KØBENHAVNS UNIVERSITET

Master's Thesis for the M.Sc. in Physics

Agent Based Modelling of Companies in a Globalized Market

Mirjam Partovi Dilami (hjc137)

Supervisor: Kim Sneppen

Niels Bohr Institut. 1. November 2021

Acknowledgements

I would like to thank my supervisor Kim Sneppen who has been an unwavering support and inspiration throughout this thesis. I am incredibly impressed with how many interesting and creative ideas he has provided - Kim Sneppen is indeed in possession of a so called *krøllet hjerne*.

My greatest gratitude befalls my friends Kim and Mikkel, who always provided me with support and guidance, and without whom I would not be achieving this degree. *Tak!*

Abstract

Background: Almost everyone everywhere wants all the things they have heard about, seen or experienced via the new technologies. This trend has pushed markets towards a new commercial reality with the emergence of global companies producing standardized consumer products on a unimagined scale of magnitude. By mass-producing a single product, global companies benefit from enormous economies of scale. For the companies to minimize their production cost further, they locate themselves in developing countries. The consequence of this is shown in unequal growth of the real Gross Domestic Product (GDP), which show a trade surplus in developing countries and a trade deficit in developed countries. **Aim of the study:** The main goal of this thesis is to contribute with agent based models that are designed to illustrate the interplay between companies in a globalized market. With a main focus on manufacturing companies, we use the parameters γ to incorporate economies of scale and σ to determine the transportation cost and with this the degree of globalization in the models. Positive and negative feedback mechanisms are also added to the models, to observe behaviour changes of the companies. The positive feedback changes the demand behaviour of companies from single customers with random distributions to being proportional to company size. The negative feedback consider labor supply deficits in areas with high company concentrations, were a lack of labor increase the wages and thus production cost of the companies. **Methods**: With statistical analyses of simulation data outputs of the variables *Company Size*, *Number of Companies*, *Company Lifetime* and *Distance between Companies*, we try to understand basic mechanisms affecting the behaviour of companies in a globalized markets. **Conclusion:** In total, we conclude that the parameters and feedback mechanisms all individually have intrinsic capabilities to adjust market mechanisms such as company size, geographical location and the number of companies. This suggest that a well considered balance between these parameters is necessary to obtain a specific desired market structure, and a change in just one parameter can completely alter the market dynamics.

Contents

1 Introduction **6**

2 Prior Work **8**

2.1 The Sneppen-Bornholdt Model (2018) 8

2.2 The Krugman-Venables Model (1995) 11

3 Research **20**

3.1 Research Strategy 20

3.2 Research Questions 21

4 Programming Tool & Method **22**

4.1 Programming Tool 22

4.2 Models 23

4.2.1 Model 0: Main Model 24

4.2.2 Model 1: One Type of Company 30

4.2.3 Model 2: Two Types of Companies 30

4.2.4 Model 3: Two Types of Companies with Positive Feedback 32

4.2.5 Model 4: Two Types of Companies with Positive and Negative Feedback . . 34

4.2.6 Model Summary 36

4.3 Model Parameters 37

4.4 Data Outputs 40

4.4.1 Variables of Interest 40

4.4.2 Steady State, Statistics and Correlation 41

5 Results **50**
5.1 Model 1 50
5.1.1 Dynamics 50
5.1.2 Steady State 53
5.1.3 Time Correlation 56
5.1.4 Company Lifetimes 64
5.1.5 Distances, Sizes and Number of Companies 68
5.2 Model 2 71
5.2.1 Model 2.a 71
5.3 Model 3 75
5.3.1 Dynamics 75
5.3.2 Distances and Sizes of Companies 77
5.3.3 Phase Diagrams 79
5.4 Model 4 81
5.4.1 Dynamics 81
5.4.2 Distances and Sizes of Companies 83
5.4.3 Phase Diagrams 85

6 Discussion **86**
6.1 Discussion of Time Correlation and Dependent Data 87
6.2 First Research Question 89
6.2.1 Discussion of τ 89
6.2.2 Discussion of σ 90
6.2.3 Discussion of γ 91
6.3 Second Research Question 92
6.4 Third Research Question 93
6.4.1 Discussion of the Positive Feedback 93
6.4.2 Discussion of the Negative Feedback 95

7 Conclusion **98**

Appendices **103**

A Additional Results **104**

A.1 Model 2.b Versus Model 3 109

B Steady State **110**

C Python Code **130**

D The Sneppen-Bornholdt Model (2018) **146**

E The Krugman-Venables Model (1995) **153**

Chapter 1

Introduction

Trade imbalances are a key feature of the latest wave of globalization. The imbalance is measured by the ratio of import and export in a country - a country with trade surplus exports more goods than it imports and a country with trade deficit imports more goods than it exports [21]. The main reason for trade imbalances starts with the powerful force of technology, which drives the world towards a converging commonality:

> *Almost everyone everywhere wants all the things they have heard about, seen, or experienced via the new technologies* (Theodore Levitt (1983), [11], page 1).

This trend has pushed markets towards a new commercial reality with the emergence of global markets for standardized consumer products on a unimagined scale of magnitude [11]. The producer of these standardized products are called global companies, and these should not be confused with multinational companies. A multinational company operates in a number of countries, and adjusts its products and practices to each of the countries. A global company operates with resolute constancy as if the entire world were a single entity, and therefore sell the same things in the same way everywhere. The two different strategies clearly give two different production costs. By mass-producing a single product and selling it the same way, global companies benefit from enormous economies of scale in production, distribution, marketing and management. Multinational companies do not achieve same benefits, as they differentiate their production to meet the needs of each country. By translating the benefits achieved by global companies into world prices, the companies decimate competitors that still live in the disabling grip of old assumptions about how the world works [11].

To further reduce their production costs, companies (from developed countries) move their production to developing countries, as these countries have low import prices (ensured by the World Trade Organization), exceedingly low wages paid to workers, lacks health and safety standards and more lenient environmental policies [21][12]. This means that an imbalance arises not only among the companies themselves, but also among the countries.

The trade imbalances between the countries are shown in the unequal growth of the real Gross Domestic Product (GDP). Figure 1.1 shows indeed that not all regions of the world recorded equal economic growth in 2018. Growth remained high, at 5.3 percent, in developing Asia and Oceania, whereas in the developing economies of America GDP increased by only 0.7 percent. The growth rate of transition and developed economies stood at 2.8 and 2.2 percent, respectively [19].

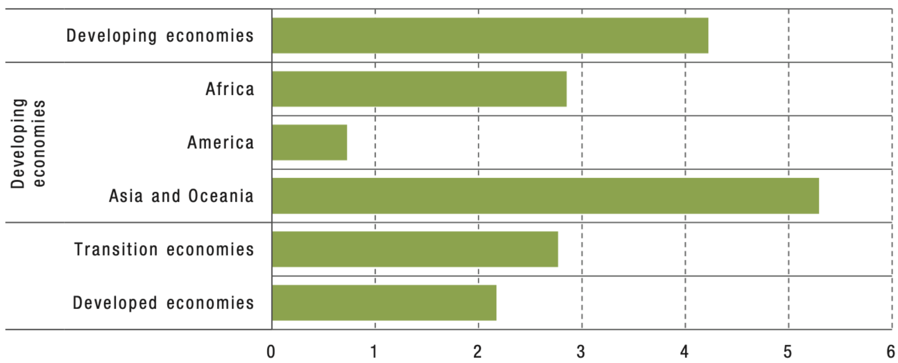


Figure 1.1: ***Percentage growth of real gross domestic product by group of economies in 2018*** *at constant 2010 United States dollars, source: United Nations Conference on Trade and Development [19]*

Although one can do nothing but indulge in the economic growth of developing countries, it may be relevant for us in the Western countries to discuss whether this growth is happening at our expense. Will our need for cheap goods cost us our prosperity and thus our welfare?

We have the impression that many economic models do not consider trade imbalances [13]. Our desire is to contribute with agent based models that are designed to illustrate the interplay between companies in a globalized market. With a main focus on manufacturing companies, we use the parameters γ to incorporate economies of scale and σ to determine the transportation cost and with this the degree of globalization in our models. From the outcome of our models, we will try to understand basic mechanisms affecting the behavior of companies in a globalized market and from this attempt to extrapolate our results to globalized market behavior. Finally, we will discuss how aspects of the globalized market affect Western welfare.

Chapter 2

Prior Work

Before presenting our models, we will do a review of previous models that we have found particularly inspiring. In the following section, we will look at the so called Sneppen-Bornholdt Model (2018) and the Krugman-Venables Model (1995).

2.1 The Sneppen-Bornholdt Model (2018)

In 2018, Kim Sneppen and Stefan Bornholdt published a paper called "Globalization in a Nutshell" (see appendix D). In this paper they discussed how global trading has emerged on a large scale due to a decreasing transportation costs and access to the Internet, which has made prices transparent to consumers beyond their regional scope. This has led to competition between different geographical regions, as optimization of manufacturing of simple products has allowed manufactures to take advantage of the economies of scale up. In their paper, Sneppen and Bornholdt were able to contribute with a simple agent based model that illustrates the interplay between transportation cost and information barriers in a model of manufacturing and trade.

In their model, they define a lattice with L identical agents at position $X = 1, 2, 3, ..., L$. The agents are interpreted as identical companies, that all produce the same type of product. The consumers of these products are the companies themselves, which define the companies as both buyers and sellers of the products. At each update, a random company at position x is chosen to make a purchase. The company receives information on the total costs offered by companies within some maximum information horizon h. The total cost, c, of one unit product produced at position y is calculated with the following equation:

$$c = s(y)^{\gamma} + |x - y|\sigma \tag{2.1}$$

Equation 2.1 consists of two separate terms. The first term calculates the production cost per unit of the company, with $s(y)$ being the company size at position y and $\gamma \leq 1$ the economies of scale exponent. The second term calculates the transportation cost, which is proportional to the linear distance between the producer at y and the consumer at x. The proportional factor is σ.

The purchasing company chooses to make a purchase from the company that can offer the lowest total cost. This company is rewarded with a one unit increase of its company size:

$$s(y) \longrightarrow s(y) + 1 \quad \text{for y with minimal c} \tag{2.2}$$

The model is executed in time steps, where each time unit consists of τ trading updates as defined above. After these τ updates, new company sizes $s(y)$ are assigned to be equal to 1 plus the accumulated orders at site y during these updates. This allows τ to be interpreted as proportional to the time it takes to rebuild the production apparatus for the considered product type.

Figure 2.1 shows the dynamics of the model with periodic boundary conditions. Each panel shows a 3D plot of the company sizes for each positions in every time step.
Panel (a) shows that companies collapse and emerge. The emergence of new companies often occurs close to previous collapsed ones, which partly pinned inheritance reflects the memory associated with the geography of surrounding companies that survive the collapse of a particular company [9]. By comparing panels (a) and (b), it is shown that a lower τ increases the number of companies collapsing and emerging and thus the degree of instability in the model. Panel (c) shows how a lower transportation cost σ can stabilize a system with low τ. Panel (d) introduces an information horizon $h = 10$. This panel uses the same other parameters as panel (c) and illustrates that a low information horizon has an effect comparable to higher transportation cost.

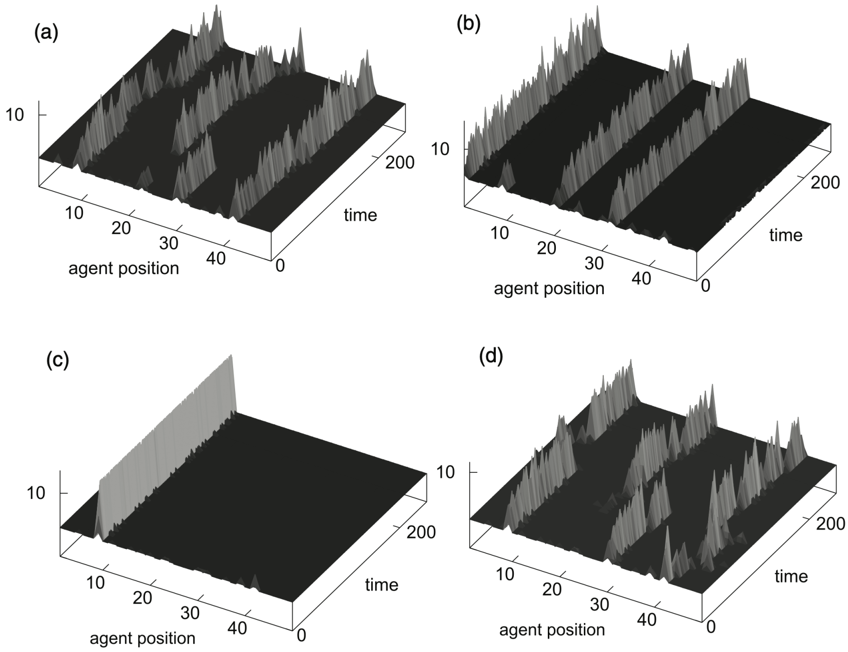


Figure 2.1: ***K. Sneppen and S. Bornholdt (2018)*** *[9]: The dynamic of one-dimensional model for $L = 50$ agents and $\gamma = -0.5$. On the x-axis they have the position, on the y-axis the time steps and on the z-axis the company sizes. (a) High noise ($\tau = 15$), high transportation cost ($\sigma = 0.05$) and information horizon $h = \infty$. (b) Low noise ($\tau = 25$), $\sigma = 0.5$ and $h = \infty$. (c) $\tau = 15$, low transportation cost ($\sigma = 0.01$) and $h = \infty$. (d) Similar to panel (c), but with $h = 10$*

Based on their model, Sneppen and Bornholdt reach the conclusion that the size of a company or its associated customer base is governed by the balance between the positive feedback of an economy of scale and the negative feedback set by transport. This means that the emergence of companies depends primarily on the economies of scale exponent γ and the transportation cost σ.

2.2 The Krugman-Venables Model (1995)

In 1995, Paul Krugman and Anthony J. Venables published a paper called "Globalization and the Inequality of Nations" (see appendix E). In this paper they discussed the impact of globalization on nations. They discussed how theorists, at first, were concerned that integration of world markets would lead to uneven development in nations, such that rich nations would rise at the expense of poor nations. However, this has shown different; Developing countries have been growing driven by an expansion on manufacturing export, and this may be at the expense of the West.
In their paper, Krugman and Venables aimed to make a model that would offer analytical proportions to the discussion of the impact of globalization on nations.

