$$\hat{n}_i = \hat{q}_i \frac{p}{w} \left( \frac{1}{1+\mu_i} \right)\,, \tag{1.2}$$

and inversely, to calculate how many goods it can produce with a given number of workers:

$$q_i = n_i \frac{w}{p} (1+\mu_i) \tag{1.3}$$

The size of a firm $i$ is defined as the number of its workers. In the following, $\hat{q}_{i,t}$ is the intended production quantity of firms, $q_{i,t}$ the actual produced quantity, and $q^s_{i,t}$ the sold quantity at iteration $t$. At each iteration, firms $i$ hire $n_i$ workers in a job market in order to produce a quantity of $q_i$ goods, which are then sold in a goods market. Firms adjust the quantity $\hat{q}_i$ they indend to produce at time $t$ based on the previously produced quantity $q_{i,t-1}$, and on their last realized profits:

$$\hat{q}_{i,t} = q_{i,t-1}(1 + \mu^n_{i,t-1})\,, \tag{1.4}$$

where $\mu^n_{i,t-1}$ is the net realized profit margin

$$\mu^n_{i,t-1} = \frac{q^s_i\, p - n_i\, w}{n_i\, w}\,. \tag{1.5}$$

It may be smaller than the expected margin $\mu_i$. In order to be able to produce this quantity, firms offer $\hat{n}_i$ jobs (1.2)

$$\hat{n}_i = \hat{q}_i\,(1 + \mu_i)\frac{p}{w}\,. \tag{1.6}$$

The job market then opens, where open positions and available workers are matched at random. Every open position has the same probability of being filled by a worker, or in the case of too few jobs, every worker has the same chance of being hired. After the job market, production takes place, where firms produce $q_i$ goods according to equation (1.3). These goods are put in a goods market, which follows the same algorithm as the job market: every produced good has the same chance of being sold, or in the case of shortage of goods, every demand has the same chance of being satisfied. This hypothesis, which is equivalent to the microcanonical ensemble in statistical physics, is also used elsewhere in the context of growth processes, albeit for internal firm structure [97, 107][1].

Throughout this chapter, the expected profits are assumed to be the same for all firms. This is the simplest case of this model.[2]. The following two cases are discussed separately:

(i) both workers and firms are consumers, who spend their salaries and their profits in the goods market

(ii) only workers spend their salaries.

Firms compete for two limited resources, workforce and purchasing power of customers. In case (i), purchasing power is sufficient and the limited resource is workforce. In case (ii) only workers are consumers, firms compete for limited purchasing power. Since the job market and the goods market are based on the same algorithm, these two settings lead to the same evolution of firm size distribution and growth rate distribution, but for clarity they are presented separately in subsections 1.3.1 and 1.3.2.

[1] The concept of a statistical equilibrium in a market has also been applied to markets with very different hypotheses. In section 2.4.3, this market is compared to the market setting by [38] which equally exhibits a statistical equilibrium.

[2] A scenario with heterogeneous $\mu_i$ of firms is discussed in chapter 2.

In case (i), aggregate demand is sufficient, firms sell their entire production, i.e. $q^s_{i,t} = q_{i,t}$ and $\mu^n_{i,t} = \mu$, and demand a discrete quantity of $\hat{n}_i$ workers, proportional to their size

$$\hat{n}_{i,t} = n_{i,t-1}(1+\mu)\,. \tag{1.7}$$

Since $\mu > 0$, firms attempt to increase their size, and the stationary state of the system corresponds to full employment. In case (ii), not all goods $q_{i,t}$ are sold, i.e. $q^s_{i,t} < q_{i,t}$ and $\mu^n_{i,t} < \mu$. As a consequence, the same number of workers will be hired as in the previous iteration, and consequently the same total number $Q_t$ of goods will be produced. In that case, firms demand less workers than available, and do not compete for workforce.

### 1.3.1 The job market

In the job market, workers do not stay at their firm but are newly placed at every iteration. Similar results would hold if only a (constant) fraction of workers of each company were newly placed. In case (i), there is a workforce shortage, and the allocation of workers gives rise to interesting dynamics: since the $N_w\,(1+\mu)$ open positions are covered at random with $N_w$ workers, the actual number $n_{i,t+1}$ of employees hired by a firm $i$ at time $t+1$, is likely to be smaller than their job offer $\hat{n}_{i,t} = n_{i,t}(1+\mu)$.

It can even be smaller than the number of employees $n_{i,t}$ in the preceding period, which includes a situation where firms may receive no worker at all and vanish. The number of active firms would decrease continuously, and workers would eventually accumulate in a monopoly, which is avoided by the introduction of new firms. To maintain $N_f$ constant, extinct firms are replaced by new ones[3], initialized with a number of workers $n^{new}_{i,t}$ drawn from a distribution $\mathcal{F}(n^{new})$. New firms contribute to the total demand for workforce in the next period with the quantity $\hat{n}^{new}_t = (1+\mu)n^{new}_{i,t}$. Analytically, the matching process in a market is described by a multinomial probability distribution. A simpler description of the evolution of the system is obtained if the number of workers is assumed to be conserved only on average: $\langle \sum_i n_{i,t} \rangle = N_w$. The probability for an open position to be filled then becomes

$$p = \frac{N_w}{\sum_i \hat{n}_i}\,, \tag{1.8}$$

where $\sum_i \hat{n}_i = (1+\mu)\,N_w + \sum_i n^{new}_i$. Since $\sum_i n^{new}_i \ll N_w$, the probability of a position to be filled is approximately $p \approx (1+\mu)^{-1}$. The probability for a firm of size $n_i$ to receive $k_i$ workers in the next period is given by the binomial distribution

$$P(k_i|n_i) = \binom{\hat{n}_i}{k_i} p^{k_i}(1-p)^{\hat{n}_i - k_i}\,, \tag{1.9}$$

[3]It is also possible to re-insert firms at a constant rate, in which case the level of active firms will become stationary after some time. In this chapter, a strictly constant number of active firms is used, since this guarantees that all presented systems will be the same size.

with mean $\langle n_{i,t} \rangle = \hat{n}_{i,t}\, p = n_{i,t-1}$, which is precisely the number of workers of firm $i$ at the previous time step. The variance is $\hat{n}_i\, p\,(1-p) = n_i \mu/(1+\mu)^2$. With probability $P(0|n_i) = (1-p)^{\hat{n}_i}$ a firm does not receive any workers and disappears.

For large $n_i$, the binomial distributions may be approximated by Gaussian distributions, whose variance exhibits the same $n$-dependence. Alternatively, another implementation uses a different rounding method by which firms determine their job offer, such that the Gaussian distribution is a good approximation even for small firms. It can be seen as growth of entirely independent subunits, where subunits are jobs (see section 1.8). For both the alternative method and for equation (1.9), the Gaussian approximation of the probability for a firm of size $n_{i,t}$ to reach size $n_{i,t+1}$ is written as

$$\mathcal{P}(n_{i,t+1} = k_i | n_{i,t}) = \frac{1}{\sigma_{i,t}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{k_i - n_{i,t}}{\sigma_{i,t}}\right)^2}\,, \tag{1.10}$$

where the mean has been replaced by its value $n_{i,t}$, and $\sigma_{i,t}^2 = n_{i,t}\frac{\mu}{(1+\mu)^2} = c\, n_{i,t}$. If the growth rate of a firm is defined as

$$g_{i,t} = \frac{n_{i,t+1}}{n_{i,t}}\,, \tag{1.11}$$

equation (1.10) yields for the growth rate probability density $\mathcal{G}$ (dropping the index $t$):

$$\mathcal{G}(g_i|n_i) = \sqrt{\frac{n_i}{2\pi\, c}} e^{-\frac{1}{2}\frac{n_i}{c}\,(g_i - 1)^2}\,, \tag{1.12}$$

where $c = \mu/(1+\mu)^2$ for the binomial approximation[4]. Thus, in the present model, the scaling exponent $\beta$ of the growth rate's standard deviation is defined through

$$\sigma(n) \propto n^{-\beta}\,, \tag{1.13}$$

and has the value $\beta = 0.5$. This value for $\beta$ is a general feature of models that explain firm growth as being the sum of the growth of independent subunits. In other published models, subunits often represent the sectors in which the firm is active [75, 97, 107]; in this model these are jobs. Other values for $\beta$ and the corresponding empirical evidence are addressed in section 1.9.

### 1.3.2 The goods market

In scenario (ii), firms do not spend their profits in the goods market. Aggregate demand $D = \sum_j d_j$ consists only of the wages which are paid to the $j$ employees. A quantity $\sum_i q_{i,t}\, p$ is produced, but the demand, i.e. the overall wages paid to workers are

$$D_t = \sum_i n_{i,t}\, w = \sum_i q_{i,t}\, p\, \frac{1}{1+\mu}\,. \tag{1.14}$$

[4]For the rounding method detailed in section 1.8, this constant is $\frac{2\mu}{(1+\mu)^2}$

It is clear that the aggregate sales $\sum_i q^s_{i,t}\, p$ is smaller than the production $Q_t = \sum_i q_{i,t}\, p$, since in this scenario workers are the only consumers. Then, the probability for a produced good to be sold becomes

$$v = \frac{D_t}{Q_t} = \frac{\sum_i q_{i,t}\, p\, \frac{1}{1+\mu}}{\sum_i q_{i,t}\, p} = \frac{1}{1+\mu}\,, \tag{1.15}$$

which is the analogon to equation (1.8). Unsold goods are lost, they cannot be stored and put in the market in following iterations. Since every good has the same chance of being sold, the allocation of demand also follows a binomial distribution

$$V(q^s_i|q_i) = \binom{q_i}{q^s_i} v^{q^s_i}(1-v)^{q_i - q^s_i}\,. \tag{1.16}$$

Again, this binomial distribution can be approximated with a Gaussian for large $q$. It becomes

$$\mathcal{V}(q^s_{i,t+1} = k_i|q_{i,t}) = \frac{1}{\sigma_i\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{k_i - \hat{q}_{i,t}}{\sigma_i}\right)^2}\,. \tag{1.17}$$

On average, each firm sells a quantity $\langle q^s_{i,t}\rangle = q_i\, v = \frac{w}{p} n_{i,t}$. Therefore, their average realized profit is $\langle \mu^{net}_{i,t}\rangle = \frac{\langle q^s_{i,t} p\rangle - n_{i,t} w}{n_{i,t} w} = 0$. In the next iteration, firms will demand $\langle \hat{n}_{i,t+1}\rangle = n_{i,t}$, and since there is no competition of workforce, $\langle n_{i,t+1}\rangle = n_{i,t}$. The average quantity of sold goods can be expressed in terms of the previously sold quantity: $\langle q^s_{i,t}\rangle = q_i\, v = \frac{w}{p} n_{i,t} = \frac{w}{p}\left\langle n_{i,t-1}(1+\mu^{net}_{i,t-1})\right\rangle = \frac{w}{p}\left\langle n_{i,t-1}(\frac{q^s_{i,t-1}}{n_{i,t-1}})\right\rangle = \langle q^s_{i,t-1}\rangle$. Thus, on average, firms stay constant in size, both measured in terms of employees, and in terms of sales. The growth rate can be written as

$$g_{i,t} = \frac{n_{i,t+1}}{n_{i,t}} \tag{1.18}$$

or equivalently as

$$g^{sales}_{i,t-1} = \frac{q^s_{i,t}}{q^s_{i,t-1}}\,. \tag{1.19}$$

This yields the growth rate probability density

$$\mathcal{G}(g_i|q^s_i) = \sqrt{\frac{q^s_i}{2\pi\, c}} e^{-\frac{1}{2}\frac{q^s_i}{c}(g_i - 1)^2}\,, \tag{1.20}$$

with $c = \frac{\mu}{(1+\mu)^2}$, where the standard deviation of the growth rate scales as

$$\sigma(q) \propto q^{-\beta}\,, \tag{1.21}$$

with $\beta = 0.5$, as in the job market competition scenario.

## 1.4 Preliminaries for analyzing the dynamics

In the following, the dynamics of the model is analyzed. Iteration of scenarios (i) or (ii) each result in the same fat-tailed size distribution, which is however not a Zipf law (i.e. a power law with exponent $\alpha$=1). In order to be able to describe it, some existing theory on the formation of power laws is introduced.

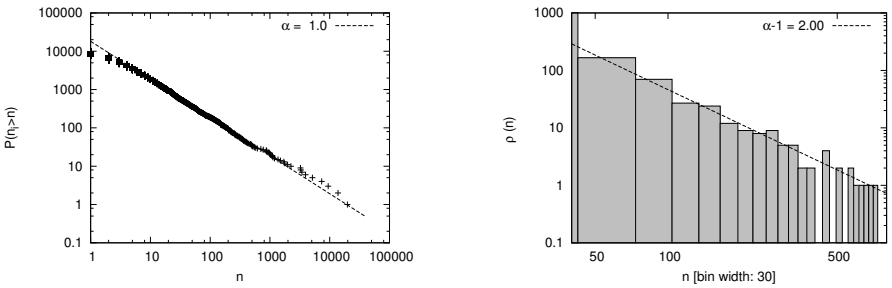
### 1.4.1 Brief introduction of existing models and power laws

Since in the literature, firm growth has often been associated with the formation of power laws, they are briefly introduced here. A power law distribution of a variable $x$ is characterized by a density

$$p(x) \propto x^{-\alpha-1} \tag{1.22}$$

Since this density diverges at the origin, its integral is often used,

$$P(x \geq x') = x^{-\alpha} \tag{1.23}$$




(a) (b)

Figure 1.1: Examples of counter-cumulative size distribution and density of the same power-law distributed data.

It is more precise to plot the counter-cumulative size distribution of data in order to verify whether it is power-law distributed, since for the density, data needs to be grouped into bins, where some information gets lost. Besides these two, a more precise method is the use of a rank-1/2-estimator [18, 73], which has not been applied yet to the presented results.

Power laws are an ubiquituous phenomenon in complex systems and are found in physical, biological and in particular social systems, which are studied more in detail in this thesis. They occur for instance in phase transitions of physical systems, in a blood vessel system of mammals, as well as in the distributions of firm sizes, city sizes, income, or stock market fluctuations (see for instance [15, 105, 95, 18, 5]). Different mechanisms exist which can generate them [70, 73], but for this thesis only explanations with some sort of multiplicative noise are relevant. A model for the formation of power laws (originally

formulated for the genera size distribution of biological species) was introduced in 1925 by U. Yule [108]. However in the following, the focus lies on explanations for systems of constant size, since these are more apt for a comparison to the model studied in this thesis.

### 1.4.2 The Langevin equation

The Langevin equation was introduced by P. Langevin in 1908 in order to describe Brownian motion of particles in a fluid [62]. In its linear form in discrete time, it can be written as

$$n_{t+1} = g_t n_t + \xi \,. \tag{1.24}$$

If $g$ is a damping constant $< 1$ and $\xi$ a stationary noise term, this equation is solved by a normally distributed function $\rho(n)$. This is for instance used to derive Boltzmann-Gibbs statistics for particles in an ideal gas: If $n$ is identified as the momentum of particles $p$, the distribution of the energy $E = \frac{p^2}{2m}$ of free particles decays exponentially. The Langevin equation has also been widely studied where $g$ is a multiplicative noise term, e.g. by [84, 98, 9, 15] who give examples from physics and chemistry, where the interpretation of the multiplicative noise are fluctuations of an external field. This explanation has also been applied to the formation of firm size and city size distributions [109, 68, 43], as well as to income distribution (e.g. [25, 58]). A variable $n$ whose evolution is described by equation (1.24) exhibits a stationary distribution with a power law tail, in the presence of a multiplicative noise term $g$ and an additive noise term $\xi$. $n$ may be a continuous variable or, as in the context of firm growth, discrete, denoting the size of a firm $i$ from an ensemble of firms. The existence of a stationary distribution with a power law decay can be derived from an equation in discrete time (1.24) or continuous time (1.31). In the following, two formalisms taking different approximations are presented.

### 1.4.3 Derivation via the master equation

Gabaix [44] shows the existence of power laws for a city size distribution based on an argument by Champernowne in 1953 [25] and developed rigorously by Kesten [56]. The counter-cumulative size distribution, which is the probability that a firm $i$ is bigger than a value $x$, is defined as $H_{t+1}(x) = P(n_{i,t+1} > x)$. Its evolution in one timestep, without additive term $\xi$, can be written as

$$H_{t+1}(x) = P(n_{i,t+1} > x) = P(g_{i,t} n_{i,t} > x) = P\left(n_{i,t} > \frac{x}{g_{i,t}}\right) \tag{1.25}$$

$$= \int_0^\infty H_t\left(\frac{x}{g}\right) G(g) d\,g \,, \tag{1.26}$$

where $G(g)$ is the distribution of the growth rate $g$. In a stationary state, $H_{t+1} = H_t$, so the relation becomes

$$H(n) = \int_0^\infty H\left(\frac{n}{g}\right) G(g) d\, g\,. \tag{1.27}$$

For the distribution $H(n)$, [44] and [84] show that a trial function $H(n) = c/n^\alpha$ ($c = const$) yields the following relation for the noise $g$: $1 = \int_0^\infty g^\alpha G(g) dg$ which is

$$E[g^\alpha] = 1\,. \tag{1.28}$$

This holds, however, only if such a stationary state exists, which is only the case if additionally some additive noise $\xi$ is present. Takayasu et al. [98] equally derive equation (1.28). He derives the fact that the stable distribution has a power law tail at its lowest order, i.e. $H(\geq n) = c/n^\alpha$, from the condition of continuity of its characteristic function (i.e. its Fourier transform). However, the additive term $\xi$ in (1.24) is needed for the system to achieve a stationary state where $H(x)$ follows a power-law distribution. For systems of constant global size, where the additive term is small, the noise distribution $\mathcal{G}(g)$ is centered around a value close to 1, so $\langle H_{t+1}\rangle = \langle g\rangle \langle H_t\rangle$, so $\langle g\rangle = 1$. For such noise, equation (1.28) yields $\alpha = 1$, which is called the Zipf law. Applied to firm growth, possible interpretations are that the additive term $\xi$ prevents firms from becoming too small, or that some firms are continuously being started and compensate firms that become smaller than a threshold and exit[5]. Since power law distributions are conserved under addition of a faster decaying distribution [44], the presence of additive noise does not affect the power law exponent.

A model for firm growth consisting only of multiplicative noise was introduced by R. Gibrat in 1931 [46]. Since a stabilising additive term is absent, it does not yield a stationary distribution. His model supposes that the noise can be written as $g_t = 1 + \epsilon_t$, where $\epsilon_t$ is uncorrelated noise. Iterating equation (1.24) over $m$ timesteps in absence of additive noise yields,

$$\frac{n_{t+m}}{n_t} = \prod_{t=0}^{m-1} (1 + \epsilon_t) \tag{1.29}$$

Taking the logarithm of equation 1.29,

$$\log \frac{n_{t+m}}{n_t} = \sum_{t=0}^{m-1} \log(1 + \epsilon_t)\,, \tag{1.30}$$

the right side of equation 1.30 can be approximated as $\sum_{t=0}^{m-1} \epsilon_t$, which is normally distributed due to the Central Limit theorem. For the growth rate $\frac{n_{t+m}}{n_t}$, this yields a lognormal firm size distribution, whose variance grows over time [43]. A number of later models have used and built on this model [71, 48, 52,

[5] If firm size is discrete, this can simply mean that firms which have reached size $n_i = 0$ are replaced.

54, 53] As derived in [27] and detailed further in chapter 3, the dynamics of the model by Gibrat are compatible with a stationary power law-decaying size distribution if new firms enter the system, such that an exponential age distribution is assumed. This amounts to assuming an additive noise term $\xi \neq 0$. This $\xi$ has a positive mean *for the firms of size 0*, but for the entire system it can have a mean of 0 or $> 0$.

### 1.4.4 Derivation via the Fokker-Planck equation

The existence of power laws in a stationary state can also be derived by solving the Fokker-Planck equation instead of the master equation [84, 17, 15]. Here the derivation by Biro and Jakovác [15], who study the Langevin equation in continuous form, is detailed:

$$\dot{n} + \gamma n = \xi \tag{1.31}$$

with a multiplicative noise term $\gamma$ and additive noise term $\xi$ [6]. In their derivation, $\gamma$ is assumed to be white noise of mean $\langle\gamma(t)\rangle = G$ and autocorrelation $\langle\gamma(t)\gamma(t')\rangle - \langle\gamma(t)\rangle\langle\gamma(t')\rangle = 2C\delta(t-t')$. The additive noise has mean $\langle\xi(t)\rangle = F$ and autocorrelation $\langle\xi(t)\xi(t')\rangle - \langle\xi(t)\rangle\langle\xi(t')\rangle = 2D\delta(t-t')$. The cross-correlation between the two noises is $\langle\gamma(t)\xi(t')\rangle - \langle\gamma(t)\rangle\langle\xi(t')\rangle = 2B\delta(t-t')$. To derive a differential equation they apply a method of multiplying the distribution $f(n,t)$ by a trial function $R(n)$, which they expand until the second order of $n$, valid in infinitesimally short timescales. Keeping only linear terms in $dt$ they derive a Fokker-Planck equation, which is solved for the stationary state, as also done by [17, 106]. The idea of this method is that in short timescales, the variable $n$ cannot vary very widely. It has been introduced by [104] originally to derive the Gaussian distribution as the stationary distribution of Brownian motion. Here, it leads to the Fokker-Planck equation

$$\frac{\partial f}{\partial t} = -\frac{\partial(F - Gn)}{\partial n} + \frac{\partial^2((D - 2Bn - Cn^2)f)}{\partial n^2}\,, \tag{1.32}$$

whose general solution for the distribution $f(n)$ is

$$ln\frac{f(n)}{f(0)} = -\left(1 + \frac{G}{2C}\right) ln\frac{(D - 2Bn + Cn^2)}{D} - \frac{G\frac{B}{C} - F}{\sqrt{CD - B^2}}\text{atn}\left(\frac{\sqrt{CD - B^2}\,n}{D - Bn}\right) \tag{1.33}$$

where atn is the inverse tangent function. This has been solved in two limits:

1. In the limit C=B=0, which means the absence of multiplicative noise $\gamma$ (and absence of the cross-correlations of the noises). The last term vanishes since the argument of the inverse tangent function (1.33) becomes zero. The derived solution for $f(n)$ is

$$f(n) = f(0)e^{-\frac{G}{2D}n^2}e^{\frac{F}{D}n}\,, \tag{1.34}$$

[6] Compared to the discrete equation (1.24), the multiplicative term $g$ corresponds to $1 - \gamma$ here.

which is the Gaussian distribution and which simplifies to $f(n) = f(0)e^{-\frac{G}{2D}n^2}$ without drift ($F = 0$). If $n$ is thought of as the momentum of particles, this solution corresponds to a Gaussian distribution of the momentum of free particles. This in turn leads to the Boltzmann-Gibbs statistics for gas particles for $n^2$.

2. In the limit D=B=0, which is the case of pure multiplicative noise and corresponds to a large argument of the inverse tangent function in equation 1.33.

$$f(n) = f(0)n^{-2-G/C}e^{\frac{F}{C|n|}}\,, \tag{1.35}$$

which is a Gamma distribution in $1/n$ and has a power law tail.

The Fokker-Planck equation (1.32) has also been derived by [81], who solve it in a different way: they show that by choosing trial functions in a specific way, the equation reduces to one that describes random walk for a variable in a specific potential field. The solution of this reduced equation is an exponential decaying distribution, as is the case the with Boltzmann formula. Through the retransformation of the continuous Langevin equation to the original variable $n$, the Gamma distribution is derived. Other approaches leading to these results are [94] and [67].

### 1.4.5 The relationship between multiplicative noise and entropy

Power law distributions in a system can be the result of long-range interactions of its elements, as is the case in phase transitions [80]. These states are not stationary states, and only with some additional additive term will the system stay in that state. However, a power law distribution can be an attractor of the dynamics, such that a system may in practice never leave it [37]. Instead of deriving a stationary distribution from the dynamics of the system, another possible way is to derive it from extremality of an entropy function, taking into account constraints on the degrees of freedom. Its idea is that in a stationary state, the system is as random as possible under these constraints [8]. For instance, the normal distribution, which is the stationary distribution of equation (1.31) with additive uncorrelated noise and a multiplicative factor $\gamma < 1$, can also be derived from the extremality of the Gibbs-Shannon entropy

$$S^{(G)}(\{p_i\}) = \sum_i p_i k \log p_i \tag{1.36}$$

where $k$ is a constant, under the constraint $\sum_i p_i = 1$ and $\sum_i H_i\, p_i = E$ are conserved, where $H_i \propto p_i^2$:

$$\frac{\partial}{\partial p_i}\left[\sum_i p_i k \log p_i - \lambda\left(\sum_i H_i p_i - E\right) - \eta\left(\sum_i p_i - 1\right)\right] = 0\,, \tag{1.37}$$

where $\lambda$ and $\eta$ are Lagrangian multipliers. It is solved by

$$p_i = e^{\left[-\frac{\lambda}{k}H_i + (\frac{\eta}{k}-1)\right]} \tag{1.38}$$

If $H_i \propto p_i^2$, $p_i$ is normally distributed, like the solution (1.34) of the Fokker-Planck equation. This derivation corresponds to the canonical ensemble. The probability $p_i$ in equation (1.38), if written as a function of energy $H_i \propto p_i^2$, is called the Boltzmann factor of a particular energy state, $B(p_i^2) = B(H_i) \propto e^{-\beta H_i}$.

Analogously, the stationary state of a system with multiplicative noise can be explained by the extremality of a different entropy function than the Gibbs-Shannon entropy, which becomes extremal for equilibrium systems.

[15, 106] showed that the solution for the Fokker-Planck equation (1.32) with large multiplicative noise yields a Gamma distribution for $n$ (equation (1.34), and, as distribution for a quantity $H_i$ ($\propto n_i^2$):

$$p_i = f_0 \left(1 + (q-1)\frac{H_i}{T}\right)^{\frac{q}{1-q}} . \tag{1.39}$$

This distribution is termed $q$-exponential distribution, since it is the analogon to an exponential distribution (1.38) resulting from the maximum of the Gibbs-Shannon entropy. In physical applications, $H$ corresponds to energy, and $T$ to the temperature of a system. In [15], the exponent $q$ is identified as $q = 1 + \frac{2C}{G}$, where $C$ is the variance of the multiplicative noise, and $G$ is its mean value. The same distribution can be derived from extremality of the Rényi entropy $S^{(R)}$ [82]:

$$S^{(R)}(\{p_i\}) = \frac{q}{q-1} \log \sum_i p_i^{1/q} , \tag{1.40}$$

which becomes extremal when $p_i$ is Gamma distributed, exhibiting a power law tail. When exhibiting this distribution, the system is as indeterminate as possible. In order to derive the distribution (1.39), the derivative of equation (1.40) with respect to $p_i$ is set to zero, again under the constraints that the quantities $1 = \sum_i p_i$ and $E = \sum H_i p_i$ are conserved (see for instance [8][7]). Alternatively, equation (1.39) can also be derived from the nonextensive Tsallis entropy, which is equal to the linear expansion of the logarithm in the Rényi entropy (1.40). Other derivations of the parameter $q$ from the dynamics of the system have been derived by [9, 106]. In the limit $q \to 0$, which corresponds to a vanishing amplitude of the multiplicative noise, both the Rényi entropy and the Tsallis entropy become the Gibbs-Shannon entropy. To summarize, extremality of the Gibbs-Shannon entropy under the constraint of a conserved quantity yields a Gaussian distribution for $n$, so the Gibbs-Shannon entropy corresponds to solution (1.34) of equation (1.33), and the Rényi entropy to solution (1.35).

Although a lot of empirical evidence has been reported, a question remains as to why the Rényi entropy is the appropriate function that becomes extremal for a system with multiplicative noise. This entropy function was introduced by Rényi

[7] [8] uses the convention $q' = 1/q$ of the parameter $q$ presented here

[82] as a function of a generalized probability distribution that satisfies an axiom of additivity for entropies of independent degrees of freedom. Bashkirov [8] gives an intuitive interpretation of its definition in the context of physics: Since a nonequilibrium system exchanges entropy with its environment, he proposes the construction of a new function, which is the cumulant generating function of the Gibbs-Shannon entropy

$$\psi_S(\lambda) = \log \sum_i e^{\lambda S_i^{(G)}} \tag{1.41}$$

Multplied by a factor $\frac{1}{1+\lambda}$, i.e. the Rényi entropy is obtained $S^R = \frac{1}{1+\lambda} \log \sum_i e^{\lambda S_i^{(G)}}$ $\equiv \log \sum_i p_i^\lambda$. The factor guarantees that for $\lambda \to -1$, the Rényi entropy becomes the Gibbs-Shannon entropy again. His reasoning follows an analogy to the way the free energy $F = -kT \log \sum_i e^{-H_i/(kT)} = \sum_i p_i H_i = U$ can be constructed from the internal energy, where $F$ is extremal in equilibrium thermodynamic systems that can exchange heat with their environment.The temperature $T$ is the parameter needed for the passage from a mechanical equilibrium to a thermodynamic equilibrium. [8] draws a comparison from $T$ to the parameter $\lambda$ which is needed to describe a system which can exchange entropy with its environment.

## 1.5 Theoretical Analysis

### 1.5.1 Discussion on additive and multiplicative noise

Additive and multiplicative noise terms in the Langevin equation lead to different stationary size distributions, which have been described in section 1.4.4, where results of the corresponding Fokker-Planck equation are given. It is illustrated with numerical simulations of both systems in figures 1.2 and 1.3. If $f(n)$ is thought of as the firm size distribution, additive fluctuations tell the absolute change in size of a firm, and multiplicative fluctuations tell the ratio by which a firm's size has changed, i.e. they are relative fluctuations. These two have different $n$-dependencies, which are both different to the $n$-dependency of the model introduced in section 1.3. $n$-dependency of the standard deviation of absolute and relative fluctuations are compared for additive noise, multiplicative noise and the introduced model in table 1.1.

For the introduced model, the time evolution of the firm size distribution cannot be described by equation (1.24), where $g$ and $\xi$ are drawn from distributions independent of $n$. Here, the standard deviation $\sigma$ of the noise depends on $n$. It may either be written as additive Gaussian noise with $\sigma_{add} \propto \sqrt{n}$

$$n_{t+1} = n_t + \xi(\sqrt{n})\,, \tag{1.42}$$

where $\xi$ is Gaussian with $\sigma \propto \sqrt{n}$, or as multiplicative Gaussian noise with $\sigma_{mult} \propto \frac{1}{\sqrt{n}}$,

$$n_{t+1} = g(\frac{1}{\sqrt{n}})n_t \tag{1.43}$$

| | $\sigma$ of relative fluctuations | $\sigma$ of absolute fluctuations | stationary size distribution |
|---|---|---|---|
| purely additive noise | $\propto 1/n$ | *const* | Gaussian |
| purely multiplicative noise | *const* | $n$ | Power law |
| studied model | $\propto 1/\sqrt{n}$ | $\propto \sqrt{n}$ | Fat-tailed |

Table 1.1: Different $n$-dependencies for additive and multiplicative noise in comparison to the one from the introduced model.

where $g$ is Gaussian with $\sigma \propto \frac{1}{\sqrt{n}}$ noise. Formulation (1.43) has been chosen for the equations (1.12) and (1.20). It corresponds to a more complicated stochastic process that has unfortunately not yet been solved analytically. Also in the literature, only few solutions for more complicated processes exist [84]. A different process whose noise has the same $n$-dependency as this model is treated in [68], and discussed further in section 1.11.

**Comparison of the noises in a different formulation.** This technical reflexion is useful for the description of the fluctuations of the studied model. Systems with multiplicative noise can not be described as maximizing the Gibbs-Shannon entropy, as described above. The probabilities of particular states are not Gaussian distributed, and their squares are not exponentially distributed, i.e. they are not described by Boltzmann factors of the form $B(H) \propto e^{-\beta H}$. However, it is possible to re-formulate equation (1.31) such that multiplicative noise is expressed in terms of some additive noise which follows a particular, non-Gaussian distribution $f(\hat{\xi})$. This equivalent additive noise then exhibits a different $n$-dependence (see table 1.1):

$$n_{n+1} = g_t n_t = n_t + (g_t - 1)n_t \equiv n_t + \hat{\xi}_n \tag{1.44}$$

This additively written noise $\hat{\xi}_n$ contains information on the *absolute* fluctuations of $n$, but it remains $n$-dependent. It is possible to express the noise distribution in terms of $B(H)$, where integration over $n$ of $f(\xi_n)$ is performed and yields $B(H) = \int \rho(n) e^{-f(\xi_n)H} dn$. This $B(H)$ is a distribution independent of $\xi_n$, which may be thought of as the analogon of firm size distribution in the studied model. Beck and Cohen [10] describe such a system by introducing a generalized Boltzmann factor $B(H)$, where the integration is done over inverse temperatures $\beta$. They call the concept superstatistics, i.e. statistics of statistics.

The concept of mapping the noise and performing an integral over a $n$-dependency is used to explain the results of the model introduced in section 1.3. The difference is that for the firm growth rate in the model, it is not *absolute* fluctuations that are relevant, but *relative* fluctuations, which is why the noise is

expressed in terms of multiplicative noise. It will be discussed further in section 1.11, after numerical results have been provided.

### 1.5.2 The growth rate probability distribution

Figure 1.8 shows that despite the normally distributed $\mathcal{G}(g|n)$, the aggregate $\mathcal{G}(g)$ exhibits a tent-shape. The $n$-dependence (equation 1.13) is the reason why the growth rate distribution of $\mathcal{G}(g|n)$ is wider for small firms and more narrow for big firms. All firms contribute to the aggregate $\mathcal{G}(g)$ and the growth rate distribution for the $N_f$ firms can be written as

$$\mathcal{G}(g) = \frac{1}{N_f} \sum_{i=1}^{N_f} \mathcal{G}(g_i|n_i)\,, \tag{1.45}$$

or, in the continuous limit:

$$\mathcal{G}(g) = \int_0^\infty dn \mathcal{G}(g|n)\rho(n)\,, \tag{1.46}$$

where $\rho(n)$ is the firms' size distribution. Since so far no analytical expression for the simulated size distribution of this model is derived, we evaluate the integral as an approximation for power-law size distributions of exponent $\alpha$. For firms' size distributions $\rho(n) \propto n^{-\alpha-1}$ and scaling exponents $\beta = 0.5$ the integral yields

$$\mathcal{G}(g) \propto \int_0^\infty n^{0.5} \frac{1}{n^{\alpha+1}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}n\,(g-1)^2} dn\,, \tag{1.47}$$

which yields $\frac{1}{\sqrt{\pi}} 2^{-\alpha}\,(g-1)^{2\alpha-1}\Gamma\big(\frac{1}{2}-\alpha\big)$ for $\alpha < 0.5$ and simplifies to $\mathcal{G}(g) = \frac{1}{|g-1|}$ for $\alpha = 0$. Since $\rho(n)$ is not a normalized probability density, (1.47) is not defined at $g = 1$ unless if integrated from a minimal size $n_0 > 0$ instead of 0. In that case, it is defined for $\alpha < 0.5$ for any $g$, albeit it is not a normalized probability density. However its tent-shaped form explains the form of the numerically obtained growth rate probability density. Numerically, the number of firms is discrete, so (1.45) is performed, which is always finite at any value of $g$, independently of $n_0$. Integral (1.47) can be generalized to values of $\beta$ other than 0.5, which is interesting since empirical values are $\alpha \approx 1$ and $\beta \approx 0.25$. The smaller is $\beta$ , the less peaked $\mathcal{G}(g)$, which is intuitive, since if $\beta = 0$, the result is a Gaussian $\mathcal{G}(g)$.

Notice that the shape of $\mathcal{G}(g)$ is not very sensitive to the underlying size distribution: equation (1.47) yields an approximate tent-shaped $\mathcal{G}(g)$ even for exponential decay of $\rho(n)$. This suggests that despite the fact that the size distribution in the proposed model deviates from a Zipf law, the idea of performing integral (1.46) explains the observed tent-shape well.

In the literature, the principle of performing this integral has been used in the model by [42] to obtain a tent-shaped growth rate distribution of a single firm. Other models [95, 63, 19, 78] do not perform the integral and do not clearly

distinguish between the growth rate probability at firm level and at aggregate level. Both are fitted with a Laplacian distribution. The necessity to perform the integral in equation (1.46) is however independent of the assumed $\mathcal{G}(g|n)$. The functional form of $\mathcal{G}(g)$ yields an approximate $1/|g-1|$ tent-shape for both Laplacian and Gaussian $\mathcal{G}(g|n)$. If it is integrated from a size cutoff, for low values of $\beta$ , Laplacian $\mathcal{G}(g|n)$ yield a more peaked aggregate $\mathcal{G}(g)$. Since the growth rates typically have values close to 1, empirical evidence can often be fitted equally well with a Laplacian (centered around 1) and a $1/|g-1|$-function. However [42] find that the tails of the tent-shape exhibit power law decay rather than exponential decay of a Laplacian, substantiating the argument presented here.

## 1.6 Numerical implementation

Some general technical details are given here. Parameters are found in the figure captions.

### 1.6.1 The sequence of events

At each iteration, the following events happen:

1. production decision: $\hat{q}_i = q_{i,t-1}(1 + \mu^{net}_{i,t-1})$
2. hiring decision: $\hat{n}_{i,t} = \frac{p}{w}\hat{q}_{i,t}\frac{1}{1+\mu_i}$
3. job market: $n_{i,t}$ is determined.
4. production: $q_{i,t} = \frac{w}{p}n_{i,t}(1 + \mu_i)$
5. goods market: $q^s_{i,t}$ is determined.
6. $\mu^{net}_{i,t}$ is calculated.
7. firms of size 0 are replaced by one of a size drawn from a distribution.

This also holds in the extended model described in chapter 2, where additional functions are present. A schema is shown on page 4 of publication 1.

### 1.6.2 Rounding method

In equation (1.4) $\hat{q}$ (and also $\hat{n}$) are not necessarily integer numbers and rounding is needed. In order to minimize rounding errors, the following method is introduced:

Let $k_{\hat{n}} \in [0, 1]$ be

$$k_{\hat{n}} = \lceil n(1 + \mu)\rceil - \lfloor n(1 + \mu)\rfloor \tag{1.48}$$

Rounding is then done using $k_{\hat{n}}$ as a probability:

$$\hat{n} = \begin{cases} \lceil n(1+\mu) \rceil & \text{with probability } k_{\hat{n}} \\ \lfloor n(1+\mu) \rfloor & \text{with probability } 1 - k_{\hat{n}} \end{cases} \tag{1.49}$$

This minimizes the rounding errors from discretization. This rounding method implies that it is possible for small firms to grow with a rate $g > (1+\mu)$: a firm of size 1, which sold its entire production, demands $(1+\mu)$ workers in the next iteration, which may be rounded to 2 with a probability of $k_{1+\mu} = \mu$.

### 1.6.3 Additive term $\xi$ in scenarios (i) and (ii)

In scenario (i) where firms consume, and (ii) where firms do not consume, newly introduced firms affect the system in slightly different ways. In both scenarios, new firms contribute to the job offer $\sum_j \hat{n}_j^{new}$

$$\hat{N} = \sum_i \hat{n}_i + \sum_j \hat{n}_j^{new} \tag{1.50}$$

- In scenario (i), the stationary state is at full employment, and only a fraction $p = \frac{N_w}{\hat{N}}$ of positions will receive a worker. This $p$ will be slightly lower than $\frac{1}{1+\mu}$ whenever a new firm is started.
- In scenario (ii), all positions are filled, so additional job offers cause an increase in workforce (i.e. decrease in unemployment) by $N^{new} \sum_j n_j^{new}$ workers. Therefore, the next production would be higher, and if iterated many times the system would tend towards full employment. In order to avoid this, the job offer of entrant firms $N^{new}$ is subtracted from existing firms. This is implemented such that every job offer of existing firms has the same chance of being eliminated. The result is that the job offer at iteration $t+1$ also equals the job offer at iteration $t$ when new firms are started. This means that the additive noise $\xi$, if considered for the entire system, has mean 0, since it does not change the size of the system, but merely shifts some workers from existing firms with $n_i > 0$ to firms with $n_i = 0$.

The latter method is also used in the simulations of the similar model by Marsili and Zhang [68], shown in figures 1.14 and 1.15. It is not used in the extended model in chapters 2 and 3, where unemployment can evolve freely.

## 1.7 Numerical Results

Although the following two approximations do not correspond to the model introduced in section 1.3, I simulated for comparison systems with purely additive noise (figure 1.2) and multiplicative noise (figure 1.3). Only a small additional term $\xi$ has been present in order to keep the system at constant global size.

Since the elements (firms) are discrete, rounding towards discrete values has the role of additive noise. In the following subsections, the results of the model are presented.

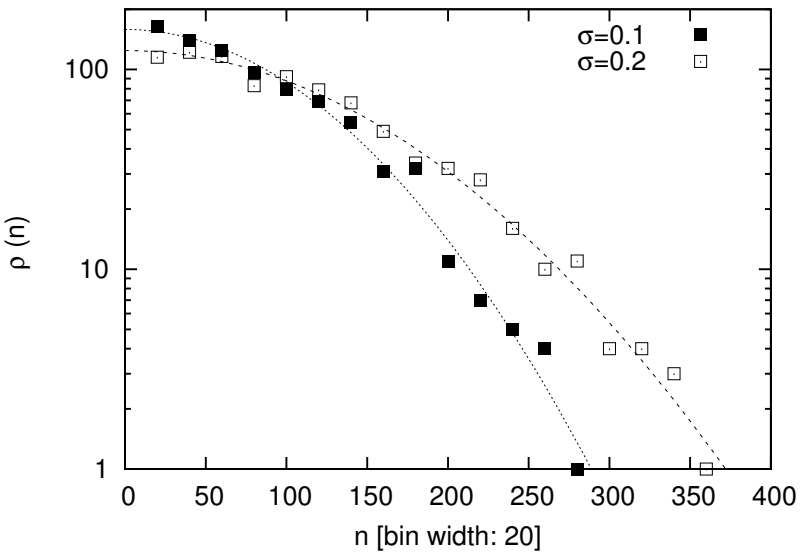



Figure 1.2: Density of the size distribution of a system with purely additive uncorrelated Gaussian noise of mean zero, and absence of a multiplicative term. Simulations of two different standard deviations, on a log-linear scale. In this scale, a parabola corresponds to a Gaussian distribution. $N_w = 10^5$, $N_f = 10^3$.