In their model, they consider a world with two economies, North and South. The two economies are allowed to differ, but are identical in endowments, preferences and technology. North and South each produce a non traded good A, and a differentiated traded good M. Good A is produced with constant returns to scale, and good M with increasing returns to scale. This gives rise to the interpretation of good A as agricultural good, and M as manufactured good - both final consumption and intermediates good.
In the following, we will describe Krugman and Venables' model - the two sectors and the trading between the two economies. We will use no index for the parameters describing the Northern economy, and index $*$ for the parameters describing the Southern economy. The description will be done in terms of North, with terms of South being identical.

North is endowed with L units of labor and wage rate w. The consumers of good A and M are the employees, and they have Cobb-Douglas preference between agriculture and manufacturing goods. This preference can be represented with the Cobb-Douglas utility function [20]:

$$u(x_A, x_M) = x_A^{(1-\delta)} x_M^{\delta} \tag{2.3}$$

Where x_A and x_M are quantities of good A and M, respectively. The value of δ gives the share of the two goods.

As the consumers receive only labor income, the budget constraint is defined as [20]:

$$p_A x_A + p_M x_M = wL \tag{2.4}$$

With p_A being the price of agricultural goods, p_M the price of manufacturing goods, and wL the amount of money available to the consumers.

For the given level of price and income, an expression for the desired amount of good A and M can be found by maximising the utility function subject to the budget constraint [20]. This is defined as the Cobb-Douglas demand for good A and M:

$$\begin{aligned} X_A(p_A, wL) &= \frac{(1-\delta)wL}{p_A} \\ X_M(p_M, wL) &= \frac{\delta wL}{p_M} \end{aligned} \tag{2.5}$$

The indirect utility function, v, can be defined by substituting x_A and x_B in equation 2.3 with the Cobb-Douglas demand.

$$v(p_A, p_M, wL) = wL\Big[\frac{1-\delta}{p_A}\Big]^{1-\delta}\Big[\frac{\delta}{p_M}\Big]^{\delta} \tag{2.6}$$

The indirect utility function is the maximum utility achievable at a given price and income. It is therefore not surprising that the value of v increases when the wage rate increases or/and the prices decreases.

The price index for good A is given with Q_A, and Q_M for good M. With the agricultural sector being the numeraire, the price of one unit of good A is:

$$Q_A = p_A = 1 \tag{2.7}$$

The agricultural good is a non traded good, and is produced under perfect competition by using one unit of labor to produce one unit of output. This leads to the following production function for agriculture:

$$\tilde{x}_A = \tilde{L}_A \tag{2.8}$$

With $\tilde{L}_A$ units of labor producing $\tilde{x}_A$ units of good A. The function implies that a one unit increase in labor allocate to one unit increase in output of agriculture.

From equations 2.7 and 2.8, the wage rate in terms of good A can be deduced to:

$$Q_A \tilde{x}_A = w\tilde{L}_A \quad \Rightarrow \quad Q_A = w, \quad \text{as} \quad \tilde{x}_A = \tilde{L}_A$$

$$w = 1 \tag{2.9}$$

If $w \neq 1$, the economy has zero agricultural production.

The manufacturing sector produces a number of varieties of different products, where it is assumed that one company produces one variety. These products are aggregated into a composite good with price index:

$$Q_M = \left[n p_M^{1-\epsilon} + n^* (p_M^* \sigma)^{1-\epsilon} \right]^{1/(1-\epsilon)} \tag{2.10}$$

Here n is the number of varieties produced in North, which are all sold at the same price p_M. Similarly, n^* is the number produced in South, and sold at price p_M^*. Southern products sold in North incur iceberg transport cost at a rate σ. For iceberg transport cost it is assumed that only a proportion $1/\sigma$ of any good shipped arrives, and with $1 - 1/\sigma$ lost in transition. The consumer has to pay for the missing goods, which results in the consumer price $p_M^* \sigma$. $\epsilon > 1$ is the elasticity of substitution between any two goods.

The manufacturing sector uses labor and a composite intermediate good to produce output. For simplicity, it is assumed that the composite intermediate good is the same as the composite consumption good. Therefore the price index of the intermediate good is (also) given in equation 2.10. This means that manufacturing goods enter the production function for other manufacturing as intermediates goods, and the utility function as final goods.

For the manufacturing sector, labor and intermediate good are combined with Cobb-Douglas technology and described with the Cobb-Douglas production function with output elasticity μ [20]:

$$\tilde{x}_M = \tilde{L}_M^{1-\mu} K^{\mu} \tag{2.11}$$

Where $\tilde{L}_M$ is labor input, K is intermediate input, and $\tilde{x}_M$ is the total output.

For a company in manufacturing producing y units of output for domestic sale and x units of output to export, the total cost of the company is:

$$CT = w^{1-\mu} Q_M^{\mu} \Big[\alpha + \beta(y + x)\Big] \tag{2.12}$$

Where $w^{1-\mu} Q_M^{\mu} \alpha$ is a fixed cost paid to start production, and $w^{1-\mu} Q_M^{\mu} \beta$ is a cost paid for each product produced.

From the total number of manufacturing product produced, X_M units of output are sold as final manufacturing goods to customers (equation 2.5). If μ is positive, a constant share μ is sold as intermediate goods to other companies. This division of sale, leads to the following expenditure function for manufacturing goods in the Northern economy:

$$E = \delta w L + \mu(y + x) p_M n \tag{2.13}$$

Here the first term is the consumer's expenditure on final manufacturing goods, and the second term is the demand of intermediate goods.

The price of a manufacturing good p_M is set as a markup over marginal cost by factor $\epsilon/(\epsilon - 1)$:

$$p_M = \epsilon/(\epsilon - 1) \times w^{1-\mu} Q_M^{\mu} \beta \tag{2.14}$$

Where the price p_M increases if the cost of production increases, and decreases if the cost of production decreases.

With the given price p_M, the Northern and Southern demand for a single variety is:

$$y = p_M^{-\epsilon} Q_M^{\epsilon - 1} E \tag{2.15}$$

$$x = p_M^{-\epsilon} \sigma^{1-\epsilon} (Q_M^*)^{\epsilon - 1} E^*$$

Where Q_M^* is the price index of manufacturing good in South, and E^* is the total expenditure on manufacturing good in South.

Companies can get in and out of the market at no cost, which in the long run will lead to all companies making zero profit. For zero profit it applies that the total earnings must be equal to the total cost:

$$p_M(y + x) = CT \tag{2.16}$$

From equation 2.16, the size of a company can be established:

$$y + x = (\epsilon - 1)\alpha/\beta \tag{2.17}$$

By choosing units of measurements such that the right hand side of equation 2.17 is equal to unity and substituting equation 2.15 in 2.17, the zero profit condition can be expressed as:

$$p_M^{-\epsilon}\Big[Q_M^{\epsilon-1}E + \sigma^{1-\epsilon}(Q_M^*)^{\epsilon-1}E^*\Big] = 1 \tag{2.18}$$

The description of Krugman and Venables' model ends here, and we would like to emphasize again that a description of the southern economy would be identical. Krugman and Venables takes many important aspects into account and incorporates these into their model. However, we choose to present only two aspects of their result, namely; the distribution of companies between the two economies and the distribution of labors between the two sectors for different values of the transport cost σ. To do so, we will in the following review 3 graphs which are included in their result section.

Figure 2.2 shows the allocation of manufacturing companies in the two economies for a relatively high transportation cost. The number of companies in North is given on the x-axis and the number of companies in South is given on the y-axis. The dashed line (NN) and the solid line (SS) indicate when companies in North and South make zero profits. In the intersection between the two lines, where $n^* = n$, an equilibrium occurs. The arrows indicate the dynamics of the model, which occurs because companies are allowed to enter the market if profits are positive and exit if they are negative.

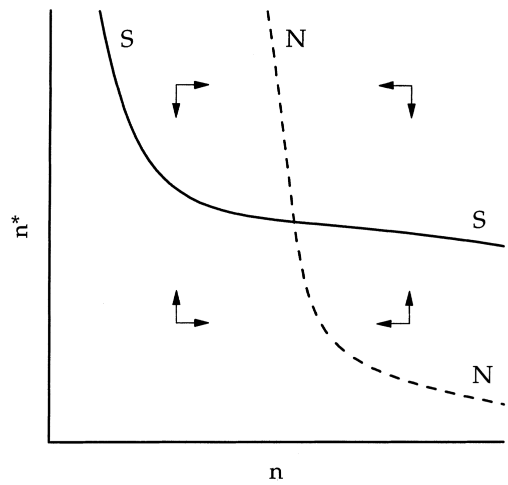


Figure 2.2: ***P. Krugman and A. J. Venables (1995)*** *[14]: Number of companies in North and South for high transport cost. The dashed line NN and the solid line SS indicate when companies in North and South earn zero profits. The arrows indicate the dynamics*

The number of manufacturing companies in North (and similar in South) has an effect on the profitability of the companies. For an increase in n and thus a decrease in n^*, the prices index Q_M is reduced (see equation 2.10). This reduction in Q_M reduces the total and marginal cost (equation 2.12 and 2.14), which increase the profit of the companies. An increase in n also increase the total expenditure on manufacturing products (equation 2.13), which raises the demand (equation 2.15) and thus the profit of each company. Krugman and Venables define the first case as cost linkage between the companies, and the second case as demand linkage between the companies [14].

Figure 2.3 shows a plot for the wage rate w as a function of labor force, where the percentage of total labor in manufacturing $\%L_M$ ($\%L_A = 100\% - \%L_M$ in agriculture) is given on the x-axis and the wage rate on the y-axis.
In equilibrium, under the assumption that agricultural good is always produced, the wage rate is $w = 1$. This is marked with the dashed line $L_A L_A$. The solid line $L_M L_M$ shows the demand for labor in manufacturing, and gives the maximum wage a company can pay and break even. At the intersection between these two curves, at point S, a symmetric stable equilibrium occurs. This means that both economies have a wage rate equal to unity i.e $w = 1$, and both economies produce agricultural and manufacturing goods. By moving some companies from one economy to the other i.e. by changing the values of n and n^*, the equilibrium can be perturbed and thus the value of $w = 1$ will change.

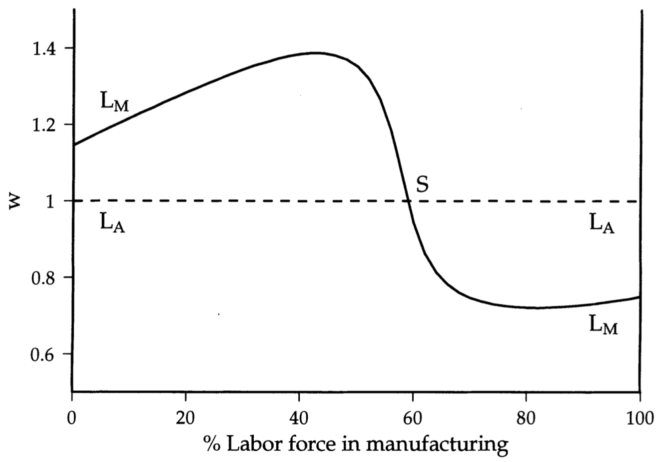


Figure 2.3: ***P. Krugman and A. J. Venables (1995)*** *[14]: Wage rate and labor demand for high transportation cost*

Figure 2.2 and 2.3 are both constructed with a high transportation cost. Figure 2.4 is the analogous diagram but with a low transportation cost. In this figure, the slope of the manufacturing labor demand curve is positive. This means that the equilibrium at point U is unstable and the equilibrium at S is stable.

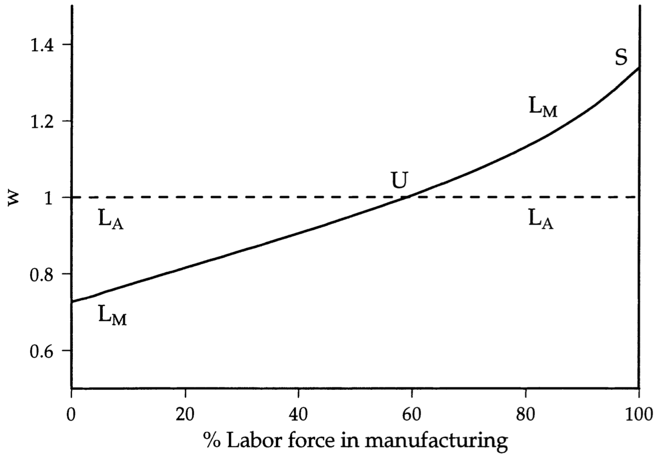


Figure 2.4: ***P. Krugman and A. J. Venables (1995)*** *[14]: Wage rate and labor demand for low transportation cost*

At point S, the manufacturing sector has a wage rate above $w = 1$ (where $w = 1$ is the wage rate offered by the agricultural sector). This high wage rate attracts labor from the agricultural sector to the manufacturing sector, which means that the economy only produces manufacturing products. This forces the other economy to produce only agricultural goods.
The reason for the reversal of the slope of the manufacturing labor demand is due to the linkages between the manufacturing companies. By relocating companies from South to North, a raise in the demand of Norther products would occur via the demand linkage. A relocation also reduces the cost of Northern companies, via the cost linkage, as varieties of intermediate good does not have to bear transportation cost. Both these linkages create forces agglomeration of manufacturing in a single location [14]. Same phenomena did not occur for high transportation cost (figure 2.2 and 2.3), as these forces are dominated by the need of companies to be near to the final consumer demand.

The overall goal of the Krugman-Venables Model and the Sneppen-Bornholdt Model were to contribute with a model that could describe the behavior of companies in a globalized market. In the Krugman-Venables Model, two types of companies where included; Agricultural and manufacturing companies. These company types were characterized by their return to scale, with agriculture companies producing goods with constant return to scale and manufacturing with increasing returns to scale. In the Sneppen-Bornholdt Model, only one type of company was included, manufacturing companies. Similar to the Krugman-Venables, these companies were defined to produce with increasing returns to scale.

In both models, the degree of globalization was determined by a transportation cost σ, where a low σ indicated a high degree of globalization.

Chapter 3

Research

In this thesis, we have an ambition to develop agent based models that can describe the interplay between companies in a globalized market. Furthermore, we intend to use aspects from the Sneppen-Bornholdt Model (2018) and the Krugman-Venables Model (1995).
From the Sneppen-Bornholdt Model, we will use the model setup but only include the parameters γ, τ and σ. From the Krugman-Venables Model, we will use the idea of two sectors or company types existing in the same market. Between these two company type, we have an intention to implement positive and negative feedback.