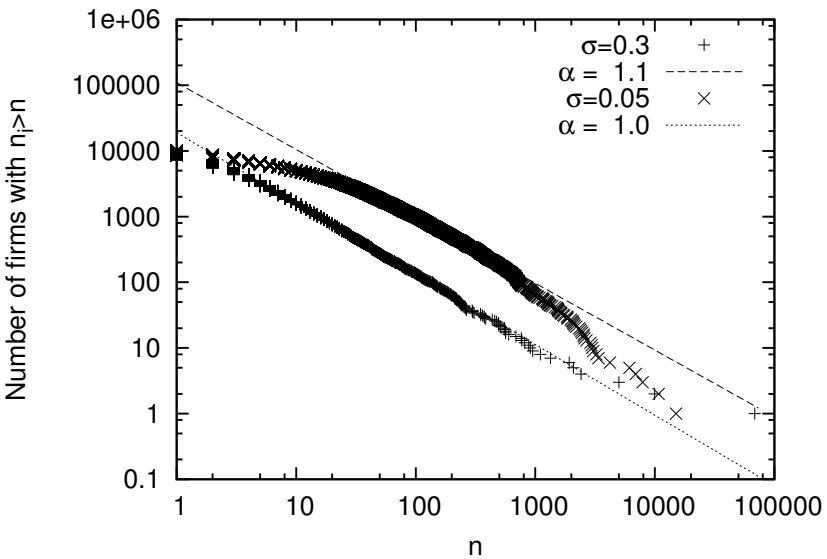
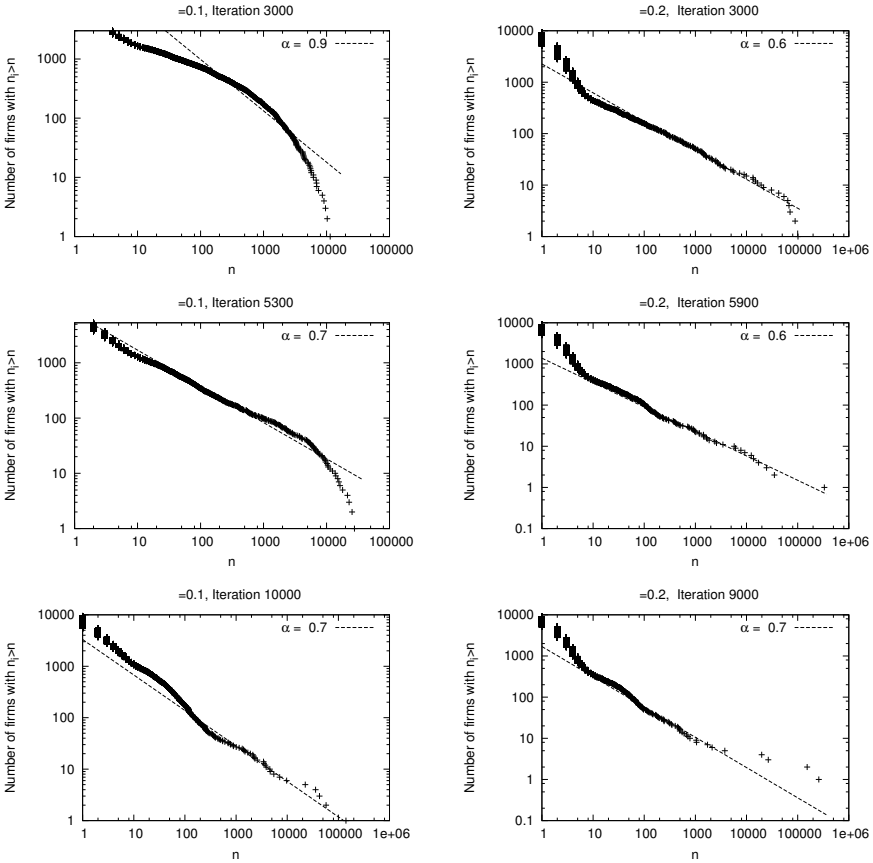



Figure 1.3: Numerical counter-cumulative size distributions of the Langevin equation (1.24), with multiplicative noise, which is uncorrelated and normally distributed. In this double logarithmic scale, a straight line corresponds to a power law. The smaller the variance of the noise distribution, the slower the convergence, and even after convergence the distribution remains concave. This is because rounding towards discrete values modifies the growth rate, such that it resembles additive noise. The smaller $\sigma$, the stronger this effect is. Results are from a system with $10^4$ firms and $5 \cdot 10^5$ workers, after 4000 iterations. Whenever an element reached size 0, it was replaced by one of average size 1.5. The result is not sensititve to the size of re-initialisation.

### 1.7.1 Size distribution

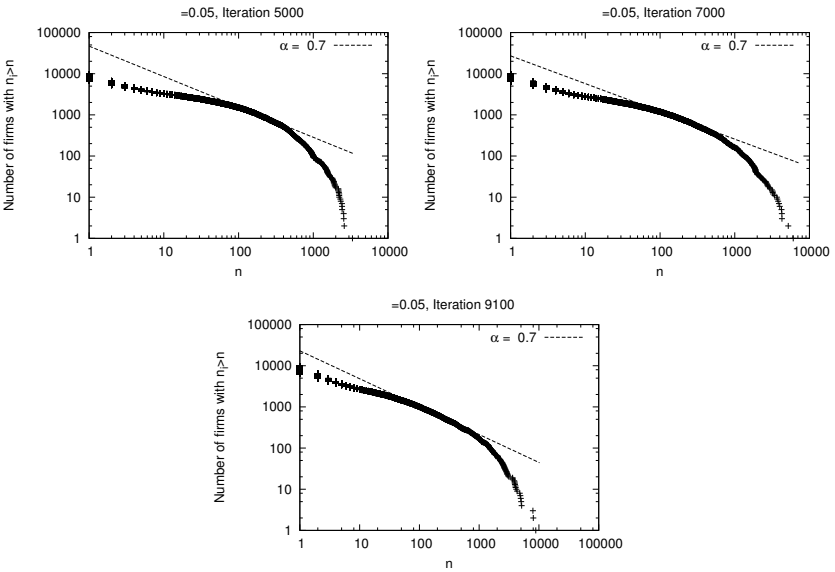
In figure 1.4, typical examples of the time evolution of a scenario where the only constraint is limited purchasing power in the goods market are shown. To keep the system at constant global size, firms that reached size 0 were re-started $n \in [1, 2]$ (1.5 on average), by the method detailed in appendix 1.6.3. Compared to simulations of systems with multiplicative noise, simulations of this model take much longer to approach a power law, whose exponent is $\alpha \approx 0.7$, which is much lower than the exponent 1 of a Zipf law. $\alpha = 1$ is the lowest exponent that is found with multiplicative noise (see figure 1.3). This flatter power law decay has been found for different values of $\mu$ in the range $0.05 \leq \mu \leq 0.2$. As shown in successive snapshots in figure 1.4, it is not stationary. The higher $\mu$ is, the faster the system converges to it, but the faster the system leaves it again too.

Typical size distributions for different values of $\mu$ are shown in figure 1.7, which remain stable over a long period. The observed effect is that the higher $\mu$ is, the closer the result approaches a power law. The interpretation is that for small $\mu$, the rounding introduced by equation (1.49) modifies the actual planned production $\hat{q}_i$ more than the multiplication by $(1 + \mu^{net}_{i,t-1})$. The modification through rounding can be regarded as additive noise rather than the noise with $\sigma \propto n^{-1/2}$- scaling. Indeed, the smaller $\mu$ is, the less the size distribution fluctuates. This effect is stronger for smaller systems, as the comparison of systems with 2000 and 10000 firms has shown (see figure 1.4).



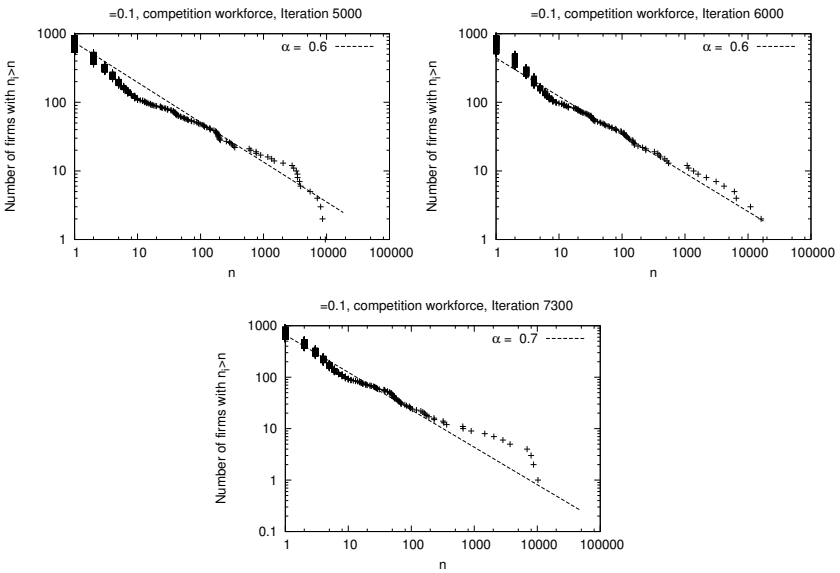

(a) (b)

Figure 1.4: Two examples of the evolution of firm size distribution, between 3000, and 10000, steps of convergence. Left column: $\mu = 0.1$, right column: $\mu = 0.2$. The power law exponent is $< 1$. Extinct firms are replaced by one of size $n \in [1, 2]$ with the method detailed in appendix 1.6.3. For a value of $\mu = 0.05$, the size distribution can be fitted with a power law of exponent $\approx 0.7$, too, but the tail stays concave and does not converge to it (see next figure).



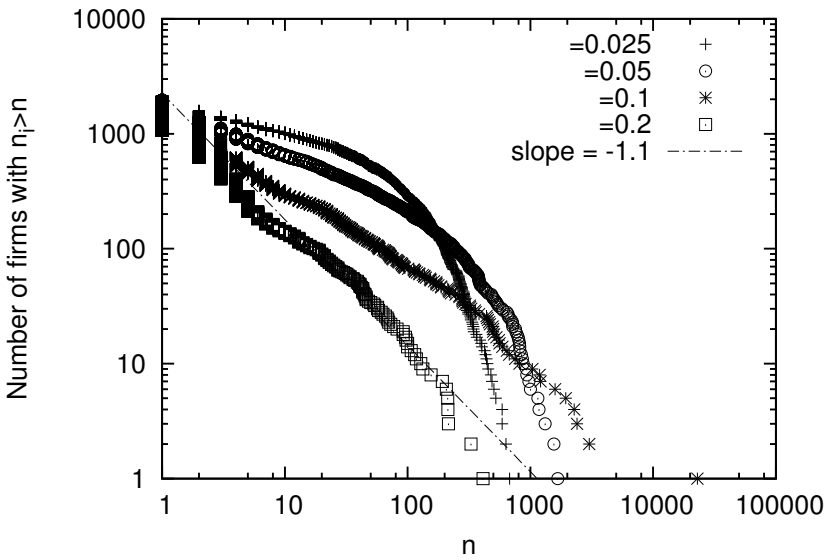

(a)

Figure 1.5: Same setting as the simulations shown in 1.4, but with $\mu = 0.05$. The size distribution stays more concave.

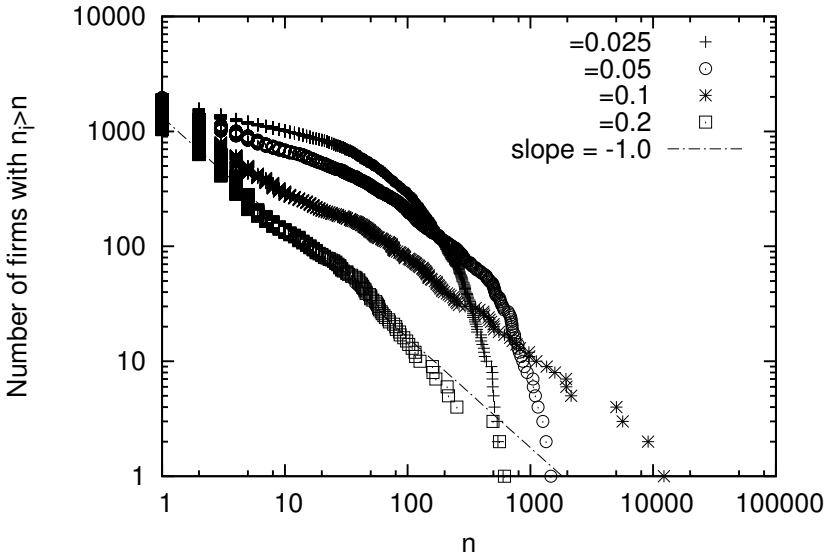



(a)

Figure 1.6: Example of the evolution of firm size distribution, between 5000 and 7300, steps of convergence. $\mu = 0.1$, for scenario (i) where the firms compete for workforce. The power law exponent is $< 1$. Extinct firms are replaced by one of size $n \in [1, 50]$. The obtained size distribution does not seem to depend on whether scenario (i) is used, or (ii) as in figure 1.4.



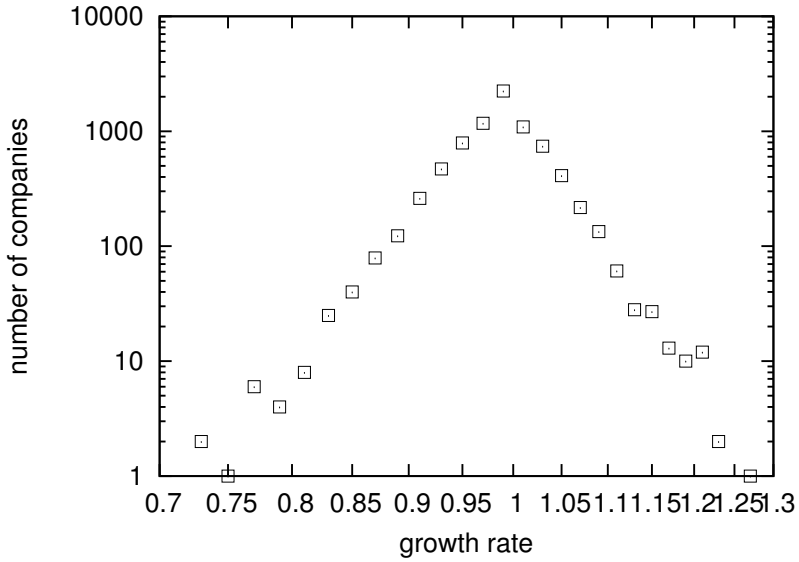

(a)




(b)

Figure 1.7: Firm size distribution after 5000 iterations for a system of $2 \cdot 10^3$ firms and $9 \cdot 10^4$ workers for different values of $\mu$. Whenever a firm has attained size 0, it is replaced by a firm (a) of size from the interval $[1, 2]$, or (b) of size from the interval $[1, 50]$, yielding very similar results. The larger $\mu$ is, the closer the distribution approaches a power law after the same time. In these simulations firms do not consume, i.e. the system's only constraint is that demand for goods is lower than production, but workforce availability is sufficient (scenario 2 detailed in section 1.3.2). Growth rates are binomial (the method in section 1.8 is not used).

### 1.7.2 Growth rate probability distribution




(a)

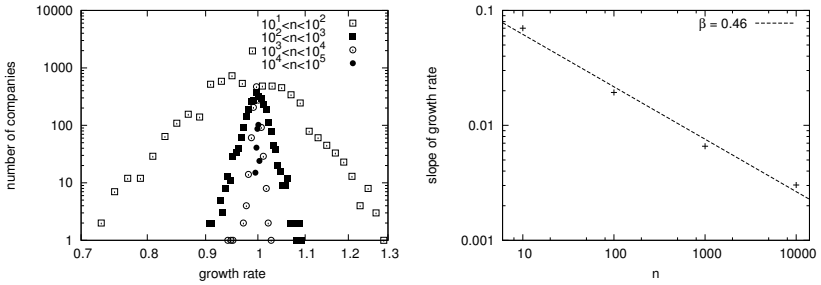
Figure 1.8: growth rate distribution with rounding detailed in 1.8, in a system with full employment, where the only constraint is workforce availability. $10^4$ firms, $10^6$ workers.

Figure 1.8 shows the tent-shaped growth rate distribution. Its explanation, detailed in 1.5.2 uses a superposition of Gaussian $\mathcal{G}(g|b)$ of $n$-dependent standard deviation. Small firms have larger $\sigma$, which is why they account for the 'fat tails', whereas big firms dominate in the peak of the growth rate distribution.

**Artefacts from binning** Some effects arising from binning data are addressed here. Empirical data in [95, 63, 97, 78] exhibits tent-shaped growth rate distributions of different widths depending on firms size (or country's size or citations respectively). For all of these, a Laplacian fit is proposed. To do this, firms are grouped according to their size in large logarithmic bins. From the slopes of the growth rate distribution on logarithmic scale, $\sigma(n)$ and its scaling exponent $\beta$ are determined.

Numerical simulations of the model show that aggregation of firm growth rates within one order of magnitude of size is sufficient to obtain a growth rate distribution that resembles a tent-shape, when $\mathcal{G}(g|n)$ is Gaussian (see figure 1.9). The reason for this is that $\rho(n) \approx 1/n^{1.7}$ and $\sigma(n) \propto n^{-0.5}$. This result implies that if the average of an ensemble of firms is used to determine the shape of $\mathcal{G}(g|n)$, its functional form is only assessed correctly if the sampled firms have precisely the same size. If $\mathcal{G}(g)$ is $n$-dependent, a size spectrum of one order of

magnitude is already enough to modify the form of $\mathcal{G}(g|n)$. The value of $\beta$ does not seem to depend on binning. By plotting the slope of the obtained $\mathcal{G}(g|n_{max})$ against $n_{max}$ of the respective bin, the found relation is again $\sigma \propto n^{-0.5}$.




(a) (b)

Figure 1.9: Histograms for growth rate distributions. (a) Aggregate growth rate distribution for companies ranging from 1 to $5 \cdot 10^4$ employees from a system with $10^7$ employees and a Zipf firm size distribution. (b) For Gaussian growth rate probability densities, the growth rate distribution of firms within one order of magnitude (its smallest size indicated in the plot) resembles a tent-shape, thus in qualitative agreement with the data shown in [95]. (c) The slopes of these approximate $1/|g-1|$-distributions follow the same scaling relation as the variance of the Gaussian growth rate probability densities $\sigma(n) \propto n^{-0.5}$.

The very simple underlying microscopic mechanisms suggest that Gaussian functions might be a simpler alternative to the commonly assumed Laplacian shape for $\mathcal{G}(g|n)$, since it also yields a tent-shaped $\mathcal{G}(g)$. Furthermore, Gaussian distributions are conjugate priors to themselves, so may result from several reasons, where each is Gaussian distributed. This idea is detailed further in section 2.3.1.

## 1.8 Alternative implementation: growth of independent subunits

The growth rate probability density of this model is binomial if there is a shortage of workforce (equation 1.9) or of purchasing power (equation 1.16). For large $n$, this distribution can be approximated with a Gaussian distribution of the same variance. An alternative rounding method is now introduced which can be used for the determination of the job offer (case (i)) or the quantity to produce (case (ii)). It is detailed using the example of the job market.

The following setup yields a discrete Gaussian growth rate probability density even for small firms. Firms demand on average a quantity of workers $\hat{n}_i = n_i\,(1+\mu)$, which can be rounded towards integer values using the method introduced in equation 1.49. Instead of rounding the quantity $\hat{n}_i$ to integers, the

rounding can also be done at the level of individual positions: for every existing position $j$, the job offer $\hat{j}$ may be 1 or 2:

$$\hat{j} = \begin{cases} 2 & \text{with probability } \mu \\ 1 & \text{with probability } (1-\mu) \end{cases} \tag{1.51}$$

Then, the job offer $\hat{n}_i$ is the sum of the job offers correponding to the positions of a firm.

$$\hat{n}_i = \sum_{j=1}^{n_i} \hat{j} \tag{1.52}$$

This $\hat{n}_i$ is the offer posted in the job market. Then, the aggregate job offer $\sum_i \hat{n}_i$ is collected. On average, it is $N(1+\mu)$, as with the standard rounding method.

**Combination with the allocation in the job market.** If available workforce $N_w$ is inferior to this offer (which is the case studied here), every open position has a probability

$$p = \frac{N_w}{\sum_i \hat{n}_i} \approx \frac{1}{1+\mu} \tag{1.53}$$

of receiving a worker. This attribution on its own would yield a binomial constraint, depending on a firm's size, as stated in equation (1.9). The growth rate has a cutoff at the upper value $(1+\mu)$, but firms can shrink to any size $\geq 0$. On the contrary, if firms determine their job offer via equations (1.51) and (1.52), the number of received workers follows a symmetric distribution between 0 and $2n$, if $n$ was the size of the firm in the previous timestep. Combining the probabilistic job offer (equation (1.51)) with the binomial allocation of workers in the job market (equation 1.9), a single job has a certain probability to double, a certain probability to reproduce itself, and a certain probability to vanish:

$$p(j=2) = q = \frac{\mu}{(1+\mu)^2} \tag{1.54}$$

$$p(j=1) = p = \frac{1+\mu^2}{(1+\mu)^2} \tag{1.55}$$

$$p(j=0) = q = \frac{\mu}{(1+\mu)^2} \tag{1.56}$$

$$\tag{1.57}$$

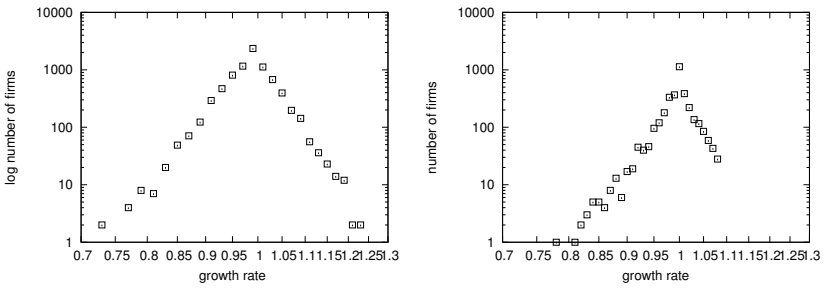
These probabilities are 'reproduction probabilities' for single positions. For a firm of size $n$, the probabilities of receiving $k$ workers can be calculated out of these probabilities $p$ and $q$, in an analogous way as the coefficients of Pascal's triangle are found. It is indeed possible to establish a recursion relation for the

coefficients $C$. The probability that a firm of size $n$ will have the size $k := 2n - l$ in the following timestep is given by

$$p(2n-l|n) = \sum_{j=0, n-l+2j>0, l-2j>0} \left( C(p^{l-2j-1} q^{n-l+2j}) + C(p^{l-2j} q^{n-l-2j-1}) \right) p^{l-2j} q^{n-l-2j} \tag{1.58}$$

In this derivation, the re-insertion of new firms has been neglected. Numerically, $G(g|n)$ is less noisy with this rounding method, compared to the case where firms offer precisely $(1+\mu)n_i$ jobs [8]. This rounding method is convenient because it yields a Gaussian $G(g|n)$ already for small firms.

This method can also be applied to the production decision of firms. In combination with a shortage of purchasing power, a growth rate probability is derived in analogy to equation (1.58). An interpretation would be that for every sold good, a firm has a probability to sell 0, 1 or 2 in the following timestep, analogously to equations (1.57).




(a) (b)

Figure 1.10: (a) and (b): Two different rounding methods for $\mu = 0.05$. Very small firms $q^s < 10$ are removed from these statistics, since their growth rates cannot take continuous values, which will distort the statistics. For instance, a firm of size 2 can only grow by $0, 0.5, 1, 1.5$ and 2, and since these small firms are numerous, peaks would be visible at these values.

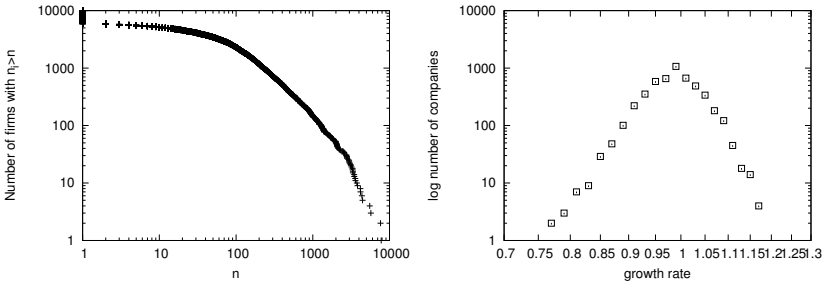
## 1.9 Results with different values of $\beta$

The empirical studies cited in the introduction find smaller values for $\beta$ than 0.5, the value in this model. These are often explained by firm-intern factors contributing to a firm's growth, as by [95, 75, 107, 97]. Intuitively, if the growth of a company was entirely dependent on the decisions of its CEO, there would be no reason to assume that a company's size should affect its growth rate variance. Under this assumption, values of $\beta$ between 0 and 0.5 are possibly

[8] Even in that case, $\lfloor (1+\mu)n_i \rfloor$ or $\lceil (1+\mu)n_i \rceil$ are offered, but no other values

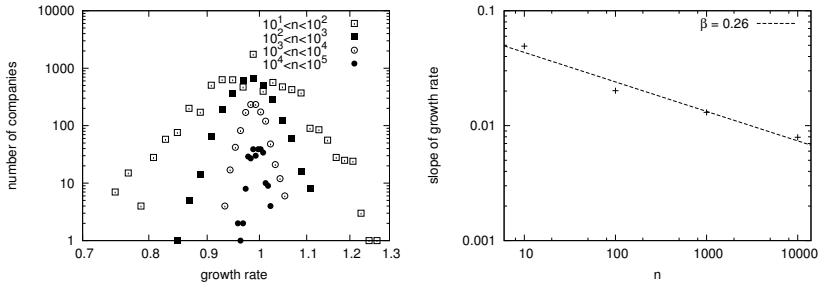
due to a contribution of both internal hierarchical structure and of a firm's size. An additional factor that can lower $\beta$ are aggregate business fluctuations, which have not been considered in the derivation above. If the total amount of availble workforce or purchasing power fluctuates, so does the probability $v$ for good to be sold, as well as the means of the Gaussian growth rate distributions (equation (1.12) and 1.20). This is independent of the firm's size.

For comparison to these empirical studies, I simulated a system with a Gaussian $\mathcal{G}(g|n)$ and a scaling exponent $\beta = 0.25$. This $\beta$ is not the result of interactions in the job market, but firm's growth consists Gaussian multiplicative noise where $\sigma \propto n^{-0.25}$, without specifying its microfoundations.




(a) (b)

Figure 1.11: (a) counter-cumulative size distribution and (b) growth rate distribution of a simulation with Gaussian multiplicative noise with a scaling relation of $\beta = 0.25$. (after 3000 iterations in a system with $10^6$ workers and $10^4$ companies)




(a) (b)

Figure 1.12: (a) Growth rate distribution where companies are clustered into size bins (system with $10^5$ companies and $10^7$ workers), (b) scaling exponent determined from the slopes in (a). It yields $\beta = 0.26$, in agreement with the scaling exponent $\sigma \propto n^{-0.25}$ of the Gaussian growth rate probability densities with which the system has been simulated.

The results of simulations shown in figures 1.11 and 1.12 are, as expected,

in between pure multiplicative noise $\beta = 0$ (figure 1.3) and the presented model with scaling exponent $\beta = 0.5$ (figure 1.8). The counter-cumulative size distribution on log-log-scale is closer to a Zipf law (i.e. $\alpha = 1$) than the one from the presented model. In contrast, the slopes of the tent-shaped growth rate distribution appear less linear than for $\beta = 0.5$.

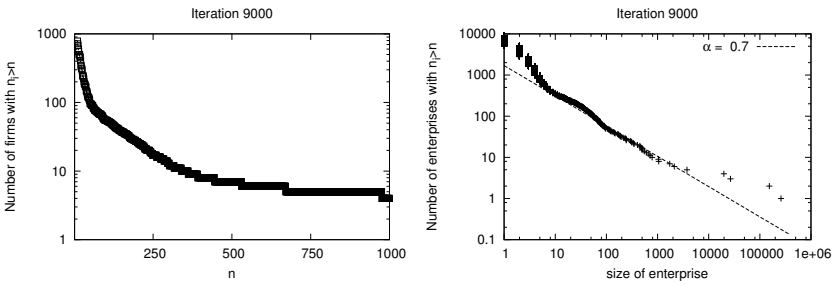
The strength of the hypothesis of Gaussian $\mathcal{G}(g|n)$ is that it would allow the explanation of heavy-tailed growth fluctuations as a collective phenomenon on aggregate level, without having to assume them on firm level, as e.g. [86]. Even if a Laplacian $\mathcal{G}(g|n)$ is assumed, as many authors do, the presence of a scaling exponent $\beta \neq 0$ does not guarantee a power law for the size distribution. As the simulation in figures 1.11 and 1.12 show (which were not simulated with the presented model), this conclusion is independent of the rationale of the model as presented in section 1.3. An alternative explanation for values of $\beta < 0.5$, which does not assume any internal structure of firms but takes into account firm's heterogeneity, is given in section 2.3.2.

## 1.10 Comparison to other models

The model can be compared to two existing models which also exhibit $\beta = 0.5$. The first is the city formation model of Marsili and Zhang [68]. They study two scenarios, of which one corresponds to pure multiplicative noise ($\beta = 0$) and yields a power law for the city size distribution, and one (which they term linear case) in which the growth rate standard deviation has scaling exponent $\beta = 0.5$. For the latter scenario, Marsili and Zhang obtained an analytical expression for the size distribution as a function of the rank $R$ of a city's size, $m(R) = m \cdot e^{-R+1}$, which is not a power law. No numerical result is shown.

The analogy to the model in this thesis is easiest for case (i) where firms compete for available workforce, and all workers are employed. Then, workers can be considered to change company at each iteration, which is why firms grow and shrink. Marsili and Zhang's setting differs from our model in that city-dwellers do not move among cities all at the same time. It corresponds to a version of this model that is simulated in sequential update, a situation where workers drawn at random can change company, and the probability of joining a particular company is proportional to its size. I have simulated these sequential dynamics for comparison (see figures 1.15 and 1.14), since the authors do not show numerical results of their 'linear case'. If the statistics of $\mathcal{G}(g|n)$ are calculated after a given number of changes at the level of workers, similar results to our tent-shaped $\mathcal{G}(g)$ are obtained, for the same reason as detailed in section 1.7.2. A difference to the presented model is that a given city may change its size within the period over which the growth rate has been evaluated, so the probabilities of receiving or loosing a city-dweller may evolve during the movements of citizens that are all represented in 1.15. In contrast, in the presented model, these probabilities remain constant during one iteration. The results obtained for the size distribution (figure 1.14) are similar to the ones from the presented model, provided that $\mu$ was chosen sufficiently large. A

conjecture is that the size distributions of the two models coincide, although it has been shown for other models that the choice of synchronous or asynchronous update does indeed influence the result [51]. Neither of the two models leads to a size distribution that can be fitted by exponential decay (see figure 1.13 where exponential and power law are compared).




(a) (b)

Figure 1.13: Plot of the counter-cumulative distribution of the same simulation as in 1.4 (b): (a) on log-linear scale, (b) on double-logarithmic scale. If decay was exponential, a the graph would be linear in (a), which is not the case.

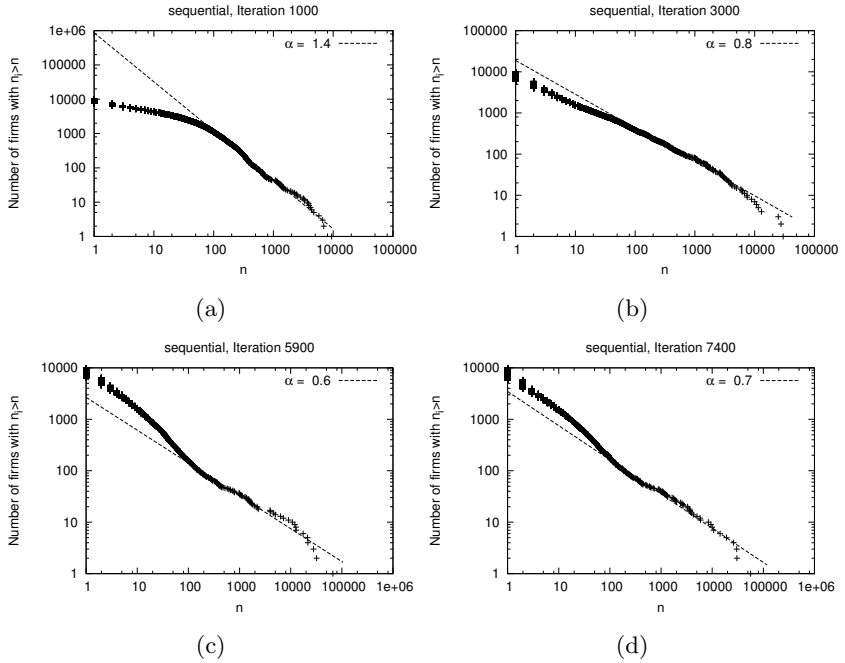



Figure 1.14: Counter-cumulative size distribution of city sizes from simulations of the linear model by Marsili and Zhang [68].

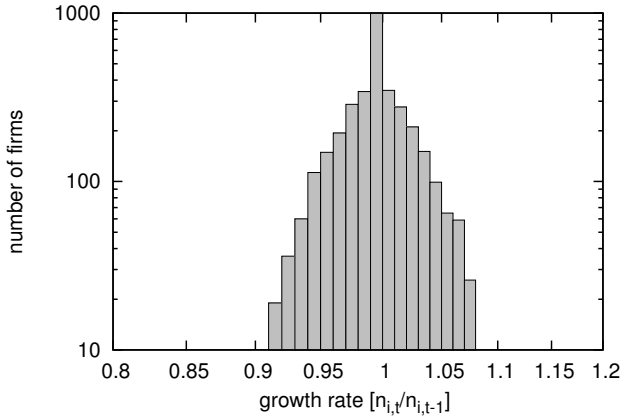



Figure 1.15: Simulation of the growth rate distribution of the linear model by [68], evaluated after 5% of workers have changed city (company).

The second interesting model with $\beta = 0.5$ is the widely used model by Yule [108] and Simon [89], which also has binomial $\mathcal{G}(g|n)$ [43]: if the constituent subunits of a firm were jobs, all existing jobs might double to two jobs with equal probability. In a particular time interval the variance of how many jobs doubled is narrower for large firms than for small firms, which may be described by a binomial distribution. (In addition, newly starting firms need to be taken into account). Yule's model leads to a Beta distribution, which exhibits a power law tail. This is due to the fact that the system is constantly growing, both in the number of employees and in the number of firms. However, it has been stated before that the power law is not found if these two assumptions are not satisfied [59].

## 1.11 Discussion

Having presented the evolution of the size distribution and the growth rate distribution of this model (which has $\beta = 0.5$), as well as simulations of systems with additive noise ($\beta = 1$), multiplicative noise ($\beta = 0$), and a system with $\beta = 0.25$ for comparison, I return to some theoretical aspects of the model. Its dynamics can be described on three levels: The noise on the elementary (i.e. job) level is the same for every element, which can double, vanish, or stay constant (see section 1.8 for the case (i)). This elementary level allows for the calculation of size evolution of companies, which are the second level. On that level, the growth rate probability density is Gaussian, with a size-dependent variance. Because of this size dependency, integral (1.46) becomes non trivial. The tent-shaped growth rate distribution of companies only holds at the aggregate system level, which is the third level.

An analogy can be drawn to a physical system, where, due to long-range interactions, the statistics of an element and the statistics of the ensemble can differ. [10] describe this by the term superstatistics, i.e. statistics of statistics, stating that in physical systems with fluctuations, the Boltzmann factor of the system is obtained by integrating the Boltzmann factors of every subsystem over their inverse temperatures. The analogy to the model presented here is the following: Instead of a Boltzmann factor, the quantity of interest is $\mathcal{G}(g|n)$, which depends on $n$. It is important to note that $\mathcal{G}(g|n)$ describes *relative* fluctuations, which are normalized by $n$, whereas additive noise (including noise mapped to additive noise as in equation (1.44)), describes absolute fluctuations. Because of the standard deviation of $1/\sqrt{n}$, the system cannot be described neither as additive nor as multiplicative noise, which leads to the conjecture that the size distribution cannot be derived from extremality neither of the Gibbs-Shannon entropy nor of the Rényi entropy. However, the concept of integrating over Boltzmann factors is the same as the integration over $n$-dependent growth rate variances in equation (1.46). This $n$-dependence may also be seen as the result of long-range interactions: the hypothesis that every job is taken with the same probability implies that every open position interacts with every available employee.

As addressed in section 1.4.5, physical systems with multiplicative noise, where the dynamics depend on the square of a Gaussian variable, exhibit the so-called q-exponential distribution [10][15], which can be derived from the extremality of the Rényi entropy. Multiplicative noise can be described as the interaction with a fluctuating external field. In contrast, in this model, the fluctuations come from competition for a limited resource, and exhibit different $n$-dependence, and different statistics as the distributions commonly found in physics. Further links of growth processes to the q-exponential distribution are presented in [92, 81] and for the case of the tent-shaped growth rate distribution [78] and [12].

## 1.12 Conclusion

In this chapter a simple agent-based model has been introduced and analyzed, in which firm growth is the result of constraints in the markets, which can be the job market or the commodity goods market. Depending on whether or not firms spend their profits in the goods market, either workforce or aggregegate demand become scarce quantities in both markets respectively. These two scenarios have been simulated separately, but yield, as expected, very similar results. A matching algorithm, which is the same in the two markets, attributes this scarce quantity, and accounts for the growth dynamics of the system. Firm growth rates are size-dependent, where the standard deviations exhibit a scaling exponent $\sigma \propto n^{-0.5}$. This feature is shared by other models that describe firm growth as the result of independent random processes. In order to keep the size of the system constant, an additive term is needed, which here corresponds to the introduction of new firms whenever a firms has attained size 0. The firm size distribution in the stationary state can be approximated by a power law of exponent $\alpha = -0.7$, which is lower than that of a Zipf law with $\alpha = 1$ found in data. This exponent is found in both scenarios and independently of profit margin $\mu$. This distribution has not yet been derived analytically, but is discussed in the context of existing results for Langevin systems with additive and multiplcative noise.

The second main result consists of the explanation of a tent-shaped growth rate probability density as a collective phenomenon. The presented model yields a growth rate probability density for firms that may be approximated by a Gaussian. Nevertheless, the aggregate growth rate pobability density of the system, for which there is empirical evidence, is tent-shaped. This tent-shaped form is also found if firms are grouped into size bins, and a tent-shaped function is fitted to the growth rates of each bin, without the need to assume a Laplacian $\mathcal{G}(g|n)$. The central idea is to take firms' size distribution into account when calculating the growth rate distribution. For comparison, simulations of a system with size-dependent Gaussian multiplicative noise ($\sigma \propto n^{-0.25}$) were carried out. Even for the latter case, which is in the range of empirical findings, the growth rate appears tent-shaped.

# Chapter 2

# The model – Economic Specification

## 2.1 Introduction

Having studied a simple model for firm growth in chapter 1, this chapter is dedicated to its extension with more features. The previously introduced stochastic growth model may be criticized because it is very abstract, and therefore difficult to interpret [96]. The extensions in this chapter provide a bridge between simple probabilistic models and more complicated agent-based macro-models. These models consist of a very complicated stochastic process that includes feedback loops, and their analysis as such is more difficult. In this chapter, the two central extensions are

(1) heterogeneity of firms' expected profit margins

(2) constraints for the availability of credit, in addition to the constraints from the markets: In order to finance production, they take out loans at a bank, which charges an interest rate $r$ (which is a constant here, i.e. the same for all firms, irrespective of their financial position or profits margin)

(3) Whenever a firm's leverage ratio has reached a threshold set by the bank, the firm is removed from the system. A constant influx of new firms compensates for these exits and ensures that the system has a number of active firms fluctuating around an average value.

The introduction of further features serves the following purposes. Firstly, an extended model combines several dynamics that are *not* happening at the same timescale. This means that results which may hold at long timescales may not explain observed phenomena at short timescales, and vice versa. Some dynamics governing longer timescales and which are being superposed on the stochastic dynamics are presented in section 2.2.1. The extended features have also been

described by equations. Firm growth is now happening at different relative speeds, since in addition to stochastic growth, firms with a high profit margin displace firms with a low profit margin. This phenomenon can be described by replicator equations. Another aspect is that the introduction of $\mu$-heterogeneity affects the scaling exponent $\beta$ of the standard deviation of growth rates (see section 2.3). In this sense, more complete explanations can also qualitatively improve those of the simple model. Secondly, the additional features provide explanations that are at a lower level of abstraction, which are easier to interpret, and take up arguments from economic theory. In particular, it links the model to macroeconomic agent-based models, as well as to stock-flow consistent models. Thirdly, along with more complicated dynamics, further possibilities for its empirical validation are given. A more complete model exhibits further distributions which can be compared to empirical data, such as firms' profit rate distribution, which can be compared to firm databases like [33, 1, 61]. This allows for a more extended validation of the model. Since the evolution of individual firms is no longer purely stochastic, it is now possible to analyze a firm's typical life cycle, and compare it to existing studies such as [103]. In addition, it is possible to study fluctuations and feedback loops, and relate them to explanation of business fluctuations and cycles.