3.1 Research Strategy

In the development of our agent based models, we have a clear intention to keep the models simple by including only few assumptions and parameters. This approach is in agreement with the supervisor of this thesis, Kim Sneppen:

"... the value of a model being high, when it is build on few assumptions and parameters. A clear wrong model is often more useful than an unclear model" (Kim Sneppen (2019), [17], page 95).

In order to do so, we will begin with a simple agent based model (the Sneppen-Bornholdt Model with $h = \infty$) to better understand how the model reacts to a change in parameter values. To maintain this understanding we will built models one by one, each time only adding one new feature. Thus the second model will be based on the first model, but with one added feature, and the third model will be based on the second model with an additional feature. This process is continued until we reach the necessary number of features that is needed to properly address and answer the following reaches questions.

3.2 Research Questions

1. What influence do the following parameters have on companies and thus the market the companies exist in?
   - *Economies of scale* (γ)
   - *Time to rebuild production apparatus* (τ)
   - *Transportation cost* (σ)

2. What impact do different returns to scale have on co-existing company types?

3. How do different feedback mechanisms affect the distance between interacting companies producing two different products?

Chapter 4

Programming Tool & Method

In order to answer the research questions (see section 3.2), we have developed four agent based models coded with the programming language Python. This section will cover the model setups, the extraction of data outputs and the data processing. The associated simulation codes can be found in appendix C.

4.1 Programming Tool

The simulation of the models and the processing of data output are both performed in Python. We have more specifically used Pycharm for the simulations and the extraction of data outputs, and Jupyter Notebook for the data processing. For the transition between the two programs, we have saved the data outputs as pickle files in Pycharm and uploaded these in Jupyter Notebook. This is clearly a matter of taste and could easily have been done differently.

Python provides various libraries that we have used. A library that has influenced our data is *multiprocessing*. With the multiprocessing module, we were able to run multiple iterations of our simulation in parallel using all cores in the computer (see appendix C.10). This can be done for all independent iterations - the second run of a simulation does not depend on the first run [16]. However, as the iterations *within* the simulation (defined as time steps) are dependent - the next iteration is dependent on the previous iteration - multiprocessing had no use. This resulted in our preference to repeat the simulations several times rather than running them for longer time steps.

Another influence that Python has had on our thesis is the use of 0-indexing. Since Python uses 0-indexing, we have chosen to do the same.

4.2 Models

The four agent based models are presented in this section. Our approach in developing these models was to start with a simple model, understand its behavior and then add an additional aspect to make the next model. In this process, our main focus was to follow the research strategy (see section 3.1) and thus to ensure the simplicity (few assumptions and parameters) of the models.

In the following, we will start by presenting the Main Model which provides a basic description of the models. Then we continue with Model 1, which is completely identical to the Sneppen-Bornholdt Model (2018) for $h = \infty$. In Model 2, we are inspired by the Krugman-Venables Model (1995) where two sectors (agriculture and manufacturing) are competing, and we thus increase the number of company types from 1 to 2. In Model 3, we introduce positive feedback between the two company types. In Model 4, we continue with the positive feedback and add a negative feedback between the companies:

Section 4.2.1 Model 0: Main Model
Section 4.2.2 Model 1: One Type of Company
Section 4.2.3 Model 2: Two Types of Companies
Section 4.2.4 Model 3: Two Types of Companies with Positive Feedback
Section 4.2.5 Model 4: Two Types of Companies with Positive and Negative Feedback

4.2.1 Model 0: Main Model

4.2.1.1 Agents

The agents of the models are companies, of which there can exist two different types Company A and Company B. The companies of type Company A all produce one type of product Product A, whereas the companies of type Company B produce Product B. The companies aim to buy and sell these products at lowest cost (see equation 4.1), thus we define the companies to be both buyers and sellers of the products. However, only the producer of a product can sell the product

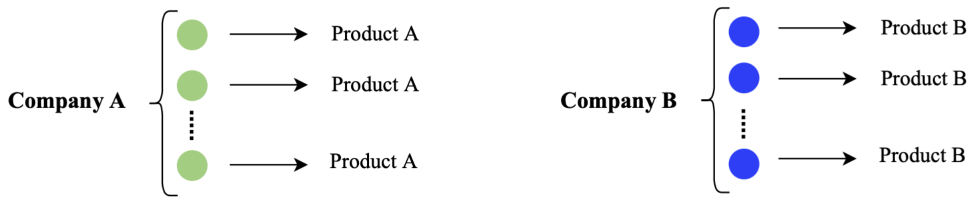


Figure 4.1: ***Company Types****: A model can contain two different company types Company A and Company B. The companies of type Company A produce Product A and the companies of type Company B produce Product B.*

The interpretation of the products must depend on the value of the *economies of scale exponent* γ, as γ determines how much the production cost is reduced with the company size i.e. S^{γ}. Here we make the same assumption as Krugman and Venables (1995); The agricultural sector produces with constant returns to scale and the manufacturing sector with increase returns to scale. This assumption leads to the following values $\gamma_{agriculture} = 0$ and $\gamma_{manufacturing} = -0.75$. In this thesis we will work with γ between these two values.
The size, S, of any company is initially, $t = 0$, set to be equal 1. With $S = 1$ a company can be considered as a "one-man business". Depending on the success or failure of the company, the size can change over time. An increase/decrease in S is considered as an increase/decrease in company size. The size of a company can change after τ interactions between the companies.

4.2.1.2 Environment

4.2.1.2.1 Agents in the Environment

The model contains L agents placed randomly on a line. Thus there will be L positions on the line, since no two agents can be placed in the same position. The positions are ordered and we will refer to these positions as $0, 1, 2, ..., L-1$. We will equip the model with periodic boundaries, having the consequence that the ends of the line are connected which forms a circle. This means that position 0 will be a neighbour to position $L-1$ (see figure 4.2). Each of the L positions will be occupied with an element from the set of available company types, which can contain Company A and Company B. The element from the set of available company types is chosen randomly with equal probabilities.

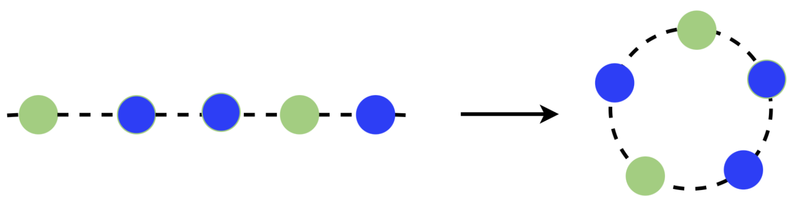


Figure 4.2: ***Periodic Boundary Conditions***

4.2.1.2.2 Transportation in the Environment

Companies can sell products to other companies. Whenever such a transaction takes place, a transportation cost is calculated. The transportation cost depends on the distance between the seller and buyer multiplied with a constant, $\sigma \geq 0$, which is encoded in the environment. Notice that this proportional dependence is markedly different from the exponential "iceberg cost" assumed in the economic literature. In fact, one may even expect modern shipping costs to increase more slowly than proportional to the distance, however we keep the linear dependence for simplicity here [9].

4.2.1.2.3 The Dynamic of the Environment

The environment can be either static or dynamic. If the environment is static it remains the same throughout the simulation, whereas a dynamic environment will change [1]. The environment is static in Model 1 and Model 2.a, and dynamic in Model 2.b, Model 3 and Model 4. In a dynamic environment, a company that has not had any customers after τ interactions i.e. one iteration of the simulation, will be defined as bankrupt. In this case all information about the company will be deleted and its position will be occupied by a new company. The type of the new company will be decided by a random generator, choosing from the set of available company types. This change in agents makes the environment dynamic (see appendix C.7).

4.2.1.3 Interaction

An iteration of the simulation can be considered as a time step, without time having any direct representation in the real world. The total number of time steps is represented with the letter T. For each time step, there will be $\tau \geq 1$ interactions between the agents. The agents are limited in who they can interact with, and each model has their interaction rules. These rules are listed in table and will be explained when each model is presented.

An interaction between the agents, must be considered as a transaction between the companies. For each count in τ, a company is randomly chosen to make a purchase from one of the available companies. An available company is a company that fulfill the conditions of the interaction rules. The randomly chosen company will choose to buy a product from the company offering the lowest total cost. For a randomly chosen company at position x, the total cost from position y at time t is set to:

$$cost = S_{t-1}(y)^{\gamma} + |x - y|\sigma \tag{4.1}$$

Equation 4.1 consists of two separate terms. The first term is the production cost per unit of the company, with $S(y)$ being the company size at position y and γ being the economies of scale exponent. This leads to $S(y)^{\gamma}$ being the economies of scale for the selling company.

The second term of the equation is the transportation cost, were $|x - y|$ is the shortest distance along the circle periphery and σ is the environment specific encoded constant (see section 4.2.1.2.2). The company size has a time index, as the size of a company can change. An increase of the company size is given as a reward to the company providing the lowest total cost (see equation 4.2). If there are more than one company offering the same low cost, a random generator will randomly choose one of the companies to receive the reward.

$$S_{t+1}(y') \longrightarrow S_{t+1}(y') + 1 \quad \text{for y' with minimal cost} \tag{4.2}$$

The size of any company at any given time $t \geq 1$ is entirely dependent on the number of sales in the previous time step. The number of sales depends on the company position and size, as these determine the total cost of interaction with the company.

The company sizes will be updated after τ interactions. For time t, after all counts in τ, the new company size $S_{t+1}(y)$ will be equal to 1 plus the accumulated orders at site y during time t. These updated company sizes will be included in the calculation of the total cost at time $t+1$. At time $t+1$, $S_{t+2}(y)$ is set to be equal one. $S_{t+2}(y)$ will again be assigned to be 1 plus the accumulated orders during time $t+1$ and used for the calculation of the total cost at time $t+2$.

By updating the company sizes after τ interactions, we interpret τ as being proportional to the time it takes to rebuild the product apparatus for the considered product type [9]. After τ interactions, the company are worn down and needs to be rebuilt. The size of the new company depends on how much the company has been rewarded in the previous time step.

The interactions at $t = 0$ between $L = 8$ agents and $\tau = 2$ are illustrated in figure 4.3. In this illustration the big circle represents the environment in which the companies are located. The companies are illustrated with small circles, where the size of the circle corresponds to the size of the company. At $t = 0$ all companies have $S = 1$. The filled yellow circle is the company who has to make a purchase. The yellow circles are the companies the buyer can interact with, and the red circles are the companies the buyer can not interact with. For each count in τ all the yellow circles offer a total cost: cost_1, cost_2 and cost_3. For the first count in τ ($\tau' = 1$) the lowest total cost is cost_1, hence we reward the company by increasing the size of its circle. Similar for $\tau' = 2$, where the company offering cost_3 is rewarded. The new circle sizes are updated after $\tau = 2$ interactions (full iteration), and these will be used in the calculation of the total cost at $t = 1$.

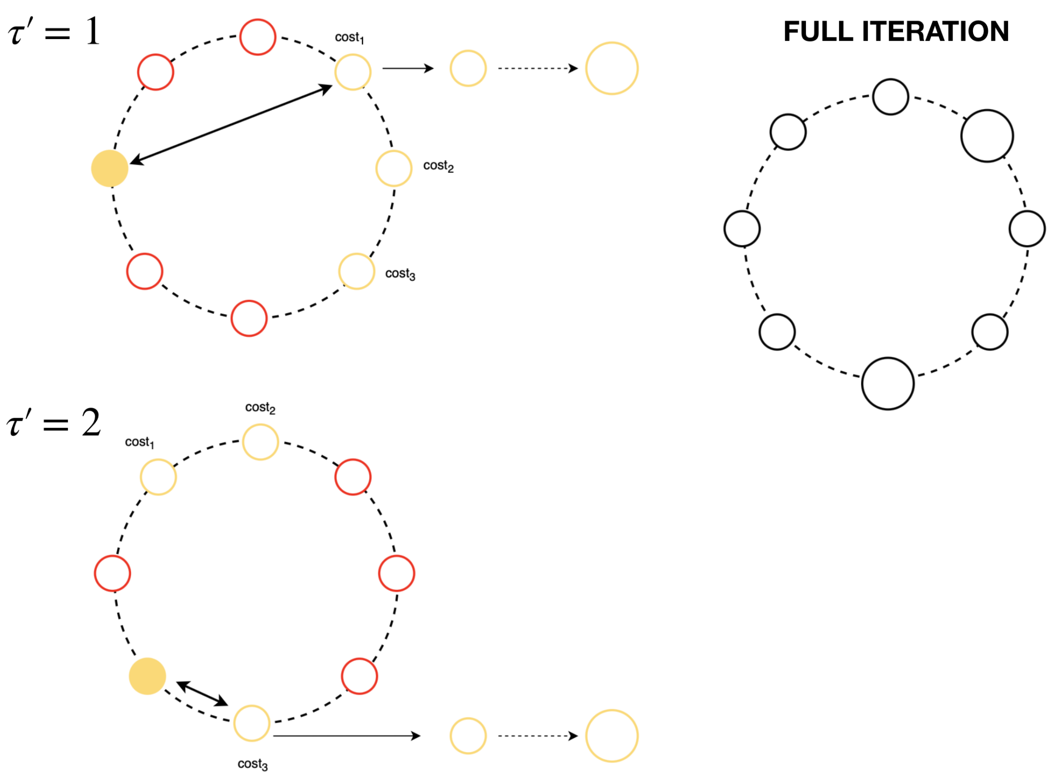


Figure 4.3: ***Interactions and Rewards***: *Illustration of interactions at $t = 0$ between $L = 8$ agents and $\tau = 2$. The environment is illustrated with a big circle, and the companies with small circles. The size of each of the small circles correspond to the company size at the given position. At $t = 0$ all the companies have size $S = 1$. The filled yellow circle is the company how has to make a purchase. The yellow circles are the companies the buyer can interact with, and the red circles are the companies the buyer can not interact with. At each count in τ all the yellow circle offer a total cost: $cost_1$, $cost_2$ and $cost_3$. The company offering the lowest total cost is rewarded with an increase in size. The company sizes are updated after the end of the iteration (full iteration)*

4.2.1.4 Remarks

A huge lack in this model is that the buying company does not gain or lose anything from a transaction. This has been an ambition to change, but it required adding new parameters. Adding new parameters will go against the research strategy and thus the principle of keeping the model simple (working with few parameters). This means that we must reformulate the goal of the agents to be:

The goal of the selling agent is to offer the lowest total cost. The goal of the buying agent is to secure a reward to the agent offering the lowest total cost.