The chapter is organized as follows: First, publication 1 is presented. It contains an introduction of the new features. In the following section 2.2, the extended model is described theoretically. In section 2.3, the effect of margin heterogeneity on the scaling exponent $\beta$ is shown numerically and theoretically. Only in section 2.4, related literature is discussed. This untypical order of sections was chosen to be able to draw comparisons with the presented model, and to keep the technical sections directly after chapter 1. Section 2.5 is the conclusion.

# Publication 1

presented at: Workshop of Economic Heterogeneous Interacting Agents 2012
(Paris, France)

# Heterogeneous Firms in a Macroeconomic Agent-Based Model
## – work in progress –

Cornelia Metzig[†] and Mirta B. Gordon[†,*]
[†]Université Joseph Fourier - Grenoble 1 / [*]CNRS
LIG (Laboratoire d'Informatique de Grenoble)
cornelia.metzig@imag.fr, mirta.gordon@imag.fr

30.07.2013

**Abstract**

We present a macroeconomic agent-based model that combines several mechanisms operating at the same timescale, while remaining mathematically tractable thanks to a low number of parameters. It comprises firms and workers who compete in a job market and a commodity goods market. The model is stock-flow consistent; a bank lends money charging interest rates, and keeps track of equities. Important features of the model are heterogeneity of firms, existence of bankruptcies and creation of new firms, as well as productivity increase. It combines probabilistic elements and deterministic dynamics, whose relative weights, determined by the parameters, may be modified according to the considered problem or the belief of the modeler. The model's evolution reproduces empirically found regularities of firms' size, growth rate and profit rate distributions. We discuss the origin and the amplitude of endogenous fluctuations of the system's stationary state for two different scenarios and analyze typical life cycles of firms.

## 1 Introduction

In economic theory, a long history of approaches exists beyond maximizing principles leading to equilibrium. These share as principles a separate treatment of households and firms and the division of income into wages and profits [1]. Within these models are agent-based models, which have been proposed as early as 1957 by G. Orcutt [2] under the name *microsimulation models*, in order to make predictions at the macroeconomic level based on the growing knowledge about decision-making units, and to avoid flawed aggregation. Such models became feasible with larger availability of computing power. First models including production, investment and consumption have been proposed by B. Bergmann [3] and G. Eliasson [4]. These models aim at predicting responses to different economic policies. To this end, the units' behaviors are calibrated with empirically based parameters, as described in [5]. More recently, agent-based models have been used to analyze the role of the units' parameters on some stylized facts of economic systems [6, 7, 8, 9, 10, 11]. In addition to earlier microsimulation models, these later models emphasize the interactions between financial and real sector, reflecting the increasing importance of finance in the economy of today. The models contain very many features, and their main purpose is the reproduction of macroeconomic emergent phenomena like business cycles. Since the number of

firms is relatively small (ranging from 10 [10], 60 [8], 100 [6], 200 [7], 250 [11]), distributions of firm characteristics are not the focus of analysis.

Many microsimulation or agent-based models conserve the flow of funds. They can serve as a tool to study the role of debt levels, interest flows, or availability of credit on economic activity. The most important implication of this stock-flow consistency is that the financial condition of agents plays a central role determining their actions. In that sense, stock-flow consistency can become the guiding principle instead of the efficient market principle, which assumes "the impact of the flow of funds and the stocks of credit and debt are fully reflected in returns and risks at the individual level"[12]. If, in contrast, agents base their decisions only on their balance sheet, and do not know the financial condition of others – as is the case in our model –, endogeneous cycles in economic activity can occur.

Several analytical stock-flow consistent models describe the money flows between sectors [13, 14], or sub-sectors [1]. These models exhibit stationary state scenarios where the money flows to and from each sector add to zero. If the model is discrete in time, in one iteration every actor needs to re-inject in the system the money he has received, be it wages, profits, interest payments or other. Otherwise, economic activity will eventually cease. Regarding firms, this phenomenon has become known as the 'paradox of monetary profits' [15], originally formulated by Marx, stating that firms can at most earn what they have paid in wages. To circumvent this, some models introduce as money flows interest payments of firms to the bank, which is a joint stock company and distributes a dividend [8], consumption of the bank [14], dividends of firms who are joint-stock companies or investment of firms[7]

Another relevant area of research for agent-based macro models are empirical studies, e.g. for firm size [16, 17], growth rate [18, 17, 19, 20, 21] and profit rate distribution [22, 19, 20].

Many of these statistical regularities have been explained by stochastic growth models, i.e. theoretical models designed to explain one or several stylized facts found in firm data. An important firm growth model has been proposed by R. Gibrat as early as [23], describing firm sizes to follow a lognormal distribution. Since then, many more have been proposed. Some describe growth as being the result of growth of independent subunits, which has been attributed to firms being active in different sectors or different markets[24, 18, 25, 26, 27]. These are often based on a model by H. Simon [28], which describes a constantly growing economy and its theoretical foundation is a model by Yule [29].

Several growth models combine an economic principle, e.g. maximization, with stochasticity [30, 21, 20]. As advocated in a survey on growth models by J. Sutton, "A proper understanding of the evolution of structure may require an analysis not only of such economic mechanisms, but also the role played by purely statistical (independence) effects, and (...) a complete theory will need to find an appropriate way of combining these two strands." [31], which relates them to agent-based macro models, since also these combine stochastic and deterministic dynamics.

The model presented here is a macroeconomic agent-based model which combines a stochastic firm growth process with stock-flow consistent principles. It abstracts from different sectors, and firms merely grow in terms of their sales and employees. Firms plan to grow depending on their last realized profits, and the actually realized growth is a result of the competition in the markets. In a simple scenario of the model analyzed theoretically in [32], a company's workplaces may be seen as the independent subunits. In the more complete model presented in this paper, the focus lies on constraints from the financial sector that influence the real sector. Interest payments lower the net realized profits of firms, and therefore influence both growth and economic activity. Once a firm has accumulated debt to a given threshold, it is declared bankrupt and exits the system, its debt being cancelled. Endogeneous fluctuations and cycles in economic activity are analyzed at the aggregate level, distributions of firm charactersitics at an intermediate level, as well as life cycles of firms may be analyzed at the individual level. The model proposed here can be analyzed theoretically

and contains fewer features than other macro-models, and may be completed successively.

The paper is organized as follows. In section 2, we introduce the model. In section 3 we show and discuss numerical results on steady states and distributions of firm characteristics. We then introduce another feature (aging), and analyze life cycles, robustness and business fluctuations with and without this feature. Finally, in section 4, we conclude and point at possible extensions and applications.

# 2 The model

We present here the main elements and the dynamics of the model, which are analyzed successively in the following sections.

## 2.1 Main elements

The system is composed of three types of agents: $N_{f,t}$ firms, $N_w$ workers, and one bank that issues credit to firms. The number of active firms fluctuates over time, since the model includes entry and exit of firms. The flow of funds is conserved.

**firms** $i$ $(1 \leq i \leq N_{f,t})$ are characterized by an expected gross profit margin $\mu_i$.

$$\mu_i = \frac{p\, q_i - n_i\, w}{n_i w}\,, \tag{1}$$

where $n_i$ is the number of hired workers, $w$ the wage, and $p\, q_i$ the expected sales. $\mu_i$ are intrinsic constant parameters drawn at random with a uniform probability density in the range $[\mu_{min} \leq \mu_i \leq \mu_{max}]$ with $0 < \mu_{min} < \mu_{max} < 1$. These expected gross profit margins are distinct from the net realized profit margin, detailed in equation (17), which can attain negative values. $\mu_i$ reflects the technological level or workforce productivity of a firm: the higher the margin, the lower the production cost per unit or the lower the number of workers needed to produce a given quantity.

Firms produce $q_i$ non-durable goods per period, which are put on the market at price $p$. This commodity good is an abstraction of purchases in the real sector, which is useful to ensure that via limited aggregate demand all firms are in competition with each other. If unemployment drops to zero, firms compete in addition for is workforce in the job market. A third constraint is the availability of credit, for which they do not compete. The number of active firms fluctuates over time.

**workers** $j$ $(1 \leq j \leq N_w)$ may be employed or unemployed. When employed they earn a wage $w$ per period. Like Godley and Lavoie [13] in their 'simplest model', and M. Kalecki [33] in his profit equation, we assume that workers try to spend all their earnings in the consumption market, unless too few goods are offered to satisfy their demand.

**the bank** represents a simple financial sector. It lends money to firms at a constant interest rate $r$, and keeps track of their equities $e_i$ $(1 \leq i \leq N_f)$, which cannot be less than a lower bound $u_i \leq 0$ :

$$e_i \geq u_i = -\gamma\, w\, n_i\,, \tag{2}$$

where $\gamma$ is some positive constant. Firms which do not satisfy this constraint are declared bankrupt, and their debts are assumed by the bank. This is the only expense of the bank, and interest its only revenue. In this simple model, the bank does not base decisions on its own balance sheet.

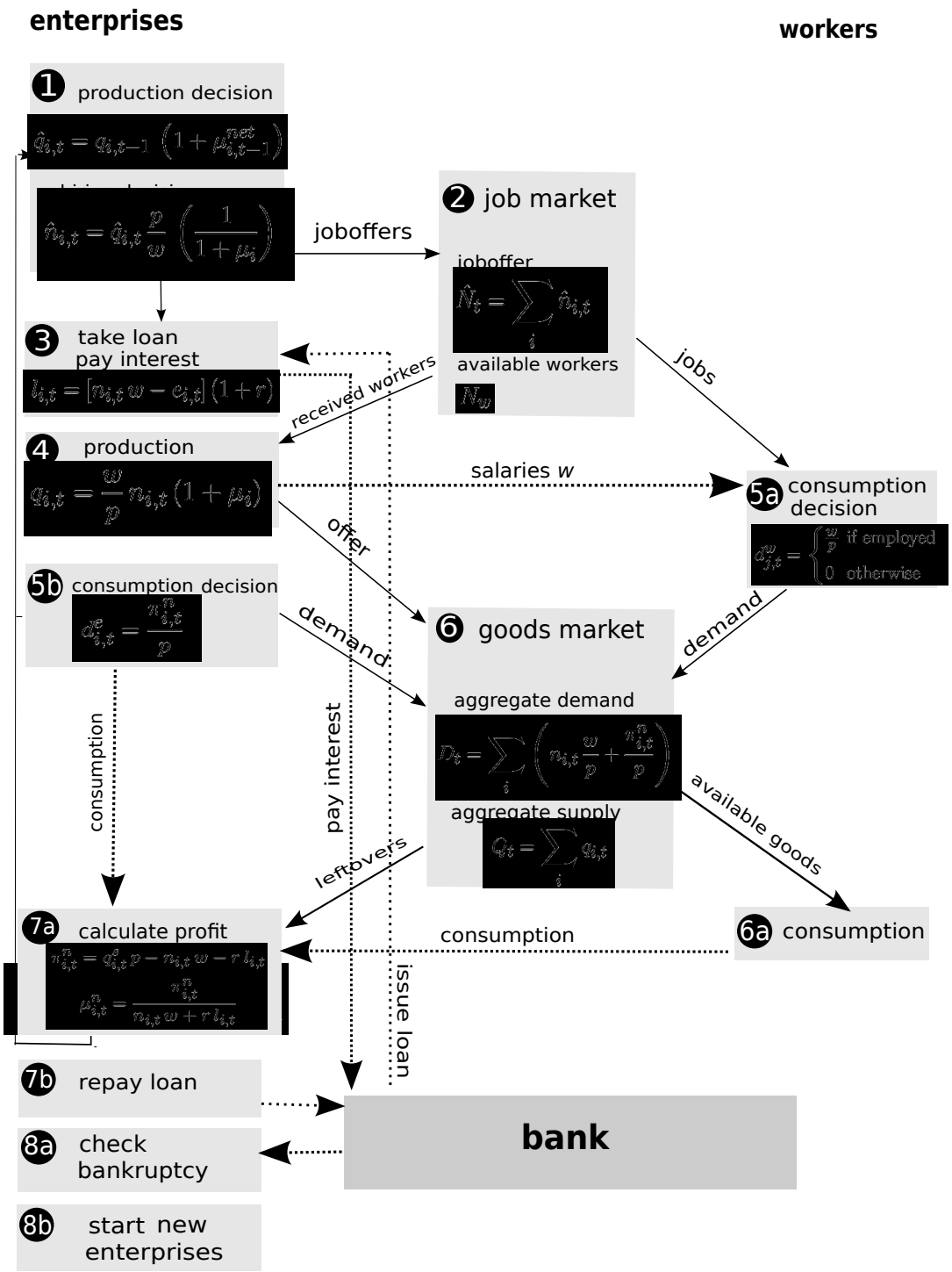



Figure 1: Scheme of the model. The arrows denote interactions, dotted lines denote money flows. The numeration corresponds to the numeration in the description.

## 2.2 Dynamics

The dynamics of the model, visualized in figure 1, is detailed below. Upper case letters stand for for aggregate quantities, lower case letters for individual quantities, both for firms $i$ or workers/households $j$. Quantities with hat ($\hat{q}_i$, $\hat{n}_i$, $\hat{d}_i$) are initially planned quantities before agents have checked constraints.

**at each time step t:**

1. **production decision of firms:** firms calculate their expected volume of production $\hat{q}_{i,t}$ according to their previous production $q_{i,t-1}$ and their net realized profit margin $\mu^n_{i,t-1}$ (which is lower than the expected profits because of interest payments and unsold goods, see equation (17)). If the net profit margin is positive (negative) they plan to produce more (less) goods proportionally:
$$\hat{q}_{i,t} = q_{i,t-1}\left(1+\mu^n_{i,t-1}\right) . \tag{3}$$
The necessary number of workers for the production of $\hat{q}_{i,t}$ goods is determined by the intrinsic margin $\mu_i$ defined in equation (1)
$$\hat{n}_{i,t} = \hat{q}_{i,t}\left(\frac{1}{1+\mu_i}\right)\frac{p}{w} . \tag{4}$$

2. **job market:** the aggregate job offers of firms
$$\hat{N}_{w,t} = \sum_i \hat{n}_{i,t} \tag{5}$$
may exceed the available number of workers $N_w$. Workers do not stay at their employer but are newly placed at each iteration through a process where job offers and available workers are matched at random[1]. If $\hat{N}_{w,t} \leq N_w$, each firm $i$ will hire exactly the desired number of workers:
$$n_{i,t} = \hat{n}_{i,t} \text{ if } \hat{N}_{w,t} \leq N_w. \tag{6}$$
Every worker has the same probability $\hat{N}_w/N_w$ of being hired by some firm but $N_w - \hat{N}_{w,t}$ of them are left unemployed. If $\hat{N}_{w,t} > N_w$, all workers are hired and each job opening has the same probability $N_w/\hat{N}_w$ of being filled and $\hat{N}_{w,t} - N_w$ positions remain vacant. On average, each firm $i$ receives
$$\langle n_{i,t}\rangle = \hat{n}_{i,t}\frac{N_w}{\hat{N}_{w,t}} \text{ if } \hat{N}_{w,t} > N_w . \tag{7}$$

3. **credit market:** once the number of hired workers $n_{i,t}$ is known, each firm calculates whether its owned equities $e_{i,t}$ are sufficient to pay the corresponding wages. If necessary, it takes out a loan of amount $l_{i,t}$ from the bank:
$$l_{i,t} = \left[n_{i,t}\, w - e_{i,t}\right](1+r) . \tag{8}$$

4. **production:** the actually produced amount of goods $q_{i,t}$ by the $n_{i,t}$ workers of firm $i$ is given by
$$q_{i,t} = \frac{w}{p} n_{i,t}(1+\mu_i) , \tag{9}$$

[1]This hypothesis is not crucial for the model, the dynamics are similar if a fixed fraction of each firm's workforce change their employer.

i.e. firms with a higher $\mu_i$ produce more goods per worker. The aggregate output $Q_t$ is

$$Q_t = \sum_i q_{i,t}. \tag{10}$$

5. **consumption decision:** workers $j$ intend to spend their wages, which allow them to buy the quantity

$$d^w_{j,t} = \begin{cases} \frac{w}{p} & \text{if employed} \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

Firms $i$ intend to spend their previous net realized profits[2]:

$$d^f_{i,t} = \frac{\pi^n_{i,t-1}}{p}\,. \tag{12}$$

The bank does not act as a consumer (as is done e.g. in [14]), it can merely lose money when firms go bankrupt. Thus, the aggregate demand is:

$$D_t = \sum_i d^f_{i,t} + \sum_j d^w_{j,t}. \tag{13}$$

6. **commodity goods market:** Aggregate $Q_t$ and $D_t$ are not known to firms and workers; they experience if their offer and demand are satisfied as result of a random matching algorithm (which is also used to match job offers and workers). If $Q_t < D_t$, all the production is sold, leaving some demands unsatisfied. Each demand – whether coming from firms or workers – has the same chance of being fulfilled:

$$q^s_{i,t} = q_{i,t} \text{ if } Q_t < D_t\,. \tag{14}$$

Conversely, if $Q_t > D_t$, firms face a random constraint due to limited aggregate demand; only a fraction $D_t/Q_t$ of all produced goods is sold. On average, each firm sells

$$\langle q^s_{i,t} \rangle = q_{i,t}\,\frac{D_t}{Q_t} \text{ if } Q_t > D_t\,. \tag{15}$$

This is the situation in the present paper, which is why realized margins (17) are generally lower than the intrinsic margins $\mu_i$ even before paying interest (see examples shown in figure 5).

The commodity good can be thought of as being a service, which provides a simple interpretation why unsold goods are lost. For other interpretations, this loss may include roughly all costs of storing and lowering future production. It is possible to reproduce size and growth rate distribution also in a situation with competition only for one single resource (see [32] where this is the job market), but competition for the non-durable commodity good adds important features to the model: It introduces a stochastic term for the last realized profits, which is the basis of firm's following production decision. Furthermore, it contains an important long-term effect: if a firm cannot sell enough goods, its net realized profits are negative, and it start the next interation with a lower financial position. This decreases with the interest payments further the next profits, and ultimately leads to a firm's bankruptcy. Even in a more complex models, such a long-term effect should be present. The principle that profits can only be estimated beforehand is in some way present in many Keynesian-inspired models, e.g. in the agent-based models [34, 7, 11].

[2]In addition to the setting analyzed in this paper, it is also possible to obtain a stationary state where firms do not consume, and their only expenses are wages and interest payments.

7. **firms' balance sheet accounting:** firms pay interests on the whole amount of their loans, i.e. on current lending and accumulated debts. If they have enough assets, they repay loans at the end of the iteration; if the liabilities exceed the assets, they start the next iteration with a negative net equity. In order to decide on the quantity to produce in the following period, firms calculate their net realized profit $\pi^n_{i,t}$, defined by

$$\pi^n_{i,t} = q^s_{i,t}\, p - n_{i,t}\, w - l_{i,t}\, r\,, \tag{16}$$

and their net realized profit margin

$$\mu^n_{i,t} = \frac{q^s_{i,t} p - n_{i,t} w - l_{i,t} r}{n_{i,t} w + l_{i,t} r}\,. \tag{17}$$

$\mu^n_{i,t}$ is generally lower than $\mu_i$ and only equals $\mu_i$ if both $r = 0$ and a firm sells all of its production.

8. **bankruptcy and new firms:** firms whose equities (after repayment of loans) fall below the threshold $u_i$ defined by eq. (2) are declared bankrupt and are removed from the system. Their (negative) equities are losses to the bank.

   In order to avoid a decline in the number of firms due to bankruptcies, $\nu$ new firms are introduced at each iteration, with an initial number of workers ($n_{i,init}$) and a margin $\mu_i$ both drawn at random from $[0, n_{max,init}]$ and $[\mu_{min}, \mu_{max}]$. With this procedure the total number of firms fluctuates over time. This differs from the approaches by Bruun [7], DelliGatti et al. [6] and Dosi et al. [11], who replace systematically each bankrupt firm by a new one, keeping thus the total number of firms constant.

## 2.3 Transaction Flow Matrix

Table 1 summarizes the money flows between sectors. A more detailed matrix would contain a separate column for the asset account and liability account of every individual agent, since the model is implemented with double-entry acoounting. Upper case letters stand for the aggregates of the quantities defined in the preceding equations. $\Delta L$ represents the changes in firms' level of loans from the bank at the considered iteration. Reimbursement of loans is conditioned by the financial position of each firm: those whose equities are lower than a threshold defined in (2) fail and their loans are never reimbursed. Consequently, $\Delta L' \leq \Delta L$.

Interest payments link stocks and flows, since their volume depends on the finiancial position of a firm, which is a stock. This interest flow influences other flows (how much firms can spend in the commodity goods market), and via the net realized margin in equation 3 the production decision of companies. Also the moment of a firm's bankruptcy depends on its money stock of debt. Stock-flow consistency determines economic activity, and itroduces a second timescale (exceeding one iteration) at which the pas influences the present economic situation.

This transaction flow matrix makes model comparable to analytical flow-of-funds models [1, 13, 14] but with the important difference that the aggregate flows are composed of the flows of individual agents, and therefore reach a *fluctuating* stationary state, i.e. the aggregated equities of each sector fluctuate around a constant. These fluctuations in the money flows are interlinked with the fluctuations in economic activity, as detailed further in section 3.5.

| **transaction** | | **sectors** | | **Σ** |
|---|---|---|---|---|
| | workers | firms | bank | 0 |
| wages | $N_t\, w$ | $-N_t\, w$ | | 0 |
| loans | | $\Delta L$ | $-\Delta L$ | 0 |
| interests | | $-L\, r$ | $L\, r$ | 0 |
| consumption | $-Q^{s,w}\, p$ | $Q^{s,w}\, p$ | | 0 |
| consumption firms | | $-Q^{s,f}\, p$ | | |
| | | $Q^{s,f}\, p$ | | 0 |
| reimbursement of loans | | $-\Delta L'$ | $\Delta L'$ | 0 |
| $\Delta$ equities | $\Delta E_w$ | $\Delta E_f$ | $\Delta E_b$ | 0 |
| **Σ** | 0 | 0 | 0 | 0 |

Table 1: Accounting matrix of the model.

# 3 Results

In this section we discuss results of simulations corresponding to different sets of parameters of the model:

- $[\mu_{min}, \mu_{max}]$, the range of expected margins,
- $r$, the interest rate,
- $\gamma$, the bankruptcy threshold,
- $\nu$, the number of new firms entering per iteration.

The parameters in the model are all relative. Roughly one month can be though of as the order of magnitude for one iteration. The type of analysis in this paper is rather suited to understand dependencies, and to address the question how much randomness a system has or should have with respect to the deterministic part of the dynamics.

## 3.1 Steady states and parameter dependence

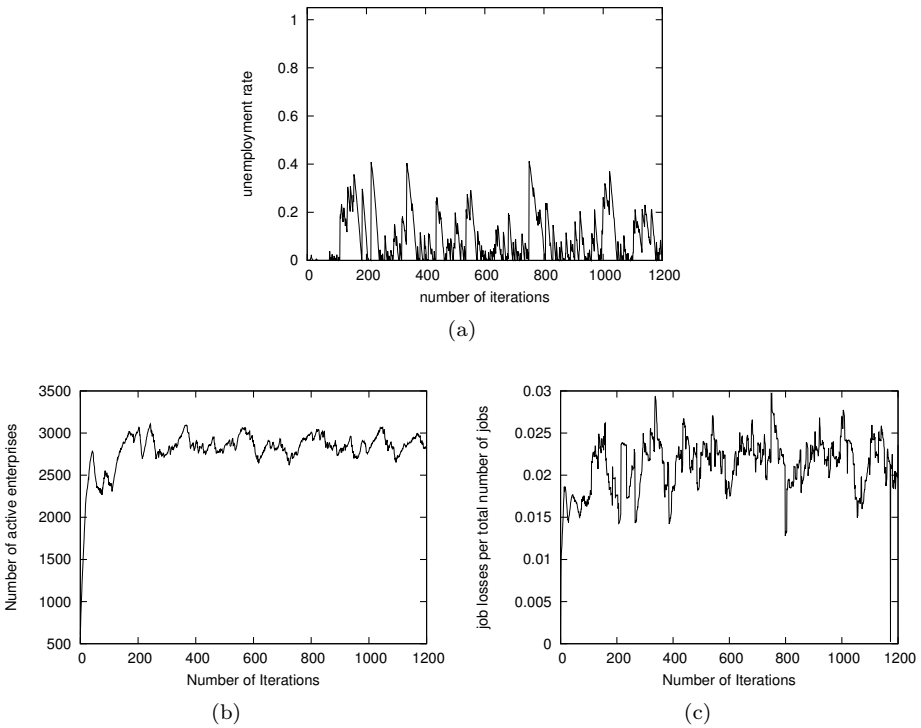



Figure 2: $N_w = 10\,000$, bankruptcy threshold $\gamma = 2$, interest rate $r$=0.01062, new firms per iteration $\nu = 120$, $[\mu_m, \mu_M] = [0, 0.1]$. Time evolution of (a) unemployment, (b) number of active firms, (c) job losses due to firm bankruptcies (averaged over 50 iterations).

Figure 2 presents results corresponding to a typical time evolution of the system. The system exhibits fluctuations also after convergence to a stationary state, since these are due to the randomness in the system. Job losses do not occur at a constant rate but fluctuate around a certain

rate (here 1.5% of the existing jobs in one iteration. In subfigure 2 (c) job losses are averaged over 50 iterations in order to see slower trends, the amplitude of actual fluctuations is heavy-tailed, as studied in [35]). The stationary state of the system is characterized by a number of active firms and a level of unemployment, which depends on the interest rate $r$, the bankruptcy threshold $\gamma$, the number $\nu$ of entering firms per iteration and the margin range $[\mu_{min}, \mu_{max}]$. The higher the interest rate, the higher the level of unemployment for a given $\nu$ and $\gamma$.

The stationary state is noisy since the dynamics has two stochastic elements, but for a given set of parameters the system converges to such a state for any initial level of unemployment, any initial size distribution of firms, and (almost) any number of initially active firms.

Unemployment fluctuates in the stationary state for the following reasons: Aggregate offer of goods exceeds aggregate demand, since firms need to service their debts and spend only their previous net realized profits in the goods market. The random allocation of demand to the produced goods implies that firms face a random constraint which lowers their profits and if these are negative, firms get into debts. Debt accumulation follows thus a random dynamics. Furthermore, the survival times depend on the intrinsic margin $\mu_i$ of each firm: the higher $\mu_i$, the longer the life time on average. Bankruptcy frequency fluctuates, and in addition firm size at bankruptcy has a fat-tailed distribution, i.e. large bankruptcies are rare events. If large firms go bankrupt, the joboffer suddenly declines, less wages are paid, and existing firms earn only lower profits. The time until unemployment diminishes depends on aggregate profits, which in turn depend on the interest rate $r$ and on the frequency of restarts $\nu$. Fluctuations and business cycles are discussed more in detail in section 3.5.

## 3.2 Distributions

For any agent-based macroeconomic model with a sufficiently large number of firms, it is possible to analyze distributions of firm characteristics. Even if the primary focus lies on aggregate variables, the arising distributions should be at least qualitatively in agreement with empirical evidence, which is one way of validating the model. Otherwise, the differences should be understood from the simplifications that the model makes. The stochastic process of a simple scenario of this model, with homogeneous $\mu$ and no interest payments has been analyzed theoretically in [32], and is the dominant dynamics also in this setting, though completed by the banktuptcy mechanism.

**Size distribution of firms.** Every single firm size varies over time, but the size distribution of firms forms a fat tailed distribution that can be approximated by a power law which remains stable (as long as firm entry and exit are possible: otherwise the stationary state is a monopoly). Most firms fail after few iterations. Typical cases for surviving firms are shown in figure 5. The stochastic process leading to this is a superposition of two effects: firstly, a stochastic process due to competition for demand and the fact that each firm plans to grow according to its realized profits (on its own described in [32]). Secondly, in the setting presented here $\mu_i$ are heterogeneous, and firms incur debts over time, which is why their net realized profits are lower than their expected ones, and why they will eventually go bankrupt. The higher the interest rate $r$ and the number of new firms per iteration $\nu$, the higher the numerically found exponent of the power law fit to the cumulative size distribution. For low interest rates, the exponents but is close to 1 (see figure 2, which is the approximate value reported in empirical data [16, 17].

**Growth rate distribution of firms.** In [32], an explanation for the stylized fact of a tent-shaped growth rate distribution is derived, both analytically and in simulations. This shape has empirically been found e.g. in [21, 18, 17], as well as for growth processes of other quantities than firms [36, 37, 38]. In our model, the necessary ingredients for this growth density are that the

size distribution is fat tailed, and that the growth rate probability density for individual firms $i$ is (approximately) Gaussian with a standard deviation $\sigma$ scaling as a power of size $n_i$: $\sigma(n) \propto n^{-\beta}$. These two features arise naturally in the presented model, but the argument holds independently of the precise reason for the Gaussian growth density, and may in principle stem from other reasons than this model suggests. In this model, three factors contribute to it, which on their own would yield a binomial growth density (that can be approximated by a Gaussian for large $n$): (i) competition in the job market for workers, (ii) competition in the goods market for purchasing power, or (iii) a probabilistic rounding method by which firms determine their joboffer (detailed in [32]). (iii) yields the same growth rate distribution as the model by H. Simon [28] based on the Yule process, although his interpretation differs. The arising size dependence of the growth rate is such that small firms have a wider growth rate density and therefore account for the tails of the distribution, whereas large firms have a more narrow growth rate distribution, and account for the middle peak.

The model offers also the possibility to compare further tendencies with empirical surveys. For instance, as A. Coad states in his survey of the literature on firm growth [39] p.12, "there is a lot of evidence that a slight negative dependence of the growth rate on size is present at various levels of industrial aggregation." Also [40] confirm this result for a database of Italian firms. This effect also occurs in our model: To reach a large size, firms need to attain a certain age. During that time, their margin diminishes due to the recentering of $\mu_{eff}$, and the level of debt will increase, so it is not possible for the net margin of big firms to range among the higher end of the distribution. Small firms, in turn, are more likely to have a net margin above average: Of the firms with negative net margin, a part goes bankrupt, and evidently cannot enter the statistics on average growth rates, so the average growth rate is shifted towards a positive value of growth.

**Profit margin distribution of firms.** Margin heterogeneity, interest payments, competition in the goods market, new firms and bankruptcies are the features accounting for the profit rate distribution shown in figures 3. The tent-shaped growth rate distribution can in principle be found without these features (see [32] where the profit rate distribution is a delta peak), and the profit rate distribution acts as a prior distribution to the growth distribution due to competition in the job market. With this prior distribution, the growth rate distribution becomes wider, though its shape is approximately the same. Depending on the scenario, simulations yield a profit rate distribution that is approximately Gaussian or Laplacian, and in some cases is asymmetric. Empirical evidence suggests that the shape of the profit rate distribution is between Gaussian and Laplacian [20, 22, 19] for databases on Italian, US and Icelandic firms. The last two publications show in addition that for the same firm data, the profit rate distribution is narrower than the growth rate distribution, which is also a result of this model (see figure 3).

The profit margin distribution is composed of firms of all stages in their life cycles. However, these are not entirely random but show some typical trajectories for the variables size, profit margin and leverage ratio over their life cycle (shown in figures 5). The dynamics of the model are crucial for the formation of the profit rate distribution, since figure 3 (c) differs clealy from a uniform distribution of the expected gross profit margin $\mu_i$ at which new firms are initialized. Firms with a low net realized profit margin are likely to incur debts and default quickly, which is why only few firms in the distribution have very low margins. Likewise, very few firms gain very high profit margins, since this is happens only under lucky coincidences: the intrinsic margin needs to be high, a firm's debt needs to be low, and no or few goods remain unsold.

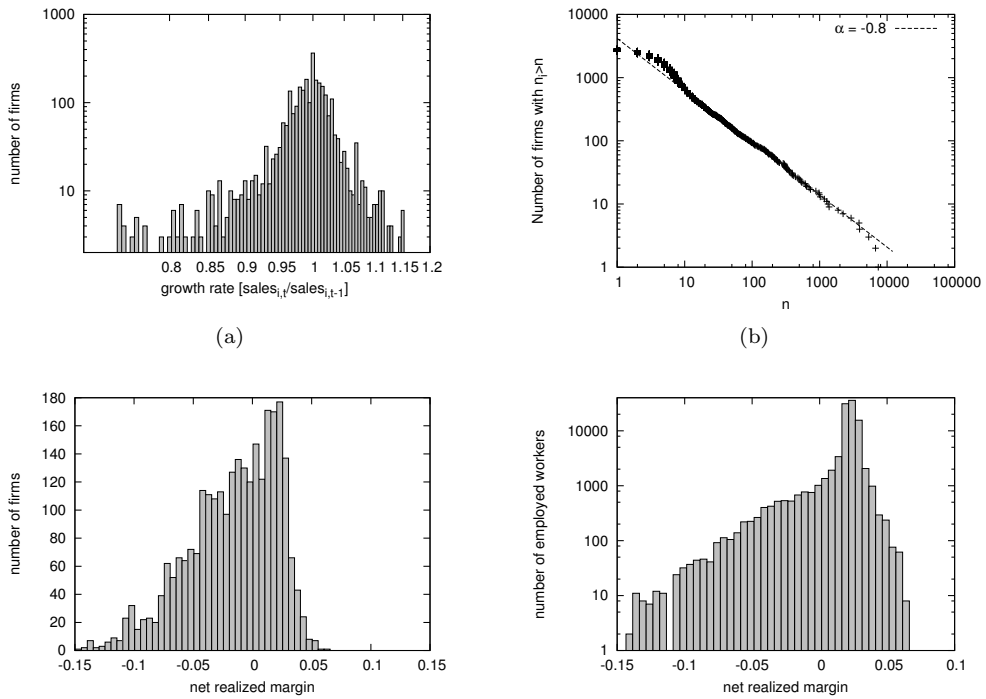



Figure 3: Interest rate 0.0162, $\nu = 120$, $\gamma = 2$, 100000 workers, $\approx 2500$ active firms, unemployment around 9% (shown in figure 2). Snapshots after 1200 iterations. (a) growth rate distribution of firms in double logarithmic scale, (b) size distribution of firms, (c) net realized profit margin distribution, (d) profit margin distribution weighted by the size of firms. If firms did not face a random constraint due to limited aggregate demand, every firm would grow according to its net realized margin, and the growth rate distribution (a) would have precisely the shape of (c), which is clearly not the case. (d) shows that the result of a uniform $\mu_i$-initialisation in combination with a dynamics with interest payments, restarts, and bankruptcies is that most workers are hired at firms whose margin is in the middle of the net margin distribution. Large firms occur mostly at the middle of the net margin distribution for two reasons: they have a lower growth rate variance from market competition, and because some time is needed to grow, in which they already incur debts, lowering $\mu_i^n$.

## 3.3 Introduction of explicit aging of firms

A valid objection to this simple model is that the only reason why older firms have a lower profitability is that they are indebted. This is certainly not the case in reality. A desired feature would be that new firms enter with a higher performing technology than incumbent ones. If this is attempted by the entry of new firms with a constantly increasing $\mu_i$, a new problem needs to be solved: the average margin should remain in a constant relation to the interest rate, which otherwise would lose its function of lowering net realized margins. Therefore, aging of firms was

added as an aditional feature. We define the *effective margin* as

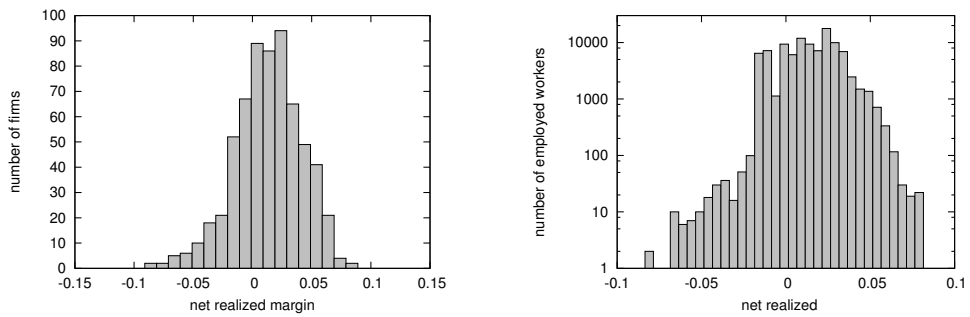
$$\mu_{eff,t} = \frac{1}{N_{w,t}} \sum_i n_{i,t}\mu_i \,. \tag{18}$$

This quantity has a natural tendency to grow, since firms with higher margins earn on average higher profits, leading to more job offers and faster growth. For comparison with the scenario presented in subsections 3.1 and 3.2, we now perform simulation with a lower interest rate, where the individual $\mu_i$ are being recentered at each iteration, such that $\mu_{eff}$ remains constant over time:

$$\mu_{i,t} = \mu_{i,t-1} - (\mu_{eff,t} - \mu_{eff,t-1}) \tag{19}$$

This has the effect that new firms are much more likely to grow, since their intrinsic margins $\mu_i$ have not diminished yet, and are therefore higher on average. The survival rate is higher and less new firms $\nu$ need to be introduced per iteration in order to obtain comparable dynamics. An additional bankruptcy criterion is needed, in that firms fail once their gross expected margin $\mu_i$ has attained zero. Also empirical studies show that several bankruptcy criteria are commonly used, e.g. [41]. Implications on distributions of these criteria are studied in detail in a forthcoming paper.

This situation may correspond to new firms entering the system with a more profitable technology. Apart from this more plausible interpretation, this leads to slightly different results: the profit margin distribution contains fewer firms with very high profit margins (see figure 4), closer in line with empirical studies [19, 20, 22]. One may go as far as to interpret the recentering as an increase productivity and in the standard of living, implying that the interpretation of one good and one salary changes. However, we encountered an inconsistency with this interpretation, since it is at odds with the interpretation that $\mu_i$ stands for the technological level or labour intensity of a firm: why should labour intensity rise over time? A possible unification of these two views is to interpret $\mu_i$ as the *relative* labour intensity of firms, indicating what fraction of the workforce is needed in order to produce a certain fraction of the total production. Another possibility to respond to this limitation could be the explicit introduction of productivity increase and inflation, at the risk of sacrificing the mathematical tractability of the model.




(a) (b)

Figure 4: Results including recentering of the margin. $r = 0.0102$, $\nu = 10$. Compared to figures 3 (c) and (d), less firms have a very high margin, and less workers accumulate in firms with a very high margin, because the intrinsic margins $\mu_i$ shrink over time.

### 3.4 Life cycles

The implications of margin aging become evident when looking at the typical time evolution of individual firms. The evolution of the leverage ratio over a firm's life cycle is closer to empirical findings: [42] report from a study on Canadian firms that firms typically start indebted (termed "liability of newness"). Firms which manage to repay them survive and grow. Later, firms incur again debts since their technology is outdated (which they term "liabilities of obsolence"). This is also visible in typical life cycles of firms if margin aging in included (see figure 5), whereas without recentering of the margin, the evolution of debt follows a linear trend. In subfigures (a,c,e) in the left column, firms are unprofitable *because* they are indebted, whereas in the situation with margin aging in subfigures (b, d, f) in the right column, firms incur debts *because* they are unprofitable, although a high leverage ratio is also self-reinforcing.

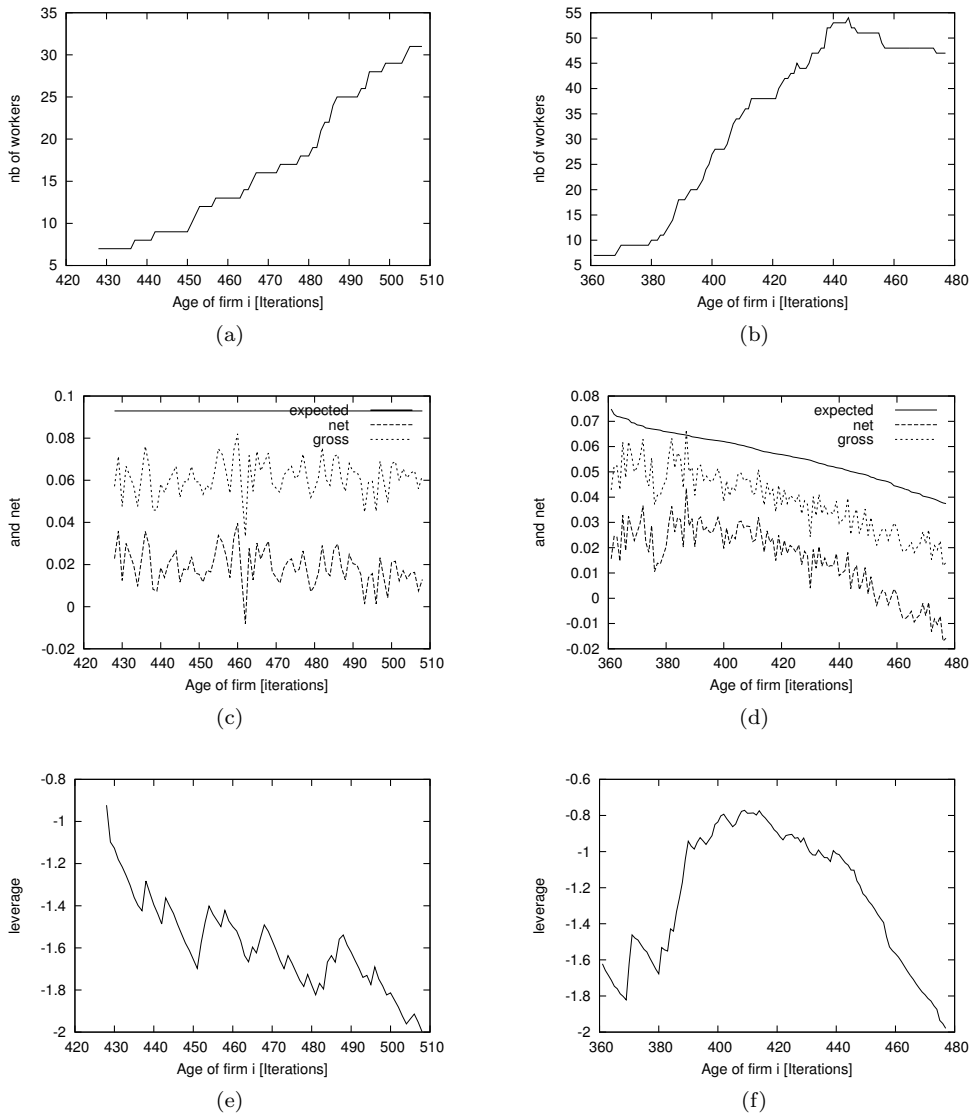



Figure 5: Comparison of typical life cycles. Without margin recentering (subfigures (a), (c), (e), $r = 0.0162$), firms' growth and debt accumulation follows a linear trend. With recentering (subfigures (b), (d), (f), $r = 0.0102$), firms start to shrink in size before failure. They repay initial debts, and only later indebt again.

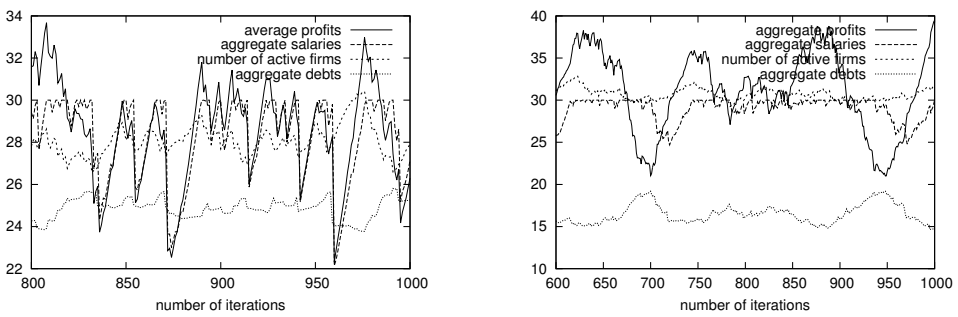
### 3.5 Fluctuations

Stability has been verified by simulating 30 times the system with same parameters and different randomization, as well as by simulating the same system with different initial conditions in successive intervals. With respect to these tests, the system is stable, though atypical behavior may in principle arise. For the observed simulations, after some time fluctuations occur always around the same mean, which is characterized by the parameters $\nu$, $\gamma$, $r$ and the range of $\mu_i$. Numerically, the stability analysis yields a value of $0.092 \pm 0.006$ for unemployment for the scenario presented in figure 2, averaged over 800 iterations, and $0.087 \pm 0.019$ for the scenario with recentering of the margin (shown in figure 4). The amplitude of business fluctuations is less stable ($0.009 \pm 0.006$ and $0.007 \pm 0.003$ respectively). The relatively few parameters of the model would allow for a Monte Carlo exploration or other estimation methods [5]. In the following we describe typical behaviour encountered when varying systematically the its parameters.

**Parameter dependence.** Simulations showed that the system is less sensitive to the number of new entries $\nu$ than to the interest rate $r$: If $\nu$ is very high, competition is high and most firms do not manage to survive the first iterations. These failures do however not affect aggregate joboffer significantly, since initial firm size is small. In contrast, the system is very sensitive to the interest rate $r$, and already a difference of 0.0001 leads robustly to higher unemployment.

A wide interval $[\mu_{min}, \mu_{max}]$ has the effect that unemployment disappears rapidly, since new profitable firms grow very fast and hire unemployed workers. In contrast, if the interval is more narrow, this may not happen fast enough and the system finds a stationary state with high unemployment.

**Business cycles** For both scenarios with and without recentering of the margin, the bankruptcy of a large firms causes a rise in unemployment (see figure 6). However, the feedback loops differ: whereas in 6 (a), lower aggregate profits cause lower employment (and therefore lower aggregate salaries) immediately, in 6 (b) unemployment occurs with some delay, and the feedback loop of such an endogeneous cycle is more complex. This is because in (b), firms first stagnate and shrink before their failure, whereas in (a) they do not shrink. Therefore, in (a) the downturns are always due to bankruptcies.




(a) (b)

Figure 6: A business cycle in the scenarios (a) without and (b) with recentering of the margin. If the salary per worker is 30, it corresponds to a system with full employment.

Roughly, this is the type of fluctuations that DelliGatti et al. [6] interpret as business fluctuations in their similar model. Since the evolution of the job losses due to bankruptcy is not

explicitely given in their book, we estimate its impact on unemployment to be at least twice as strong as in the case presented here, based on their bankruptcy frequency and the relation of firms to workers.

# 4 Discussion and Conclusion

In this paper we proposed an agent-based model linking statistical regularities of macroeconomic systems to characteristics of single firms like their typical life cycles. This is done by combining a well understood stochastic model [32] with further features. The new elements are heterogeneous margins, interest payments, bankruptcies and recentering of the margin of firms, reflecting an 'ageing process'. Altogether, the system exhibits a fat tailed size distribution (which can be approximated by a power law), a tent-shape growth rate distribution and an approximately Gaussian profit rate distribution, as well as fluctuations in unemployment, in aggregate debt level and in the number of active firms. We show and explain the feedback loops that reproduce endogeneous business cycles.

Finally we would like to point out some interesting problems for which this model can be useful: (1) This model may be developed further and calibrated such that it serves for the study of policy implications. A possibility, similarly to [8], is to vary some parameters that are currently fixed, i.e. to generate an external shock, e.g. in the interest rate, in order to reproduce a specific situation observed in the real world. (2) In the model presented here, the margin of one firm is cannot be influenced by its own behaviour throughout an firm's lifetime. Productivity increase happens thus only via the creation of new firms. A possible extension of the model would be to let the margin of an firm depend on its investments, as is done in [6], or by purchasing production goods that depreciate over time as [7]. This would allow for firms to take strategic choices and to have more variable lifecycles. (3) The financial sector is very simplified in the model. Decisions of the bank depending on its own balance sheet are not present, as well as variable interest rates or more than one bank, which seems a promising extension in order to study more in detail the role of credit. By banks' decisions whom to grant credit, it would also be possible to allow for a pro-cyclical entry rate of firms in times of booms, as present in the MOSES model [34]. Particularly, this may have a further impact on profit, growth and size distribution of firms, since investment may contribute to accumulation of capital [1].

# References


[1] Duncan K Foley and Lance Taylor. 4. a heterodox growth and distribution model. *Economic Growth and Distribution*, page 75, 2006.

[2] Guy H Orcutt. A new type of socio-economic system. *The Review of Economics and Statistics*, 39(2):116–123, 1957.

[3] Barbara R Bergmann. A microsimulation of the macroeconomy with explicitly represented money flows. In *Annals of Economic and Social Measurement, Volume 3, number 3*, pages 29–44. NBER, 1974.

[4] Gunnar Eliasson. Competition and market processes in a simulation model of the swedish economy. *The American Economic Review*, 67(1):277–281, 1977.

[5] Elisa Baroni and Matteo Richiardi. Orcutt's vision, 50 years on. *LABORatorio R. Revelli, Center for Employment Studies Working Paper No*, 65, 2007.

[6] Domenico Delli Gatti. *Macroeconomics from the Bottom-up*, volume 1. Springer, 2011.

[7] Charlotte Bruun. Rediscovering the economics of keynes in an agent-based computational setting. Conferencepaper, 2008.

[8] Pascal Seppecher. Un modèle macroéconomique multi-agents avec monnaie endogène. 2009.

[9] Domenico Delli Gatti. *Emergent macroeconomics: an agent-based approach to business fluctuations*. Springer, 2008.

[10] Silvano Cincotti, Marco Raberto, and Andrea Teglio. Credit money and macroeconomic instability in the agent-based model and simulator eurace. *Economics: The Open-Access, Open-Assessment E-Journal*, 4, 2010.

[11] Giovanni Dosi, Giorgio Fagiolo, and Andrea Roventini. Schumpeter meeting keynes: A policy-friendly model of endogenous growth and business cycles. *Journal of Economic Dynamics and Control*, 34(9):1748–1767, 2010.

[12] Dirk J Bezemer. Understanding financial crisis through accounting models. *Accounting, Organizations and Society*, 35(7):676–688, 2010.

[13] Wynne Godley and Marc Lavoie. *Monetary economics*. Palgrave Macmillan Basingstoke, 2007.

[14] Steve Keen. Solving the paradox of monetary profits. *Economics Discussion Paper*, (2010-2), 2010.

[15] Charlotte Bruun and Carsten Heyn-Johnsen. The paradox of monetary profits: an obstacle to understanding financial and economic crisis? *Economics Discussion Paper*, (2009-52), 2009.

[16] Robert L Axtell. Zipf distribution of us firm sizes. *Science*, 293(5536), 2001.

[17] Kazumi Okuyama, Misako Takayasu, and Hideki Takayasu. Zipf's law in income distribution of companies. *Physica A: Statistical Mechanics and its Applications*, 269(1):125–131, 1999.

[18] Giulio Bottazzi and Angelo Secchi. Explaining the distribution of firm growth rates. *The RAND Journal of Economics*, 37(2):235–256, 2006.

[19] Einar Jón Erlingsson, Simone Alfarano, Marco Raberto, and Hlynur Stefánsson. On the distributional properties of size, profit and growth of icelandic firms. *Journal of Economic Interaction and Coordination*, pages 1–18, 2013.

[20] Simone Alfarano, Mishael Milaković, Albrecht Irle, and Jonas Kauschke. A statistical equilibrium model of competitive firms. *Journal of Economic Dynamics and Control*, 36(1):136–149, 2012.

[21] Michael HR Stanley, Luis AN Amaral, Sergey V Buldyrev, Shlomo Havlin, Heiko Leschhorn, Philipp Maass, Michael A Salinger, and H Eugene Stanley. Scaling behaviour in the growth of companies. *Nature*, 379(6568):804–806, 1996.

[22] Marco Lamieri and Ilaria Sangalli. Economia e finanza dei distretti industriali. Report, sanpaolo intesa, 2011.

[23] Robert Gibrat. *Les inégalités économiques*. Recueil Sirey, 1931.

[24] John Sutton. The variance of firm growth rates: the 'scaling'puzzle. *Physica a: statistical mechanics and its applications*, 312(3):577–590, 2002.

[25] Dongfeng Fu, Fabio Pammolli, Sergey V Buldyrev, Massimo Riccaboni, Kaushik Matia, Kazuko Yamasaki, and H Eugene Stanley. The growth of business firms: Theoretical framework and empirical evidence. *Proceedings of the National Academy of Sciences of the United States of America*, 102(52):18801–18806, 2005.

[26] Misako Takayasu, Hayafumi Watanabe, and Hideki Takayasu. Generalized central limit theorems for growth rate distribution of complex systems. *arXiv preprint arXiv:1301.2728*, 2013.

[27] Boris Podobnik, Davor Horvatic, Alexander M Petersen, Branko Urošević, and H Eugene Stanley. Bankruptcy risk model and empirical tests. *Proceedings of the National Academy of Sciences*, 107(43):18325–18330, 2010.

[28] Herbert A Simon. On a class of skew distribution functions. *Biometrika*, 42(3/4):425–440, 1955.

[29] G Udny Yule. A mathematical theory of evolution, based on the conclusions of dr. jc willis, frs. *Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character*, 213:21–87, 1925.

[30] Boyan Jovanovic. Selection and the evolution of industry. *Econometrica: Journal of the Econometric Society*, pages 649–670, 1982.

[31] John Sutton. Gibrat's legacy. *Journal of economic Literature*, 35(1):40–59, 1997.

[32] Cornelia Metzig and Mirta B Gordon. A model for scaling in firms' size and growth rate distribution. *arXiv preprint arXiv:1304.4311*, 2013.

[33] Michal Kalecki. A theory of profits. *The Economic Journal*, 52(206/207):258–267, 1942.

[34] Gunnar Eliasson. Modeling the experimentally organized economy: Complex dynamics in an empirical micro-macro model of endogenous economic growth. *Journal of Economic Behavior & Organization*, 16(1):153–182, 1991.

[35] Cornelia Metzig and Mirta B Gordon. Cedit constraints and heterogeneity in a macroeconomic agent-based model. *Workshop on Economic Heterogeneous Interacting Agents*, 2013.

[36] Hernan Mondani. *Statistical Mechanics of Organizational Growth Processes*. PhD thesis, KTH, 2012.

[37] S Picoli Jr, RS Mendes, LC Malacarne, and EK Lenzi. Scaling behavior in the dynamics of citations to scientific journals. *EPL (Europhysics Letters)*, 75(4):673, 2006.

[38] Yonathan Schwarzkopf, Robert Axtell, and J Farmer. An explanation of universality in growth fluctuations. 2010.

[39] Alex Coad. Firm growth: A survey. 2007.

[40] Giorgio Fagiolo and Alessandra Luzzi. Do liquidity constraints matter in explaining firm size and growth? some evidence from the italian manufacturing industry. *Industrial and Corporate Change*, 15(1):1–39, 2006.

[41] Amir F Atiya. Bankruptcy prediction for credit risk using neural networks: A survey and new results. *Neural Networks, IEEE Transactions on*, 12(4):929–935, 2001.

## 2.2 Theoretical description of the extended model

In publication 1, an extended model is presented via equations and a visual scheme. In this section, some consequences of the features introduced in publication 1 are analyzed more theoretically. The central extension is the introduction of heterogeneous margins $\mu_i$. Margin heterogeneity complicates the dynamics of the model, in that now, high-margin firms grow faster on average than firms with a lower margin. This is described theoretically and discussed in 2.2.1. The introduction of heterogeneous margins affects the scaling exponent, which is discussed theoretically and numerically in 2.3.

### 2.2.1 Replicator dynamics

In this section, the relative growth of firms is addressed. This feature did not exist in the setting of chapter 1, where all firms have the same $\mu$ and differ merely in size, which evolves stochastically. Margin heterogeneity accounts for a slow relative growth by which high-margin firms outperform firms with lower margins. For its description, the concept of replicator dynamics is briefly introduced. Replicator dynamics is a concept from evolutionary biology which describes growth in terms of populations, rather than in terms of individuals [50]. It has been introduced by [100]. Replicator dynamics describe the evolution of a sub-population of size $n_i$ by the following deterministic nonlinear equation:

$$\dot{n}_i = n_i[a_i(n_i) - \phi(n)]\,, \tag{2.1}$$

where $n = \sum_i n_i$ is the size of the system. A typical application is to explain the population growth of biological species, where the evolution of one species $i$ reflects its reproductive success, $a_i$ is a function denoting the fitness of a particular species $i$. The 'average success' of the system is described by the second term $\phi(n) = \sum_{j=1}^{N_f} n_j a(n_j)$. Through this term, the success of a strategy depends on the success of all other strategies. Beyond biology, the replicator equation has various applications in evolutionary game theory. Here the interpretation of the terms is slightly different: generally a $n_i$ corresponds to the number of individuals that play a strategy in a game[1], and $a_i$ denotes its payoff. Replicator dynamics describe a game that is carried out repeatedly in an entire population of players, not merely by individuals [50]. Instead of biological reproduction, players may switch to a more rewarding strategy, due to imitation, learning or inheritance [88]. The so-called 'folk theorem' proves the existence of Nash equilibria for replicator dynamics, i.e. a set of strategies that no player will benefit from leaving unilaterally.

### 2.2.2 Application to the firm growth model

In the model as detailed in chapter 1, the dynamics was purely stochastic, and different paths of the firms depended exclusively on the stochasticity of the

[1] In this description, strategies are assumed to be pure strategies, i.e. they are not combintations of basic strategies

dynamics, since all the firms had the same margin. Introduction of margin heterogeneity adds a deterministic term to a firms' evolution, such that the mean of growth rates $\mathcal{G}(g)$ follow replicator dynamics. Equation (2.1) written in discrete time becomes

$$n_{i,t+1} = [a_i(n_{i,t}) - \phi(n) + 1]n_{i,t}\,, \tag{2.2}$$

where $n_{i,t}$, $1 \leq i \leq N_f$ are firm i's number of employees at time $t$. In the following the dynamics of $\langle n_{i,t}\rangle$ in the model is derived, but it would equally be possible to express the dynamics in terms of the produced quantity $\langle q_{i,t}\rangle$. To describe the effect of a shortage of workforce and aggregate demand, it is useful to define the effective margin $\mu_{eff,t}$, i.e. the weighted sum of firms' margins $\mu_{j,t}$

$$\mu_{eff,t} = \frac{1}{N_{w,t}} \sum_{i=1}^{N_f} \mu_i n_{i,t}\,. \tag{2.3}$$

where $N_f$ is the number of firms, and $N_{w,t}$ the total number of workers hired at iteration $t$ (since there is full employment, it equals the total amount of workers, $N_w$). (2.3) is the 'effective margin' in publication 1. The success (i.e. average growth) of a firm $i$ can be derived by expressing $\langle n_{i,t+1}\rangle$ in terms of $\langle n_{i,t}\rangle$. As in chapter 1, two cases are discussed separately: (i) firms consume, and the only limited resource is the workforce, and (ii) only workers consume, and the only limited resource for firms is the aggregate demand. In case (i), firms intend to hire

$$\hat{n}_{i,t} = n_{i,t-1}(1 + \mu^n_{i,t-1}) = n_{i,t-1}(1 + \mu_i)\,, \tag{2.4}$$

since all output is sold. The available amount of workforce is $N_w$, such that only a fraction $p$ of job offers can be filled, where $p$ is

$$p = \frac{N_w}{\sum_i \hat{n}_{i,t}} = \frac{N_w}{N_w + \sum_i \mu_i n_{i,t}} = \frac{1}{1 + \mu_{eff,t}} \tag{2.5}$$

The number of workers which a firm actually receives is on average

$$\langle n_{i,t+1}\rangle = p\,\hat{n}_{i,t+1} = n_{i,t}(1 + \mu_i)\frac{1}{1 + \mu_{eff,t}} \tag{2.6}$$

In case (ii) where firms do not consume, firms only sell a fraction $v = \frac{D_t}{Q_t}$ of their production, since demand comes only from the money which has been spent as wages in the same production cycle. Using the relationship between $n_i$ and $q_i$ (equation (1.3)), it can be written as

$$v = \frac{D_t}{Q_t} = \frac{\sum_i n_{i,t} w}{\sum_i q_{i,t} p} = \frac{\sum_i n_{i,t} w}{\sum_i (1 + \mu_i) n_{i,t} \frac{w}{p} p} = \frac{N_w}{N_w + \sum_i \mu_i n_{i,t}} = \frac{1}{1 + \mu_{eff,t}} \tag{2.7}$$

$$\langle \mu^n_{i,t}\rangle = \left\langle \frac{q^s_{i,t} p - n_{i,t} w}{n_{i,t} w} \right\rangle = \frac{q_{i,t}\, p \frac{D_t}{Q_t} - n_{i,t} w}{n_{i.t} w} \tag{2.8}$$

The number of jobs offered by firms is smaller than in scenario (i). On average, each firm demands

$$\langle \hat{n}_{i,t+1} \rangle = n_{i,t}(1 + \langle \mu^n_{i,t} \rangle) = n_{i,t} \left\langle \frac{\frac{D_t}{Q_t} q_{i,t} p}{n_{i,t} w} \right\rangle = n_{i,t} \left\langle \frac{q_{i,t} p}{n_{i,t} w} \frac{1}{1 + \mu_{eff,t}} \right\rangle \tag{2.9}$$

which is

$$\langle \hat{n}_{i,t+1} \rangle = n_{i,t}(1 + \mu_i) \frac{1}{1 + \mu_{eff,t}} \tag{2.10}$$

Since in case (ii), the workforce is sufficient, each firm is able to employ the desired number of workers, i.e.

$$\langle n_{i,t+1} \rangle = n_{i,t}(1 + \mu_i) \frac{1}{1 + \mu_{eff,t}} \,, \tag{2.11}$$

which is the same result as equation (2.6), that has been derived for case (i) where the constraint is a limited workforce.

$$a_i(n_i) = \frac{1 + \mu_i}{1 + \mu_{eff,n,t}} \,, \tag{2.12}$$

and the 'average sucess' $\phi(n)$ can then be written as

$$\phi(n) = \frac{1}{(1 + \mu_{eff,t})} \left( 1 + \frac{1}{N_w} \sum_{i=1}^{N_f} \mu_{i,t} n_{i,t} \right) = \frac{(1 + \mu_{eff,t})}{(1 + \mu_{eff,t})} = 1 \,. \tag{2.13}$$

Since other firms' sizes appear via (2.5) or (2.7) in the denominator of expression (2.12), the expression $a_i$ for a firm's success cannot be written in terms of a payoff matrix. The dynamics cannot be thought of as the repetition of two randomly drawn firms competing with each other: all firms compete *simultaneously*. The evolution then goes as follows: Based on their profit (payoff), firms adjust the planned production, anf their $n_i$ increases or decreases respectively. Other common interpretations in evolutionary game theory are that $\frac{n_i}{N_w}$ corresponds to a ratio of a particular strategy $i$, and $a_i(n_i)$ to strategy $i$'s payoff. In the presented case, equation (2.2) converges to a state where all employees are hired in firms whose $\mu_i$ equals $\mu_{eff}$, and this is the case if all employees are hired at the firm with the highest margin, which is the Nash equilibrium of the system.

### 2.2.3 Combined dynamics

The replicator dynamics with stochastic terms has been widely studied [50]. [2]. Having described how the average of a given firm size evolves according to deterministic replicator dynamics, the original problem of combining these

[2]For instance, Lotka-Volterra equations, which are equivalent to replicator dynamics (of one dimension higher), have been combined with the linear Langevin equation by [81]

dynamics with a stochastic term will now be addressed. These dynamics can be combined with the dynamics from the stochastic market constraints. Inserted into the linear Langevin equation (1.24), a firm's size evolves according to

$$n_{i,t+1} = [a_{i,t}(n_{i,t})\, g_{i,t}]n_{i,t} + f_t \tag{2.14}$$

The deterministic replicator dynamics are dominant for large firms which exhibit relatively small stochastic growth fluctuations. Their distribution $g_i$ is more narrowly centered around a mean value $m$, with standard deviation $\sigma \propto n^{-1/2}$. The value $m$ is heterogeneous among firms, since $a_{i,t}$ are heterogeneous. In contrast, for small firms, the standard deviation of $g_i$ is far larger, and their evolution in size is governed more by the stochastic term of (2.14). In the next subsection, an element is introduced that prevents the emergence of a monopoly, which would be the evolutionary stable state of the dynamics of equation (2.14)

### 2.2.4 The recentering of the effective margin

The above dynamics alone would lead on the long term to a situation where all workers accumulate in the firm with the highest margin. A possibility to avoid it is introduced in publication 1: $\mu_{eff,n}$ is recentered at each iteration, by subtracting its difference from the margins of all active firms.

$$\mu_{i,t} = \mu_{i,t-1} - (\mu_{eff,n,t} - \mu_{eff,n,t-1}) \tag{2.15}$$

This has two effects. Firstly, the margins of existing firms slowly decline, which may be interpreted as an aging of technology. Secondly, new firms are started at a fix rate with a margin drawn at random from a fixed interval. Since their margin has not been lowered by the recentering, they have some probability to have a margin higher than existing firms, and grow faster. This simple method has however some limitations. One, discussed in publication 1, is that labour productivity for a single firm declines over time. The second is that this largely deterministic evolution of the margin limits the maximum lifetime of firms. This aspect is discussed in chapter 3.

### 2.2.5 Interest payments, Entry and Exit of Firms

Replicator equations can equally be derived if firms pay interest $l_{i,t}$. In order to finance production, firms can take loans from the bank at an interest rate. At the end of each production cycle, they repay it if possible, otherwise debts accumulate and the firm starts the next iteration with lower equity. Interest is charged for the loans, whether from the actual timestep or accumulated in previous timesteps, given by equation (8) in publication 1:

$$l_{i,t} = [n_{i,t}\, w - e_{i,t}]\,(1 + r)\,. \tag{2.16}$$

This has the effect of lowering the net expected profit of a firm by the interest payments due to this amount, in addition to the fact that over time, the margin

of a firm declines due to the recentering. Firms' expected $\mu_i$ is now lowered, and becomes the time-dependent 'net expected' $\mu_i$, which can be written as

$$\mu'_{i,t} = \frac{q_i p - n_i w - l_i r}{n_i w + l_{i,t} r} \tag{2.17}$$

This affects the overall constraints in the system, and $p$ and $v$ (equations (2.5) and (2.7)) need to be rewritten with these $\mu'_{i,t}$.

One implication of interest payments are that now, constraints in the job market and in the goods market can be present simultaneously. If the scenario is (i) where firms spend their net realized profits in the goods market of the following iteration, the profits are lowered by the interest payments, and therefore $D_t < Q_t$ even if firm spending is included.

Interest payments are also the feature which links the role of the margin to the entry and exit dynamics. Lower net profits further reduce a firm's chances to repay the loan, so that at some point in time, its leverage will reach a threshold at which it is declared bankrupt by the bank. This additional exit mechanism is a consequence of loans and interest. In the simple setting of chapter 1, firms could only 'die out' by not receiving any workers in the job allocation process. Also in that case, exit of firms was a naturally arising feature, and entry of firms was a necessary ingredient to maintain a constant number of firms and a stationary size distriution. In that sense, new entry and exit conditions may be seen as a modification of the additive noise term in the Langevin equation (1.24).

To summarize, the term $a_{i,t}(n_{i,t})$ in equation (2.14) as it arises in the setting with $\mu$-recentering, interest payments and bankruptcies, becomes time-dependent. For this reason, the stationary state is not attained via replicator dynamics, which typically describes a population that reproduces itself, and in which a nontrivial evolutionary stable state may exist due to the feedback of the term $\phi(n)$. Here, in contrast, firm sizes do evolve according to equation (2.14), but the evolutionary stable state (which would be a monopoly) is never reached, since new firms are entering with a possibly higher margin. As shown in figure 3 (b) of publication 1 and figure 2.3, the size distribution of the model containing all these features is still heavy tailed, because the stochastic element is still dominant.

The possibility of describing the dynamics by a replicator equation also provides a bridge from the model towards evolutionary economics. It has also been used for the same purpose in other agent-based models inspired by Schumpeter [29]. As addressed in section 2.4, replicator dynamics contain an economic foundation: they apply whenever agents (here firms) do not have perfect information about other agents' strategies (here margins). In contrast, a scenario where firms act upon knowing other firms' margins would be described by best reponse dynamics.

## 2.3 Effect of heretogeneous $\mu_i$ on the scaling exponent $\beta$

The simple model in chapter 1 describes growth as the sum of the growth of its independent subunits, i.e. of its jobs. As other models sharing this assumption [97], it yields a scaling exponent for the growth rate standard deviation $\sigma \propto n^{-\beta}$ with $\beta = 0.5$. This result is reportedly far from empirical studies which find smaller values in the range of $0.1 < \beta < 0.3$. Furthermore, it is based on the assumption that firms all have the same expected profit margin $\mu$. In this section, the effect of the new features on this result is investigated theoretically and numerically in 2.3.1 – 2.3.2.

### 2.3.1 The prior distribution

For the sake of clarity, the following reasoning is detailed for scenario (i) where firms consume, and where the competition is for workforce in the job market. It holds equally in the goods market. To recall: In the simple setting in chapter 1, all firms could maximally earn the profits $\mu$. If on top of this margins are distributed, firms multiply their current size by heterogeneous factors to determine their demand for workers. The probability for a job to be taken depends on the total shortage of workforce in the system, and so does the average shortage for one firm. Firms with a margin inferior to $\mu_{eff}$ will therefore shrink on average. The net realized margin, which is the quantity on which firms base their production decision, is therefore a prior probability distribution to the growth rate distribution. The question becomes how to calculate $P(g)$ whose mean $m$ is distributed according to $p(m)$.

$$P(g|n) = \int_{-\infty}^{\infty} P(g|m,n)p(m)dm\,. \tag{2.18}$$

As shown in chapter 1, $P(g|m,n)$ is Gaussian for our model. If $p(m)$ is assumed to be Gaussian, $P(g|n)$ can be written as

$$P(g|n) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_1}} e^{-\frac{(g-m)^2}{2\sigma_1^2}} \frac{1}{\sqrt{2\pi\sigma_2}} e^{-\frac{m^2}{2\sigma_2^2}} dm\,, \tag{2.19}$$

where now the $n$-dependency is contained in the standard deviation $\sigma_1$. This can be written, using completing-the-square

$$P(g|n) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_1}} \frac{1}{\sqrt{2\pi\sigma_2}} e^{-\frac{g^2}{2(\sigma_1^2+\sigma_2^2)}} e^{-\left[\frac{g\sigma_3}{\sqrt{2}\sigma_1^2} - \frac{m}{\sqrt{2}\sigma_3}\right]^2} dm \tag{2.20}$$

where the abbreviation $\sigma_3^2 = \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}$ was used. It yields

$$P(g) \propto e^{-\frac{1}{2(\sigma_1^2+\sigma_2^2)} g^2} \tag{2.21}$$

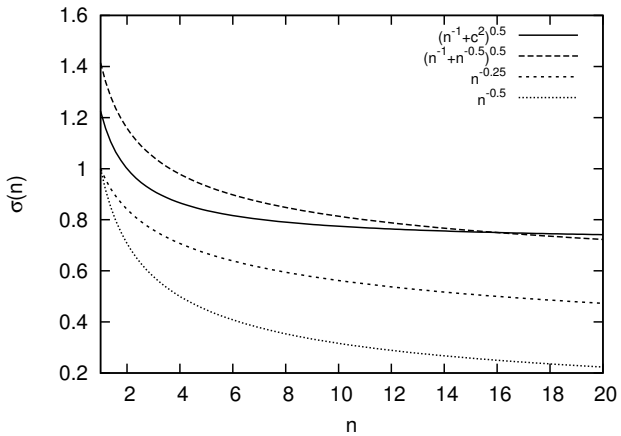



Figure 2.1: $n$-dependence of different scaling variance relations in the growth rate probability $P(g|n,m)$. The continuous line shows the scaling relation where the prior distribution is Gaussian with variance independent of firms' size, i.e. the growth rate scaling relation is $\sigma \propto \sqrt{1/n + c^2}$. If the prior distribution is itself $n$-dependent and scales like $\sigma_2 \propto n^{-0.25}$ (long-dashed line), $\sigma \propto \sqrt{n^{-1} + n^{-0.5}}$, both cannot be written in terms of a (constant) scaling exponent $\sigma \propto n^{-\beta}$, but their $n$-dependence is closer to $n \propto n^{-0.25}$ than $n \propto n^{-0.5}$. This is numerically verified, and is closer to empirical results than the simple setting with $\beta = 0.5$.

$P(g|m)$ is equal to the Gaussian $G(g|n,m)$ derived in chapter 1, whose $\sigma_1 \propto n^{-0.5}$. If this $n$-dependence is assumed for $\sigma_1$, the $n$-dependence of $P(g|n)$ in equation 2.21 becomes $\sigma \propto \sqrt{1/n + \sigma_2^2}$, which cannot be expressed by a constant scaling exponent $\beta$ : for firms with large $n$, the correction by a (constant) $\sigma_2$ is more dominant than for small $n$. A comparison is shown in figure 2.1.

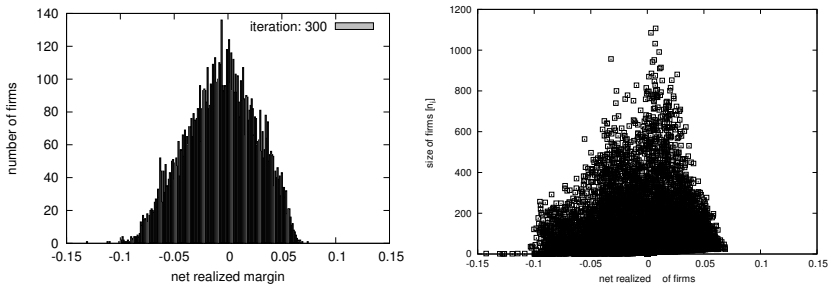
### 2.3.2 The conjugate prior distribution for job market constraints in the model

The analytic expression for the prior distribution $p(m)$ is hard to assess: the newly introduced firms have a uniform distribution of their gross expected margin, but this is not the case for the population of active firms. Over time, firms become indebted, which lowers the net expected margin. The lower the margin, the faster firms indebt. These firms do not live long, and therefore only few firms with very low net realized margins are active. Some 'lucky' firms may not accumulate debts for a very long time (allowing them to become large). Figure 2.3 (c) shows a snapshot of the numerically obtained prior distribution at the stationary state: It shows the frequency of firms as a function of their

net realized margin (repeated here from equation (20) in publication 1),

$$\mu^n_{i,t} = \frac{q_{i,t}p - n_{i,t}w - l_{i,t}r}{n_i w + l_{i,t}r} \,. \tag{2.22}$$

$\mu^{net}_{i,t}$ contains their intrinsic margin $\mu_i = \frac{q_i p - n_i w}{n_i w}$ appearing in the produced quantity $q_{i,t}$ and their debt $l_{i,t}r$. Simulations showed that the distribution in figure 2.2 (a) becomes stationary. From its shape, a Gaussian distribution is assumed as prior distribution. Based on $\mu^{net}_{i,t}$, firms decide on the quantity to produce, so they will demand $\hat{n}_{i,t} = n_{i,t-1}(1 + \mu^n_{i,t-1})$. The distribution $P(g|m)$ comes from the shortage of workforce in the job market, and can be approximated Gaussian. This can be compared to the scenario (i) presented in chapter 1, where firms consume and all output is sold. Then, $\mu^{net}_{i,t} = \mu = const$, and no prior distribution exists.




(a) (b)

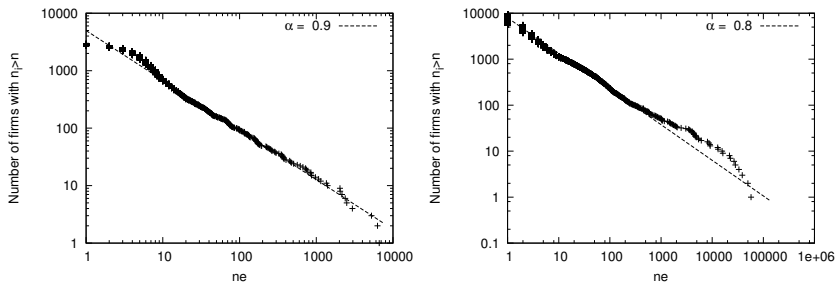
Figure 2.2: (a) Frequency of firms as a function of their net realized margins, which is the numerically determined $p(m)$. Based on this distribution, $p(m)$ has been assumed Gaussian. (b) Scatterplot of $\mu^{'}_{i,t}$ vs. size. This shows that for big firms, the distribution of the net realized margin has a smaller variance, and that $p(m)$ in equation 2.18 is $n$-dependent.

**$n$-dependence of $\sigma_2$.** An additional effect which is addressed only qualitatively is the following: the variance of $\mu^{net}_{i,t}$ can depend in addition on the firm's size. This is shown in figure 2.2 (b).

**Prior distribution in the goods market.** The idea that margin heterogeneity introduces a prior distribution for the growth rate holds also if the allocation of demand in the goods market is considered. Firms produce $q_i = \frac{w}{p} n_i(1+\mu_i)$, so the offer is $Q_t = \sum_i \frac{w}{p} n_i(1+\mu_i)$, but aggregate demand is only $D_t = \sum_i n_i w$. In this case, it is the distribution of the expected $\mu_i$ which is the prior distribution, not the distribution of $\mu^{net}_i$, as was the case for the job market.

**Other explanations for $\beta < 0.5$** This explanation for a scaling exponent $\beta < 0.5$ was based on the idea that firms of a certain size $n$ have different

averagre growth rates. $\sigma(n)$ is therefore wider than the standard deviation $\sigma(n_i)$ of a particular firm $i$. An alternative explanation or $\beta < 0.5$ has been detailed in section 1.9, where the internal structure of a firm – not detailed with microfoundations – was assumed to account for a different scaling exponent. This idea also has been suggested in [95, 97]. It is possible that both of these explanations hold true simultaneously.




(a) (b)

Figure 2.3: (a) Counter-cumulative size distribution of a complete setting including interest rate ($r = 0.0162$), margin heterogeneity ($0 \leq n_i \leq 0.1$), restarts ($\nu = 120$ per iteration) and bankruptcies, without recentering of the margin. (b) for comparison: counter-cumulative size distribution of the simple setting from chapter 1, where the only constraint is limited purchasing power. The fact that firms remain slightly smaller is a result of the recentering of the margin and the bankruptcy mechanism (detailed in 2.2.4 and 2.2.5), since this limits the lifetime of firms, and therefore also the time in which they can grow. For the simple setting shown in (b), there is no prior distribution of the net margins $p(m)$. The equivalent to the distribution shown in figure 2.2 (a) would be a delta peak. Snapshots at the stationary state after 2200 (a) and 7500 (b) iterations.

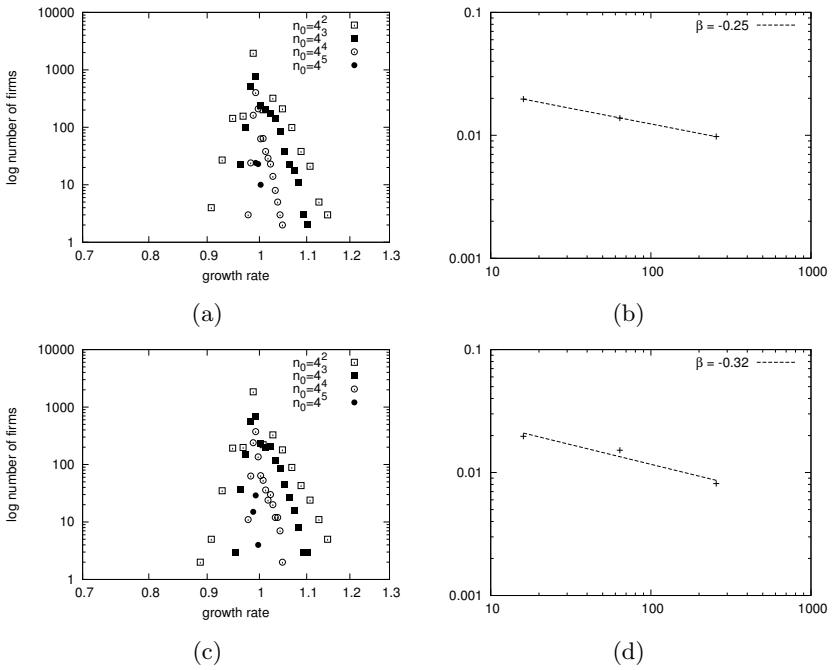
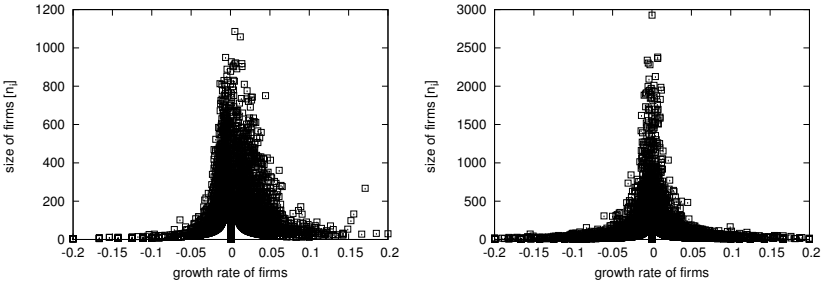



Figure 2.4: Two snapshots of firm growth rates, where the firms are grouped into size bins. Snapshots at iteration 300 (a) and (b) and 310 (c) and (d). (b) and (d) show the slopes fitted to the 'tent-shaped' growth rate distribution as a function of the smallest firm within the bin. The reason why the scaling exponent $\beta$ is not 0.5 is that the mean of the firm growth rates are distributed as a result of $\mu$-heterogeneity. The exponent $\beta$ may vary slightly over time, and its precision is not very high. Simulations with $10^6$ workers, $\approx 10^4$ active firms, interest $r = 0.01$, initial $\mu$ uniformly distributed between 0 and 0.1. The fact that more firms with positive growth rates exist results from the bankruptcy threhold: The same plot for the simple setting without margin heterogeneity is shown in figure 1.12, yielding $\beta = 0.5$.




(a) (b)

Figure 2.5: Scatterplots of firm sizes vs. growth rates. (a) complete setting with same parameters as in figure 2.4 with scaling exponent $\beta \approx 0.3$, (b) simple setting from chapter 1 with $\beta = 0.5$. As expected, in (a) the growth rate variance shrinks to a lesser extent for bigger firms than in (b).

## 2.4 The context of economic models

Since the model has assumptions at short and long term, and combines stochastic and deterministic dynamics, it can also incorporates argument from different theories, which can be weighted differently depending on the choice of the parameters. For instance, the choice of the width of the margin distribution influences the speed at which the effective margin increases, and therefore allows to weight the deterministic dynamics with respect to the stochastic dynamics. The smaller this range, the more important the stochastic dynamics is.

The level of abstraction at which the economy is analyzed is very high. Although many other models explain firm growth by emphasizing the role of the number of sectors in which a firm is active [107] [19], in this model no sectors are present. Here, firms produce an abstract good, which can be thought of as a service, since it is non durable. This simple hypothesis may be plausible, given that in modern industrialized countries the service sector accounts for more than 2/3 of the emloyment [34]. An abstract good is convenient since it allows the approximation that all firms are in competition with all other firms in the goods market. The introduction of sectors would make the competition more complex: firms would compete with firms of their sector to sell their goods, but would also compete for the limited purchasing power of workers with *all* other firms. The microfoundations of the model should be handled with care. It is not possible to make any statement about features which the model has abstracted from.

The model describes an industrial economy of firms and employees. Firms have more possible choices than employees: depending on their margin, they offer workplaces, produce goods, sell them, take loans. Workers, in contrast, are very simple in this model. They are homogeneous and cannot compete. This separation has been described by economists since the work David Ricardo (1772 – 1823) and is a central feature of so-called heterodox models [91, 40].