A consequence of letting the agents buy without any loss (payment) means that there are unlimited resources in the environment of the simulation. With unlimited resources and no end-event, the simulation can keep going in the same way indefinitely. This means that the model must be defined as a *non-terminal simulation model* [3]. For such a simulation model it is needed to determine the steady state behavior of the model. With the steady state we can determine how long we should run the simulation and when to start saving data outputs. This is all explained in more detail in section 4.4.2.1.
The model is also defined as a stochastic model, due to the use of random generators through the simulation. We use random generators in three case:

- *Each of the L positions is occupied with a randomly chosen agent.*
- *An agent is randomly chosen to make a purchase.*
- *If more than one agent are offering the lowest total cost, one is randomly chosen to be rewarded.*

An alternative, to avoid randomness in the model, could be to systematically distribute and select the agents. However, we believe that a random choice is more realistic. This has the consequence that data outputs of the simulation are random, making runs of the simulation program *realization* of the system performance [2]. This requires that we repeat the simulation and make statistics on the outputs. The statistical approach is described in section 4.4.2.2.
These conditions also apply to all variations of the Main Model.

4.2.2 Model 1: One Type of Company

Model 1 is identical to the Sneppen-Bornholdt Model (2018) for $h = \infty$. The model consist of one company type (Company A) and therefore contains L identical companies, where any two companies can interact. The environment is static, meaning that a non-rewarded company continues in the next time step.

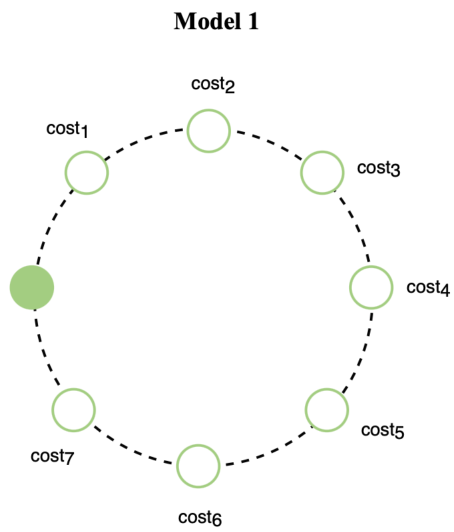


Figure 4.4: ***Illustration of Model 1****: The big circle illustrates the environment, and the small circles the companies. The small circles are all colored in green, as the model only consist of one company type Company A. The filled green circle is the company that has to make a purchase. Since any two companies can interact, all other companies announce their total cost*

4.2.3 Model 2: Two Types of Companies

Model 2 consist of two company types Company A and Company B. The companies of type Company A can only interact with companies of type Company B, in the same way companies of type Company B can only interact with companies of type Company A.

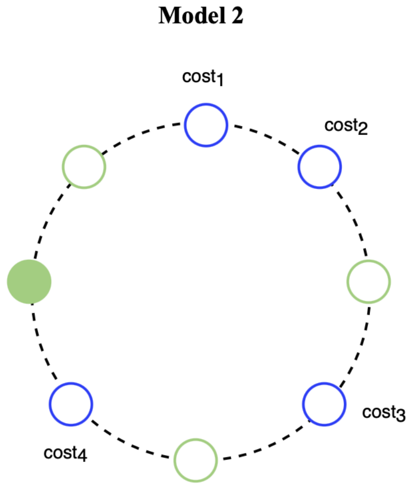


Figure 4.5: ***Illustration of Model 2:*** *The big circle illustrates the environment, and the small circles the companies. As the model consist of both company types, the companies of type Company A is colored in green and the companies of type Company B in blue. The filled green circle is a company of type Company A, that has to make a purchase. Since companies of type Company A can only interact with companies of type Company B, only the blue circles announce their total cost. The companies of type Company A are silence*

As the total number of companies is equal to L, the number of companies of each type is initially $\sim L/2$. This number can change during the simulation if the environment is dynamic. We divide Model 2 into two sub-models, where the environment of Model 2.a is static and the environment of Model 2.b is dynamic. A dynamic environment is illustrated in figure 4.6, where non-rewarded companies are replaced with new companies.

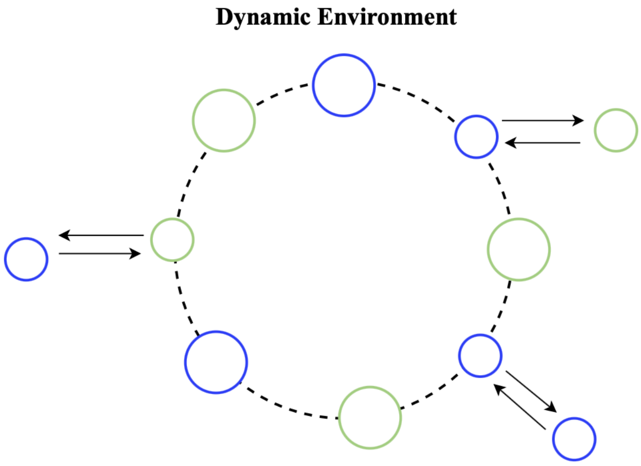


Figure 4.6: ***Illustration of Dynamic Environment:*** *At the end of each time step, non-rewarded companies are replaced with new companies. The type of a new company is determined by a random generator with equal probability, which means that the chance that the new company is type Company A/Company B is* 50%

4.2.4 Model 3: Two Types of Companies with Positive Feedback

Model 3 consists of two company types Company A and Company B, where only companies with opposite type can interact (see figure 4.5). The environment is dynamic, which means that at the end of each time step, non-rewarded companies are replaced with new companies (see figure 4.6). In this model, we add a positive feedback between the two company types. This is obtained by changing the way of choosing the purchasing company (see figure 4.7 and appendix C.8). In the Main Model (Model 1 and Model 2), a random generator select the position of the purchasing company (position X) by choosing between all the company positions $[0, 1, 2, .., L-1]$ with equal probability. In Model 3, a random generator starts by choosing the type of the purchasing company with equal probability. When the type is determined (type Company X), a list containing the company sizes for all companies with type Company X is given to a random generator. This generator will choose the position of the purchasing company with probability that is proportional to the size of a company. This means that larger companies have a greater chance of being selected to make a purchase. This setup is based on the assumption that larger companies have a higher production and therefore their need for intermediates are higher.

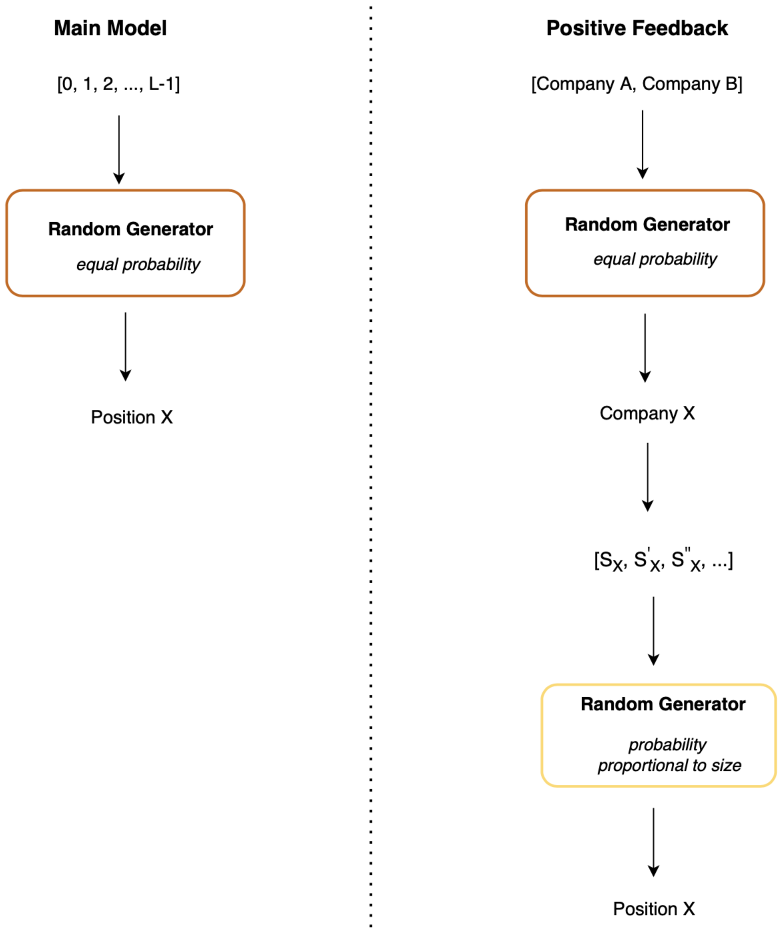
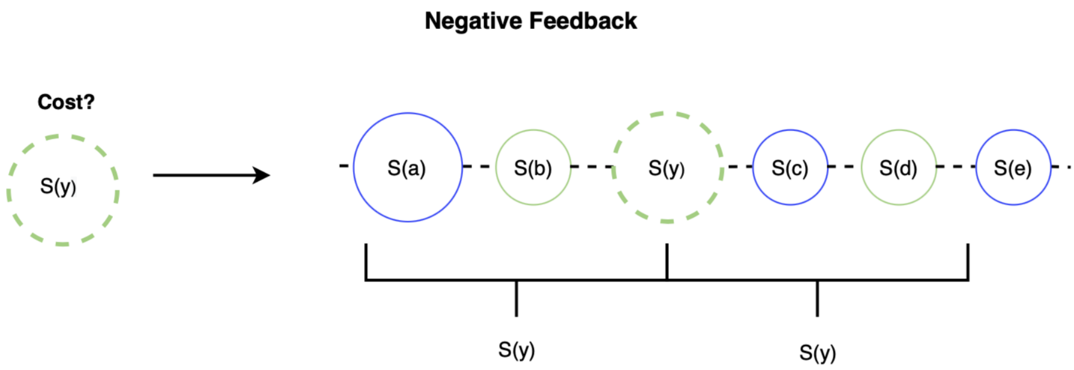


Figure 4.7: ***Positive Feedback (The Purchasing Company)***: *The flowcharts show the different ways to select the purchasing company.* ***Main Model***: *A random generator choose the position of the purchasing company (position x) by choosing between all company positions $[0, 1, 2, .., L-1]$ with equal probability.* ***Model 3***: *A random generator choose the type of the purchasing company with equal probability. When the type is determined (Company X), a random generator is given a list containing all sizes of companies with type Company X. This generator will choose the position of the purchasing company with probability that is proportional to the size of a company*

4.2.5 Model 4: Two Types of Companies with Positive and Negative Feedback

Model 4 is similar to Model 3 with a negative feedback. Therefore, there exist both positive and negative feedback between the companies. The positive feedback is illustrated in figure 4.7, and the negative feedback in figure 4.8 (see appendix C.9).



$$S_{factor} = S(a) + S(b) + S(c) + S(d)$$

$$cost = \big(log_{10}(S_{factor}) + 1\big) \cdot S(y)^{\gamma} + |x - y|\sigma$$

Figure 4.8: ***Negative Feedback (The Total Cost):*** *The total cost of a selling company at position y with size $S(y)$ is dependent on company sizes of companies (with any type) within the distance $S(y)$. In the illustration, the companies at positions a, b, c and d with sizes $S(a)$, $S(b)$, $S(c)$ and $S(d)$ are within this limit, thus $S_{factor} = S(a) + S(b) + S(c) + S(d)$. The production cost for the company at position y is then $cost = \big(log_{10}(S_{factor}) + 1\big) \cdot S(y)^{\gamma} + |x - y|\sigma$*

The negative feedback occurs in the calculation of the total cost for each of the selling companies. The total cost for a company at position y with size $S(y)$ at time steps t is:

$$cost = \big(log_{10}(S_{factor}) + 1\big) \cdot S_{t-1}(y)^{\gamma} + |x - y|\sigma \tag{4.3}$$

The negative feedback is given as a factor multiplied by the production cost. This factor depends on the company sizes of companies within the distance $S(y)$ from the company at position y. The factor is the company sizes added:

$$S_{factor} = \sum_{i=y-S_{t-1}(y)}^{y+S_{t-1}(y)} S_{t-1}(i) \tag{4.4}$$

Where $S_{t-1}(i)$ is the company size for the company at position i and time step t.

When the total number of companies is equal L, the company positions are ranged from 0 to $(L-1)$ (0-indexing (see section 4.1)), and as the model is equipped with periodic boundary conditions (see figure 4.2), the following is used for the sum in equation 4.3:

$$if \quad 0 > y - S(y) \quad : \quad y - S(y) \rightarrow L - |y - S(y)|$$

$$if \quad y + S(y) > (L-1) \quad : \quad y + S(y) \rightarrow |L - \big(y + S(y)\big)|$$

The negative feedback is obtained since areas with large companies or a large number of companies will increase the demand of labor and thus force an increase in the wage rate of labor [6]. This increase in wage rate, will increase the cost of production, hence the factor in equation 4.3 is multiplied with the production cost S^{γ}.
In *the Classical Theory of Employment* the supply of labor is determined by the real wage rate (the wage rate divided by the price of a good) [10].

$$N = F(w/p) \quad \Rightarrow \quad w/p = F^{-1}(N) \tag{4.5}$$

Where N is the labor supply, w is the wage rate and p is the price of a good. The equation is assumed for full employment, which means that all who wants to work at the given real wage rate can find a job.

We assume that $p = 1$ and no inflation or deflation exist in the model, thus the price of a good is always constant. We also assume that the size of a company is equivalent to the number of company employees, and for this reason we will set:

$$S(y) = N(y) \tag{4.6}$$

Where $S(y)$ and $N(y)$ respectively is the company size and the number of company employees at position y.

This means that the labor supply N within the distance $y \pm S(y)$ is:

$$N = S_{factor} \tag{4.7}$$

and therefore:

$$w = F^{-1}(S_{factor}) \tag{4.8}$$

The function F is unspecified in economic literature [5]. However, it seems plausible to assume that a small increase in wage rate will yield a relatively large increase in labor supply. We will obtain this relation with the function:

$$F^{-1}(S_{factor}) = log_{10}(S_{factor}) + 1 \tag{4.9}$$

1 is added to achieve $log_{10}(S_{factor}) + 1 > 1$. This is necessary to guarantee that we always have a negative feedback, and therefore an increased cost; Alternatively scenarios could arise where $log_{10}(S_{factor}) < 1$ and the feedback would be positive with the consequence of lowering the total cost.