### 2.4.1 Keynesian arguments

As is the case other macroeconomic agent-based models [21, 30, 87], the model incorporates several of Keynes's ideas. The common points to Keynesian arguments refer mainly to production:

- Monetary profits are the driving force for economic activity[90]
- Firms base their decisions on the expectations of profits [39]
- Competitiom is present in quantity and not price[90, 91]
- Firms focus on the 'short run'. In Keynes's view, "there was only a succession of short runs"[39] rather than a long run. This also holds true in the model, in that firms plan only one iteration in advance, and because new firms are constantly arriving. If described by a differential equation, boundary conditions would change at every timestep (see figure 2.6).

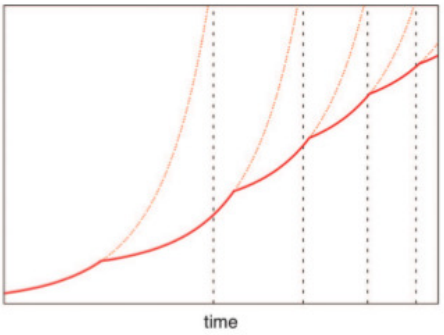



Figure 2.6: Scheme in [13] to show how changing boundary conditions influence the evolution of some quantity. The dynamics can be seen as a succession of short run developments. In a similar way, productivity evolves in the presented model. Without the arrival of new firms, productivity (see equation (21) in publication 1) would converge to a constant value.

As to Keynes's arguments on investment, the comparison is hardly possible, since at this stage the model has a very simple financial sector, and firms have access to credit in a 'binary' mode: below a certain leverage threshold they do, above they don't. The financial sector has potential to be described more realistically and has so far not been the focus of the analysis of the model (see chapter 5). However the comparison cannot be too in depth, since Keynes himself did not specify microfoundations for his macroeconomic theory. As states [39] p.195, "The problem of linking Keynesian macroeconomics to a coherent and persuasive theory of competition among individual firms remains a central unresolved issue in contemporary economics."

### 2.4.2 Neoclassical arguments

Few common points can be stated, since the neoclassical axioms of perfect information and market-clearing prices are completely missing. The absence of perfect information of firms is also addressed in 2.2.1. The only way that firms can know their own competitiveness is through their sales in the market. Past experience is also their only way to experience the aggregate demand: for instance, if a large firm goes bankrupt, unemployment rises and aggregate demand drops suddenly, but other firms will only notice this once they have incurred negative profits due to overproduction. What distinguishes this model most from neoclassical assumptions is that there is no asset *market* that takes the function of allocating risk. For instance via variable interest rates, a negative feedback loop is generated that keeps the system in an equilibrium and does not allow for systemic risk [39],[14].

### 2.4.3 Framework of statistical market equilibrium [38]

D. Foley describes a framework of a market which exhibits a statistical equilibrium, which is methodologically weaker than a Walrasian equilibrium. It applies to the markets in the studied model. It considers a general market where agents' offer and demand of a traded good are posted, Both quantity and offer/demand price can vary among agents. All feasible outcomes (called *transactions*) are equally likely. (Only one will actually happen and clear the market). The results that can be achieved by several transactions have higher probabilities, as in this model. For the distribution of results, the Gibbs-Shannon entropy is calculated, and maximized under 2 constraints, the conservation of probability of outcomes, and that only transactions take place which on average leave every agent in an advantageous position, using Langrangian multipliers. This setting is equivalent to the canonical ensemble in statistical physics. The Langrangian multiplier of the second constraint is interpreted as "market temperature". When it is high, many transactions can possibly take place, and "agents are spread out widely in terms of the gains from trade". When it is low, only few transactions are possible. Applied to this model, a high temperature corresponds to a high parameter $\mu_{eff}$, where many transactions are possible (although only one will be realized).

### 2.4.4 Evolutionary Economics

Evolutionary economics is a field inspired by concepts of evolutionary biology. It was termed so by Thorstein Veblen (1857 – 1929) who was inspired by Ch. Darwin's *The Origin of Species* [49]. In evolutuionary economics often the role of competition for resources and growth are emphasized. In the perspective of evolutionary economics, Schumpeter's approach to this was to describe the economy as being in a macroeconomic equilibrium on the short term, which is being transformed by the arrival of innovations. This is assumed to happen discontinuously and shakes apparently calm periods [85]. These arguments are taken up by the fact that new firms are entering continuously, which are potentially more productive than existing firms, and displace them. This displacement follows replicator dynamics, as detailed in 2.2.1. While this is happening, the system may be regarded to be in a short-term equilibrium, since the job market and the goods market are both designed in a way that every realization has the same probability, which guarantees that the system exhibits a stationary firm size distribution (see figure 2.3).

### 2.4.5 Stock-flow consistent models

The guiding principle of stock-flow consistent models is the conservation of money within the system [47]. This conservation imposes some constraints on macroeconomic outcomes that do not need to be explained by other arguments, e.g. behavioral arguments. From this follows that if a model should reach a stationary state, every economic actor has to spend his earnings at each timestep

(see also the introduction of publication 1). Stock-flow consistent models share the assumption of an agent-based approach, in that economic actors do not have foresight, and therefore base their decisions on simple rules, which for stock-flow consistent models are based on their own balance sheet. So does this model, which is stock-flow consistent. How accounting is precisely done is considered to be crucial in stock-flow consistent models, as well as in this model.

**The balance sheet transaction matrix.** The flows of any stock-flow consistent model can be described in a matrix. Some of these models are agent-based, such as [21, 87], or analytical [40, 47, 55]. As also discussed in publication 1, the cited stock-flow consistent models include different money flows. All models have in common that profits earned by firms need to be spent in some way in order to obtain a stationary state. Otherwise, firms would either not earn any profit, or economic activity would eventually cease, which is Marx's "paradox of monetary profits" [55, 22]. Firm spending can either happen through investment in capital goods [21, 87], through dividends if they are public companies [21] or through high interest payments to a bank, which then spends its earnings. This may be because the bank itself is a public company [87], because the bank acts as a consumer in the goods market [55], or because the bank cancels debt of bankrupt firms, which is the case in the presented model. The "simplest model" by Godley and Lavoie [47] has instead a state who collects taxes from firms, and re-distributes them in the form of unemployment benefits. One alternative to firm spending is present in the model by J. Mimkes [69], who describes the production cycle as an analogon to a heat pump, which creates mechanical energy (profits) by exploiting two different temperature levels (wage levels): applied to the production cycle, production needs to take place at a lower wage level than consumption. This solution however is not possible in a closed system as described here, where the recipients of wages coincide with the customers. Table 2.1 shows all the money flows of the presented model. For the sake of simplicity only the aggregate flows of this model are shown, but the matrix has in fact as many rows and columns as there are agents.

In publication 1, an effect of stock-flow consistency is manifested in the coupling of aggregate wages and spending in the next iteration, since the modeled economic system is closed. In a situation with high unemployment, aggregate demand is low, and therefore firms are constrained in how much they can grow. Another feature where stock-flow-consistency is important is the role of credit availability: If firms did not face any constraint and could take infinite debts, the job losses due to bankruptcies would be non existent, since the only reason for a firm to vanish is because it doesn't sell anything (or doesn't receive any workers), which only occurs for very small firms.

| **transaction** | | **sectors** | | **Σ** |
|---|---|---|---|---|
| | workers | firms | bank | 0 |
| wages | $N_t\, w$ | $-N_t\, w$ | | 0 |
| loans | | $\Delta L$ | $-\Delta L$ | 0 |
| interests | | $-L\, r$ | $L\, r$ | 0 |
| consumption | $-Q^{s,w}\, p$ | $Q^{s,w}\, p$ | | 0 |
| consumption firms | | $-Q^{s,f}\, p$ | | |
| | | $Q^{s,f}\, p$ | | 0 |
| reimbursement of loans | | $-\Delta L'$ | $\Delta L'$ | 0 |
| $\Delta$ equities | $\Delta E_w$ | $\Delta E_f$ | $\Delta E_b$ | 0 |
| **Σ** | 0 | 0 | 0 | 0 |

Table 2.1: Accounting matrix of the model, aggregated into sectors. $L$ denotes aggregate loans, $E$ is the financial position of an agent.

### 2.4.6 Macroeconomic Agent-Based Modelling

Recently, the neoclassical approach of using a representative agent has been widely criticised, in that it makes what is called a "fallacy of composition" (e.g. [28], [90]) Agent-based modelling as a methodology can circumvent this [32],[102]. For the same reason, the use of agent-based models in economic science has been advocated as early as 1957 by Orcutt [77]. Among the earlier of these models, which were at the time called microsimulation models, are [11, 31]. Later models, which are discussed below, detail more the interaction between the financial and the real sector.

Despite their diversity, agent-based macro-models have several features in common. Typically, the agents are households, firms, one or many banks, and possibly a government. They model the production cycle, as well as the money flows, which is why some of them are stock-flow consistent. Agents act according to relatively simple rules, even though some models may differ in this point, and may take ideas from different economic arguments, e.g. contain elements of neoclassical optimization [45]. They all typically exhibit feedback loops, which may arise from constraints in the markets, from long-term growth dynamics, and possibly imitation of other agents' behaviour. Feedback may be positive (reinforcing) or negative (stabilizing). There are always some stochastic elements in the markets. To understand more precisely business fluctuations in a model, the crucial elements are the timescales at which these feedback loops operate, and the amplitude of stochastic fluctuations admitted in the model. Depending on these, business fluctuations may result from feedback loops of the underlying stochastic differential equations, or merely from fluctuations.

#### Existing models

Here some approaches of macroeconomic agent-based models are briefly presented. They differ from the model studied in this thesis in that they contain more features, and that they are not analyzed theoretically as in chapter 1 and sections 2.2.1 and 2.3.1. Other differences are that all of these simulate a much smaller number of agents, and that the focus lies more on predictions and policy implications, which is not the case of the model in this thesis. Common results with the presented model are that some of these models find fat tailed size distribution and a tent-shaped growth rate distribution [45, 30]. Since they contain many features, the dynamics is necessarily some type of multiplicative noise. Furthermore, superposition of many features may lead to a Gaussian growth rate probability density of individual firms, so that the explanation for a tent-shaped $\mathcal{G}(g)$ derived in 1.7.2 might approximately hold in these models, although this is not explicited.

**The model by G. Dosi et al. [30, 29]** The model captures elements from the theories of Keynes (short term demand fluctuations) and of Schumpeter (technology-driven growth), hence the name K+S (Keynes and Schumpeter). It consists of two types of firms, workers (which are not simulated explicitly), and

a state. One type of firm produces investment goods, another one consumption goods, as in [21]. Investment good producers can both innovate, and imitate other firms' technology. Firms exit if their size has attained very small values, or if their net assets become negative. Every exit is followed by the entry of a small new firm, so the number of active firms is strictly constant (not fluctuating as in the case of the presented model). In the labor market, wage adjustments take place, but not to an extent that unemployment disappears. As in the model of this thesis, firms' market shares are reported to evolve according to replicator dynamics, such that profitable firms outperform unprofitable ones, which are however not specified as done for this model in 2.2.1. The system has 22 parameters; the number of firms is 250. Workers are not simulated explicitly, which is why their model is only partly comparable to the one presented in this chapter. In contrast to the model in this thesis, the main purpose of the K+S model is to test different policy scenarios. The key variables that govern economic activity are tax rates and unemployment benefits. Validation is done by exploiting statistical regularities of empirical data. For this aim the authors compute averages over many simulations in order to compensate for its small size. A tent-shaped growth rate distribution is reported, as well as a size distribution, which decays faster than lognormal.

**The 'Eurace' model by S. Cincotti et al. [26]** The Eurace model is stock-flow consistent and focuses on the interrelation of financial and real sectors. It models a job market, a goods market and a financial market. Agents are several commercial banks, a central bank, a state, firms which are joint-stock companies, and households. The goods market takes place at malls to which firms deliver their production. A homogeneous good is traded, but at varying prices. Households try to spend all their money at one mall and split their puchase only if the mall does not have the desired quantity. After the second attempt, money is kept, so not necessarily all potential purchasing power is spent, unlike the case in the model in this thesis. In the job market, firms offer jobs at different wages. A matching algorithm assigns workers to open jobs. If not all positions are filled, the desired quantity cannot be produced. The malls act as a buffer, since they may dispose of stock, and because shortage in workforce at one time step can be compensated by higher production in the following time step. This is not so in the model of this thesis, where shortage in workforce immediately repercutes in a change of company size and its next production. In the financial market, the government, firms and the households participate for the purpose of saving and speculation. Separately, firms can take loans from commercial banks in a credit market to finance their production. Besides this, they invest in capital. A central bank sets the maximum leverage of commercial banks. It may buy government bonds and therefore influence the amount of credit money in the system, in addition to loans issued by commercial banks to firms (as in the model in this chapter). Both explicitely simulated external shocks, and endogeneous feedback loops exist in the Eurace model, which are identified as reasons for fluctuations in economic activity. The reasons

are "the coordination failure between demand and supply of consumption goods, strong fluctuations in the investment in physical capital, and disruptions in the supply chain as well as mass layoffs due to firms' bankruptcies"[26].

The Eurace model's dynamics are governed by 10 parameters. Its analysis is uniquely done for time series of aggregate quantities. Neither distributions nor individual life cycles are addressed.

**The model by Ch. Bruun [21].** This model is an implementation of the theory of Keynes. It comprisies workers, firms, a simple bank, and a state which collects taxes and pays unemployment benefits. As in our model, wages and prices are fixed, and competition happens only in quantities, except in a stock market where stock prices may change. Firms need a certain stock of capital in order to be able to produce, which they buy, and whose value depreciates over time. However, capital and labour cannot be substituted as in neoclassical production functions. Firms adjust their production quantity depending on last period's profits, albeit not linearly dependent as in our model. In the model, two types of firms exist: producers of consumption goods, and producers of capital goods. The firms are joint-stock companies, who pay dividends to shareholders (which are workers and other firms). Firms also make an investment decision, depending on experienced capital constraints, stock prices, and their financial position. From their intrinsic parameters, firms are homogeneous, but the many degrees of freedom introduce heterogeneity. Workers are placed on a spatial network. They do not spend all of their wages but their consumption depends on their neighbours' consumption. Randomness is present in the job market, the goods market and in the stock market. Analysis is mainly via aggregate variables and their correlation in a business fluctuation. As a distributional feature, the Gini index, which measures skewness of the wealth distribution of households, is calculated. Despite its similar assumptions to the presented model, the comparison is difficult, since it has 17 parameters.

**The model by D. Delli Gatti et al. [45].** In their book, the authors have analyzed several scenarios. I will refer to their central model, which they term 'basic agent-based model' (BAM). The agents of this model are households, firms and banks. Despite using an agent-based approach, this model uses some principles from neoclassical economics: In the goods market, the job market and the credit market, prices can adjust because agents demand several offers. However, the agents are adaptive and do not perform real optimization. The markets are decentralized, without a clearing mechanism, which is why the outcomes are characterized by unemployment and unsold goods .Inflation is present, which is a result of the price mechanisms in the markets. In a simple scenario, no productivity increase is allowed. If over indebted, firms can go bankrupt, and are replaced by new firms. Business fluctuations are explained as originating from waves of bankruptcies (and restarts). The model has 14 parameters. Results of the model have been compared to a firm database of Italian firms, which does however not contain bankruptcies and birth of firms. The model re-

produces a fat-tailed size distribution and a Laplacian size distribution, whose parameters are compared to data. Since the shown size bins are very large, it is difficult to compare the size distribution with the one in the presented model. Other comparisons to data are the evolution of average interest rates, and wage distributions.

**The model by P. Seppecher [87].** This model is equally stock-flow consistent and describes a system of households, firms, and a bank. It bases itself monetary circuit theory, i.e. the endogeneous creation of money. This means that once a firm takes a loan, an amount of money is being created positive on the firm's account, and negative on the bank's account. Additionally, the bank earns some interest. After firms earn revenues from the production financed by the loan, they pay it back. This is termed 'destruction of money'.No phyical capital good exists in this model, the only costs for firms are labour and interest payments. In contrast to the model of this thesis, a bankruptcy mechanism does not exist Seppecher's model. Over-indebted firms are merely charged a higher interest rate, but continue to exist even once they are declared bankrupt. New firms are not introduced. The absence of bankruptcies (i.e. absence of money losses to the bank) poses the problem that the bank accumulates money from interest payments, which would eventually stop economic activity. This situation is prevented in P. Seppecher's model by the introduction of high dividend payments from both firms and the bank, which is a stock company, to shareholders (except some own funds that are kept by the bank). Other differences to the presented model are the existence variable prices and wages, and therefore inflation. The markets are search algorithms for individual agents, which end after a certain number of search steps. The evolution of distributions has not been presented. The model has 27 parameters, but how many of them are important for the dynamics is not discussed .

**The presented model in this context.** The model presented in this thesis differs in that it is much simpler. $\mu_{min}$ is always 0, which is why it has not

| $[\mu_{min}, \mu_{max}]$ | range of expected gross profit margins of firms |
|---|---|
| $r$ | interest rate |
| $\gamma$ | bankruptcy threshold |
| $\nu$ | number of restarts per iteration |
| $n_{init}$ | initial size of a firm |
| $l_{init}$ | initial debts of a firm |

Table 2.2: List of parameters

been listed separately. The model depends sensitively on the $\mu$ and $r$. The bankruptcy threshold $\gamma$ and number of restarts per iteration $\nu$ affect the size of the system, as well as the average lifetime of firms. The dynamics is less

sensitive to the itilialisation parameters $l_{init}$ and $n_{init}$. Further dependence on parameters is discussed in publication 1.

Furthermore, the presented model focuses less on the financial markets than the mentioned approaches, at the present stage of its development. Its main contribution is to study the evolution of distributions, not merely the mean values of aggregate quantities. This provides a link between mathematically tractable simpler growth models and the macroeconomic agent-based models.

**Comparison of the business cycle mechanisms.** One object of study of the models by [45] and [21] is to describe business cycles. In economic theory, business cycles are explained as being either reactions to exogeneous shocks, or emerge endogeneously, i.e. they are positive feedback loops. In the cited models, these are endogeneous cycles, or simulated reactions to a shock, as in [87]. In [21], the cyclic behaviour seems to be governed by a herding effect of household's decision to invest in stocks or to consume, which then influences the investment and production decision of firms. In [45], downturns of economic activity are released by waves of bankruptcies. Since they may occur abruptly, they may also be seen as *endogeneous* shocks due to randomness in the system.

For the model in this thesis, some business fluctuations have been presented in publication 1 (figure 6). It is shown that the more features are included, the less fluctuations are due to randomness, and the more they are indeed feedback loops. In figure 2.7, the feedback loop is driven by the entry and exit of firms, firms' consumption in the goods market, and the recentering of $\mu_{eff}$. It goes as follows: While firms' net realized profits are increasing, spending in the goods market also increases. High profits incite firms to produce too much, which is why not all goods can be sold and their profits saturate [3]. Since $\mu_{eff}$ is recentered, firms' expected margins slowly decline, and the indebtedness increases, which further lowers their net profits. Firms who intend to spend their last realized profits as consumption in the goods market, consume less. Due to their low profits, firms do not plan to grow quickly, and at some point they cannot absorb the workforce which is set free due to large bankruptcies. Unemployment is a consequence (visible here in the decline of the aggregate salaries, since the salary is a constant). However, since the failing firms are unprofitable and indebted, a wave of bankruptcies lowers aggregate debt in the system. This causes aggregate profits to rise again, because new unindebted firms dominate who can earn higher profits. Due to their high profits, they offer more jobs, which is why unemployment declines again.

Further features causing endogeneous feedback-loops could readily be added to this model, for instance a feature in order to create a feedback loop from credit availability. Since for example the stationary state of the level of unemployment depends on interest rate $r$ and bankruptcy threshold $\gamma$, it is clear that if these parameters depend on economic prospects, and these in turn on unemployment, the system would exhibit more complex endogeneous cycles. The introduction

[3] Note that in the presence of interest payments, $D_t < Q_t$ even if firms spend their last realized profits.

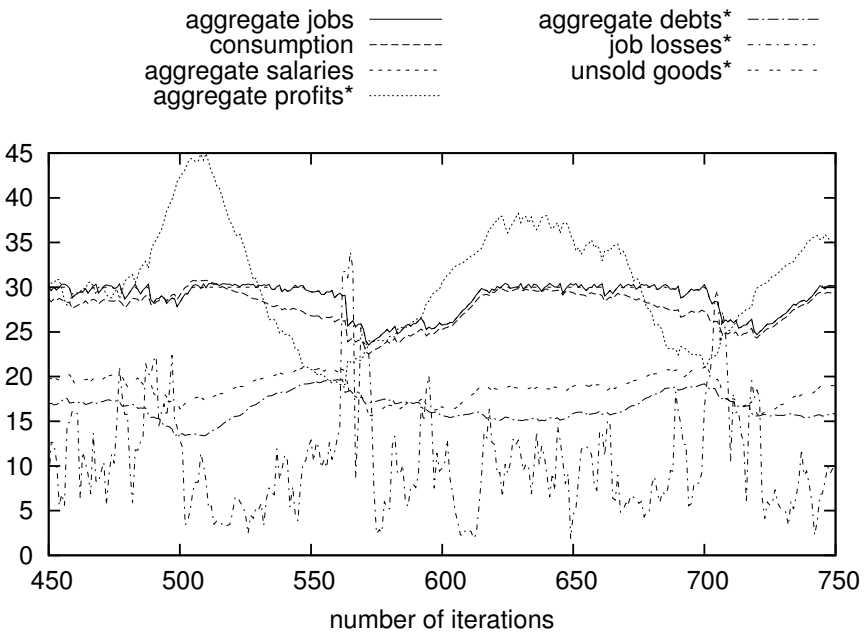



Figure 2.7: Feedback loop of several quantities. Aggregate salaries of 30 correspond to full employment. Quantities denoted with $*$ are rescaled, but no offset from 0 is added.

of further features can also have a stabilizing effect on economic activity, i.e. they may introduce negative feedback loops.

## 2.5 Concluding Remarks

In this chapter, a simple stochastic firm growth model has been extended with several features: margin heterogeneity of firms, interest payments and a limited volume of credit. As in the simple scenario of chapter 1, the dynamics of the model is stochastic, and generates a power law decay of the size distribution, as well as a tent-shaped growth rate distribution. In addition, the profit rate distribution is analyzed, as well as life cycles of individual firms.

It has shown that with the introduction of margin heterogeneity on its own, the system would evolve towards a monopoly. However, if additional, 'counteracting' features exist, a stationary state may be obtained. In the presented case, these are interest payments which lower the gross expected margin, 'ageing' of margins due to recentering of the effective margin $\mu_{eff,t}$, as well as bankruptcies and entry of firms. These features, however, represent only one possibility of further specification of a stochastic firm growth model, and other specifications may be closer to reality. Even though comparison to empirical studies look promising, what can really be stated is merely that they does not seem to falsify the model.

The presented scenario is much simpler and more tractable than existing agent-based macro models. It can be seen as a baseline scenario for further development towards model as comprehensive as other macro models. The necessity of some extension depends of course on the purpose for which the model shall be used, and care needs to be taken not to loose track of the causality. Possible elements are a more cognitive bank, which takes into account the profits, margin or the financial position of a firm when deciding on the volume of credit a firm receives. Another possible feature is the introduction of investment goods. In view of globalization, a further possible scenario is to study a system which is not closed, e.g. where customers do not coincide with employees. In addition to new functions, it is possible to study different behavioral rules for firms and workers, for instance by comparing different levels of information when taking decisions.

# Chapter 3

# Study of Bankruptcies in the model

## 3.1 Introduction

This chapter consists of a working paper, containing the sequel of the analysis carried out in chapter 2. The theoretical aspects in section 2.2 and 2.3, as well as the comparison to models in section 2.4 apply equally for the results of this chapter. Without complexifying the model, a number of new results are analyzed. These center around bankruptcies. For their description, the focus lies on the dynamics of ageing and debt accumulation, as well as on the criteria by which the bank declares firms bankrupt. In the following working paper, different bankruptcy criteria are tested, and how this choice affects the distribution of active and bankrupt firms is shown. Results are the age distribution and size distribution of active firms, and those of firms at the moment of bankruptcy. They allow the bankruptcy probability as a function of age and size to be calculated. Further results are a more thorough analysis of life cycles, the distribution of the liabilities, and the relation between several variables. These results have been compared with a number of empirical studies of firm characteristics [35, 1, 33, 61, 60, 58, 20], and to stochastic models for bankruptcy [79].

# Publication 2

presented at: Workshop of Economic Heterogeneous Interacting Agents 2013
(Reykjavik, Iceland)

# Credit Constraints and Heterogeneity in a Macroeconomic Agent-Based Model
# – work in progress –

Cornelia Metzig, Mirta B. Gordon

March 5, 2014

**Abstract**

This paper presents results of a macroeconomic agent based model comprising households, firms and a bank. Firms statistics of active firms are linked to bankruptcy statistics by focusing on the bankruptcy criteria and firms life cycles. Bankruptcy probabilities as a function of age and size are calculated using Bayesian inference. The model builds upon a simpler stochastic model [36] and is mathematically tractable. Several distributions are compared to empirical studies, such as size, growth rate, age, margin and debt distribution, as well as correlations of variables.

From the complexity perspective, a goal is to explain jointly arising stylized facts with coherent microfoundations. This is one of the purposes of macroeconomic agent-based models, which typically have several types of agents: households, firms, and one or many banks. They attempt to relate a simple theory of the firm to macroeconomic outcomes. This type of model has as a common feature that they avoid what is called "fallacy of composition" since the sectors are not aggregated to representative sectors. Besides that, they may incorporate different economic ideas, e.g. Keynesian [10], neoclassical [26] or Post-Keynesian as the model by [41] which focuses on monetary circuit theory, and does not include bankruptcies. All of the cited models are also stock-flow consistent.

Any agent-based model is based on a stochastic process, which can be described more or less explicitly [19]. Many relatively simple models exist, which describe only one or two stylized facts of an economic system. These may be city size or firm size formation [34, 49], firms' growth rate distribution [8]; growth rate distribution and scaling of the size-dependent growth rate variance [39, 22, 46]; or scaling laws for firms' bankruptcy risk [40]. These models rely on the stochastic evolution of firms' sizes. However, these simple models can be difficult to interpret in economic theory, since they typically do not model the production cycle [45, 24]. Another approach are macroeconomic agent-based models, which have more features, and in which mechanisms from these simple models may be present to some extent, even if they dynamics are not uncovered and analyzed explicitly. We showed in [37] that the construction from simpler firm growth models is a means to keep the model mathematically tractable. more than one timescale is relevant for the dynamics,

The strength of macroeconomic agent-based models is that they not only allow to study aggregate quantities like unemployment and GDP, but also distributions [25], which is a crucial intermediary step to validate plausibility of certain microfoundations. To do so, it is crucial to dispose of empirical evidence. Important stylized facts are the heavy-tailed firm size distribution [6][38]; the tent-shaped growth rate distribution and the scaling exponent of the growth rate variance [44][39][46][38]. Other interesting empirical results are the profit rate distributin [20, 2, 32], the effect of credit constraint on age and growth [21][11], the margin distribution [32], debt distribution [31, 40], bankruptcy probability of firms [40] and debt at bankruptcy [4].

The focus of this paper is the analysis of credit constraints and bankruptcies, in a framework that has been introduced and analyzed in other respects

in [36, 37, 35]. Since different factors are reported to contribute to firm failure, the model needs some sophistication. For the problem of corporate failure prediction, different techniques have been used in the relevant literature: neural networks [5, 17], data mining [33], and support vector machines [42]. The typically studied quantities in these approaches are as in [3]: working capital/total assets, retained earnings/total assets, earnings before interest and taxes/total assets, market capitalization/total debt, sales/total assets.

The idea of this paper is to link distributions of active and bankrupt firms in an agent-based model, and to exploit empirical studies for its qualitative validation. We use the previously introduced model [35] but focus on the lifespan of firms and the relation between distributions of active firms and bankrupt firms. Bankruptcy statistics depend crucially on typical life cycles of firms, which in turn depend on the rules of the model, as well as on the parameters, which weight the speed of the different dynamics. In the following we analyze the model along three axes: Distributions at a certain point in time, the evolution of single firms, and bankruptcy statistics. We refer to literature and empirical findings on these three aspects.

The paper is organized as follows. In section 1, we recall the equations of the model. In section 2, we study the age distribution of firms, and discuss how it is influenced by the used bankruptcy criterion. In section 4 we discuss with the aid of typical examples how firms evolve over time in their size, debt and profit margin. In section 5, further distributions of firm characteristics are shown, at this stage without an in-depth discussion. In section 6, we discuss and conclude.

# 1 The model

We recall the macroeconomic agent-based model that has been introduced in [35]. It comprises a number $N_w$ of available workers, a fluctuating number $N_{f,t}$ of firms and one bank representing a simplified financial sector. Firms are heterogeneous in their gross expected profit margins $\mu_i$. What differs significantly from the models by [26, 10, 41, 18, 12] is that the number of agents isone to two orders of magnitude larger, and that the validation is done via the stationary distributions of characteristic firm variables. To a part, these may be described theoretically, since the model is a mathematically tractable extension [37] of a stochastic model for firm growth [36]. Since the dynamics generates scaling phenomena, it is important that the number of firms and workers is big enough to be able to produce these phenomena[1]. The basic concept is that firms are in competition for limited resources, which in this paper are aggregate purchasing power in the goods market and workforce in the job market. As soon as there is a shortage of one of these resources, a stochastic element appears in the dynamics, because every demand has the same chance of being satisfied. That is, every offered unit of commodity goods has the same chance of being sold, and every offered position has the same chance of being filled. Competition happens thus only via quantities and no price or wage adjustment mechanisms are present in the model. A second important ingredient is heterogeneity of firms profitability, which leads to relative replicator-type growth of firms. A third aspect, which

[1] typically, the number of firms is $10^4$ (in this preliminary version $10^3$), and the number of workers $10^6$ (or $10^5$)

comes from the model's stock-flow consistency, is that firms finance production. Firms can demand loans at an interest rate, which at this point serves as abstraction for various kinds of capital costs. Over longer lapses of time, firms can accumulate debt until a threshold where they go bankrupt. The option of bankruptcy limits the life span of firms, and interest payments limit firms' net realized profits, on which they base the decision of how much to produce in the next iteration. New firms are being introduced at a constant rate, such that the system finds a (fluctuating) level of active firms rather than forming a monopoly. Since debt accumulation depends on the (heterogeneous) margins as well as on the stochastic dynamics, the life time is also heterogeneous.

## 1.1 The Sequence of Events – Equations

We recall the model. One iteration is composed of the following steps:

1. **Production decision of firms.** Firms decide a quantity $\hat{q}_{i,t}$ of goods to produce at time $t$. It is based on their production in the previous iteration $q_{i,t-1}$, adjusted by a factor that grows linearly with their net realized profit margin $\mu^n_{i,t-1}$ (defined in equation (15)).

$$\hat{q}_{i,t} = q_{i,t-1}\left(1 + \mu^n_{i,t-1}\right), \tag{1}$$

where $p$ is the price of a unit of commodity goods. To be able to produce this quantity, their demand for workforce $\hat{n}_{i,t}$ is given by

$$\hat{n}_{i,t} = \hat{q}_{i,t}\,\frac{p}{w}\,\frac{1}{1+\mu_i}\,, \tag{2}$$

where $w$ denotes the (constant) wage.

2. **Job market.** All $N_{f,t}$ firms post their job offers, which sum to

$$\hat{N}_{w,t} = \sum_{i=1}^{N_{f,t}} \hat{n}_{i,t}\,. \tag{3}$$

If the number $N_w$ of available workers is sufficient, firms hire the planned quantity:

$$n_{i,t} = \hat{n}_{i,t} \text{ if } \hat{N}_{w,t} \leq N_w\,. \tag{4}$$

If $\hat{N}_{w,t} > N_w$, workers are attributed with equal probability to every open position, so that the actual received quantity by a firm follows a binomial distribution, which has the mean value $\langle n_{i,t}\rangle = \hat{n}_{i,t}\,\frac{N_w}{\hat{N}_{w,t}}$ (see [36]). Since margis are heterogeneous, after the placement of workers in the job market, it is useful to calculate the effective margin of the economy:

$$\mu_{eff,t} = \frac{1}{N_{w,t}} \sum_{i=1}^{N_{f,t}} n_{i,t}\,\mu_i\,, \tag{5}$$

where $N_{w,t}$ denotes the number of employed workers. The competitiveness of a firm in the system depends on its margin compared to $\mu_{eff,t}$.

3. **Possibility to take a loan.** Firms check if they dispose of enough equity $e_{i,t}$ to be able to pay wages. If not, they demand a loan $l_{i,t}$ at interest rate $r$:

$$l_{i,t} = [n_{i,t}\, w - e_{i,t}]\,(1 + r)\,. \tag{6}$$

All agents do double-entry accounting, and technically all transactions are booked in a way that guarantees stock-flow consitency.

4. **Production and payment of salaries.** With a number $n_{i,t}$ of workers, a firm produces $q_{i,t}$ units of consumption goods:

$$q_{i,t} = \frac{w}{p}\, n_{i,t}(1 + \mu_i)\,, \tag{7}$$

Hired workers receive a (constant) salary $w$.

5. **Consumption decision.** Workers $j$ are assumed to spend their salary on a number of $d$ consumption goods:

$$d_{j,t} = \begin{cases} \frac{w}{p} & \text{if employed} \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

Scenarios have been simulated where also firms spend their net realized profits from the previous iteration. These are discussed in [35]. However, in the simulations presented here, workers are the only consumers, so that total demand is

$$D_t = \sum_{j=1}^{N_w} d_{j,t}\,. \tag{9}$$

6. **Commodity goods market.** Analogously to the job market, if aggregate demand $D_t$ is sufficient, all goods

$$Q_t = \sum_{i=1}^{N_{f,t}} q_{i,t} \tag{10}$$

are sold:

$$q^s_{i,t} = q_{i,t} \text{ if } Q_t \leq D_t\,. \tag{11}$$

If $Q_t > D_t$, the allocation of demand is again probabilistic, such that the actual sold quantity follows a binomial distribution, whose mean is $\langle q^s_{i,t} \rangle = q_{i,t}\, \frac{D_t}{Q_t}$. After the transactions, firms calculate their sales. Since unsold goods cannot be stored and are lost, the abstract good can also be interpreted as a service.

7. **Repayment of the loan and possible bankruptcy.** Firms which dispose of enough assets repay their loan to the bank. If their assets are not sufficient, they will start the next iteration with a more negative financial position. The bank declares a firm bankrupt according to two criteria:

(1) A firm's margin $\mu_i$ reaches

$$\mu_{i,t} = \mu_{crit} = 0 \tag{12}$$

This can happen since firm margins diminish due to an ageing process introduced in equation (16). As (16), bankruptcy criterion 1 is an optional feature.

(2) The debt ratio reaches a threshold $\gamma_{crit}$, and a firm's debts reaches a level $l_{i,crit}$:

$$l_{i,t}^{crit} = -\gamma_{crit}\, w\, n_{i,t} \tag{13}$$

Once a firm is declared bankrupt, its debts are losses to the bank. Implications of these two criteria are analyzed separately in this paper.

8. **Calculation of profits.** Firms net realized profits are

$$\pi_{i,t}^{n} = q_{i,t}^{s}\, p - n_{i,t}\, w - l_{i,t}\, r\,, \tag{14}$$

where $l_{i,t}$ may have accumulated over many iterations. This means that the net realized profit margin is

$$\mu_{i,t}^{n} = \frac{\pi_{i,t}^{n}}{(n_{i,t}\, w + l_{i,t}\, r}) \tag{15}$$

9. **Recentering of firms' margins** An optional feature is that at the end of each iteration, each of the intrinsic margins $\mu_i$ is lowered by an amount such that $\mu_{eff,t}$ remains constant.

$$\mu_{i,t} = \mu_{i,t-1} - \Delta_t \tag{16}$$

where

$$\Delta_t = \mu_{eff,t} - \mu_{eff,t-1}\,. \tag{17}$$

$\mu_{eff}$ has a natural tendency to grow at every iteration, since the dynamics of the job market has the effect that on average, firms with a high margin will offer more jobs and attract more workers. This recentering is a means to describe ageing of the technology of a firm and is important for the long-term dynamics, if new firms are introduced. $\Delta_t$ describes the productivity increase within iteration $t$.

10. **Introduction of new firms.** At the end of each iteration, a constant number $\nu$ of new firms are introduced. They are initialized as follows:

- $\mu_i \in [\mu_{min}, \mu_{max}]$. Therefore, on average their margins are higher than those of older firms, whose margins have already been shifted to lower values due to recentering.
- $n_{i,t} \in [0, n_{max}^{new}]$. The simulations presented in this paper use $n_{max}^{new} = 4$.
- $q_{i,t}$ and $\mu_{i,t}^{net}$ of new firms is calculated from these two values, which is used by firms for making their first production decision $\hat{q}_{i,t+1}$.
- assets $a_i$, liabilities $l_i$ are initialized, as well as equities $e_i = a_i - l_i$. Here, initial $l_i$ liabilities are distributed between $\gamma n_i w \leq l_i \leq 0$.

## 2 Results I – Age distribution

As a starting point we analyze the age distribution of active firms in empirical studies and in existing stochastic growth models. We then show the age distribution of purely stochastic version of this model [36]. This serves as a theoretical foundation for the results presented in subsection 2.2 where hetergeneity, debt accumulation and bankruptcy are added: the age distribution for different bankruptcy criteria. From these results follow all further results of this paper.

**Empirical Evidence.** In empirical studies different functions have been fitted to the age distribution of firms: [21] and [30] find that firm age follows a lognormal distribution in a database of Italian and Irish firms respectively. [15] finds firm age to be power-law distributed in a database of 6 million U.S. firms, [9, 7, 14] find an exponential age distribution (for Italian, Spanish and Indian firms respectively)[2].

**Two existing models.** A commonly used growth model is the model by H. Simon [43], which describes a constantly growing system. It has a constant influx of firms but no exit, so the age distribution of active firms exhibits linear decay. Another famous model is the model by Gibrat[27], which describes a multiplicative noise process that leads to a lognormal distribution for firm sizes $n_t$,

$$P(n_t) = \frac{1}{n_t\sqrt{2\pi\sigma^2 t}} e^{\frac{(\ln n_t - \epsilon_t)^2}{2\sigma^2 t}} , \tag{18}$$

whose width increases over time. Gibrat's model is based on the hypothesis that all firms have the same age, which limits its applicability for real-world data. [1] has solved this problem by working backwards: if an exponential age distribution

$$A(t) = \lambda e^{-\lambda t} \tag{19}$$

of firms is assumed, and if firms follow Gibrat's dynamics, the size distribution becomes

$$P(n_t) = \int \lambda e^{\lambda t} \frac{1}{n_t\sqrt{2\pi\sigma^2 t}} e^{-\frac{(\ln n_t - \epsilon_t)^2}{2\sigma^2 t}} dt \tag{20}$$

This firm size distribution is not lognormal any more, but exhibits a power-law decay in the first approximation:

$$P(n_t) \propto C\, n_t^{-\alpha-1} \tag{21}$$

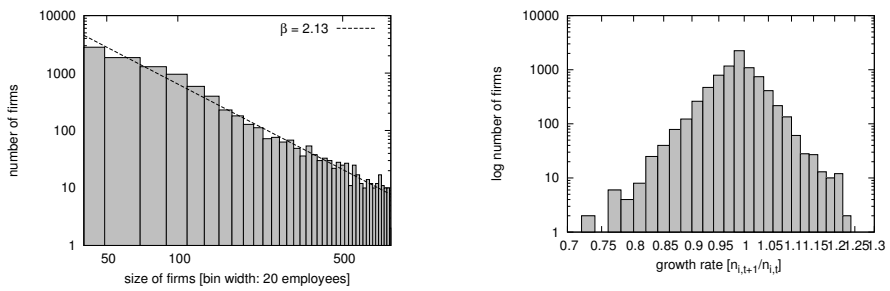
For this result there is empirical evidence [6, 38]. This coincides with the size distribution of models with a multiplicative noise term in the Langevin equation, where entry and exit are neccessary elements [47].

### 2.1 A purely stochastic scenario.

We now turn to the introduced model and recall a purely stochastic scenario analyzed in [36] where firms have a homogeneous $\mu$, with $r = 0$ and consequently

[2] The scaling reported here is only for firms above the median age, i.e. we do not consider differences reported for small firms here.

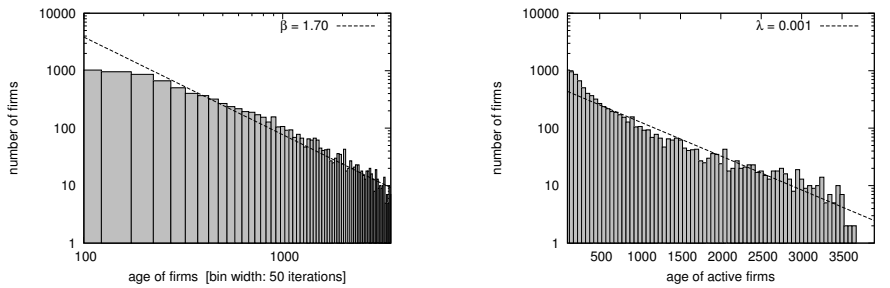
no bankruptcy, where firms only exit when they shrink to size zero. Its resulting size and growth rate distribution (figure 1) are close to widely reported empirical results (e.g. [6][44]). They correspond to the nonequilibrium stationary state of the stochastic firm growth which is generated by the competition in the markets. It yields a counter-cumulative size distribution which is not a Zipf law but decays approximately like a power law of exponent $\alpha \approx 0.7$, (see [36]). The reasoning from equations (20) and (21) applies to this result and suggests that for the age distribution, exponential decay may be a good approximation for our model, which has been fitted in figure 2. Since entry and exit of firms as




(a) (b)

Figure 1: Results discussed in detail in [36]: (a) Heavy-tailed size distribution of active firms in a simulation of homogeneous firms in double-logarithmic scale. (b) tent-shaped growth rate distribution.

necessary elements to reach this stationary state, an age distribution of active firms arises naturally.