4.2.6 Model Summary

From the four models presented above we now have sufficient information to address the research questions given in section 3.2. Our aim is that Model 1 is able to answer Question 1, Model 2 to answer Question 2 and Model 3 + Model 4 are able to answer Question 3.
To get an overview over the models, some of the conceptions are summarised in table 4.1.

Model	Number of Company Types	Interaction Rule	Environment	Feedback
Model 1	1	all	static	
Model 2.a	2	opposite type	static	
Model 2.b	2	opposite type	dynamic	
Model 3	2	opposite type	dynamic	positive
Model 4	2	opposite type	dynamic	positive and negative

Table 4.1: ***Models:*** *Model summary*

4.3 Model Parameters

The model includes five different parameters γ, τ, σ, L and T. Definitions and limitations of the parameters are given in table 4.2.

Symbol	Definition	Limitation	Group
γ	economies of scale exponent	$-1 \leq \gamma \leq 0$	$\mathbb{Q}_-$
τ	time it takes to rebuild production apparatus	$\tau \geq 1$	$\mathbb{N}$
σ	transportation cost	$\sigma > 0$	$\mathbb{Q}_+$
L	total number of companies	$L \gg \gamma, \tau, \sigma$	$\mathbb{N}$
T	total number of time steps	$T \gg t_0$	$\mathbb{N}$

Table 4.2: ***Model Parameters****: Definitions and limitations of the model parameters*

The limit of γ is set to meet the definition of economies of scale. With $\gamma < 0$, a company is defined to produce products with increasing return to scale. This means that an increase in production (company size) lowers the production cost of one unit product S^γ. A decrease in the value of γ (to a higher negative value) increases the efficiency of the company.

With $\gamma = 0$, the production cost is not reduced at all, which is why the company is defined to produce with constant return to scale. This case may be true for the agricultural sector, whereas the first case may be realistic for the manufacturing sector [14][8].

The value of τ must be a natural number, as it determines the number of interactions between the companies at each time step. Since new sizes are assigned to the companies after the τ interaction, τ can be interpreted as proportional to the time is takes to rebuild the production apparatus.

The transportation cost is proportional to the value of σ. Since there do not exist negative transportation cost, the value of σ must be positive.

The value of L determines the total number of companies existing in the simulated environment. The value of L is irrelevant, as long as it is much larger than the domain scale set by the other parameters γ, τ and σ [9].

The value of T determines the total number of iteration of the simulation. This value must be much larger than the length of the warm up state t_0 (t_0 is defined in section 4.4.2.1).

The parameters can obviously have different values, and these can be combined in different ways. A model with specific parameter choices will be referred to as a *system*. The size of a system is determined with the values of L and T, whereas the dynamic of the system depends on the values of γ, τ and σ.

The values of the dynamic parameters are chosen from 3D plots showing the dynamics of a system. Two examples of non-proper parameter choices are given in figure 4.9. The dynamics of the system shown in panel (a) are too high and we therefore do not see any clear phenomenon. The dynamics of the system shown in panel (b) are too low and we therefore define the dynamics of the system as frozen.

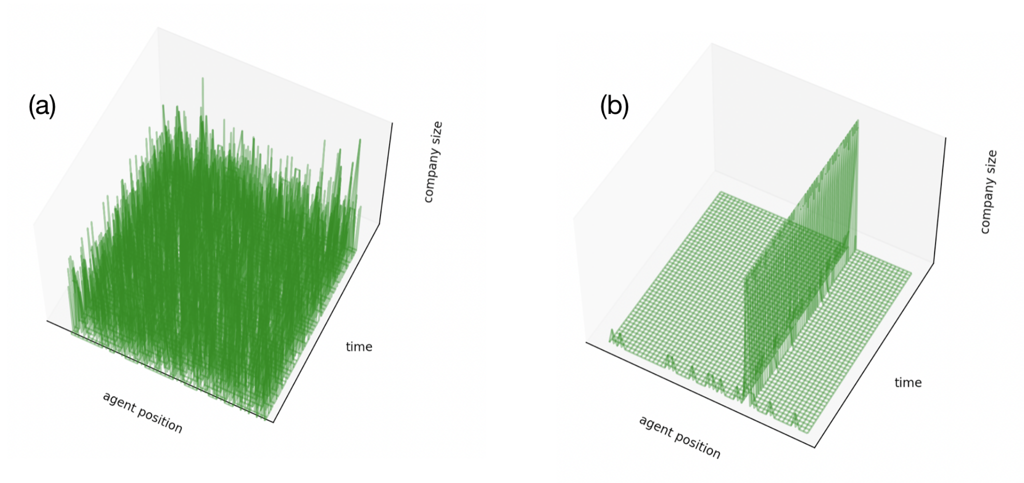


Figure 4.9: ***Non-Proper Values of the Dynamic Parameters****: 3D plots showing the dynamics of two systems. The company positions are given on the x-axis, the number of time step on the y-axis and the company sizes on the z-axis. The dynamics of the system shown in panel (a) are too high, whereas the dynamics of the system shown in panel (b) are too low*

The values of γ and σ are independent on the system size, however the value of τ must be changed when the system size is changed. We therefore use the following expression when referring to the value of τ:

$$\tau = p \times L \tag{4.10}$$

where τ approximately expresses the chances $p \leq 1$ for each company to make at least one purchase between each reassignment of the company sizes.

Figure 4.10 shows two systems with two different system sizes. The values of γ, σ and T are kept fixed, whereas the value of L and τ are changed in the two panels. In panel (a) $L = 50$ and $\tau = 0.4 \times 50 = 20$. In panel (b) $L = 200$ and $\tau = 0.4 \times 200 = 80$. Although the value of L and τ is 4 times greater for panel (b), the percentage is the same; each company has maximum 40% chance to make a purchase between each reassignment of the company sizes.

The dynamics of the two panels are the same. The picture in panel (a) is just replicated 4 times in panel (b), as the system size is 4 times larger.

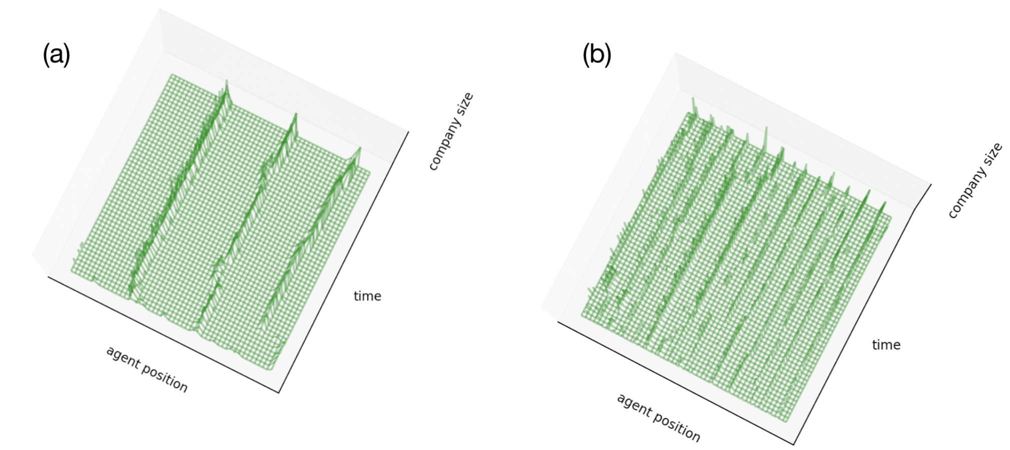


Figure 4.10: ***System Size****: 3D plots showing the dynamics of two systems. The company positions are given on the x-axis, the number of time step on the y-axis and the company sizes on the z-axis. Panel (a):* $L = 50$ *and* $\tau = 0.4 \times 50 = 20$*. Panel(b):* $L = 200$ *and* $\tau = 0.4 \times 200 = 80$*. The dynamics of the two system are the same*

In this thesis we will use two different system sizes $L = 50$ and $L = 200$. The $L = 50$ system size is used to show 3D plots of the system dynamics, whereas the $L = 200$ system size is used for data extraction to increase the number of data points.

4.4 Data Outputs

Agent Based Modeling is interesting because of the phenomenons that emerges as a consequence of the interactions between the agents. In order to understand the phenomenons, it requires that we extract data outputs that describe different aspects of a system. With graphical and statistical analysis of the data, we can understand and discuss the phenomenons that occurs in the models.
We divide this section into two parts; In section 4.4.1 we present the variables that we believe describe the phenomenons that occur in our models and on which we perform statistical analysis. In section 4.4.2, we describe how data outputs of the variables are extracted and processed.

4.4.1 Variables of Interest

Company Size — The size of a company is given with the value of S. The size indicates how much the company has been rewarded, and thus how significant the company is in the system.
The mean and distribution of company sizes for companies with $S > 1$ allow us to compare the sizes of each company type. This will get us an understanding of which type being the most dominant in the system.

Number of Companies — The number of companies will be given as a count of the number of companies with $S > 1$. By comparing the count of each company type, we get an estimation of the distribution of each company type within the simulation environment.

Company Lifetime — The lifetime of a company is defined as the number of time steps a company can remain a size $S > 1$. The lifetime of each company is initially, $t = 0$, set to be equal 0. If the size of a company increases to $S > 1$, a count will start. For each time step the company can remain a size $S > 1$ one count will be added to its lifetime. If the size of the company drops to $S = 1$, the lifetime will be saved and again set to be equal 0.
The lifetimes of companies will give an indication of the dynamics and stability of the system. If the lifetimes are short, the dynamics of the system are high and thus the system is unstable. If the lifetimes are long, the dynamics are low and the system is stable.

Distance between Companies — The distance between companies refers to the shortest distance between a company of type X with $S > 1$ and a company of type Y with size $S > 1$. X and Y can be equal to or different from each other.

The distance between companies indicates how close the companies prefer to be. We assume that if the distances between companies are relatively small, then there is a positive feedback between the companies. If the distances are relatively high, then there is a negative feedback between the companies.

4.4.2 Steady State, Statistics and Correlation

4.4.2.1 Steady State

In section 4.2.1.4, we defined the models as non-terminal simulation models. For such models, we must be able to distinguish between the warm up state and the steady state behavior for each system of interest. The transition between the two states will be given as a simulation time step t_0, where there for any given time step $t \in T$ applies:

$$\begin{aligned} t < t_0 &\quad \text{simulation in warm up state} \\ t \geq t_0 &\quad \text{simulation in steady state} \end{aligned} \tag{4.11}$$

The value of t_0 is used to determine two things; the time step at which we can start extracting simulation data outputs, and the total number of simulation time steps T. The extraction of data outputs must start after the simulation has completed the first t_0 iterations. By extracting data before the simulation has reached a steady state behavior (in the warm up state of the simulation), transient data will be included in the data set. This will introduce bias to the results [3]. The total number of time steps must be much larger than the value of t_0, i.e. $T \gg t_0$.

For this use of t_0 it is not necessary to know the exact time step for the transition, but only to ensure that the system has reached a steady state behavior. We can therefore easily settle with an *estimation* of t_0. For this estimation we use *Welch's Method*. This method is a simple graphical procedure, that uses the moving average of a variable of interest to look for the point in a time series at which the moving average flattens out [3]. This point in the time series gives the time step for the transition between the warm up state and the steady state, and thus the value of t_0.

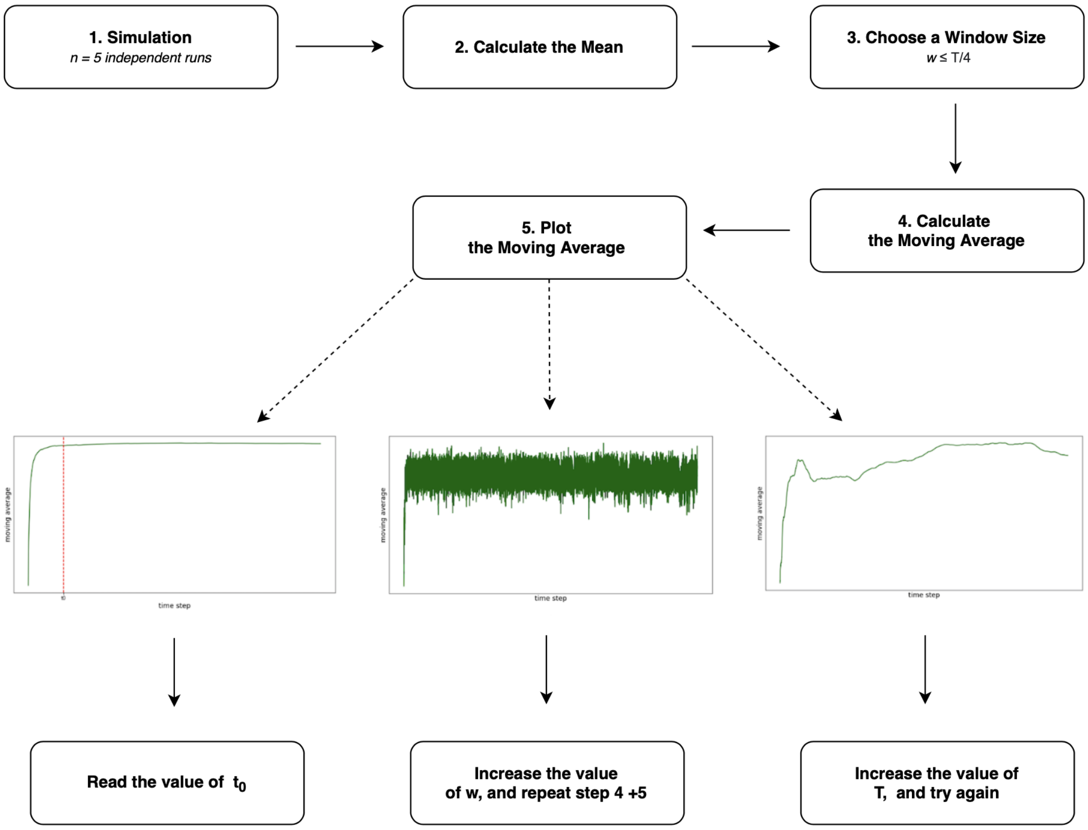


Figure 4.11: ***Flowchart showing Welch's Method:*** *Welch's method is divided into 5 steps. Step 1, retrieving $n = 5$ time series from simulations. Step 2, calculating the mean. Step 3 + 4, calculating the moving average for a given window size. Step 5, plotting the moving average. The plot of the moving average can have one of three outcomes. Left plot: The value of t_0 can be read on the x-axis. Middle plot: The value of w must be increased. Right plot: The value of T must be increased.*

We use the moving average of the sum of company sizes of companies with $S > 1$. The flowchart of the procedure is shown in figure 4.11. In step 1, we run $n = 5$ independent simulations for T time steps, where the value of T is much higher than the expected value of t_0. By repeating the simulation $n = 5$ times, we can compensate for possible variations in the output [3]. In step 2, we use equation 4.12 to calculate the mean value for each time step.

$$\bar{S}_t = \frac{1}{n} \sum_{i=1}^{n} S_{ti} \quad \text{for} \quad t = 0, ..., (T-1) \tag{4.12}$$

Here S_{ti} is the sum of company sizes at time step t in the i'th run (repetition).