(a) (b)

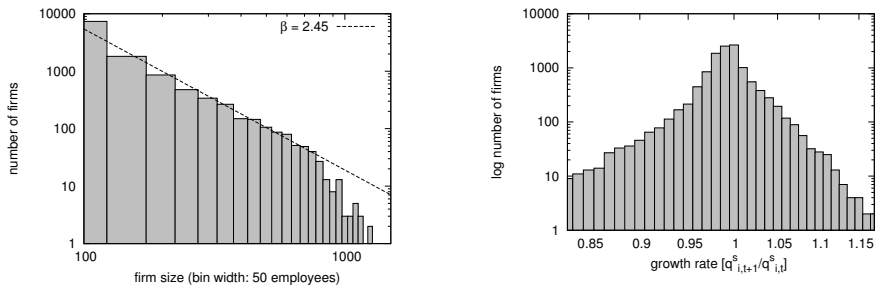
Figure 2: age distribution of active firms in a simulation of homogeneous firms discussed in [36]. (a) power law fit with exponent $-\beta$ (in log-log scale), (b) exponential fit with decay constant $\lambda$ (in log-linear scale).

A strong limitation of the model is that this age distribution arises for an economically unplausible reason: The only criterion by which firms exit the system is when they have shrunk to size 0. This highlights the need to complexify the model, which has been done by the introduction of heterogeneous profit margins, interest payments and, most importantly, bankruptcies. These

features superpose another relative growth dynamics that can be described by replicator equations [37] and is much slower than the stochastic growth. How these features modify the results is discussed in the following.

## 2.2 Bankruptcy criteria in the complete model

The model as detailed in subsection 1.1 equally exhibits a fat-tailed size distribution and a tent-shaped growth rate distribution, since these distributions are the result of the firms' stochastic growth (figure 3).




(a) (b)

Figure 3: (a) Heavy-tailed size distribution of active firms in a simulation of homogeneous firms in double-logarithmic scale. (b) tent-shaped growth rate distribution. $10^6$ workers, $10^4$ firms, interest $r = 0.005$, $\nu = 120$ new firms per iteration, $0 \leq \mu_{i,init} \leq 0.1$, $0 \leq l_{i,t,init} \leq \gamma n_i\, w$ bankruptcy criterion 2 (equation 13), $\gamma_{crit} = 2$. Firms' lifetimes are further limited by interest payments, bankruptcy and recentered margins, so the size distribution decays faster than in the simple setting, especially for large firms the decay is faster. $\beta$ is the exponent of a power law fit to the size frequency. This value depends on the chosen bin size. The deviation shows that the power law hypothesis is only approximately fulfilled. It corresponds to $\beta - 1$ for the cumulative size distribution.

Firms' evolution is not only probabilistic, but in addition $\mu$-heterogeneity and $\mu$-evolution create a typical 'life cycle' for firms' size, net realized margin, and debt ratio[3], defined as

$$\gamma_{i,t} = \frac{l_{i,t}}{n_{i,t} w} \,. \tag{22}$$

The dynamics of the expected values of size, margin and debt ratio is however slower than the stochastic market dynamics, i.e. it is weak on the short term, where the stochastic dynamics dominate. Whereas already for a purely stochastic model, entry and exit of firms are necessary to reach a stationary state, in the setup with interest payments and bankruptcies, the choice of the entry and exit criteria is economically more plausible. The aim is to understand how the arising distributions depend on this choice, and to compare the results qualitatively with data. These variables may be size, growth rate, age of active firms,

[3]Since in the model firm's margins $\mu_i \in [0, 0.1]$, the results are approximately the same if the debt ratio is normalized by a firm's sales $q^s p$ instead of wage payments. These would correspond to firm's assets before the repayment of a loan, i.e. $\gamma = \frac{l_{i,t}}{a_{i,t}}$

debt of active firms, net realized margin of active firms, as well as statistics of these variables at the moment of firm's bankruptcy, averaged over a certain time span. Two different bankruptcy criteria have been tested:

1. The expected margin $\mu_i$ reaches 0. If a firm does not even have a theoretical chance to earn profit, the bank stops financing it.

2. The debt ratio reaches $\gamma_{crit}$. This is also used in the agent-based models by [10] and [26].

Despite the simplicity of the model, these conditions contain similar information to the features traced by data-mining approaches enunciated in section **??**, where the *realized* profits are used.

Depending on which bankruptcy criterion is used, the size distribution of active firms is different, since margin and debt are only slightly correlated in the model (see figure 4).

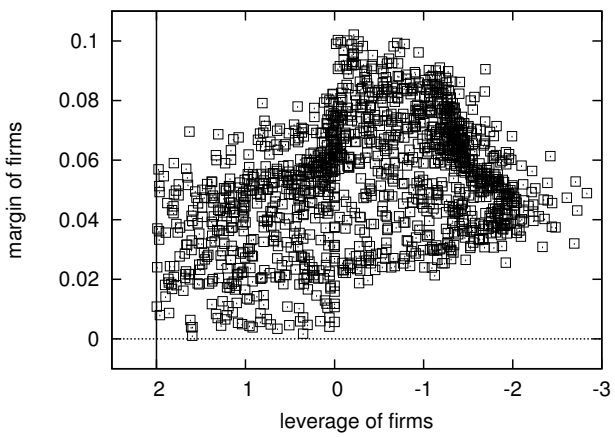
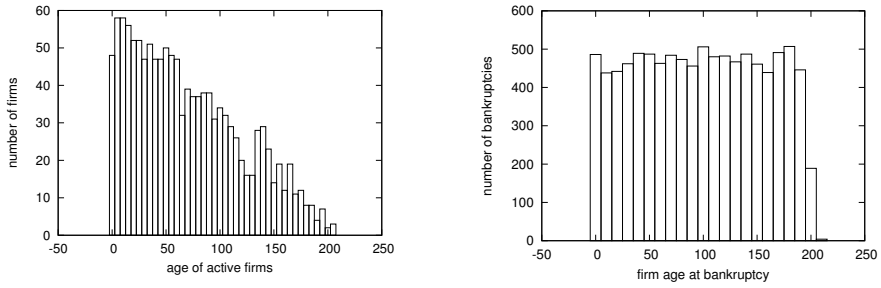



Figure 4: Typical snapshot of the 2-dimensional plane showing expected margin vs. debt of a company. Each firm corresponds to one point. Bankruptcy criteria 1 (dotted line) and 2 (continuous line) are shown. In this setting, the margin and debt ratio are only slightly correlated. Their correlation depends on the $\mu$-distribution and on the resource constraints which are responsible for the randomness of the dynamics. $10^5$ workers, $10^3$ firms, $r = 0.012$, iteration $= 1000$, new firms initialized with $0 \leq \mu_i \leq 0.1$ and debt $0 < l_i < \gamma$ (values $< 0$ correspond to positive equity). Bankruptcy criterion 2 has been used for the simulation

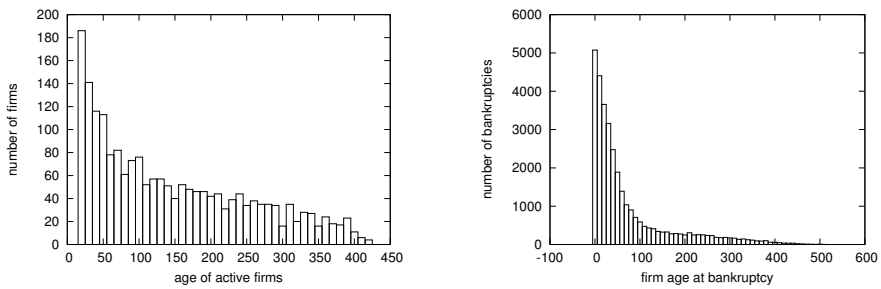
**Bankruptcy criterion 1: margin threshold $\mu_{crit} = 0$.** In this scenario, lifetime of firms is determined by their initial $\mu_i$ and system variables (number of entrants per iteration and possibly an interest rate $r \neq 0$). Firms are initialized with $\mu_i$ from a uniform distribution $\in [0, \mu_{max}]$. The rate how recentering lowers margins is the same for all firms, so life times are (approximately) linearly distributed (this is shown numerically in figure 5(a)). The age distribution at bankruptcy is a constant, shown in figure 5(b).




(a) (b)

Figure 5: Age distribution of active firms after 100 iterations for bankruptcy criterion 1 ($\mu_i = \mu_{crit} = 0$). In (a) a firm's lifetime depends only on its initial margin, and if this is drawn from a uniform distribution, the age distribution is linear.

**Bankruptcy criterion 2 (equation (13)): debt threshold $\gamma = \gamma_{crit}$.** In this scenario, firms are equally initialized with a margin $\mu_i$ and debt ratio drawn from a uniform distribution. They fail once their debt ratio has attained $\gamma = \gamma_{crit}$. In contrast to the margin $\mu_i$, the debt ratios evolve heterogeneously, depending on $\mu_i$, debt ratio itself and the stochastic market outcomes. The arising age distribution of active firms decays slower and is closer to an exponential distribution found in empirical observations [13] (see figure 6(a)). It differs from the purely stochastic model in that firms are still large when they fail, so their lifetime is shorter. If in addition recentering of $\mu_i$ is used, the maximal time in which a firm is profitable is lowered additionally. Therefore, even well-performing firms will eventually incur debt and fail, which is why the age distribution is cut at a certain value, depending on $r$ and the frequency of new entries. Also the age distribution for bankrupt firms has been simulated numerically. It is decaying approximately exponentially (see figure 6 (b)).




(a) (b)

Figure 6: Age distribution of active firms after 100 iterations in log-linear representation for bankruptcy criterion 2 (a firms' debt reaches the threshold $l_{crit}$). The age distribution of active firms depends on their debts, which in turn accumulate over time and diminsh the net realized margin, and therefore growth.

For both bankruptcy criteria, size and growth rate distributions are similar (see further discussion plots in appendix A). Compared to the purely stochastic model, for both bankruptcy criteria firms remain smaller and the tail of the size distribution decays much faster, since a limited lifetime also limits the maximum size to which firms can grow.

Analysis of the arising age distribution shows that in order to obtain an approximately exponential age distribution, the variable that determines bankruptcy should follow stochastic dynamics. In principle this is possible for the evolution of $\mu_i$ if more sophisticated rules for $\mu$-evolution are introduced. In order to keep the model simple, in this paper we limit our study to bankruptcy criterion 2 (debt threshold), where debts accumulate heterogeneously, which accounts for a slower decaying age distribution. This is analyzed in detail in the remainder of the paper.

# 3 Results II – Distributions of active and bankrupt firms

We start by analyzing the bankruptcy probability

$$p(b) = \frac{n_{bankrupt,t}}{N_{f,t}}\,, \tag{23}$$

i.e. the fraction of firms that fail in one timestep divided by the number of active firms. Some authors [13] study instead the *survival* probability, which is $p(s) = 1 - p(b)$. This probability relates statistics on active firms to statistics on bankrupt firms and provides further evidence as to whether the dynamics produce reasonable results. Unsurprisingly, on average, young firms are smaller, (e.g. [21] for Italian firms), which is also a result of this model. However, as figure 8 shows, the size distribution is fat-tailed for all age groups, which implies that size and age are not proxies for each other, and that a separate analysis of thhe bankruptcy probability of both variables is necessary.

## 3.1 Bankruptcy probability depending on age

The theorem of Bayes describes the relationship between the age distribution of active firms $p(a)$, the age distribution of firms at their bankruptcy $p(a|b)$, the probability of bankruptcy $p(b)$, and the probability of bankruptcy as a function of a firm's age, $p(b|a)$:

$$p(b|a) = \frac{p(a|b)p(b)}{p(a)} \tag{24}$$

Distributions $p(a)$, $p(a|b)$, and the rate $p(b)$ are direct results of the model. In order to obtain a higher precision, for the age distribution of active firms $p(a)$ and the age distribution at bankruptcy $p(a|b)$, both distributions were aggregated over 800 iterations. Summing over all ages of these two distributions, and dividing the total number of bankruptcies by the total number of active firms, a division yields the probability of bankruptcy $p(b)$. Using (24), $p(b|a)$, the bankruptcy probability depending on age, can be calculated.

For bankruptcy criterion 1 ($\mu_i = \mu_{crit} = 0$), bankruptcy probability $p(b|a)$ increases with age, which is unplausible. For bankruptcy criterion 2 ($\gamma = \gamma_{crit}$),

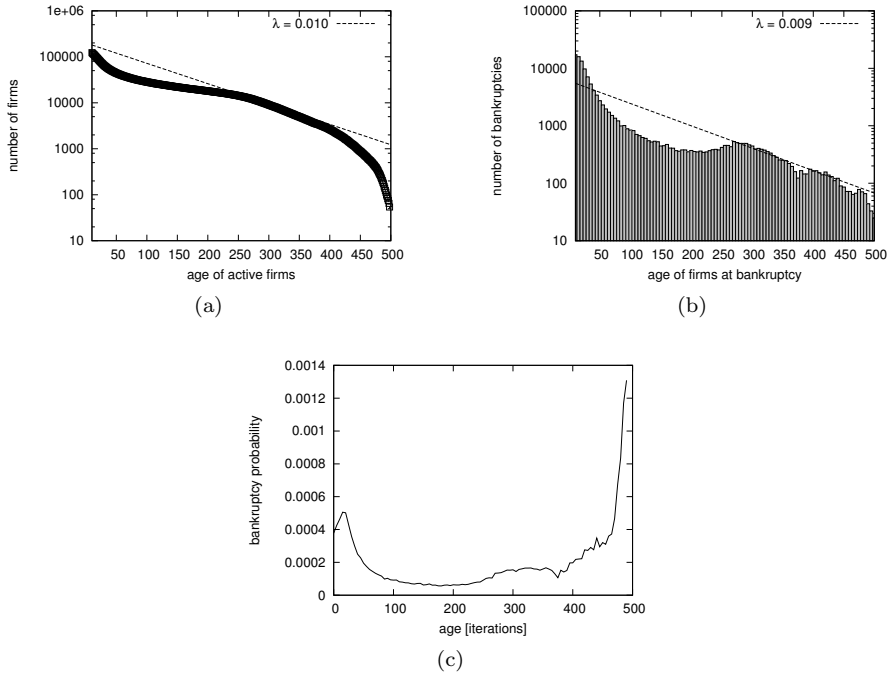



Figure 7: (a) age distribution (bin width 5) (snapshot at iteration 2000) (b) size distribution of bankrupt firms, averaged over 1300 iterations. (c) bankruptcy probability as a function of age, averaged over 1300 iterations. Total bankruptcy probability $p(b)$ is $1.52 \cdot 10^{-2}$. $10^6$ workers, $8 \cdot 10^3$ firms, interest $r = 0.012$. This is very low and means that firms have time to diminish their size before failure. (c) bankruptcy probability as a function of age, averaged over 1300 iterations. Firms are initialized with $0 \leq \mu_i \leq 0.1$ and debt $0 < l_i < \gamma$.

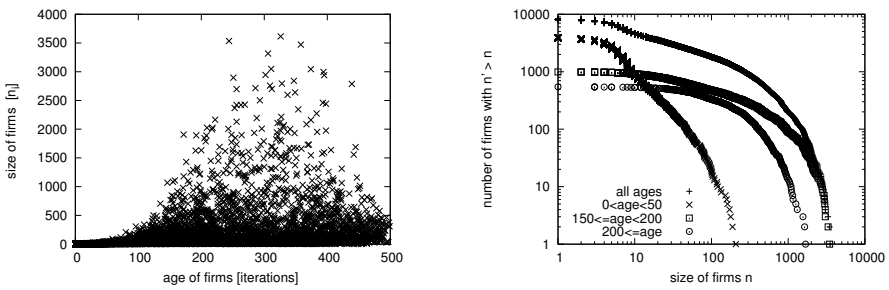
the probabilities $p(a)$, $p(a|b)$, and $p(b|a)$ are shown in figure 17 (a)-(c). For young firms, this probability $p(b|a)$ decreases for two reasons: young firms are small and therefore experience stronger random fluctuations. They start with debts that not all manage to repay. Older firms have already grown and repaid initial debts, and their bankruptcy probability is lower. This has also been observed empirically: [13] finds that young firms' bankruptcy rates decrease in their first years, by analyzing bankruptcy (or survival) probabilities separately for 6 datasets on Portuguese, Swedish and US firms. This lowered bankruptcy probability has been reported to be highly statistically significant, and is present independently of the year of firm entry.

However, the strong increase of $p(b|a)$ for old firms is an unwanted artefact that is due to the method of recentering $\mu_{eff}$. It limits the life time of firms, since even firms with a high $\mu_i$ will become unprofitable, incur debts, and eventually fail. The number of firms that reach this age and contribute to the peak is low, since most fail at an earlier age.

The same results for another simulation with a very low interest rate $r$ are presented in appendix B for comparison. In this scenario, the effect of $\mu_i$-recentering is more dominant. It is the reason why firms become unprofitable,

eventually incur debt and fail. The bankruptcy probability exhibits a higher peak for old firms than in figure 17 (c), where the interest payments were more dominant in reducing $\mu_i^{net}$ and led to different speeds at which firms become indebted.

In his empirical analysis, [13] show that the age distribution of different sectors may however deviate significantly from the aggregate age distribution: in the example of the international airline sector, he shows that multimodal behaviour exists, and that the corresponding age distribution decays faster than exponential. The total age distribution is a superposition of all sectors. Relating these results to our model suggests that more heterogeneity is needed in order to obtain an age distribution without cutoff. Introducing sectors with different ageing constants may be one of several possibilities.

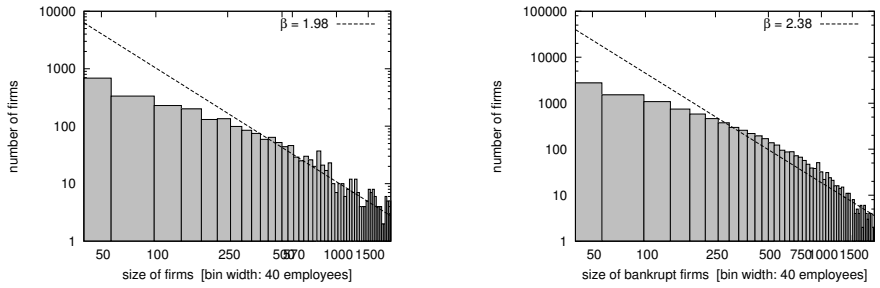



(a) (b)

Figure 8: (a) scatterplot size vs. age of firms, (b) rank-size distribution for selected age groups. It shows that every age group individually follows a fat-tailed distribution. Very large firms are only present in older groups, but the dynamics of the model is such that at a certain age, firms do not grow any more. Therefore the oldest age group does not contain the biggest firms. Snapshot after 2000 iterations, $10^6$ workers, $8 \cdot 10^3$ firms, $r = 0.012$. (same parameters as in figure 17).

## 3.2 Bankruptcy probability depending on size

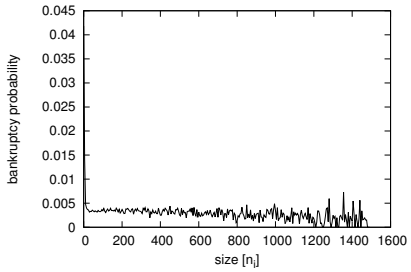
Bayes' theorem can also be used to calculate the bankruptcy probability as a function of firms' size, $p(b|n)$, as a function of the the size distributions of active firms $p(n)$, the size distribution of bankrupt firms $p(n|b)$, and the total bankruptcy probability $p(b)$. The result is shown in figure 9: The size distribution for bankrupt firms, measured in terms of assets, has been found by [40] to follow a power law. The same is found for the liabilities distribution, since they are linked by the debt ratio threshold $\gamma$. A power-law size distribution has been found empirically by [4, 23] for Japanese firms, measured in terms of their debt. The latter found an exponent of $\alpha = 0.911$ for the cumulative size distribution.by [15] for US firms, and by [28] for OECD countries.




(a)




(b)




(c)

Figure 9: (a) size distribution at iteration 1000. $10^6$ workers, $10^4$ firms, interest $r = 0.012$. This is very low and means that firms have time to diminish their size before failure. (b) size distribution of bankrupt firms, averaged over 800 iterations (c) bankruptcy probability as a function of size, averaged over 800 iterations.) Same parameters as in figure 17.

# 4 Results II – Lifecycles

In the model, a firm's evolution depends on its margin, and the three constraints which it faces: credit constraints, workforce constraints and constraints in demand. Over its lifespan, it evolves in size, debt, $\mu_i$, and earns fluctuating profits. These variables are traced for every firm, and should fit empirical studies on firm evolution. Their analysis provides another possibility to validate the model. Empirically, life cycles have been analysed not only regarding the exit probabilities which depend on the size and age of a firm, but rather focus on the underlying reasons that lead to an observed pattern of bankruptcies. An empirical study on life cycles of Canadian firms [48] uses as criteria (other than size and age) the firm level resources, a firms capabilities, as well as the competitive environment. Their results about firms level resources are that young and small firms may have liabilities of "adolescence". Those who cannot repay them because of a lack of managerial experience will fail once their internal assets are expired. In contrast, old firms who survived this period may fail for another reason: they are inable to adapt to a changed environment. In our model, a firm's resources are the equity on which a firm is started. Its capabilities are reflected in the margin $\mu_i$. The competitive environment is reflected in its net expected margin with respect to the $\mu_{eff,t}$ of the system.

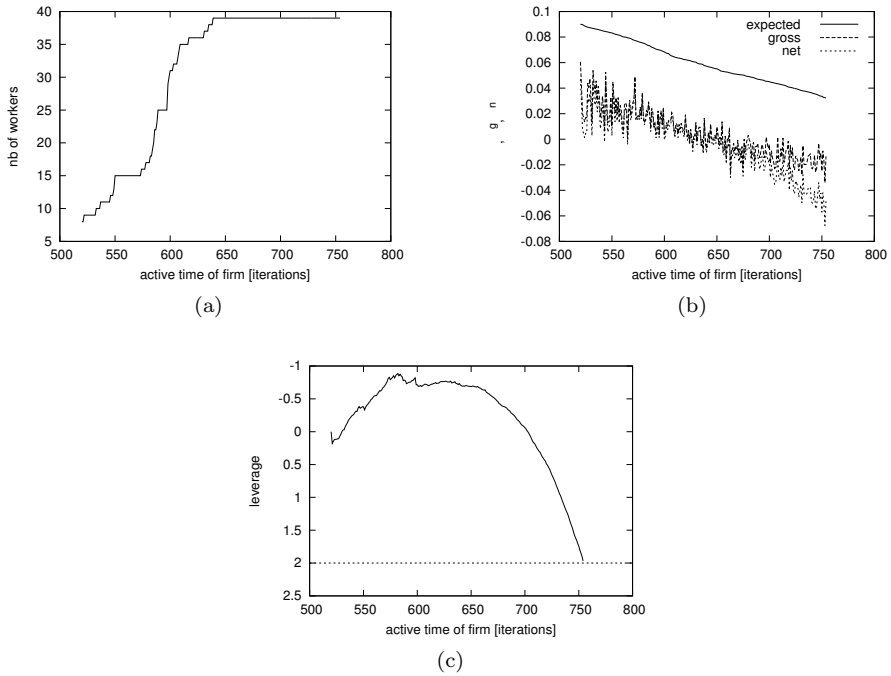



Figure 10: (a) size evolution (b) margin evolution (c) debt evolution of a firm that is active over 160 iterations. The dotted line represents $\gamma_{crit}$, $r = 0.012$. During this period, the system is has unemployment. Despite competition in the goods market, firms grow on average, and negative returns are not enough to shrink for these firms.

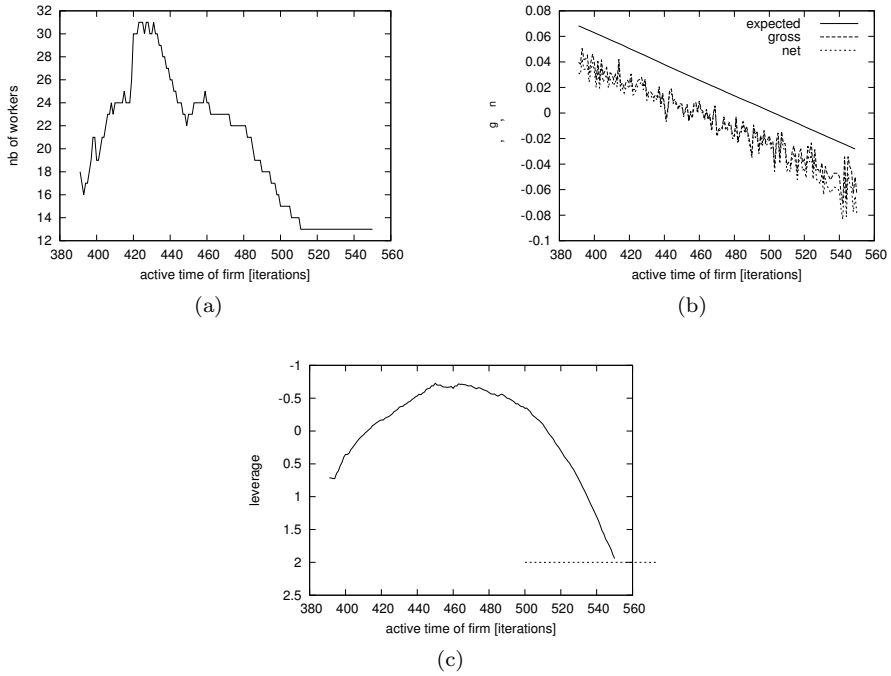



Figure 11: (a) size evolution (b) margin evolution (c) debt evolution of a firm that is active over 235 iterations. The dotted line represents $\gamma_{crit}$, $r = 0.005$. Firms compete both for purchasing power and workforce. Due to the low $r$, the firm has time to shrink before reaching $\gamma_{crit}$. From the same parameters as figure17.

**Qualitative results.** We present two typical examples of firms' life cycles for different parameter settings in figures 10 and 11. As expected from the age distribution, a wide range of trajectories is found. Many of the firms fail while still small and young, depending on the random constraints, their initial debt ratio itself, and their margin $\mu_i$. Profitable firms manage to survive and repay initial loans, but eventually they will also become indebted and fail because their profit margin has become too low. The correlation among variables depends on parameters:

- If the interest rate is very high: firms go bankrupt when they are still big. The lower the interest rate, the more time firms have to shrink, and then go bankrupt once they are already smaller.
- In the examples in figures 10 and 11 firms start with low debt, and it is mainly shortly before bankruptcy that they become indebted in a self-reinforcing way. If firms start with high debt, their net realized margin may be lowered significantly.
- If recentering is used: Firms start with debt, then they diminish them, and only when their margin has lowered due to ageing, they start spending their assets, and their debts increase again.

These results may be used to adjust parameters of the model according to empirical observations. For instance, a possible comparison would be to the empirically observed fraction of the profits needed for interest payments, dividends or other sources of capital. Empirical studies may also help to find parameters that weight the reasons why a firm becomes unprofitable, i.e. how strong the ageing of technology should be with respect to a lower net margin due to debts. This type of analysis also raises the question of whether the model captures dominant patterns or whether it needs more heterogeneity, e.g. by introducing adaptive interest rates.

# 5 Results IV – Further distributions

Distributions of the debt and margins of active firms (in addition to size and age), as well as correlations between variables, open further possibilities to control the dynamics of the model and check its validity. These results are in turn strongly influenced by the life cycles, since the firms of all ages contribute to these statistics, although the variables for individual firms are evolving. In the following we present observations, which will be further compared to data in ongoing work.

**Leverage distribution.** A distribution of debt, which in this simple model is the same as liabilities normalized by the size of a firm, of active firms is shown in figure 12. It qualitatively reproduced the fat-tailed distributions found in [40] and [31]. Generally, firms with a high debt ratio are overrepresented, because in the current model firms do not have any strategy to lower their debt. They can only do so if their parameters and the constraints allow it. An introduction of variable interest rates could also change this. A scatterplot debt vs. size of

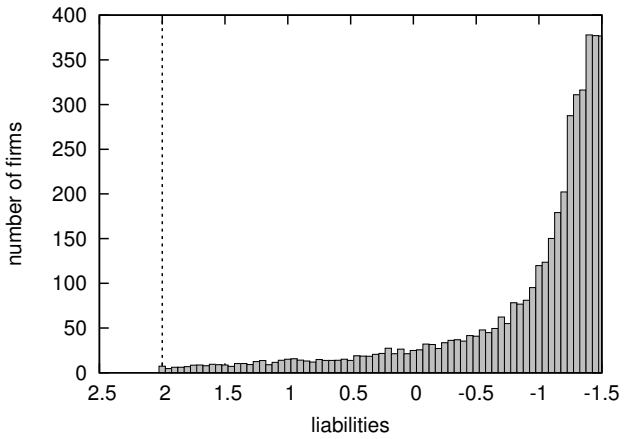
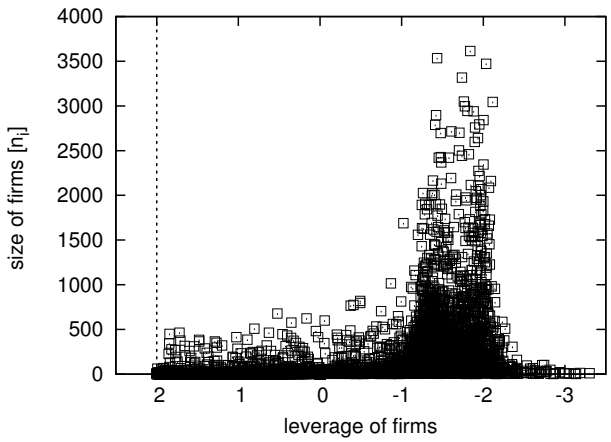



Figure 12: $10^6$ workers, $1.2 \cdot 10^4$ firms, $r = 0.005$, iteration = 1000. (Negative debt correponds to positive equity).

firms is presented in figure 13. The debt vs. the age of firms is fat-tailed in our results. [13] and [21] find in their database on Italian firms that younger firms are more liquidity-constrained, which is also the case of these results.

**Margin distribution.** The distribution shown in figure 14 has, roughly approximated, Gaussian shape. It exhibits asymmentry: no firm has a very high net realized margin, and the distribution decays much slower for lower margins. Interestingly, both the approximate Gaussian shape, as well as the same asymmetrical shape, have been found empirically by [32] for Italian firms. Other empirical studies [20] have found a distribution which is in between Gaussian and Laplacian for Icelandic firms, and [2] for US firms a distribution which is close to a Laplacian distribution.




(a)

Figure 13: scatterplot of debt vs size (snapshot at iteration 2000, same parameters as simulation in figure 17).

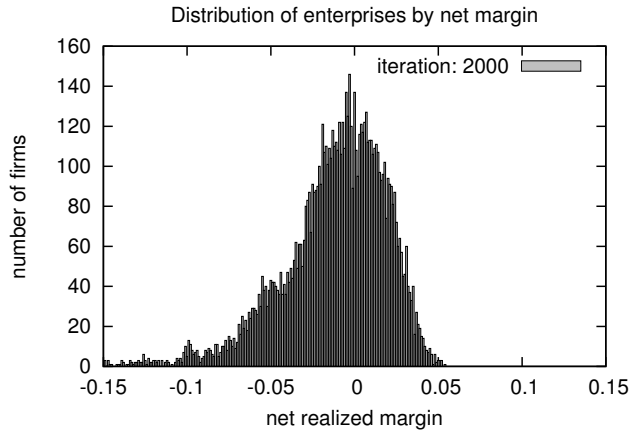



Figure 14: Distribution of the net realized margin. same parameters as simulation in figure 17, iteration = 2000.

**The relation between growth and size of firms.** The effect of liquidity constraints has been studied by [21] for Italian firm data (which also exhibits other distributions that this model reproduces qualitatively, like size and growth rate distribution). Besides stating that liquidity constraints have a negative effect on growth, they also find that that smaller firms are less affected than bigger firms by liquidity constraints. This finding is also a result from our simulation: Since small firms have a much bigger growth rate variance in the presence of high interest payments many will go bankrupt quickly. Then these firms are removed from the statistics, rather than entering the statistics with negative growth rates. That is, a high interest rate 'clears up' among small firms. Large firms fluctuate much less and typically incur debt for an extended period

before failure. Since interest payments lower their net margin and therefore their growth rate, they exhibit lower growth rates during that period.

**The relation between growth and age of firms.** [21] find that on average, younger firms grow more, which is also an outcome of our model (see figure 15). It can be obtained with and without recentering of the margin, i.e. the subtraction of $\Delta_t$ from the margins after each iteration (see equation (16)). If this is done, old firms whose margin has been recentered many times cannot range among the highest margins, whereas new firms are initialized with a margin $\in [\mu_{min}, \mu_{max}]$, typically $[0, 0.1]$. This important fact is also the reason why only very few firms become very old. Also the study by [16] find that the age dependence of growth cannot solely be explained by technological obsolence and selection, but there must be a contribution of liquidity constraints. This is the case in our model, too, since interest payments (which represent a larger sum for indebted businesses) are subtracted from the gross profit, and firms base their production decision on the net profit. The model by [16] proposes that the equity of firms affects firms' financing decision, and firms' finance patterns depend on their size. This approach, which shares a number of results with ours, is an interesting yet complex alternative: in our model, interest payments are an abstraction of the entire capital cost. The negative relationship between age and growth has also been found in a study by [30] on Irish firms.

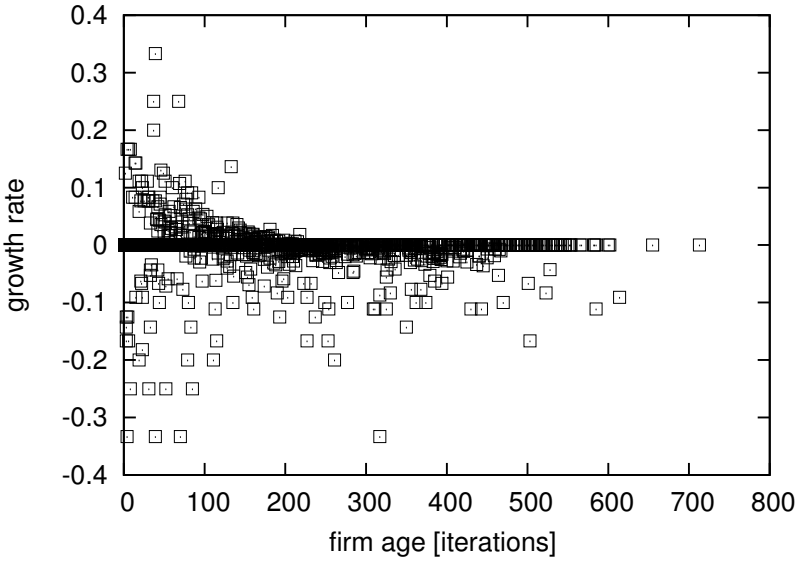



Figure 15: Scatterplot of age vs. growth of firms. Although growth is stochastic, it is visible that firms with very high growth rates are concentrated at young ages.

# 6 Discussion and Conclusion

We have presented the role of financial constraints in a macroeconomic agent-based model. It has been studied with focus on distributions of active and bankrupt firms, as well as on life cycles that firms undergo. The relation between these distributions is described theoretically, and verified with numerical examples for different parameter settings. A number of empirical studies have been used for comparison and have shown that the model, despite its simplicity, reproduces qualitatively a considerable number of empirical features. Nevertheless, certain limitations have been revealed. Clearly, comparison of further features to data is needed. A project is to validate it with a database, and possibly extend it.

A possible critique to our approach may be that firms do not have sufficient strategic choice. This could be increased, for instance if firms could modify their profit margin. However, stochastic growth is compatible with the interpretation of deliberate decisions made by firms: The model does not describe *why* a firm acts in a particular way, but merely the outcome of their choice.

# References


[1] Lada Adamic and Bernardo Huberman. The nature of markets in the world wide web. 1999.

[2] Simone Alfarano, Mishael Milaković, Albrecht Irle, and Jonas Kauschke. A statistical equilibrium model of competitive firms. *Journal of Economic Dynamics and Control*, 36(1):136–149, 2012.

[3] Edward I Altman. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The journal of finance*, 23(4):589–609, 1968.

[4] Hideaki Aoyama, Wataru Souma, Yuichi Nagahara, Mitsuhiro P Okazaki, Hideki Takayasu, and Misako Takayasu. Pareto's law for income of individuals and debt of bankrupt companies. *Fractals*, 8(03):293–300, 2000.

[5] Amir F Atiya. Bankruptcy prediction for credit risk using neural networks: A survey and new results. *Neural Networks, IEEE Transactions on*, 12(4):929–935, 2001.

[6] Robert L Axtell. Zipf distribution of us firm sizes. *Science*, 293(5536), 2001.

[7] Agustí Segarra Blasco and Mercedes Teruel. *Dinámica empresarial, creación de empleo y productividad en las manufacturas españolas*. Ministerio de Industria, Turismo y Comercio, Centro de Publicaciones, 2008.

[8] Giulio Bottazzi and Angelo Secchi. Explaining the distribution of firm growth rates. *The RAND Journal of Economics*, 37(2):235–256, 2006.

[9] Giulio Bottazzi, Angelo Secchi, and Federico Tamagni. Productivity, profitability and financial performance. *Industrial and Corporate Change*, 17(4):711–751, 2008.

[10] Charlotte Bruun. Rediscovering the economics of keynes in an agent-based computational setting. Conferencepaper, 2008.

[11] Luis MB Cabral and Jose Mata. On the evolution of the firm size distribution: Facts and theory. *American Economic Review*, pages 1075–1090, 2003.

[12] Silvano Cincotti, Marco Raberto, and Andrea Teglio. Credit money and macroeconomic instability in the agent-based model and simulator eurace. *Economics: The Open-Access, Open-Assessment E-Journal*, 4, 2010.

[13] Alex Coad. Investigating the exponential age distribution of firms. *Economics: The Open-Access, Open-Assessment E-Journal*, 4, 2010.

[14] Alex Coad and Jagannadha Pawan Tamvada. The growth and decline of small firms in developing countries. Technical report, Papers on economics and evolution, 2008.

[15] William Cook and Paul Ormerod. Power law distribution of the frequency of demises of us firms. *Physica A: Statistical Mechanics and its Applications*, 324(1):207–212, 2003.

[16] Thomas F Cooley and Vincenzo Quadrini. Financial markets and firm dynamics. *American Economic Review*, pages 1286–1310, 2001.

[17] AI Dimitras, R Slowinski, R Susmaga, and C Zopounidis. Business failure prediction using rough sets. *European Journal of Operational Research*, 114(2):263–280, 1999.

[18] Giovanni Dosi, Giorgio Fagiolo, and Andrea Roventini. Schumpeter meeting keynes: A policy-friendly model of endogenous growth and business cycles. *Journal of Economic Dynamics and Control*, 34(9):1748–1767, 2010.

[19] Joshua M Epstein. Agent-based computational models and generative social science. *Generative Social Science: Studies in Agent-Based Computational Modeling*, pages 4–46, 1999.

[20] Einar Jón Erlingsson, Simone Alfarano, Marco Raberto, and Hlynur Stefánsson. On the distributional properties of size, profit and growth of icelandic firms. *Journal of Economic Interaction and Coordination*, pages 1–18, 2013.

[21] Giorgio Fagiolo and Alessandra Luzzi. Do liquidity constraints matter in explaining firm size and growth? some evidence from the italian manufacturing industry. *Industrial and Corporate Change*, 15(1):1–39, 2006.

[22] Dongfeng Fu, Fabio Pammolli, Sergey V Buldyrev, Massimo Riccaboni, Kaushik Matia, Kazuko Yamasaki, and H Eugene Stanley. The growth of business firms: Theoretical framework and empirical evidence. *Proceedings of the National Academy of Sciences of the United States of America*, 102(52):18801–18806, 2005.

[23] Yoshi Fujiwara. Zipf law in firms bankruptcy. *Physica A: Statistical Mechanics and its Applications*, 337(1):219–230, 2004.

[24] Mauro Gallegati, Steve Keen, Thomas Lux, and Paul Ormerod. Worrying trends in econophysics. *Physica A: Statistical Mechanics and its Applications*, 370(1):1–6, 2006.

[25] Domenico Delli Gatti. *Emergent macroeconomics: an agent-based approach to business fluctuations*. Springer, 2008.

[26] Domenico Delli Gatti. *Macroeconomics from the Bottom-up*, volume 1. Springer, 2011.

[27] Robert Gibrat. *Les inégalités économiques*. Recueil Sirey, 1931.

[28] Corrado Di Guilmi, Mauro Gallegati, and Paul Ormerod. Scaling invariant distributions of firms' exit in oecd countries. *Physica A: Statistical Mechanics and its Applications*, 334(1):267–273, 2004.

[29] Byoung Hee Hong, Kyoung Eun Lee, and Jae Woo Lee. Power law in firms bankruptcy. *Physics Letters A*, 361(1):6–8, 2007.

[30] Stephen Kinsella. The age distribution of firms in ireland, 1961–2009. *Department of Economics, University of Limerick, Ireland*, 2009.

[31] Marco Lamieri. Debt maturity structure of italian firms. *Available at SSRN 1373752*, 2009.

[32] Marco Lamieri and Ilaria Sangalli. Economia e finanza dei distretti industriali. Report, sanpaolo intesa, 2011.

[33] Feng Yu Lin and Sally McClean. A data mining approach to the prediction of corporate failure. *Knowledge-Based Systems*, 14(3):189–195, 2001.

[34] Matteo Marsili and Yi-Cheng Zhang. Interacting individuals leading to zipf's law. *Physical Review Letters*, 80(12):2741–2744, 1998.

[35] Cornelia Metzig and Mirta Gordon. Heterogeneous firms in a macroeconomic agent-based model. *arXiv preprint arXiv:1211.5575*, 2012.

[36] Cornelia Metzig and Mirta B Gordon. A model for scaling in firms' size and growth rate distribution. *Physica A*, 2014.

[37] journal=in preparation year=2014 Metzig, Cornelia. forthcoming.

[38] Kazumi Okuyama, Misako Takayasu, and Hideki Takayasu. Zipf's law in income distribution of companies. *Physica A: Statistical Mechanics and its Applications*, 269(1):125–131, 1999.