In step 3 and 4, we use equation 4.13 to calculate the moving average for a value of the window size $w \leq T/4$ (see appendix C.11).

$$\bar{S}_t(w) = \begin{cases} \bar{S}_0 \quad \text{for} \quad t = 0 \\ \frac{1}{2t}\sum_{m-t}^{t} \bar{S}_{t+m} \quad \text{for} \quad 1 \leq t \leq w \\ \frac{1}{2w+1}\sum_{m=0}^{2w} \bar{S}_{t-w+m-1} \quad \text{for} \quad w+1 \leq t \leq (T-1) - w \end{cases} \tag{4.13}$$

In step 5 the moving average is plotted as a function of time steps. This plot can have one of three outcomes: The plot to the left, the plot in the middle or the plot to the right. Left plot: The curve for the moving average increases and then flattens out. The red dashed line illustrates the transition between the warm up state and the steady state, with the warm up state on the left side of the line and steady state on the right side. The value of t_0 can be read from the x-axis. Middle plot: The curve of the moving average is fluctuating, and it is therefore hard to determine when the system is in steady state. In such situations, we increase the value of w which will give us a more smooth curve. We increase the value of w slowly to ensure that we use the smallest possible value of w. Right plot: For this case, the window size used in the calculation of the moving average is $w = T/4$ (see equation 4.13). The curve is slightly fluctuating, but not completely smooth. We increase the value of simulation time steps, and repeat the procedure.

From Welch's Method, we are able to determine the value of t_0. The saving of data outputs starts after the simulation has completed the first t_0 iteration and thus after the system has reached a steady state behavior. We will refer to the time step at which we start saving data as $t_{data} \geq t_0$. This means that for each variable of interest, a time series with sample size $k = (T-1) - t_{data}$ is returned at the end of the simulation:

$$z^k = (z_0, z_1, ..., z_k) \tag{4.14}$$

with z_0 being the data point at time step t_{data}, and z_t being the t'th data point in the time series z^k.

4.4.2.2 Statistics

In section 4.2.1.4, we defined the models as stochastic simulation models as some of the input processes are driven by random generators. Having these random elements in our simulation, leads to data outputs being random. Therefore runs of the simulation will be *realizations* of the system performance [2]. In order to analyse the realizations statistically, we will use the Law of Large Numbers and thus we need a lot of data. We will obtain this by running the simulation 50 times longer than the length of the warm up before steady state occurs [1], and repeat this n times:

$$\begin{aligned} z_1^k &= (z_{10}, z_{11}, ..., z_{1k}) \\ z_2^k &= (z_{20}, z_{21}, ..., z_{2k}) \\ &... \\ z_n^k &= (z_{n0}, z_{n1}, ..., z_{nk}) \end{aligned} \tag{4.15}$$

Where $n = 10$ is used to calculate the mean value of a variable, and $n = 100$ for histograms showing the distribution of values of the variable.

The n time series are independent from each other, however each time step in a single time series depends on the previous time step in the same time series. In order to infer statistical conclusions from the data, we will consider the dependent data as independent. Viewing this data as independent is necessary to better perform statistical analysis, however the quality of the statistical result will not be optimal. If it turns out that correlation (see section 4.4.2.3) of the data within a run is large, another approach to the statistical analysis could be to increase the number of runs and decrease the number of time steps. However, the scope of this thesis will be limited to the initial approach.

4.4.2.2.1 Mean of Variable

The mean of a variable is used to create phase diagrams that must be able to describe a system with one point. The mean of a variable is calculated by taking the mean of the time steps in a single

[1] Since the models have different values of t_0, we choose that T should be 50 times greater than t_0. This is only chosen to have a way of choosing the value of T

time series and then taking the mean of the mean of the time series (see appendix C.12):

$$\bar{z} = \langle\langle z\rangle_t\rangle_i = \frac{1}{n}\frac{1}{k}\sum_{i=1}^{n}\sum_{t=0}^{k} z_{it} \tag{4.16}$$

Where $k = t_{data}(50-1)$ and $n = 10$.

For our use of the mean of a variable, we do not find it necessary to define the standard deviation.

4.4.2.2.2 Histogram

Histogram for a variable is created by using all time steps in all time series as one big data set with size nk (see appendix C.13):

$$Z^{nk} = (z_{10}, z_{11}, ..., z_{1k}, z_{20}, z_{21}, ..., z_{nk}) \tag{4.17}$$

Where $k = t_{data}(50-1)$ and $n = 100$.

Alternatively, to avoid the dependence within a run, we could have used the mean value for each time series as a point in the data set Z^{nk} [3]. However, this would lead to loss of information as the output values are often discrete and taking the average would often give a decimal result.
For histograms, we will reveal the number of entries, the mean and the standard deviation. The mean and the standard deviation is calculated as if Z^{nk} was a set with nk independent data points. If a histogram is exponentially distributed we will, per definition, set the standard deviation to be equal to the mean.

$$\mu = \frac{1}{nk}\sum_{j=1}^{nk} Z_j \tag{4.18}$$

$$\sigma_\mu = \sqrt{\frac{1}{nk}\sum_{j=1}^{nk}(Z_j-\mu)^2} \tag{4.19}$$

Where Z_j is the j'th element in the data set Z^{nk}.
To avoid a misrepresentation of the mean due to the dependency of the data points, we will present the mean as $\mu \pm \sigma_\mu/\sqrt{n}$ instead of $\mu \pm \sigma_\mu/\sqrt{nk}$.

Histogram of Lifetimes

Figure 4.12 shows two histograms of the same company lifetime distribution, but with two different binning. The top histogram shows that the lifetimes are power-law distributed, which means that the chosen bin size at 25 is not very useful. The problem is indeed the power-law distribution of the data, where many of the companies have short lifetimes between 1 and 200 (the peak of the histogram) and the remaining lifetimes are distributed between lifetime 200 and 1771. The latter is a quite large interval. This means that a bin size of 25 is too small for the short lifetimes, where we have good statistics, and too large for the long lifetimes. This problem can be solved with logarithmic binning with bin edges a^i, where a is the basis for the bins and i is the bin number [18]. For the bottom histogram in figure 4.12, we have chosen the basis $a = 2$ for the bins (see appendix C.14). The bin edges will therefore be 2^0, 2^1, 2^2,..., which is 1, 2, 4,...

The lifetimes of the companies in our models are all power-law distributed hence the histograms for the lifetimes are logarithmic binned.

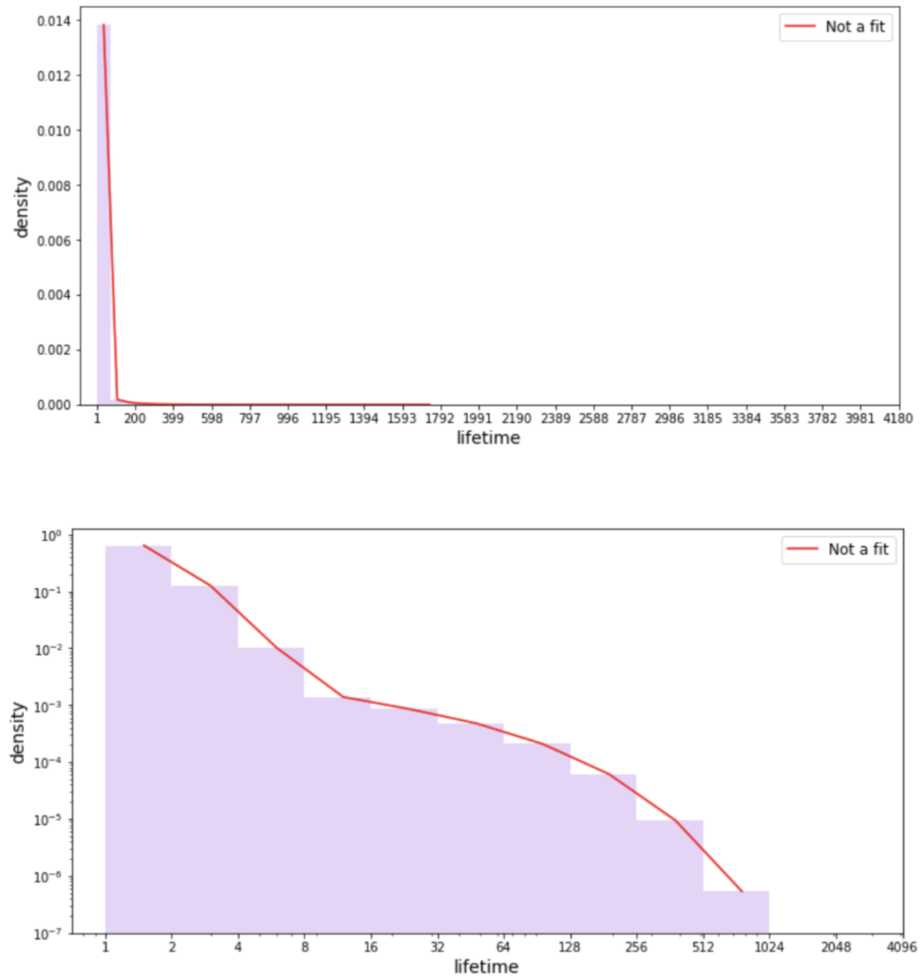


Figure 4.12: ***Binning of histogram for the distribution of the lifetimes****: Data output of the lifetime is shown with two histograms with two different binning. The upper histogram has bin size* 25 *and the lower histogram has logarithmic binning.*

4.4.2.3 Time Correlation

As mentioned previously, dependent data is treated as independent. The degree of dependency between each time step in a time series can we examined from the time correlation of a system. By plotting the time correlation as a function of time lag, we can get an understanding of how great an impact this can have on the statistical analysis.

The interest in the time correlation of a system is to have a mathematical representation of the degree of similarity between a given time series and a lagged version of itself over successive time intervals. In our calculation of the time correlation, we use a time series for each position of the company sizes (see figure C.15):

$$C(\Delta t) = \frac{\sum_{x,t} S(t,x)S(t+\Delta t,x)}{\sum_{x,t}(S(t,x))^2} \tag{4.20}$$

With $S(t,x)$ being the company size at position x for time t, and Δt being the time lag. By summing over time and position, we include the positions in our calculation of the time correlation. The denominator has the purpose of normalizing the value of $C(\Delta t)$. This means that the value of $C(\Delta t)$ is between 0 and 1 i.e. $0 < C(\Delta t) < 1$.
The time series contains company sizes after the system has reached steady state. This means that the time series is stationary (or weakly stationary), which is characterized by a constant mean. Thus the correlation depends only on the time lag between two time steps, but independent of the value of the time step. This is the reason why Δt is the only variable for the time correlation [15].
We will calculate the time correlation for the following intervals of Δt: $[0;600[$ with step 1 and $[600;(T-t_{data})[$ with step 200. We have chosen to change the number of steps for the intervals of Δt as the correlation has ceased for the systems of interest for time lag much smaller than 600.
The time correlation is calculated for systems with size $L = 200$ and $T = 10,000$. Due to the randomness of our model, we repeat the simulation $n = 10$ times, and calculate an average value of $C(\Delta t)$ (however, we will still use the notation $C(\Delta t)$ instead of $\langle C(\Delta t)\rangle$):

$$\langle C(\Delta t)\rangle = \frac{1}{n}\sum_{i=1}^{n} C(\Delta t)_i \tag{4.21}$$

Figure 4.13 shows an example of the time correlation as a function of Δt. The correlation decreases exponentially, and approaches asymptotically a specific value. Here it is important to notice that due to our definition of the function for the time correlation, $C(\Delta t)$ never actually reaches 0, even though the correlation has ceased to be. Instead, no correlation should be assumed when the graph has flattened out.

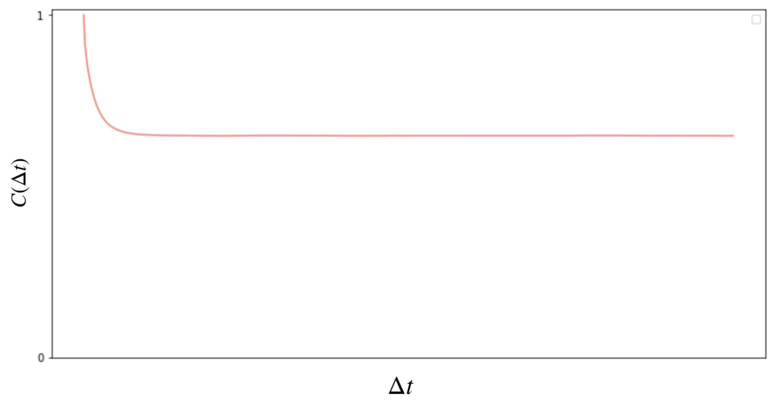


Figure 4.13: ***Example*** *of a graph showing the time correlation as a function of* Δt*.*

In order to the better understand the development of the graph, we fit the data of the time correlation with the following fitting function:

$$f(\Delta t) = a + be^{-\Delta t/\alpha} \tag{4.22}$$

Where a, b and α are fitting parameters. The value of α gives the time interval at which the correlation has decayed with $1/e$, and this is the parameter we are interested in. The fit is made with *iminuit*, which is a tool created for Python by CERN [4] (see appendix C.16).