[39] S Picoli Jr, RS Mendes, LC Malacarne, and EK Lenzi. Scaling behavior in the dynamics of citations to scientific journals. *EPL (Europhysics Letters)*, 75(4):673, 2006.

[40] Boris Podobnik, Davor Horvatic, Alexander M Petersen, Branko Urošević, and H Eugene Stanley. Bankruptcy risk model and empirical tests. *Proceedings of the National Academy of Sciences*, 107(43):18325–18330, 2010.

[41] Pascal Seppecher. Un modèle macroéconomique multi-agents avec monnaie endogène. 2009.

[42] Kyung-Shik Shin, Taik Soo Lee, and Hyun-jung Kim. An application of support vector machines in bankruptcy prediction model. *Expert Systems with Applications*, 28(1):127–135, 2005.

[43] Herbert A Simon. On a class of skew distribution functions. *Biometrika*, 42(3/4):425–440, 1955.

[44] Michael HR Stanley, Luis AN Amaral, Sergey V Buldyrev, Shlomo Havlin, Heiko Leschhorn, Philipp Maass, Michael A Salinger, and H Eugene Stanley. Scaling behaviour in the growth of companies. *Nature*, 379(6568):804–806, 1996.

[45] John Sutton. Gibrat's legacy. *Journal of economic Literature*, 35(1):40–59, 1997.

[46] John Sutton. The variance of firm growth rates: the 'scaling'puzzle. *Physica a: statistical mechanics and its applications*, 312(3):577–590, 2002.

[47] Hideki Takayasu, Aki-Hiro Sato, and Misako Takayasu. Stable infinite variance fluctuations in randomly amplified langevin systems. *Physical Review Letters*, 79(6):966–969, 1997.

[48] Stewart Thornhill and Raphael Amit. Learning about failure: bankruptcy, firm age, and the resource-based view. *Organization Science*, 14(5):497–509, 2003.

[49] Damián H Zanette and Susanna C Manrubia. Role of intermittency in urban development: a model of large-scale city formation. *Physical Review Letters*, 79(3):523–526, 1997.

# Appendices

## A Size and growth rate for the two bankruptcy criteria – images

In figure 16, size and growth rate distributions of simulations are shown, (a) and (c) corresponding to bankruptcy criterion 1 ($\mu_i = \mu_{crit} = 0$), (b) and (d) to criterion 2 ($\gamma_i = \gamma_{crit}$).The results of the simple setting [36] are widely conserved: they exhibit a tent-shaped aggregate growth rate distribution, a fat-tailed size distribution which decays slighlty faster than power law. For criterion 1, where bankruptcy is more deterministic, a sharp bend in the size distribution is visible. The fact that the decay is faster than power law is related to the fact that the system exhibits (for both bankruptcy criteria) in addition a size-dependent growth rate variance. Whereas in the simple setting, the width of a (Gaussian) growth rate distribution was characterized by $\sigma n \propto n^{-\beta}$ with $\beta = 0.5$, in current work shows that the scaling exponent in these simulations is lower (numerically $\beta \approx 0.3$). A lower $\beta$ is the result of the fact that margin heterogeneity widens the growth rate variance fo a certain size class, and this effect is comparatively stronger for larger firms. As a result, the curve of the the cumulative size distribution is less concave as the simple setting, and therefore closer to empirical results [6][38].

## B Age and size distribution for active and bankrupt firms in a simulation with low interest rates

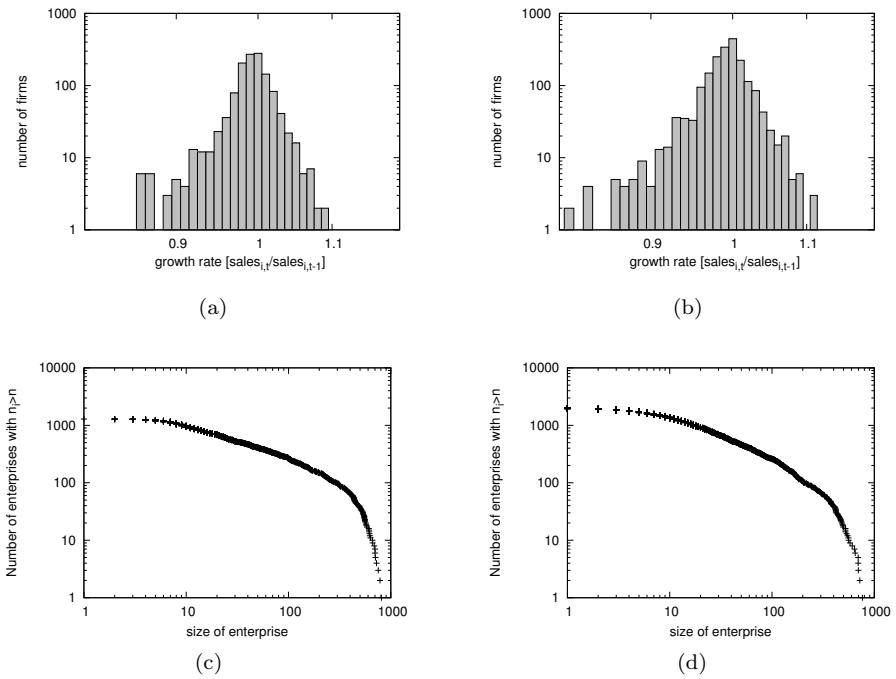



Figure 16: Comparison of size and growth rate distribution for bankruptcy criterion 1 (a) and (c) and 2 (b) and (d). Age distribution needs to be studied to understand the difference which is not obvious from these distributions. In (d), the age distribution decays slower, which is why some firms have time to get bigger, and the cutoff of sizes is less pronounced. In a setting without margin recentering in [35] in figure 3, the effect is absent. $10^5$ workers, $10^3$ firms, $r = 0.011$.

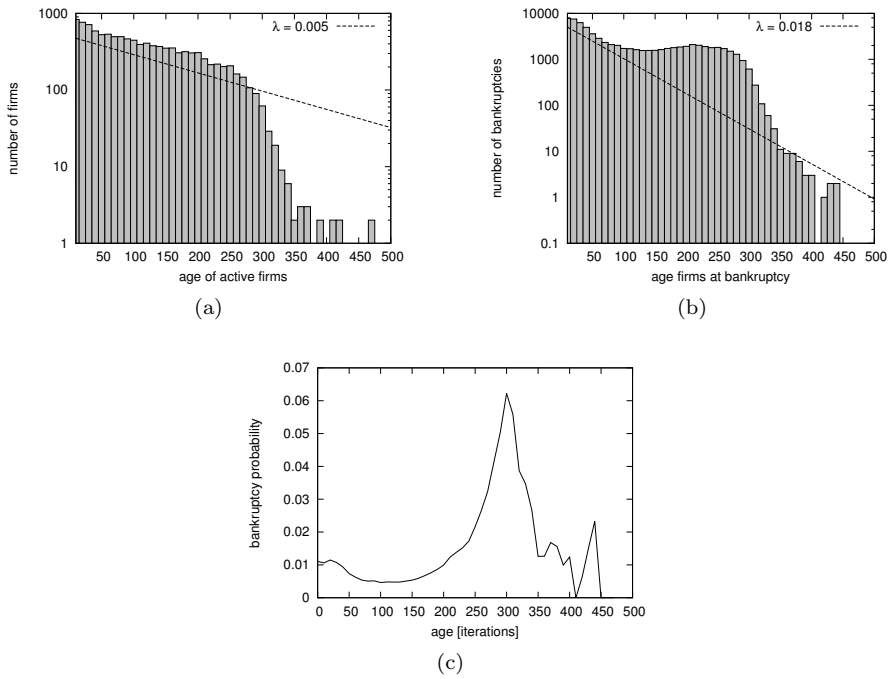



Figure 17: (a) age distribution $p(a)$ (snapshot at iteration 2000). (b) age distribution $p(a|b)$ of bankrupt firms, averaged over 1300 iterations. (c) bankruptcy probability as a function of age $p(b|a)$, averaged over 1300 iterations. $10^6$ workers, $8 \cdot 10^3$ firms, interest $r = 0.005$. The same plots for a higher interest rate are shown in figure 9. Here, margin recentering is dominant with respect to interest payments in lowering the margin. Therefore, the peak of bankruptcy probability at the maximal life time is very pronounced.

## 3.2 Concluding Remarks and Outlook

In this chapter, which is thematically the continuation of chapter 2, bankruptcies have been studied. The model is the same as introduced in the two previous chapters. A situation where firms do not spend their profits has been analyzed in this chapter. The comparison with data suggests some future improvements:

- The recentering should be changed because it is the reason why no firm is active longer than a certain maximum age. The maximum age is the one at which firm which starts without debt and a very high margin can exist, before finally going bankrupt. Next, size distribution is expected to decay more linearly than e.g. in figure 8 (b) of the working paper. More specifically, the comparison to a system where firms merely age due to debt accumulation, but without a deterministic margin recentering, shows that this leads to a size distribution with slower decay (see figure 2.3) and figure 3 in publication 1). A possible rationale could be that firms can modify their margin via investment into physical capital, as e.g. [45, 21] do.

- If firms $\mu_i$ can evolve stochastically, the use of bankruptcy criterion 1 of the working paper ($\mu_i = 0$) would yield more realistic results. The criterion was proposed because it corresponds to one of the bank's criteria reported in studies [4], and because it is unplausible for firms to get financed when they do not have a theoretical chance to earn profits. However, with the method of recentering, firms' lifetimes are almost linearly dependent on their initial $\mu_i$, which is the reason why the age distribution obtained when using criterion 1 is not heavy-tailed at all (see figure 5 in the working paper). This is far from the reported empirical evidence.

- One hypothesis, which is difficult to justify, is the credit policy of the bank: each firm is given credit, independently of their margin and of their leverage. There are neither different interest rates, nor a limited volume of credit for less trustworthy firms. Simulations have shown that this has an unplausible effect: the larger the volume of debt in the system (i.e. the larger $\gamma$), the higher unemployment is, since unprofitable firms continue to be active. Simulations for different interest rates $r$ showed that the higher $r$ is, the higher unemployment is, which seems plausible. As already stated in section 2.5, bank rules should be found such that these two parameters yield results that reproduce qualitatively empirical data.

Even if these three modifications are taken into account, the fact that there is no explicit capital good in the model is a strong limitation for some purposes. In particular, it makes it difficult to compare the results to data where, for instance, the leverage is defined as liabilities/assets. In this model, the assets are merely positive money holdings. Currently, the model is being tested on databases, which will suggest further ideas for future work. Since many variables can be compared to the same dataset, it is a much more thorough validation than the comparison of the separate results to various empirical studies. The results

in the working paper are only qualitative, and even if some empirically found distributions are reproduced, the rationale should be further tested, to prevent false conclusions being drawn. One contribution of this work is however the method of validation of an agent-based model, which differs from the agent-based macroeconomic models described in section 2.4.

# Chapter 4

# Fat-tailed distributions in other Social Systems

## 4.1 Introduction

Fat tailed distributions exist in various complex systems beyond economics. In this section, results are shown from a project where a stochastic growth model was designed to describe power law formation in large web taxonomies, with a future application to machine learning purposes. Besides the application also the theoretical model differs from the previous chapters. It is based on a twofold preferential attachment growth, since a constantly growing system is described. The common concept with the economic model of the previous chapters is that empirical evidence shows again several jointly arising fat-tailed distributions, which can be described in a joint model in which they are intrinsically linked.

**Formation of fat-tailed distributions in web databases.** Large web taxonomies are databases of websites, which are organized in classes, according to their content. The classification of websites into these classes is done by a large number ($10^5$) of voluntary referees. This taxonomy is hierarchical. For example, the classes may correspond to different topics like sports and music, and the subclasses may be football, tennis, skiing, pop music, classical music, which may be again subdivided on the following level. The arising size distribution of subclasses (measures in the number of websites it contains) is fat tailed, in every class and for the total system. It can be very approximately fitted by a power law decay. Within a subclass at the lowest level of the hierarchy, the formation of this power law can be explained from the way the database is generated. Since the database is constantly growing, both in the number of classes and in the number of websites, a candidate for the explanation of class size distribution is the model by Yule [108, 73].

The fact that a power law is also found for the sizes of categories, as well as for the number of children categories, has been explained by a catenated preferetial

attachment process, which is in line with the instructions for the voluntary referees how to update the taxonomy. Other fat tailed distributions result from the formation process, such as for the number of features per category, and the categoriy size distribution at a given level of the hierarchy[43, 37].

# Publication 3

ACM SIGKDD Explorations Newsletter - Special issue on big data, Volume 16
Issue 1, June 2014, Pages 47-56

# On Power Law Distributions in Large-scale Taxonomies

Rohit Babbar

Cornelia Metzig, Ioannis Partalas
University of Grenoble Alpes, LIG
F-38000 Grenoble, France
firstname.lastname@imag.fr

Eric Gaussier, Massih-Reza Amini

## ABSTRACT

In many of the large-scale physical and social complex systems phenomena fat-tailed distributions occur, for which different generating mechanisms have been proposed. In this paper, we study models of generating power law distributions in the evolution of large-scale taxonomies such as Open Directory Project, which consist of websites assigned to one of tens of thousands of categories. The categories in such taxonomies are arranged in tree or DAG structured configurations having parent-child relations among them. We first quantitatively analyse the formation process of such taxonomies, which leads to power law distribution as the stationary distributions. In the context of designing classifiers for large-scale taxonomies, which automatically assign unseen documents to leaf-level categories, we highlight how the fat-tailed nature of these distributions can be leveraged to analytically study the space complexity of such classifiers. Empirical evaluation of the space complexity on publicly available datasets demonstrates the applicability of our approach.

## 1. INTRODUCTION

With the tremendous growth of data on the web from various sources such as social networks, online business services and news networks, structuring the data into conceptual taxonomies leads to better scalability, interpretability and visualization. Yahoo! directory, the open directory project (ODP) and Wikipedia are prominent examples of such web-scale taxonomies. The Medical Subject Heading hierarchy of the National Library of Medicine is another instance of a large-scale taxonomy in the domain of life sciences. These taxonomies consist of classes arranged in a hierarchical structure with parent-child relations among them and can be in the form of a rooted tree or a directed acyclic graph. ODP for instance, which is in the form of a rooted tree, lists over 5 million websites distributed among close to 1 million categories and is maintained by close to 100,000 human editors. Wikipedia, on the other hand, represents a more complicated directed graph taxonomy structure consisting of over a million categories. In this context, large-scale hierarchical classification deals with the task of automatically assigning labels to unseen documents from a set of target classes which are represented by the leaf level nodes in the hierarchy.

In this work, we study the distribution of data and the hierarchy tree in large-scale taxonomies with the goal of modelling the process of their evolution. This is undertaken by a quantitative study of the evolution of large-scale taxonomy using models of preferential attachment, based on the famous model proposed by Yule [33] and showing that throughout the growth process, the taxonomy exhibits a fat-tailed distribution. We apply this reasoning to both category sizes and tree connectivity in a simple joint model. Formally, a random variable $X$ is defined to follow a power law distribution if for some positive constant $a$, the complementary cumulative distribution is given as follows:

$$P(X > x) \propto x^{-a}$$

Power law distributions, or more generally fat-tailed distributions that decay slower than Gaussians, are found in a wide variety of physical and social complex systems, ranging from city population, distribution of wealth to citations of scientific articles [23]. It is also found in network connectivity, where the internet and Wikipedia are prominent examples [27; 7]. Our analysis in the context of large-scale web-taxonomies leads to a better understanding of such large-scale data, and also leveraged in order to present a concrete analysis of space complexity for hierarchical classification schemes. Due to the ever increasing scale of training data size in terms of the number of documents, feature set size and number of target classes, the space complexity of the trained classifiers plays a crucial role in the applicability of classification systems in many applications of practical importance.

The space complexity analysis presented in this paper provides an analytical comparison of the trained model for hierarchical and flat classification, which can be used to select the appropriate model a-priori for the classification problem at hand, without actually having to train any models. Exploiting the power law nature of taxonomies to study the training time complexity for hierarchical Support Vector Machines has been performed in [32; 19]. The authors therein justify the power law assumption only *empirically*, unlike our analysis in Section 3 wherein we describe the generative process of large-scale web taxonomies more concretely, in the context of similar processes studied in other models. Despite the important insights of [32; 19], space complexity has not been treated formally so far.

The remainder of this paper is as follows. Related work on reporting power law distributions and on large scale hierarchical classification is presented in Section 2. In Section 3, we recall important growth models and quantitatively justify the formation of power laws as they are found

in hierarchical large-scale web taxonomies by studying the evolution dynamics that generate them. More specifically, we present a process that jointly models the growth in the size of categories, as well as the growth of the hierarchical tree structure. We derive from this growth model why the class size distribution at a given level of the hierarchy also exhibits power law decay. Building on this, we then appeal to Heaps' law in Section 4, to explain the distribution of features among categories which is then exploited in Section 5 for analysing the space complexity for hierarchical classification schemes. The analysis is empirically validated on publicly available DMOZ datasets from the Large Scale Hierarchical Text Classification Challenge(LSHTC)[1] and patent data (IPC) [2] from World Intellectual Property Organization. Finally, Section 6 concludes this work.

## 2. RELATED WORK

Power law distributions are reported in a wide variety of physical and social complex systems [22], such as in internet topologies. For instance [11; 7] showed that internet topologies exhibit power laws with respect to the in-degree of the nodes. Also the size distribution of website categories, measured in terms of number of websites, exhibits a fat-tailed distribution, as empirically demonstrated in [32; 19] for the Open Directory Project (ODP). Various models have been proposed for the generation power law distributions, a phenomenon that may be seen as fundamental in complex systems as the normal distribution in statistics [25]. However, in contrast to the straight-forward derivation of normal distribution via the central limit theorem, models explaining power law formation all rely on an approximation. Some explanations are based on multiplicative noise or on the renormalization group formalism [28; 30; 16]. For the growth process of large-scale taxonomies, models based on preferential attachment are most appropriate, which are used in this paper. These models are based on the seminal model by Yule [33], originally formulated for the taxonomy of biological species, detailed in section 3. It applies to systems where elements of the system are grouped into classes, and the system grows both in the number of classes, and in the total number of elements (which are here documents or websites). In its original form, Yule's model serves as explanation for power law formation in any taxonomy, irrespective of an eventual hierarchy among categories. Similar dynamics have been applied to explain scaling in the connectivity of a network, which grows in terms of nodes and edges via preferential attachment [2]. Recent further generalizations apply the same growth process to trees [17; 14; 29]. In this paper, describe the approximate power-law in the child-to-parent category relations by the model by Klemm et al. [17]. Furthermore, we combine this formation process in a simple manner with the original Yule model in order to explain also a power law in category sizes, i.e. we provide a comprehensive explanation for the formation process of large-scale web taxonomies such as DMOZ. From the second, we infer a third scaling distribution for the number of features per category. This is done via the empirical Heaps's law [10], which describes the scaling relationship between text length and the size of its vocabulary.

Some of the earlier works on exploiting hierarchy among target classes for the purpose of text classification have been studied in [18; 6] and [8] wherein the number of target classes were limited to a few hundreds. However, the work by [19] is among the pioneering studies in hierarchical classification towards addressing web-scale directories such as Yahoo! directory consisting of over 100,000 target classes. The authors analyse the performance with respect to accuracy and training time complexity for flat and hierarchical classification. More recently, other techniques for large-scale hierarchical text classification have been proposed. Prevention of error propagation by applying *Refined Experts* trained on a validation set was proposed in [4]. In this approach, bottom-up information propagation is performed by utilizing the output of the lower level classifiers in order to improve classification at top level. The deep classification method proposed in [31] first applies hierarchy pruning to identify a much smaller subset of target classes. Prediction of a test instance is then performed by re-training Naive Bayes classifier on the subset of target classes identified from the first step. More recently, Bayesian modelling of large-scale hierarchical classification has been proposed in [15] in which hierarchical dependencies between the parent-child nodes are modelled by centring the prior of the child node at the parameter values of its parent.

In addition to prediction accuracy, other metrics of performance such as prediction and training speed as well as space complexity of the model have become increasingly important. This is especially true in the context of challenges posed by problems in the space of Big Data, wherein an optimal trade-off among such metrics is desired. The significance of prediction speed in such scenarios has been highlighted in recent studies such as [3; 13; 24; 5]. The prediction speed is directly related to space complexity of the trained model, as it may not be possible to load a large trained model in the main memory due to sheer size. Despite its direct impact on prediction speed, no earlier work has focused on space complexity of hierarchical classifiers.

Additionally, while the existence of power law distributions has been used for analysis purposes in [32; 19] no thorough justification is given on the existence of such phenomenon. Our analysis in Section 3, attempts to address this issue in a quantitative manner. Finally, power law semantics have been used for model selection and evaluation of large-scale hierarchical classification systems [1]. Unlike problems studied in classical machine learning sense which deal with a limited number of target classes, this application forms a blue-print on extracting hidden information in big data.

## 3. POWER LAW IN LARGE-SCALE WEB TAXONOMIES

We begin by introducing the complementary cumulative size distribution for category sizes. Let $N_i$ denote the size of category $i$ (in terms of number of documents), then the probability that $N_i > N$ is given by

$$P(N_i > N) \propto N^{-\beta} \tag{1}$$

where $\beta > 0$ denotes the exponent of the power law distribution.[3] Empirically, it can be assessed by plotting the rank of a category's size against its size (see Figure 1) The derivative of this distribution, the category size probability

[1] http://lshtc.iit.demokritos.gr/

[2] http://web2.wipo.int/ipcpub/

[3] To avoid confusion, we denote the power law exponents for in-degree distribution and feature size distribution $\gamma$ and $\delta$.

| Level | # Categories |
|---|---|
| 1 | 11 |
| 2 | 343 |
| 3 | 3670 |
| 4 | 13255 |
| 5 | 18169 |

Table 1: Number of categories at each level in the hierarchy of the LSHTC2-DMOZ database.

density $p(N_i)$, then also follows a power law with exponent $(\beta+1)$, i.e. $p(N_i) \propto N_i^{-(\beta+1)}$.
Two of our empirical findings are a power law for both the complementary cumulative category size distribution and the counter-cumulative in-degree distribution, shown in Figures 1 and 2, for LSHTC2-DMOZ dataset which is a subset of ODP. The dataset[4] contains 394,000 websites and 27,785 categories. The number of categories at each level of the hierarchy is shown in Table 1.

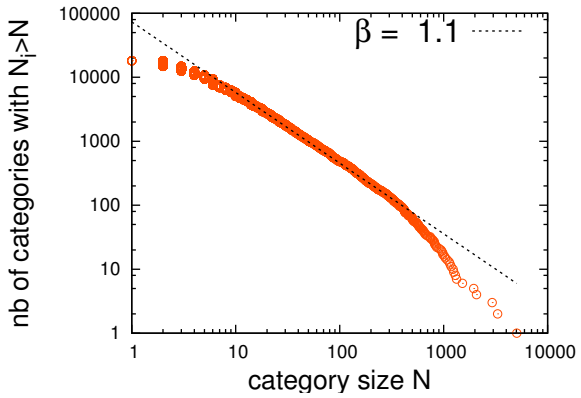



Figure 1: Category size vs rank distribution for the LSHTC2-DMOZ dataset.

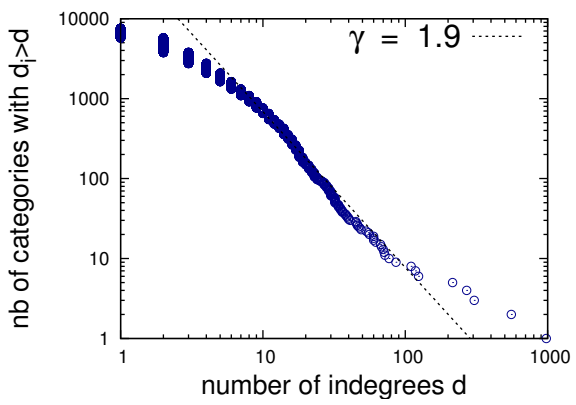



Figure 2: Indegree vs rank distribution for the LSHTC2-DMOZ dataset.

We explain the formation of these two laws via models by Yule [33] and a related model by Klemm [17], detailed in sections 3.1 and 3.2, which are then related in section 3.3.

[4]http://lshtc.iit.demokritos.gr/LSHTC2_datasets

| **Variables** | |
|---|---|
| $N_i$ | Number of elements in class $i$ |
| $d_i$ | Number of subclasses of class $i$ |
| $f_i$ | Number of features of class $i$ |
| $\kappa$ | Total number of classes |
| $D$ | Total number of in-degrees (=subcategories) |
| $p_{N,\kappa}$ | Fraction of classes having $N$ elements when the total number of classes is $\kappa$ |
| **Constants** | |
| $m$ | Number of elements added to the system after which a new class is added |
| $w$ | $\in]0,1]$ Probability that attachment of subcategories is preeferential |
| **Indices** | |
| $i$ | Index for the class |

Table 2: Summary of notation used in Section 3

## 3.1 Yule's model

Yule's model describes a system that grows in two quantities, in elements and in classes in which the elements are assigned. It assumes that for a system having $\kappa$ classes, the probability that a new element will be assigned to a certain class is proportional to its current size,

$$p(i) = \frac{N_i}{\sum_{i'=1}^{\kappa} N_{i'}} \tag{2}$$

It further assumes that for every $m$ elements that are added to the pre-existing classes in the system, a new class of size 1 is created[5].
The described system is constantly growing in terms of elements and classes, so strictly speaking, a stationary state does not exist [20]. However, a stationary distribution, the so-called Yule distribution, has been derived using the approach of the master equation with similar approximations by [26; 23; 17]. Here, we follow Newman [23], who considers as one time-step the duration between creation of two consecutive classes. From this follows that the average number of elements per class is always $m+1$, and the system contains $\kappa(m+1)$ elements at a moment where the number of classes is $\kappa$. Let $p_{N,\kappa}$ denote the fraction of classes having $N$ elements when the total number of classes is $\kappa$. Between two successive time instances, the probability for a given pre-existing class $i$ of size $N_i$ to gain a new element is $mN_i/(\kappa(m+1))$. Since there are $\kappa\, p_{N,\kappa}$ classes of size $N$, the expected number such classes which gain a new element (and grow to size $(N+1)$) is given by :

$$\frac{mN}{\kappa(m+1)}\,\kappa\, p_{N,\kappa} = \frac{m}{(m+1)}\,N\, p_{N,\kappa} \tag{3}$$

The number of classes with $N$ websites are thus fewer by the above quantity, but some which had $(N-1)$ websites prior to the addition of a new class have now one more website. This step depicting the change of the state of the system from $\kappa$ classes to $(\kappa+1)$ classes is shown in Figure 3. Therefore, the expected number of classes with $N$ documents when the

[5]The initial size may be generalized to other small sizes; for instance Tessone et al. consider entrant classes with size drawn from a truncated power law [29] .

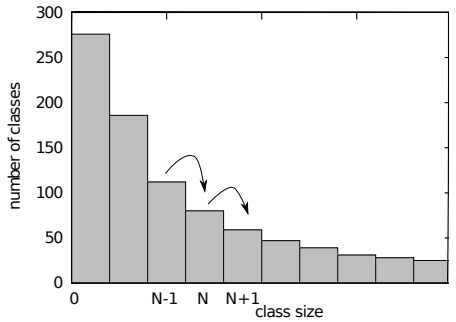



Figure 3: Illustration of equation 4. Individual classes grow constantly i.e., move to the right over time, as indicated by arrows. A stationary distribution means that the height of each bar remains constant.

number of classes is ($\kappa$+1) is given by the following equation:

$$(\kappa+1)p_{N,(\kappa+1)} = \kappa\, p_{N,\kappa} + \frac{m}{m+1}[(N-1)(p_{(N-1),\kappa}) \\ -Np_{N,\kappa}] \quad (4)$$

The first term in the right hand side of equation 4 corresponds to classes with $N$ documents when the number of classes is $\kappa$. The second term corresponds to the contribution from classes of size $(N-1)$ which have grown to size $N$, this is shown by the left arrow (pointing rightwards) in Figure 3. The last term corresponds to the decrease resulting from classes which have gained an element and have become of size $(N+1)$, this is shown by the right arrow (pointing rightwards) in Figure 3. The equation for the class of size 1 is given by:

$$(\kappa+1)p_{1,(\kappa+1)} = \kappa\, p_{1,\kappa} + 1 - \frac{m}{m+1}p_{1,\kappa} \quad (5)$$

As the number $\kappa$ of classes (and therefore the number of elements $\kappa(m+1)$) in the system increases, the probability that a new element is classified into a class of size $N$, given by Equation (3), is assumed to remain constant and independent of $\kappa$. Under this hypothesis, the stationary distribution for class sizes can be determined by solving equation (4) and using equation (5) as the initial condition. This is given by

$$p_N = (1+1/m)B(N, 2+1/m) \quad (6)$$

where $B(.,.)$ is the beta distribution. (6) has been termed *Yule distribution* [26]. Written for a continuous variable $N$, it has a power law tail:

$$p(N) \propto N^{-2-\frac{1}{m}}$$

From the above equation the exponent of the density function is between 2 and 3. Its cumulative size distribution $P(N_k > N)$, as given by equation (1), has an exponent given by

$$\beta = (1 + (1/m)) \quad (7)$$

which is between 1 and 2. The higher the frequency $1/m$ at which new classes are introduced, the bigger $\beta$ becomes, and the lower the average class size. This exponent is stable over time although the taxonomy is constantly growing.

## 3.2 Preferential attachment models for networks and trees

A similar model has been formulated for network growth by Barabási and Albert [2], which explains the formation of a power law distribution in connectivity degree of nodes. It assumes that the networks grow in terms of nodes and edges, and that every newly added node to the system connects with a fixed number of edges to existing nodes. Attachment is again preferential, i.e. the probability for a newly added node $i$ to connect to a certain existing node $j$ is proportional to its number of existing edges of node $j$.

A node in the Barabási-Albert (BA) model corresponds a class in Yule's model, and a new edge to two newly assigned element. Every added edge counts both to the degree of an existing node $j$, as well as to the newly added node $i$. For this reason the existing nodes $j$ and the newly added node $i$ grow always by the same number of edges, implying $m = 1$ and consequently $\beta = 2$ in the BA-model, *independently* of the number of edges that each new node creates.

The seminal BA-model has been extended in many ways. For hierarchical taxonomies, we use a preferential attachment model for trees by [17]. The authors considered growth via directed edges, and explain power law formation in the *in-degree*, i.e. the edges directed from children to parent in a tree structure. In contrast to the BA-model, newly added nodes and existing nodes do not increase their in-degree by the same amount, since new nodes start with an in-degree of 0. Leaf nodes thus cannot attract attachment of nodes, and preferential attachment alone cannot lead to a power-law. A small random term ensures that some nodes attach to existing ones independently of their degree, which is the analogon to the start of a new class in the Yule model. The probability $v$ that a new node attaches as a child to the existing node $i$ of with indegree $d_i$ becomes

$$v(i) = w\frac{d_i - 1}{D} + (1-w)\frac{1}{D}\,, \quad (8)$$

where $D$ is the size of the system measured in the total number of in-degrees. $w \in [0,1]$ denotes the probability that the attachment is preferential, $(1-w)$ the probability that it is random to any node, independently of their numbers of indegrees. As it has been done for the Yule process [26; 23; 14; 29], the stationary distribution is again derived via the master equation (4). The exponent of the asymptotic power law in the in-degree distribution is $\beta = 1 + 1/w$.This model is suitable to explain scaling properties of the tree or network structure of large-scale web taxonomies, which have also been analysed empirically, for instance for subcategories of Wikipedia [7]. It has also been applied to directory trees in [14].

## 3.3 Model for hierarchical web taxonomies

We now apply these models to large-scale web taxonomies like DMOZ. Empirically, we uncovered two scaling laws: (a) one for the size distribution of leaf categories and (b) one for the indegree (child-to-parent link) distribution of categories (shown in Figure 2). These two scaling laws are linked in a non-trivial manner: a category may be very small or even not contain any websites, but nevertheless be highly connected. Since on the other hand (a) and (b) arise jointly, we propose here a model generating the two scaling laws in a simple generic manner. We suggest a combination of the two processes detailed in subsections 3.1 and 3.2 to describe the growth process: websites are continuously added to the system, and classified into categories by human ref-

erees. At the same time, the categories are not a mere set, but form a tree structure, which grows itself in two quantities: in the number nodes (categories) and in the number of in-degrees of nodes (child-to-parent links, i.e. subcategory-to-category links). Based on the rules for voluntary referees of the DMOZ how to classify websites, we propose a simple combined description of the process. Altogether, the database grows in *three* quantities:

(i) *Growth in websites.* New websites are assigned into categories $i$, with probability $p(i) \propto N_i$ (Figure 4). This assignment happens independently of the hierarchy level of category. However, only leaf categories may receive documents.

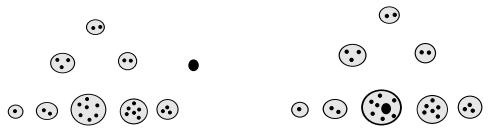


Figure 4: (i): A website is assigned to existing categories with $p(i) \propto N_i$.

(ii) *Growth in categories.* With probability $1/m$, the referees assign a website into a newly created category, at any level of the hierarchy (Figure 5).

This assumption would suffice to create a power law in the category size distribution, but since a tree-structure among categories exists, we also assume that the event of category creation is also attaching at particular places to the tree structure. The probability $v(i)$ that a category is created as the child of a certain parent category $i$ can depend in addition on the *in-degree* $d_i$ of that category (see equation 9).

Figure 5: (ii): Growth in categories is equivalent to growth of the tree structure in terms of in-degrees.

(iii) *Growth in children categories.* Finally, the hierarchy may also grow in terms of levels, since with a certain probability $(1 - w)$, new children categories are assigned independently of the number of children, i.e. its in-degree $d_i$ of the category $i$. (Figure 6). Like in [17], the attachment probability to a parent $i$ is

$$v(i) = w\frac{d_i - 1}{D} + (1 - w)\frac{\epsilon_i}{D}\,. \tag{9}$$

Equation (8), where $\epsilon_i = 1$, would suffice to explain power law in-degrees $d_i$ and in category sizes $N_i$.

To link the two processes more plausibly, it can be assumed that the second term in equation (9) denoting

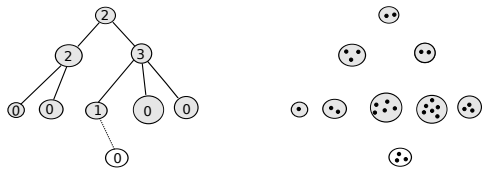



Figure 6: (iii): Growth in children categories.

assignment of new 'first children' depends on the size $N_i$ of parent categories,

$$\epsilon_i = \frac{N_i}{N}\,, \tag{10}$$

since this is closer to the rules by which the referees create new categories, but is not essential for the explanation of the power laws. It reflects that the bigger a leaf category, the higher the probability that referees create a child category when assigning a new website to it.

To summarize, the central idea of this joint model is to consider two measures for the size of a category: the number of its websites $N_i$ (which governs the preferential attachment of new websites), and its in-degree, i.e. the number of its children $d_i$, which governs the preferential attachment of new categories. To explain the power law in the category sizes, assumptions (i) and (ii) are the requirements. For the power law in the number of indegrees, assumptions (ii) and (iii) are the requirements. The empirically found exponents $\beta = 1.1$ and $\gamma = 1.9$ yield a frequency of new categories 1/m=0.1 and a frequency of new indegrees $(1 - w) = 0.9$.

## 3.4 Other interpretations

Instead of assuming in Equations (9) and (10) that referees decide to open a single child category, it is more realistic to assume that an existing category is *restructured*, i.e. one or several child categories are created, and websites are moved into these new categories such that the parent category contains less websites or even none at all. If one of the new children categories inherits all websites of the parent category (see Figure 7), the Yule model applies directly. If the websites are partitioned differently, the model contains effective shrinking of categories. This is not described by the Yule model, and the master Equation (4) considers only growing categories. However, it has been shown [29; 21] that models including shrinking categories also lead to the formation of power laws. Further generalizations compatible with power law formation are that new categories do not necessarily start with one document, and that the frequency of new categories does not need to be constant.

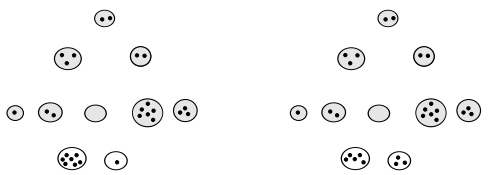


Figure 7: Model without and with shrinking categories. In the left figure, a child category inherits all the elements of its parent and takes its place in the size distribution.

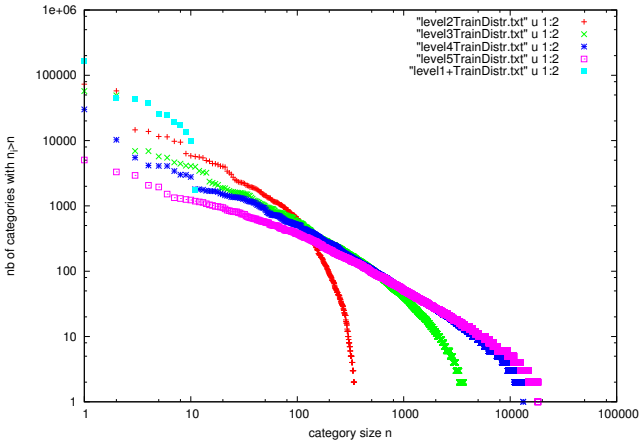



Figure 8: Category size distribution for each level of the LSHTC1-large dataset.

### 3.5 Limitations

However, Figures 1 and 2 do not exhibit perfect power law decay for several reasons. Firstly, the dataset is limited. Secondly, the hypothesis that assignment probability (2) depends uniquely on the size of a category might be too strong for web directories, neglecting the change in importance of topics. In reality, big categories can exist which receive only few new documents or none at all. Dorogovtsev and Mendes [9] have studied this problem by introducing an assignment probability that decays exponentially with age. For a low decay parameter they show that the stronger this decay, the steeper the power law; for strong decay, no power law forms. A last reason might be that referees re-structure categories in ways strongly deviating from the rules (i) - (iii).

### 3.6 Statistics per hierarchy level

The tree-structure of a database allows also to study the sizes of class belonging to a given level of the hierarchy. As shown in table 1, the DMOZ database contains 5 levels of different size. If only classes on a given level $l$ of the hierarchy are considered, we equally found a power law in category size distribution as shown in Figure 8. Per-level power law decay has also been found for the in-degree distribution. This result may equally be explained by the model introduced above: Equations (2), and (9) respectively, are valid also if instead of $p(k)$ one considers the conditional probability $p(l)p(i|l)$, where $p(l) = \frac{\sum_{i'=1,l}^{\kappa} N_{i',l}}{\sum_{i'=1}^{\kappa} N_{i'}}$ is the probability of assignment to a given level, and $p(i|l) = \frac{N_{i,l}}{\sum_{i'=1,l}^{\kappa} N_{i',l}}$ the probability of being assigned to a given class within that level. The formation process may be seen as a Yule process *within* a level if $\sum_{i'=1,l}^{\kappa} N_{i',l}$ is used for the normalization in equation (2), and this formation happens with probability $p(l)$ that a website gets assigned into level $l$. Thereby, the rate at $m_l$ at which new classes are created need not be the same for every level, and therefore the exponent of the power law fit may vary from level to level. Power law decay for the per-level class size distribution is a straightforward corollary of the described formation process, and will be used in Section 5 to analyse the space complexity of hierarchical classifiers.

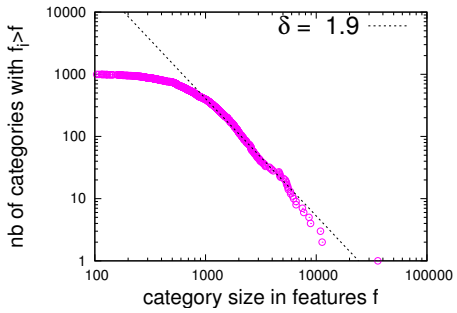



Figure 9: Number of features vs number of documents of each category.

## 4. RELATION BETWEEN CATEGORY SIZE AND NUMBER OF FEATURES

Having explained the formation of two scaling laws in the database, a third one has been found for the number of features $f_i$ in each category, $G(f)$ (see Figures 9 and 11). This is a consequence of the power law in category size distribution, shown (in Figure 1) in combination with another power law, termed Heaps' law [10]. This empirical law states that the number of distinct words $R$ in a document is related to the length $n$ of a document as follows

$$R(n) = Kn^{\alpha}\,, \tag{11}$$

where the empirical $\alpha$ is typically between 0.4 and 0.6. For the LSHTC2-DMOZ dataset, Figure 10 shows that for the collection of words and the collection of websites, similar exponents are found. (An interpretation of this result is that the total number words in a category can be measured approximately by the number of websites in a category, although not all websites have the same length.)

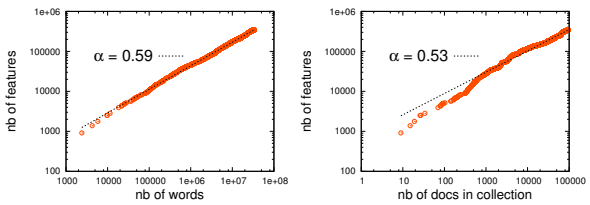



Figure 10: Heaps' law: number of distinct words vs. number of words, and vs number of documents.

Figure 10 shows that bigger categories contain also more features, but this increase is weaker than the increase in websites. This implies that less very 'feature-rich' categories exist, which is also reflected in the high decay exponent $\delta = 1.9$ of a power-law fit in figure 9, (compared to the slower decay of the category size distribution shown in figure 1 where $\beta = 1.1$). Catenation of the size distribution measured in features and Heaps' law yields again dize distribution measured in websites: $P(i) = R(G(f_i))$, i.e. multiplication of the exponents yields that $\delta \cdot \alpha = 1.1$ which confirms our empirically found value $\beta = 1.1$.