Chapter 5

Results

In this chapter, we will present the results of the four developed agent based models (see section 4.2). The results are obtained from simulation data outputs of the variables: *Company Size*, *Number of Companies*, *Company Lifetime* and *Distance between Companies* (see section 4.4.1). The data is retrieved after each system of interest have reached a steady state behavior, which was determined with the moving average of time series of the sum of company sizes (see section 4.4.2.1).
The most essential findings of the results are summarized in chapter 6 (the Discussion chapter) as indented paragraphs, where they will be discussed in depth.

5.1 Model 1

5.1.1 Dynamics

The dynamics of Model 1 are shown in figure 5.1 for four different systems. The panels each show a 3D plot of the company sizes for each positions in every time step. A peak indicates a company with size $S > 1$, whereas no peak indicates a company with $S = 1$. Recall that Model 1 is identical to the Sneppen-Bornholdt Model for $h = \infty$. Therefore, some of the dynamics shown in figure 5.1 are similar to those shown in figure 2.1.
Panel (a) in figure 5.1 illustrates that a given company may collapse while other emerge. The emergence of new companies often occurs close to the position of the previously collapsed ones. This shows that when a company disappears, it leaves open a wide business niche because of the cost associated with the distance from local customers to deal with companies farther away in the larger

neighborhood [9].

The difference between panel (a) and (b) is the value of τ. By comparing these, one notice that a low τ destabilizes companies as more companies collapse and emerge. A low τ corresponds to the case where it only costs a few product unit to build a new production facility. Therefore, a higher startup cost will tend to stabilize existing companies [9].

In the case of low τ, the company sizes are in general small. For small company sizes, the production cost S^γ is not reduced much with economies of scale. This means that the difference in production costs for the companies is small, which means that the competition between them is more significant. This is shown in panel (a), where larger companies collapse and new companies emerge.

In panel (b) a high τ allows the companies to grow in size. Thus, the production cost is reduced much with economies of scale, which makes competition between companies less significant. This is shown, as large companies are able to survive throughout the simulation.

By changing the value of the economies of scale exponent from $\gamma = -0.5$ to $\gamma = -0.75$, we can increase the influence of economies of scale on the production cost. This is shown in panel (c), which is similar to panel (a) but with $\gamma = -0.75$. As companies benefit more from economies of scale, competition between them is small and the system is therefore more stable.

In panel (d), we see that a low transportation cost also can stabilize a system with low τ. We see that the system simply needs one large company that can cheaply transport products along the entire circle periphery. This company have monopoly on the supply of Product A.

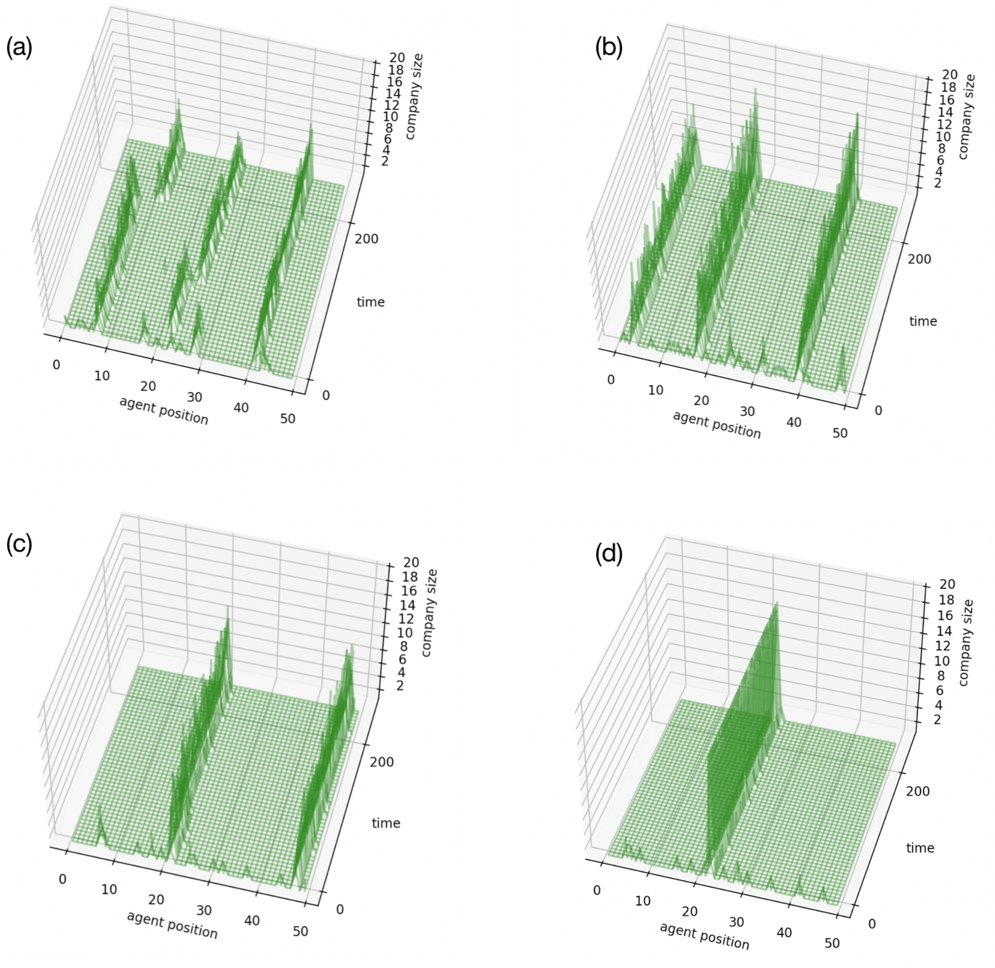


Figure 5.1: ***Dynamics of Model 1***: *3D plots showing the dynamics of Model 1 for* $L = 50$ *companies and* $T = 250$ *time steps.* ***Panel (a)***: $\gamma = -0.5$ *with high noise and high transportation cost* ($\tau = 0.3 \times 50 = 15$, $\sigma = 0.05$). ***Panel (b)***: $\gamma = -0.5$ *with low noise and high transportation cost* ($\tau = 0.5 \times 50 = 25$, $\sigma = 0.05$). ***Panel (c)***: $\gamma = -0.75$ *with high noise and high transportation cost* ($\tau = 0.3 \times 50 = 15$, $\sigma = 0.05$). ***Panel (d)***: $\gamma = -0.5$ *with high noise and low transportation cost* ($\tau = 0.3 \times 50 = 15$, $\sigma = 0.01$)

5.1.2 Steady State

The steady state for each system of interest is found with Welch's Method (see section 4.4.2.1). The results are shown for 48 systems in figure 5.2 and figure 5.3. The remaining are shown in appendix B. Figure 5.2 shows the moving average for systems with $\tau = 0.05 \times L$, whereas figure 5.3 shows for $\tau = 0.5 \times L$. These are the extreme values of τ used in this thesis. In each figure, the moving average is plotted as a function of time steps for 8 different values of σ and 3 different values of γ, where panel (a), (b) and (c) show plots for $\gamma = -0.25$, $\gamma = -0.5$ and $\gamma = -0.75$, respectively.

Each panel has a left and a right plot. The left plot shows the moving average for the first $10,000$ time steps and the right plot for the first 200 time steps. The purpose of the left plot is to show that there only exist one steady state, where the purpose of the right plot is to get a close up of the moving average. Note that the curves for $\sigma = 0.01$ are removed from the left plot to get a better close up.

The values of the moving average depend on the sum of company sizes. Therefore, we see that the curves are shifted along the y-axis. The curves for $\tau = 0.5 \times L$ are higher on the y-axis, as the company sizes are larger.

The curves in figure 5.2 appear more fluctuated compared to those in figure 5.3. The reason is that lower τ destabilize the system and thus fluctuated the value of the moving average (this was also discussed in the discussion of figure 5.1 panel (a) and (b)).

The time step at which a system reaches a steady state behavior seem to depend on the value of τ, σ and γ. It appears that a system with low σ, high τ and high negative γ increases the length of the warm up state t_0 [1]. We assess that all systems (of interest) for Model 1 have achieved a steady state behavior at time step $t_0 \leq 200$ (see appendix B.1, B.2, B.3). We therefore start saving data outputs in time step $t_{data} = 200$.

[1] If one means to observe differently, it is due to the scales on the y-axis; The scales on the y-axis for the left plots are smaller in figure 5.2 than in figure 5.3. This difference in scaling has been forced, as the curves in figure 5.2 are closer together than they are in figure 5.3. We therefore recommend to look at the graphs for the moving average in the appendix B.

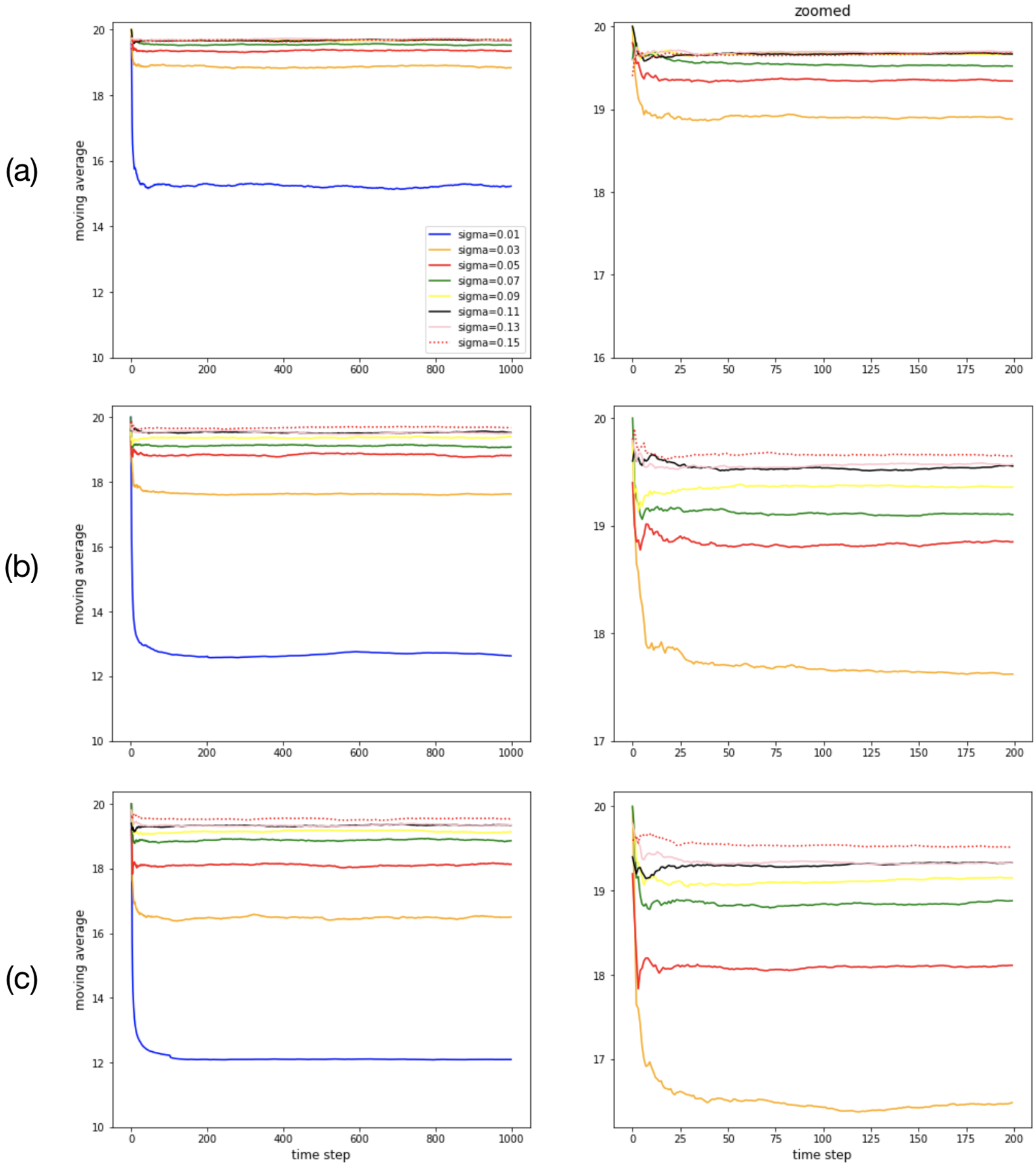


Figure 5.2: ***Model 1: Steady State (low τ):*** *Plots of moving average with $w = 100$ as a function of time steps. All system have size $L = 200$ and $T = 20,000$, and $\tau = 0.05 \times 200 = 10$. Each panel shows system with $\sigma = 0.01$, $\sigma = 0.03$, $\sigma = 0.05$, $\sigma = 0.07$, $\sigma = 0.09$, $\sigma = 0.11$, $\sigma = 0.13$ and $\sigma = 0.15$. Panel (a): $\gamma = -0.25$, panel (b): $\gamma = -0.5$ and panel (c): $\gamma = -0.75$.*