## 5. EXPLOITING POWER LAW DISTRIBUTION IN LARGE-SCALE CLASSIFICATION

Fat-tailed distributions in large-scale web taxonomies highlight the underlying structure and semantics which are useful to visualize important properties of the data especially in big data scenarios. In this section we focus on the applications in the context of large-scale hierarchical classification, wherein the fit of power law distribution to such taxonomies can be leveraged to concretely analyze the space complexity of large-scale hierarchical classifiers in the context of a generic linear classifier deployed in top-down hierarchical cascade.

In the following sections we first present formally the task of hierarchical classification and then we proceed to the space complexity analysis for large-scale systems. Finally, we empirically validate the derived bounds.

## 5.1 Hierarchical Classification

In single-label multi-class hierarchical classification, the training set can be represented by $S = \{(\mathbf{x}^{(i)}, y^{(i)})\}_{i=1}^{N}$. In the context of text classification, $\mathbf{x}^{(i)} \in \mathcal{X}$ denotes the vector representation of document $i$ in an input space $\mathcal{X} \subseteq \mathbb{R}^d$.

The hierarchy in the form of rooted tree is given by $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where $\mathcal{V} \supseteq \mathcal{Y}$ denotes the set of nodes of $\mathcal{G}$, and $\mathcal{E}$ denotes the set of edges with parent-to-child orientation. The leaves of the tree which usually form the set of target classes is given by $\mathcal{Y} = \{u \in \mathcal{V} : \nexists v \in \mathcal{V}, (u, v) \in \mathcal{E}\}$. Assuming that there are $K$ classes, the label $y^{(i)} \in \mathcal{Y}$ represents the class associated with the instance $\mathbf{x}^{(i)}$. The hierarchical relationship among categories implies a transition from generalization to specialization as one traverses any path from root towards the leaves. This implies that the documents which are assigned to a particular leaf also belong to the inner nodes on the path from the root to that leaf node.

## 5.2 Space Complexity of Large-Scale Classification

The prediction speed for large-scale classification is crucial for its application in many scenarios of practical importance. It has been shown in [32; 3] that hierarchical classifiers are usually faster to train and test time as compared to flat classifiers. However, given the large physical memory of modern systems, what also matters in practice is the size of the trained model with respect to the available physical memory. We, therefore, compare the space complexity of hierarchical and flat methods which governs the size of the trained model in large scale classification. The goal of this analysis is to determine the conditions under which the size of the hierarchically trained linear model is lower than that of flat model.

As a prototypical classifier, we use a linear classifier of the form $\mathbf{w}^T\mathbf{x}$ which can be obtained using standard algorithms such as Support Vector Machine or Logistic Regression. In this work, we apply one-vs-all $L2$-regularized $L2$-loss support vector classification as it has been shown to yield state-of-the-art performance in the context of large scale text classification [12]. For flat classification one stores weight vectors $\mathbf{w}_y, \forall y$ and hence in a $K$ class problem in $d$ dimensional feature space, the space complexity for flat classification is:

$$Size_{Flat} = d \times K \tag{12}$$

which represents the size of the matrix consisting of $K$ weight vectors, one for each class, spanning the entire input space.

We need a more sophisticated analysis for computing the space complexity for hierarchical classification. In this case, even though the total number of weight vectors is much more since these are computed for all the nodes in the tree and not only for the leaves as in flat classification. Inspite of this, the size of hierarchical model can be much smaller as compared to flat model in the large scale classification. Intuitively, when the feature set size is high (top levels in the hierarchy), the number of classes is less, and on the contrary, when the number of classes is high (at the bottom), the feature set size is low.

In order to analytically compare the relative sizes of hierarchical and flat models in the context of large scale classification, we assume power law behavior with respect to the number of features, across levels in the hierarchy. More precisely, if the categories at a level in the hierarchy are ordered with respect to the number of features, we observe a power law behavior. This has also been verified empirically as illustrated in Figure 11 for various levels in the hierarchy, for one of the datasets used in our experiments. More formally, the feature size $d_{l,r}$ of the $r$-th ranked category, according to the number of features, for level $l$, $1 \leq l \leq L-1$, is given by:

$$d_{l,r} \approx d_{l,1} r^{-\beta_l} \tag{13}$$

where $d_{l,1}$ represents the feature size of the category ranked 1 at level $l$ and $\beta > 0$ is the parameter of the power law. Using this ranking as above, let $b_{l,r}$ represent the number of children of the $r$-th ranked category at level $l$ ($b_{l,r}$ is the branching factor for this category), and let $B_l$ represents the total number of categories at level $l$. Then the size of the entire hierarchical classification model is given by:

$$Size_{Hier} = \sum_{l=1}^{L-1} \sum_{r=1}^{B_l} b_{l,r} d_{l,r} \approx \sum_{l=1}^{L-1} \sum_{r=1}^{B_l} b_{l,r} d_{l,1} r^{-\beta_l} \tag{14}$$

Here level $l = 1$ corresponds to the root node, with $B_1 = 1$.

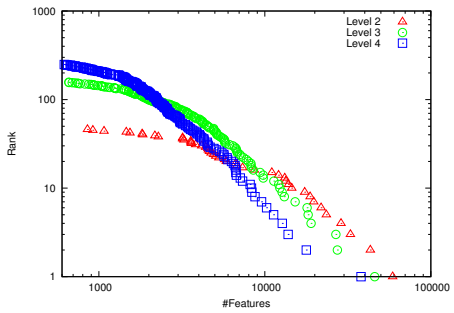



Figure 11: Power-law variation for features in different levels for LSHTC2-a dataset, Y-axis represents the feature set size plotted against rank of the categories on X-axis

We now state a proposition that shows that, under some conditions on the depth of the hierarchy, its number of leaves, its branching factors and power law parameters, the size of a hierarchical classifier is below that of its flat version.

PROPOSITION 1. *For a hierarchy of categories of depth $L$ and $K$ leaves, let $\beta = \min_{1 \leq l \leq L} \beta_l$ and $b = \max_{l,r} b_{l,r}$. Denoting the space complexity of a hierarchical classification model by $Size_{hier}$ and the one of its corresponding flat ver-*

*sion by $Size_{flat}$, one has:*

$$\text{For } \beta > 1, \text{if } \beta > \frac{K}{K - b(L-1)}(> 1), \text{ then } \quad Size_{hier} < Size_{flat} \tag{15}$$

$$\text{For } 0 < \beta < 1, \text{if } \frac{b^{(L-1)(1-\beta)} - 1}{(b^{(1-\beta)} - 1)} < \frac{1-\beta}{b}K, \text{then} \quad Size_{hier} < Size_{flat} \tag{16}$$

PROOF. As $d_{l,1} \leq d_1$ and $B_l \leq b^{(l-1)}$ for $1 \leq l \leq L$, one has, from Equation 14 and the definitions of $\beta$ and $b$:

$$Size_{hier} \leq bd_1 \sum_{l=1}^{L-1} \sum_{r=1}^{b^{(l-1)}} r^{-\beta}$$

One can then bound $\sum_{r=1}^{b^{(l-1)}} r^{-\beta}$ using ([32]):

$$\sum_{r=1}^{b^{(l-1)}} r^{-\beta} < \left[\frac{b^{(l-1)(1-\beta)} - \beta}{1-\beta}\right] \text{ for } \beta \neq 0, 1 \tag{17}$$

leading to, for $\beta \neq 0, 1$:

$$\begin{aligned} Size_{hier} &< bd_1 \sum_{l=1}^{L-1} \left[\frac{b^{(l-1)(1-\beta)} - \beta}{1-\beta}\right] \\ &= bd_1 \left[\frac{b^{(L-1)(1-\beta)} - 1}{(b^{(1-\beta)} - 1)(1-\beta)} - (L-1)\frac{\beta}{(1-\beta)}\right] \end{aligned} \tag{18}$$

where the last equality is based on the sum of the first terms of the geometric series $(b^{(1-\beta)})^l$.
If $\beta > 1$, since $b > 1$, it implies that $\frac{b^{(L-1)(1-\beta)}-1}{(b^{(1-\beta)}-1)(1-\beta)} < 0$. Therefore, inequality (18) can be re-written as:

$$Size_{hier} < bd_1(L-1)\frac{\beta}{(\beta-1)}$$

Using our notation, the size of the corresponding flat classifier is: $Size_{flat} = Kd_1$, where $K$ denotes the number of leaves. Thus:

$$\text{If } \beta > \frac{K}{K - b(L-1)}(> 1), \text{then } Size_{hier} < Size_{flat}$$

which proves Condition (15).
The proof for Condition (16) is similar: assuming $0 < \beta < 1$, it is this time the second term in Equation 18 ($-(L-1)\frac{\beta}{(1-\beta)}$) which is negative, so that one obtains:

$$Size_{hier} < bd_1 \left[\frac{b^{(L-1)(1-\beta)} - 1}{(b^{(1-\beta)} - 1)(1-\beta)}\right]$$

and then:

$$\text{If } \frac{b^{(L-1)(1-\beta)} - 1}{(b^{(1-\beta)} - 1)} < \frac{1-\beta}{b}K, \text{then } Size_{hier} < Size_{flat}$$

which concludes the proof of the proposition. □

It can be shown, but this is beyond the scope of this paper, that condition 16 is satisfied for a range of values of $\beta \in ]0, 1[$. However, as is shown in the experimental part, it is condition 15 of Proposition 1 that holds in practice.
The previous proposition complements the analysis presented in [32] in which it is shown that the training and test time of hierarchical classifiers is importantly decreased with respect to the ones of their flat counterpart. In this work we show that the space complexity of hierarchical classifiers is also better, under a condition that holds in practice, than the one of their flat counterparts. Therefore, for large scale taxonomies whose feature size distribution exhibit power law decay, hierarchical classifiers should be better in terms of speed than flat ones, due to the following reasons:

1. As shown above, the space complexity of hierarchical classifier is lower than flat classifiers.

2. For $K$ classes, only $O(\log K)$ classifiers need to be evaluated per test document as against $O(K)$ classifiers in flat classification.

In order to empirically validate the claim of Proposition 1, we measured the trained model sizes of a standard top-down hierarchical scheme (TD), which uses a linear classifier at each parent of the hierarchy, and the flat one.
We use the publicly available DMOZ data of the LSHTC challenge which is a subset of Directory Mozilla. More specifically, we used the large dataset of the LSHTC-2010 edition and two datasets were extracted from the LSHTC-2011 edition. These are referred to as LSHTC1-large, LSHTC2-a and LSHTC2-b respectively in Table 3. The fourth dataset (IPC) comes from the patent collection released by World Intellectual Property Organization. The datasets are in the LibSVM format, which have been preprocessed by stemming and stopword removal. Various properties of interest for the datasets are shown in Table 3.

| Dataset | #Tr./#Test | #Classes | #Feat. |
|---|---|---|---|
| **LSHTC1-large** | 93,805/34,880 | 12,294 | 347,255 |
| **LSHTC2-a** | 25,310/6,441 | 1,789 | 145,859 |
| **LSHTC2-b** | 36,834/9,605 | 3,672 | 145,354 |
| **IPC** | 46,324/28,926 | 451 | 1,123,497 |

Table 3: Datasets for hierarchical classification with the properties: Number of training/test examples, target classes and size of the feature space. The depth of the hierarchy tree for LSHTC datasets is 6 and for the IPC dataset is 4.

Table 4 shows the difference in trained model size (actual value of the model size on the hard drive) between the two classification schemes for the four datasets, along with the values defined in Proposition 1. The symbol $\bigtriangledown$ refers to the quantity $\frac{K}{K-b(L-1)}$ of condition 15.

| Dataset | TD | Flat | $\beta$ | $b$ | $\bigtriangledown$ |
|---|---|---|---|---|---|
| **LSHTC1-large** | 2.8 | 90.0 | 1.62 | 344 | 1.12 |
| **LSHTC2-a** | 0.46 | 5.4 | 1.35 | 55 | 1.14 |
| **LSHTC2-b** | 1.1 | 11.9 | 1.53 | 77 | 1.09 |
| **IPC** | 3.6 | 10.5 | 2.03 | 34 | 1.17 |

Table 4: Model size (in GB) for flat and hierarchical models along with the corresponding values defined in Proposition 1. The symbol $\bigtriangledown$ refers to the quantity $\frac{K}{K-b(L-1)}$

As shown for the three DMOZ datasets, the trained model for flat classifiers can be an order of magnitude larger than for hierarchical classification. This results from the sparse

and high-dimensional nature of the problem which is quite typical in text classification. For flat classifiers, the entire feature set participates for all the classes, but for top-down classification, the number of classes and features participating in classifier training are inversely related, when traversing the tree from the root towards the leaves. As shown in Proposition 1, the power law exponent $\beta$ plays a crucial role in reducing the model size of hierarchical classifier.

# 6. CONCLUSIONS

In this work we presented a model in order to explain the dynamics that exist in the creation and evolution of large-scale taxonomies such as the DMOZ directory, where the categories are organized in a hierarchical form. More specifically, the presented process models jointly the growth in the size of the categories (in terms of documents) as well as the growth of the taxonomy in terms of categories, which to our knowledge have not been addressed in a joint framework. From one of them, the power law in category size distribution, we derived power laws at each level of the hierarchy, and with the help of Heaps's law a third scaling law in the features size distribution of categories which we then exploit for performing an analysis of the space complexity of linear classifiers in large-scale taxonomies. We provided a grounded analysis of the space complexity for hierarchical and flat classifiers and proved that the complexity of the former is always lower than that of the latter. The analysis has been empirically validated in several large-scale datasets showing that the size of the hierarchical models can be significantly smaller that the ones created by a flat classifier.
The space complexity analysis can be used in order to estimate beforehand the size of trained models for large-scale data. This is of importance in large-scale systems where the size of the trained models may impact the inference time.

# 7. ACKNOWLEDGEMENTS

This work has been partially supported by ANR project Class-Y, BioASQ European project and Persyval Lab.

# 8. REFERENCES

[1] R. Babbar, I. Partalas, C. Metzig, E. Gaussier, and M.-R. Amini. Comparative classifier evaluation for web-scale taxonomies using power law. In *European Semantic Web Conference*, 2013.

[2] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.

[3] S. Bengio, J. Weston, and D. Grangier. Label embedding trees for large multi-class tasks. In *Neural Information Processing Systems*, pages 163–171, 2010.

[4] P. N. Bennett and N. Nguyen. Refined experts: improving classification in large taxonomies. In *Proceedings of the 32nd international ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 11–18, 2009.

[5] L. Bottou and O. Bousquet. The tradeoffs of large scale learning. In *Advances In Neural Information Processing Systems*, pages 161–168, 2008.

[6] L. Cai and T. Hofmann. Hierarchical document categorization with support vector machines. In *Proceedings of the thirteenth ACM international conference on Information and knowledge management*, pages 78–87, 2004.

[7] A. Capocci, V. D. Servedio, F. Colaiori, L. S. Buriol, D. Donato, S. Leonardi, and G. Caldarelli. Preferential attachment in the growth of social networks: The internet encyclopedia wikipedia. *Physical Review E*, 74(3):036116, 2006.

[8] O. Dekel, J. Keshet, and Y. Singer. Large margin hierarchical classification. In *Proceedings of the twenty-first international conference on Machine learning*, ICML '04, pages 27–34, 2004.

[9] S. N. Dorogovtsev and J. F. F. Mendes. Evolution of networks with aging of sites. *Physical Review E*, 62(2):1842, 2000.

[10] L. Egghe. Untangling herdan's law and heaps' law: Mathematical and informetric arguments. *Journal of the American Society for Information Science and Technology*, 58(5):702–709, 2007.

[11] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology. *SIGCOMM*.

[12] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin. LIBLINEAR: A library for large linear classification. *Journal of Machine Learning Research*, 9:1871–1874, 2008.

[13] T. Gao and D. Koller. Discriminative learning of relaxed hierarchy for large-scale visual recognition. In *IEEE International Conference on Computer Vision (ICCV)*, pages 2072–2079, 2011.

[14] M. M. Geipel, C. J. Tessone, and F. Schweitzer. A complementary view on the growth of directory trees. *The European Physical Journal B*, 71(4):641–648, 2009.

[15] S. Gopal, Y. Yang, B. Bai, and A. Niculescu-Mizil. Bayesian models for large-scale hierarchical classification. In *Neural Information Processing Systems*, 2012.

[16] G. Jona-Lasinio. Renormalization group and probability theory. *Physics Reports*, 352(4):439–458, 2001.

[17] K. Klemm, V. M. Eguíluz, and M. San Miguel. Scaling in the structure of directory trees in a computer cluster. *Physical review letters*, 95(12):128701, 2005.

[18] D. Koller and M. Sahami. Hierarchically classifying documents using very few words. In *Proceedings of the Fourteenth International Conference on Machine Learning*, ICML '97, 1997.

[19] T.-Y. Liu, Y. Yang, H. Wan, H.-J. Zeng, Z. Chen, and W.-Y. Ma. Support vector machines classification with a very large-scale taxonomy. *SIGKDD*, 2005.

[20] B. Mandelbrot. A note on a class of skew distribution functions: Analysis and critique of a paper by ha simon. *Information and Control*, 2(1):90–99, 1959.

[21] C. Metzig and M. B. Gordon. A model for scaling in firms' size and growth rate distribution. *Physica A*, 2014.

[22] M. Newman. Power laws, pareto distributions and zipf's law. *Contemporary Physics*, 46(5):323–351, 2005.

[23] M. E. J. Newman. Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, 2005.

[24] I. Partalas, R. Babbar, É. Gaussier, and C. Amblard. Adaptive classifier selection in large-scale hierarchical classification. In *ICONIP*, pages 612–619, 2012.

[25] P. Richmond and S. Solomon. Power laws are disguised boltzmann laws. *International Journal of Modern Physics C*, 12(03):333–343, 2001.

[26] H. A. Simon. On a class of skew distribution functions. *Biometrika*, 42(3/4):425–440, 1955.

[27] C. Song, S. Havlin, and H. A. Makse. Self-similarity of complex networks. *Nature*, 433(7024):392–395, 2005.

[28] H. Takayasu, A.-H. Sato, and M. Takayasu. Stable infinite variance fluctuations in randomly amplified langevin systems. *Physical Review Letters*, 79(6):966–969, 1997.

[29] C. J. Tessone, M. M. Geipel, and F. Schweitzer. Sustainable growth in complex networks. *EPL (Europhysics Letters)*, 96(5):58005, 2011.

[30] K. G. Wilson and J. Kogut. The renormalization group and the  expansion. *Physics Reports*, 12(2):75–199, 1974.

[31] G.-R. Xue, D. Xing, Q. Yang, and Y. Yu. Deep classification in large-scale text hierarchies. In *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '08, pages 619–626, 2008.

[32] Y. Yang, J. Zhang, and B. Kisiel. A scalability analysis of classifiers in text categorization. In *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval*, SIGIR '03, pages 96–103, 2003.

[33] G. U. Yule. A mathematical theory of evolution, based on the conclusions of dr. jc willis, frs. *Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character*, 213:21–87, 1925.

# Chapter 5

# Conclusion

In this thesis, a macroeconomic agent-based model, which has been studied numerically and theoretically, is proposed. It is constructed successively. In chapter 1 a simple scenario is analyzed, where firms have homogeneous expected profit margins $\mu$. Since firms are in competition either for workforce, or for purchasing power of customers, the dynamics of firm growth is stochastic. Results are: firm-size dependent scaling of the growth rate's standard deviation $\sigma(n) \propto n^{-0.5}$, a tent-shaped growth rate distribution, and a fat-tailed size distribution, whose counter-cumulative distribution can be approximated by a power law of exponent $\approx 0.7$.

In chapter 2, the simple scenario of chapter 1 is extended with heterogeneous margins, interest payments, ageing, and firm birth-death dynamics. This makes the results easier to interpret in the context of economics, and relates them to stock-flow consistent models, to evolutionary economics, and to macroeconomic agent-based modelling. The previously found results of a tent-shaped growth rate distribution and fat-tailed size distribution are largely conserved, since the stochastic dynamics that generate them is still present. Margin heterogeneity accounts for relative growth of firms, for which replicator equations have been derived. Another new result, which has been shown numerically and theoretically, is that margin heterogeneity modifies the growth rate's scaling exponent towards lower values. This result is closer to empirical studies than the scenario of chapter 1.

In chapter 3, the model, as it has been introduced in chapter 2, is further analyzed. Financial constraints and firms' age distribution are discussed, as well as distributions of age and size at the moment of firms' bankruptcy. These analyses allow the results of the model to be compared with further empirical evidence like [35, 20, 23, 3, 1]. Since the studies provide evidence from many different countries (Italy, US, Ireland, Japan, Portugal), they might contain universal features of industrial organization, and therefore suggest further possible improvements. In current work, further validation on databases is being carried out.

Chapter 4 presents an additional work, the exploitation of fat-tailed distri-

butions for the purpose of hierarchical classification. This work is not directly related to the previous three chapters.

The thesis concludes with some critical reflexions on the approach taken. Firstly, the work can be criticized for a number of reasons. Apart from technical shortcomings, the methodology has some flaws:

The model needs to be kept simple for its mathematical tractability, so potentially the same mistake is repeated which models do:

- As models with representative agent do, the model might assume some average behaviour in order to explain some statistical regularities, but the reality is far too complex to be cast into a model[24].

- The fact that prices and wages are constants is doubtful. Even if it is used in this way in this model in order to reason in terms of profit margins, sales, and 'wagebill' of firms, it is unlikely to be accepted by the economics community.

- The model omits some important features which impact its results, e.g. physical capital, variable interest rates, household saving, and herding effects, but also the fact that it is a closed system, and the trade balance is not considered. Some of these features may possibly be included in future work.

- As other macroeconomic models, its level of abstraction is not equal for all sectors. This highlights the difficulty of the approach.

Even if the model has been kept simple, it may already be too complex for many applications. The model is abstract and not designed for prediction or policy analysis, which is a central task of macroeconomic models. Another limitation is that at this stage, the size distribution resulting from the stochastic process of the simple scenario has not been solved analytically.

Nevertheless, there are some strong points, which might be worth pursuing.

Several effects exist simultaneously in reality, so the attempt to reproduce several of them simultaneously can in principle work. For instance, some effects of financial constraints on growth have been found empirically [20, 35]. This can be attempted to be explained in a joint framework with other results resulting from stochastic growth, such as size, growth rate, and profit rate distribution [18, 95, 1, 61]. The large availability of firm growth databases can serve as a guideline. Its apparent success to reproduce jointly arising distributions on industrial dynamics might back the arguments being used to explain. Caution is however needed, since this can be a coincidence, and it may be easily falsified.

The fact that market allocations are modelled in a way where every outcome is equally likely is a very general principle that is central in the description of complex systems. In physics, it is associated with the principle of entropy maximization, which Bashkirov described as the "maximum honesty principle" [8], since no further information is assumed. These market allocations account for the binomial growth rate distribution, and the size distribution found in

the model. Other economic models have taken up the idea of a statistical equilibrium in a market, e.g. [38, 1].

Even if some of the hypotheses of this model may not be convincing, some results will hold true even under different assumptions. For instance, if many factors contribute to the growth of a firm, from the Central limit theorem, it seems plausible that the growth rate probability density is Gaussian, or close to Gaussian. If in addition, for large firms, the growth rate variance is narrower than for small firms, it is necessary to perform integral (1.46), which will be more peaked than the growth rate probability density for individual firms, if in addition a fat-tailed size distribution is assumed. Furthermore, the model is stock-flow consistent, which is a framework that allows to study systemic risk [14]. The model incoporates the Keynesian viewpoint that firms base their decisions on their financial position [90, 21].

Two timescales are considered, a short and a long one, at which firms are in competition. The long one relates the model to evolutionary economics, since it can be described by replicator equations. The argument takes up evolutionary arguments. In particular, for given parameters ($[\mu_{min}, \mu_{max}]$, $r$, $\nu$), a *speed* of the dynamics with respect to the lifetime of firms, and with respect to random events, can be set.

Since the market algorithms can be described theoretically, the model is more mathematically tractable than macroeconomic agent-based models. Finally, the theoretical growth model presented in chapter 1 is very general and might serve to explain growth process in various applications of complex systems.

The presented work is still ongoing, so no premature overall conclusion is drawn here. While conceiving, implementing and studying this model, I gained many insights, practical experience and learned very valuable knowlege, and also identified many interesting questions which might be addressed in the future.

# Bibliography


[1] Simone Alfarano, Mishael Milaković, Albrecht Irle, and Jonas Kauschke. A statistical equilibrium model of competitive firms. *Journal of Economic Dynamics and Control*, 36(1):136–149, 2012.

[2] Luiz GA Alves, Haroldo V Ribeiro, and Renio S Mendes. Scaling laws in the dynamics of crime growth rate. *Physica A: Statistical Mechanics and its Applications*, 2013.

[3] Hideaki Aoyama, Wataru Souma, Yuichi Nagahara, Mitsuhiro P Okazaki, Hideki Takayasu, and Misako Takayasu. Pareto's law for income of individuals and debt of bankrupt companies. *Fractals*, 8(03):293–300, 2000.

[4] Amir F Atiya. Bankruptcy prediction for credit risk using neural networks: A survey and new results. *Neural Networks, IEEE Transactions on*, 12(4):929–935, 2001.

[5] Robert L Axtell. Zipf distribution of us firm sizes. *Science*, 293(5536), 2001.

[6] Per Bak, Maya Paczuski, and Martin Shubik. Price variations in a stock market with many agents. *Physica A: Statistical Mechanics and its Applications*, 246(3):430–453, 1997.

[7] Albert-László Barabási. Scale-free networks: a decade and beyond. *Science*, 325(5939):412–413, 2009.

[8] Andrei Grigor'evich Bashkirov. Renyi entropy as a statistical entropy for complex systems. *Theoretical and mathematical physics*, 149(2):1559–1573, 2006.

[9] Christian Beck. Dynamical foundations of nonextensive statistical mechanics. *Physical Review Letters*, 87(18):180601, 2001.

[10] Christian Beck and EGD Cohen. Superstatistics. *Physica A: Statistical Mechanics and its Applications*, 322:267–275, 2003.

[11] Barbara R Bergmann. A microsimulation of the macroeconomy with explicitly represented money flows. In *Annals of Economic and Social Measurement, Volume 3, number 3*, pages 29–44. NBER, 1974.

[12] Luis Bettencourt, Jose Lobo, and Hyejin Youn. The hypothesis of urban scaling: formalization, implications and challenges. *arXiv preprint arXiv:1301.5919*, 2013.

[13] Luís MA Bettencourt, José Lobo, Dirk Helbing, Christian Kühnert, and Geoffrey B West. Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the National Academy of Sciences*, 104(17):7301–7306, 2007.

[14] Dirk J Bezemer. "no one saw this coming": Understanding financial crisis through accounting models. 2009.

[15] Tamás S Biró and Antal Jakovác. Power-law tails from multiplicative noise. *Physical review letters*, 94(13):132302, 2005.

[16] Richard A Blythe and Alan J McKane. Stochastic models of evolution in genetics, ecology and linguistics. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(07):P07018, 2007.

[17] Lisa Borland. Ito-langevin equations within generalized thermostatistics. *Physics Letters A*, 245(1):67–72, 1998.

[18] Giulio Bottazzi, Davide Pirino, and Federico Tamagni. Zipf law and the firm size distribution: a critical discussion of popular estimators. Technical report, Laboratory of Economics and Management (LEM), Sant'Anna School of Advanced Studies, Pisa, Italy, 2013.

[19] Giulio Bottazzi and Angelo Secchi. Explaining the distribution of firm growth rates. *The RAND Journal of Economics*, 37(2):235–256, 2006.

[20] Giulio Bottazzi and Angelo Secchi. Financial fragility and the distribution of firm growth rates. Technical report, Laboratory of Economics and Management (LEM), Sant'Anna School of Advanced Studies, Pisa, Italy, 2013.

[21] Charlotte Bruun. Rediscovering the economics of keynes in an agent-based computational setting. Conferencepaper, 2008.

[22] Charlotte Bruun and Carsten Heyn-Johnsen. The paradox of monetary profits: an obstacle to understanding financial and economic crisis? *Economics Discussion Paper*, (2009-52), 2009.

[23] Luis MB Cabral and Jose Mata. On the evolution of the firm size distribution: Facts and theory. *American Economic Review*, pages 1075–1090, 2003.

[24] Claudio Castellano, Santo Fortunato, and Vittorio Loreto. Statistical physics of social dynamics. *Reviews of modern physics*, 81(2):591, 2009.

[25] David G Champernowne. A model of income distribution. *The Economic Journal*, 63(250):318–351, 1953.

[26] Silvano Cincotti, Marco Raberto, and Andrea Teglio. Credit money and macroeconomic instability in the agent-based model and simulator eurace. *Economics: The Open-Access, Open-Assessment E-Journal*, 4, 2010.

[27] Alex Coad. Investigating the exponential age distribution of firms. *Economics: The Open-Access, Open-Assessment E-Journal*, 4, 2010.

[28] David Colander, Peter Howitt, Alan Kirman, Axel Leijonhufvud, and Perry Mehrling. Beyond dsge models: toward an empirically based macroeconomics. *The American Economic Review*, 98(2):236–240, 2008.

[29] Giovanni Dosi, Giorgio Fagiolo, Mauro Napoletano, and Andrea Roventini. Income distribution, credit and fiscal policies in an agent-based keynesian model. *Journal of Economic Dynamics and Control*, 2013.

[30] Giovanni Dosi, Giorgio Fagiolo, and Andrea Roventini. Schumpeter meeting keynes: A policy-friendly model of endogenous growth and business cycles. *Journal of Economic Dynamics and Control*, 34(9):1748–1767, 2010.

[31] Gunnar Eliasson. Competition and market processes in a simulation model of the swedish economy. *The American Economic Review*, 67(1):277–281, 1977.

[32] Joshua M Epstein. Agent-based computational models and generative social science. *Generative Social Science: Studies in Agent-Based Computational Modeling*, pages 4–46, 1999.

[33] Einar Jón Erlingsson, Simone Alfarano, Marco Raberto, and Hlynur Stefánsson. On the distributional properties of size, profit and growth of icelandic firms. *Journal of Economic Interaction and Coordination*, pages 1–18, 2013.

[34] Eurostat. news release 142 2012, 2012.

[35] Giorgio Fagiolo and Alessandra Luzzi. Do liquidity constraints matter in explaining firm size and growth? some evidence from the italian manufacturing industry. *Industrial and Corporate Change*, 15(1):1–39, 2006.

[36] J Doyne Farmer and Duncan Foley. The economy needs agent-based modelling. *Nature*, 460(7256):685–686, 2009.

[37] J Doyne Farmer and John Geanakoplos. Power laws in economics and elsewhere. *Santa Fe Institute, Santa Fe*, 2008.

[38] Duncan K Foley. A statistical equilibrium theory of markets. *Journal of Economic Theory*, 62(2):321–345, 1994.

[39] Duncan K Foley. *Adam's fallacy: a guide to economic theology.* Harvard University Press, 2008.

[40] Duncan K Foley and Lance Taylor. 4. a heterodox growth and distribution model. *Economic Growth and Distribution*, page 75, 2006.

[41] Santo Fortunato, Michael Macy, and Sidney Redner. Editorial: Statistical mechanics and social sciences. *arXiv preprint arXiv:1304.1171*, 2013.

[42] Dongfeng Fu, Fabio Pammolli, Sergey V Buldyrev, Massimo Riccaboni, Kaushik Matia, Kazuko Yamasaki, and H Eugene Stanley. The growth of business firms: Theoretical framework and empirical evidence. *Proceedings of the National Academy of Sciences of the United States of America*, 102(52):18801–18806, 2005.

[43] Xavier Gabaix. Zipf's law for cities: an explanation. *The Quarterly Journal of Economics*, 114(3):739–767, 1999.

[44] Xavier Gabaix. Power laws in economics and finance. Technical report, National Bureau of Economic Research, 2008.

[45] Domenico Delli Gatti. *Macroeconomics from the Bottom-up*, volume 1. Springer, 2011.

[46] Robert Gibrat. *Les inégalités économiques*. Recueil Sirey, 1931.

[47] Wynne Godley and Marc Lavoie. *Monetary economics*. Palgrave Macmillan Basingstoke, 2007.

[48] Peter E Hart and Nicholas Oulton. Growth and size of firms. *The Economic Journal*, pages 1242–1252, 1996.

[49] Geoffrey M Hodgson. On the evolution of thorstein veblen's evolutionary economics. *Cambridge Journal of Economics*, 22(4):415–431, 1998.

[50] Josef Hofbauer and Karl Sigmund. Evolutionary game dynamics. *Bulletin of the American Mathematical Society*, 40(4):479–519, 2003.

[51] Bernardo A Huberman and Natalie S Glance. Evolutionary games and computer simulations. *Proceedings of the National Academy of Sciences*, 90(16):7716–7718, 1993.

[52] Stephen Hymer and Peter Pashigian. Firm size and rate of growth. *The Journal of Political Economy*, 70(6):556–569, 1962.

[53] Atushi Ishikawa. Pareto index induced from the scale of companies. *Physica A: Statistical Mechanics and its Applications*, 363(2):367–376, 2006.

[54] Atushi Ishikawa. The uniqueness of firm size distribution function from tent-shaped growth rate distribution. *Physica A: Statistical Mechanics and its Applications*, 383(1):79–84, 2007.

[55] Steve Keen. Solving the paradox of monetary profits. *Economics Discussion Paper*, (2010-2), 2010.

[56] Harry Kesten. Random difference equations and renewal theory for products of random matrices. *Acta Mathematica*, 131(1):207–248, 1973.

[57] Stephen Kinsella. The age distribution of firms in ireland, 1961–2009. *Department of Economics, University of Limerick, Ireland*, 2009.

[58] Stephen Kinsella, Matthias Greiff, and Edward J Nell. Income distribution in a stock-flow consistent model with education and technological change. *Eastern Economic Journal*, 37(1):134–149, 2011.

[59] Paul R Krugman and Paul Krugman. *The self-organizing economy*. Blackwell Publishers Cambridge, Massachusetts, 1996.

[60] Marco Lamieri. Debt maturity structure of italian firms. *Available at SSRN 1373752*, 2009.

[61] Marco Lamieri and Ilaria Sangalli. Economia e finanza dei distretti industriali. Report, sanpaolo intesa, 2011.

[62] Paul Langevin. *On the theory of Brownian motion*. 1908.

[63] Youngki Lee, Luís A Nunes Amaral, David Canning, Martin Meyer, and H Eugene Stanley. Universal features in the growth dynamics of complex organizations. *Physical Review Letters*, 81(15):3275, 1998.

[64] Moshe Levy and Sorin Solomon. Power laws are logarithmic boltzmann laws. *International Journal of Modern Physics C*, 7(04):595–601, 1996.

[65] Thomas Lux. Stochastic behavioral asset pricing models and the stylized facts. *Handbook of Financial Markets: Dynamics and Evolution. North-Holland: Amsterdam*, pages 161–216, 2009.

[66] Alfred Marshall. *Principles of economics: unabridged eighth edition*. Cosimo, Inc., 2009.

[67] Matteo Marsili, Sergei Maslov, and Yi-Cheng Zhang. Dynamical optimization theory of a diversified portfolio. *Physica A: Statistical Mechanics and its Applications*, 253(1):403–418, 1998.

[68] Matteo Marsili and Yi-Cheng Zhang. Interacting individuals leading to zipf's law. *Physical Review Letters*, 80(12):2741–2744, 1998.

[69] Jürgen Mimkes. Introduction to macro-econophysics and finance. *Continuum Mechanics and Thermodynamics*, 24(4-6):731–737, 2012.

[70] Michael Mitzenmacher. A brief history of generative models for power law and lognormal distributions. *Internet mathematics*, 1(2):226–251, 2004.

[71] Hernan Mondani. *Statistical Mechanics of Organizational Growth Processes*. PhD thesis, KTH, 2012.

[72] Richard Münch and Neil J Smelser. Relating the micro and macro. *The micro-macro link*, pages 356–388, 1987.

[73] Mark EJ Newman. Power laws, pareto distributions and zipf's law. *Contemporary physics*, 46(5):323–351, 2005.

[74] MEJ Newman. Complex systems: A survey. *arXiv preprint arXiv:1112.1440*, 2011.

[75] Luís A Nunes Amaral, Sergey V Buldyrev, Shlomo Havlin, Philipp Maass, Michael A Salinger, H Eugene Stanley, and Michael HR Stanley. Scaling behavior in economics: the problem of quantifying company growth. *Physica A: Statistical Mechanics and its Applications*, 244(1):1–24, 1997.

[76] Kazumi Okuyama, Misako Takayasu, and Hideki Takayasu. Zipf's law in income distribution of companies. *Physica A: Statistical Mechanics and its Applications*, 269(1):125–131, 1999.

[77] Guy H Orcutt. A new type of socio-economic system. *The Review of Economics and Statistics*, 39(2):116–123, 1957.

[78] S Picoli Jr, RS Mendes, LC Malacarne, and EK Lenzi. Scaling behavior in the dynamics of citations to scientific journals. *EPL (Europhysics Letters)*, 75(4):673, 2006.

[79] Boris Podobnik, Davor Horvatic, Alexander M Petersen, Branko Urošević, and H Eugene Stanley. Bankruptcy risk model and empirical tests. *Proceedings of the National Academy of Sciences*, 107(43):18325–18330, 2010.

[80] Linda E Reichl and Ilya Prigogine. *A modern course in statistical physics*, volume 71. University of Texas press Austin, 1980.

[81] Peter Richmond and Sorin Solomon. Power laws are disguised boltzmann laws. *International Journal of Modern Physics C*, 12(03):333–343, 2001.

[82] ALFRPED RRNYI. On measures of entropy and information. In *Fourth Berkeley Symposium on Mathematical Statistics and Probability*, pages 547–561, 1961.

[83] Thomas C Schelling. Dynamic models of segregation†. *Journal of mathematical sociology*, 1(2):143–186, 1971.

[84] A Schenzle and H Brand. Multiplicative stochastic processes in statistical physics. *Physical Review A*, 20(4):1628, 1979.

[85] Joseph Alois Schumpeter. *Capitalism, socialism and democracy.* Harper & Row New York, 1962.

[86] Yonathan Schwarzkopf, Robert Axtell, and J Farmer. An explanation of universality in growth fluctuations. 2010.

[87] Pascal Seppecher. Un modèle macroéconomique multi-agents avec monnaie endogène. 2009.

[88] Karl Sigmund. *The calculus of selfishness*. Princeton University Press, 2010.

[89] Herbert A Simon. On a class of skew distribution functions. *Biometrika*, 42(3/4):425–440, 1955.

[90] Robert Skidelsky. *Keynes: the return of the master*. PublicAffairs, 2010.

[91] Mark Skousen. *The Big Three in Economics: Adam Smith, Karl Marx, and John Maynard Keynes*. ME Sharpe Inc, 2007.

[92] Danyel JB Soares, Constantino Tsallis, Ananias M Mariz, and Luciano R da Silva. Preferential attachment growth model and nonextensive statistical mechanics. *EPL (Europhysics Letters)*, 70(1):70, 2005.

[93] Didier Sornette. Multiplicative processes and power laws. *Physical Review E*, 57(4):4811, 1998.

[94] Didier Sornette and Rama Cont. Convergent multiplicative processes repelled from zero: power laws and truncated power laws. *Journal de Physique I*, 7(3):431–444, 1997.

[95] Michael HR Stanley, Luis AN Amaral, Sergey V Buldyrev, Shlomo Havlin, Heiko Leschhorn, Philipp Maass, Michael A Salinger, and H Eugene Stanley. Scaling behaviour in the growth of companies. *Nature*, 379(6568):804–806, 1996.

[96] John Sutton. Gibrat's legacy. *Journal of economic Literature*, 35(1):40–59, 1997.

[97] John Sutton. The variance of firm growth rates: the 'scaling'puzzle. *Physica a: statistical mechanics and its applications*, 312(3):577–590, 2002.

[98] Hideki Takayasu, Aki-Hiro Sato, and Misako Takayasu. Stable infinite variance fluctuations in randomly amplified langevin systems. *Physical Review Letters*, 79(6):966–969, 1997.

[99] Misako Takayasu, Hayafumi Watanabe, and Hideki Takayasu. Generalized central limit theorems for growth rate distribution of complex systems. *arXiv preprint arXiv:1301.2728*, 2013.

[100] Peter D Taylor and Leo B Jonker. Evolutionary stable strategies and game dynamics. *Mathematical Biosciences*, 40(1):145–156, 1978.

[101] Leigh Tesfatsion. Agent-based computational economics: modeling economies as complex adaptive systems. *Information Sciences*, 149(4):262–268, 2003.

[102] Leigh Tesfatsion and Kenneth L Judd. *Handbook of computational economics, Volume 2: Agent-based computational economics*, volume 2. North-Holland, 2006.

[103] Stewart Thornhill and Raphael Amit. Learning about failure: bankruptcy, firm age, and the resource-based view. *Organization Science*, 14(5):497–509, 2003.

[104] Ming Chen Wang and George Eugène Uhlenbeck. On the theory of the brownian motion ii. *Reviews of Modern Physics*, 17(2-3):323–342, 1945.

[105] Geoffrey B West, James H Brown, and Brian J Enquist. The origin of universal scaling laws in biology. *Scaling in biology*, pages 87–112, 2000.

[106] G Wilk and Z Włodarczyk. Interpretation of the nonextensivity parameter q in some applications of tsallis statistics and lévy distributions. *Physical review letters*, 84(13):2770–2773, 2000.

[107] Matthieu Wyart and Jean-Philippe Bouchaud. Statistical models for company growth. *Physica A: Statistical Mechanics and its Applications*, 326(1):241–255, 2003.

[108] G Udny Yule. A mathematical theory of evolution, based on the conclusions of dr. jc willis, frs. *Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character*, 213:21–87, 1925.

[109] Damián H Zanette and Susanna C Manrubia. Role of intermittency in urban development: a model of large-scale city formation. *Physical Review Letters*, 79(3):523–526, 1997.