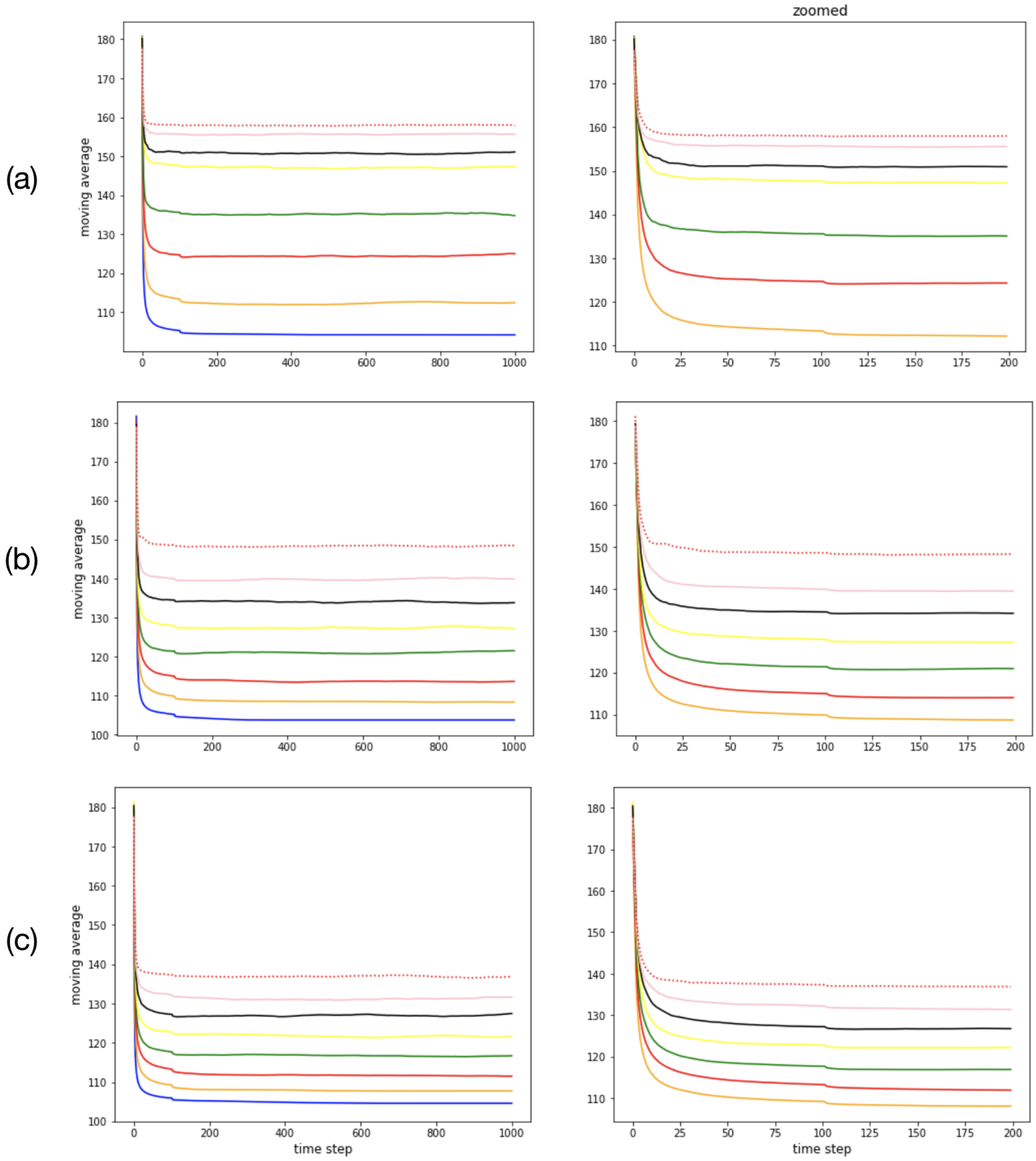


Figure 5.3: ***Model 1: Steady State (high*** τ***):*** *Plots of moving average with* $w = 100$ *as a function of time steps. All system have size* $L = 200$ *and* $T = 10,000$*, and* $\tau = 0.5 \times 200 = 100$*. Each panel shows system with* $\sigma = 0.01$*,* $\sigma = 0.03$*,* $\sigma = 0.05$*,* $\sigma = 0.07$*,* $\sigma = 0.09$*,* $\sigma = 0.11$*,* $\sigma = 0.13$ *and* $\sigma = 0.15$*. Panel (a):* $\gamma = -0.25$*, panel (b):* $\gamma = -0.5$ *and panel (c):* $\gamma = -0.75$*.*

5.1.3 Time Correlation

In this section we will look at the time correlation for different systems. More specifically, we will present graphs of α, the time interval at which the correlation has decayed with $1/e$, as a function of different values of the model parameters.

The time correlation for systems in steady state are shown in figure 5.4 for different values of σ. The figure consists of two plots, where the plot to the left shows the correlation for all time lags and the plot to the right shows the correlation for $0 \leq \Delta t \leq 200$.

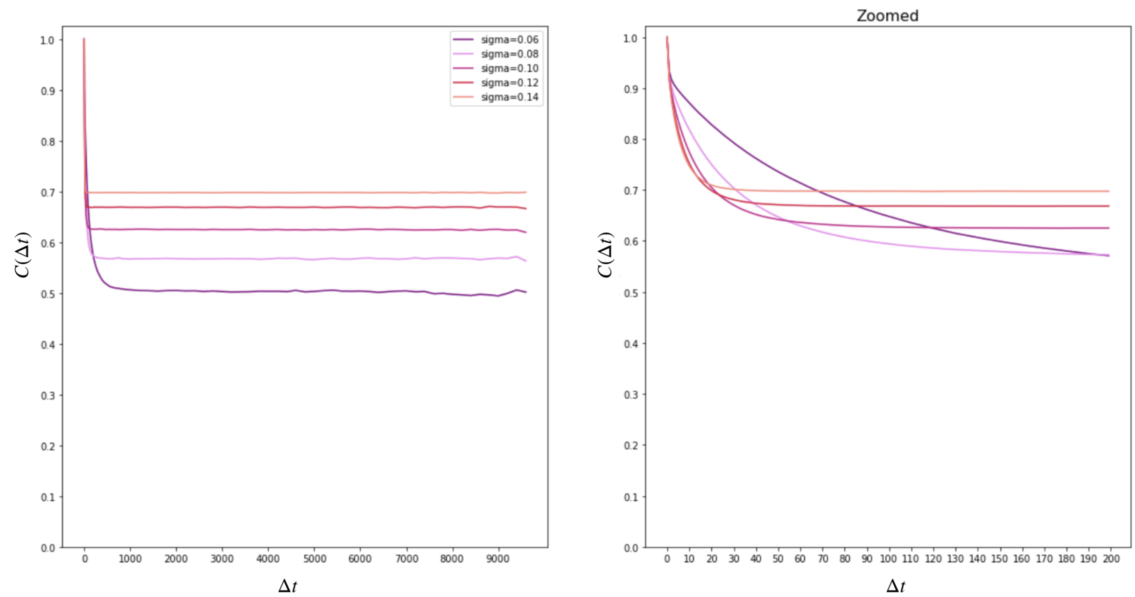


Figure 5.4: ***Model 1: Time Correlation****: Time correlation as a function of Δt for systems with size $L = 200$ and $T = 10,000$. The correlation is shown for systems with $\gamma = -0.75$ and $\tau = 0.3 \times 200 = 60$ for $\sigma = 0.06$, $\sigma = 0.08$, $\sigma = 0.10$, $\sigma = 0.12$ and $\sigma = 0.14$*

The figure shows that the curves of the time correlation all decay exponentially and reach a horizontal asymptote. Usually for correlation functions, the horizontal asymptote is at $C(\Delta t) = 0$, which indicates no correlation. However, due to our definition of the time correlation (see equation 4.21) the horizontal asymptote is between $C(\Delta t) = 0.5$ and $C(\Delta t) = 0.7$ depending on the company sizes included in the calculation of the correlation. The system should be assumed as uncorrelated when the curve reaches the asymptotic behavior.

To understand the behavior of the time correlation, the curves in figure 5.4 are fitted with the exponential function given in equation 4.22. An example of a fit is given for $\sigma = 0.12$ in figure 5.5.

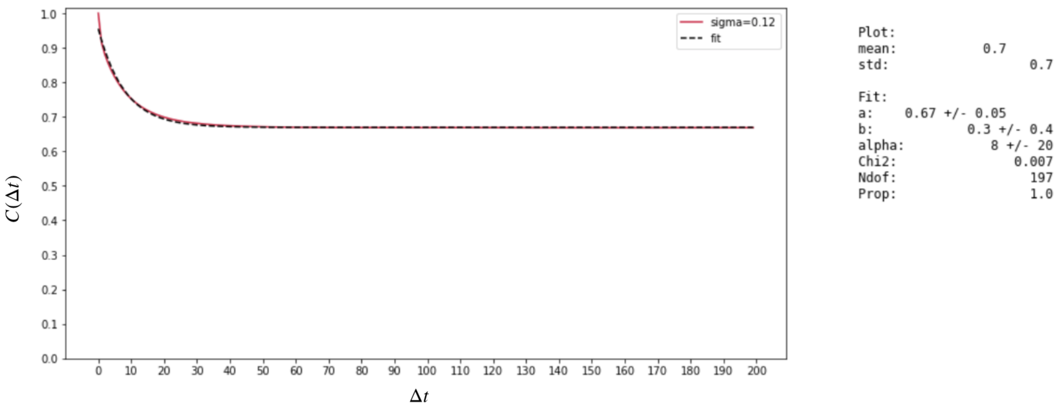


Figure 5.5: ***Model 1: Fit Curve of Time Correlation**: The curve of the time correlation for $\sigma = 0.12$ from figure 5.4 is fitted with the exponential function given in equation 4.22. The curve is plotted with a red line and the fit is plotted with a black dashed line.*

Iminuit estimates the fitting parameters to be $a = 0.67 \pm 0.05$, $b = 0.3 \pm 0.4$ and $\alpha = 8 \pm 20$. The parameter of interest is α, the time interval at which the correlation has decayed with $1/e$. We notice that the uncertainty of α exceeds the value - and this problem recurs for all fits. However, we choose to continue with the method as the fitted line appears to be a good fit for the correlation curve.

We perform a fit for each curve in figure 5.4 to be able to determine the value of α. The values of α are plotted in figure 5.6 for different values of σ. In this figure, we notice that the values of α for $\sigma = 0.01$ and $\sigma = 0.02$ exceed the value of simulation time steps $\alpha > T = 10,000$. This result suggests that the correlation of the two systems did not decay with $1/e$ before the end of the simulation, and thus we should have run the simulations longer than $T = 10,000$. These two values of α must be discarded.

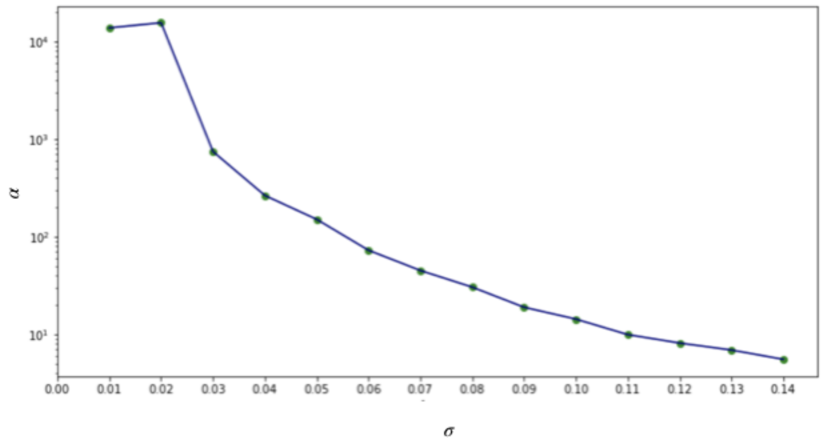


Figure 5.6: ***Model 1:*** α ***as a function of*** σ: *Plot of the correlation time interval as a function of* σ. *The values of* α *are found by fitting the curves in figure 5.4 with the exponential function given in equation 4.22.*

In figure 5.6, we see that the value of α decreases with an increase in σ. To understand this behavior, we show 3D plots for two systems in figure 5.7. The left 3D plot shows the dynamics of a system with low transportation cost ($\sigma = 0.06$) and the right of a system with high transportation cost ($\sigma = 0.14$). The plots are marked with red lines, indicating when the correlation of a system has decreased with $1/e$, i.e. the value of α. The value of α is ~ 72 for the system with $\sigma = 0.06$, and ~ 6 for the system with $\sigma = 0.14$.

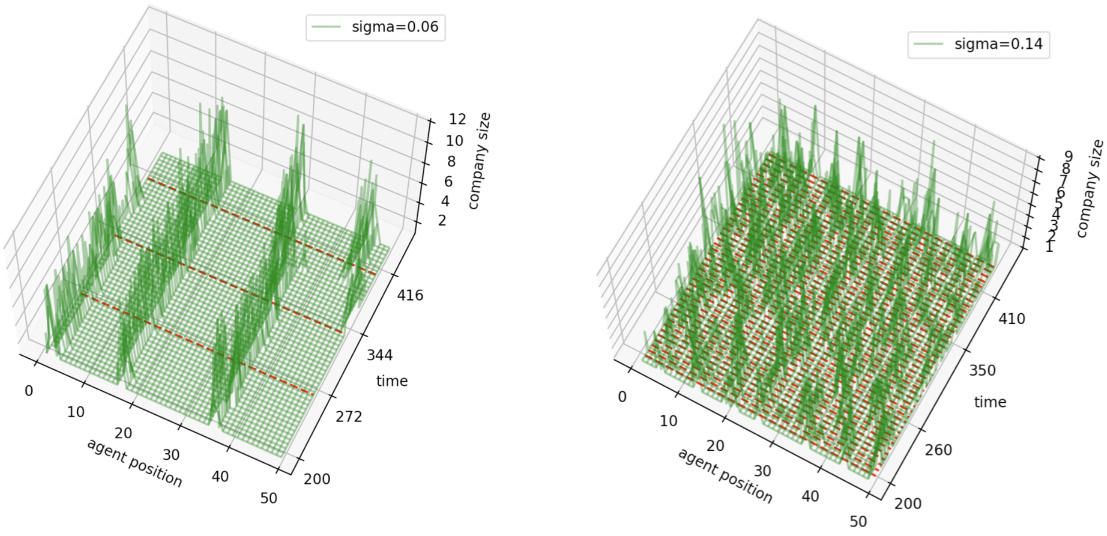


Figure 5.7: ***Model 1: Dynamics for different values of*** σ: *3D plots of the dynamics for systems with* $\sigma = 0.06$ *and* $\sigma = 0.14$. *The remaining parameters are kept fixed:* $L = 50$, $\gamma = -0.75$, $\tau = 0.3 \times 50 = 15$. *The red dashed lines mark the value of the time interval* α, *where the correlation has decreased with* $1/e$

The correlation between each time step for the system with low transportation cost appears higher, as the changes in company sizes and positions are more slow. This is compared to the dynamics of the system with high transportation cost, where the changes appear more rapid. It is therefore a matter of course that α is lower for $\sigma = 0.14$ than for $\sigma = 0.06$, as the correlation decreases faster. The result indicates that for high values of σ, the dynamics appear more disordered thus the time interval of the correlation α is lower, but for decreasing σ the dynamics become more ordered making α higher.

Figure 5.8 shows plots of α as a function of σ for systems with $\tau = 0.15 \times L$ and for $\tau = 0.6 \times L$, thus we have halved and doubled the value of τ we had in figure 5.6. The plot for the system with $\tau = 0.15 \times L$ is shown in the upper plot, and $\tau = 0.06 \times L$ in the lower plot.

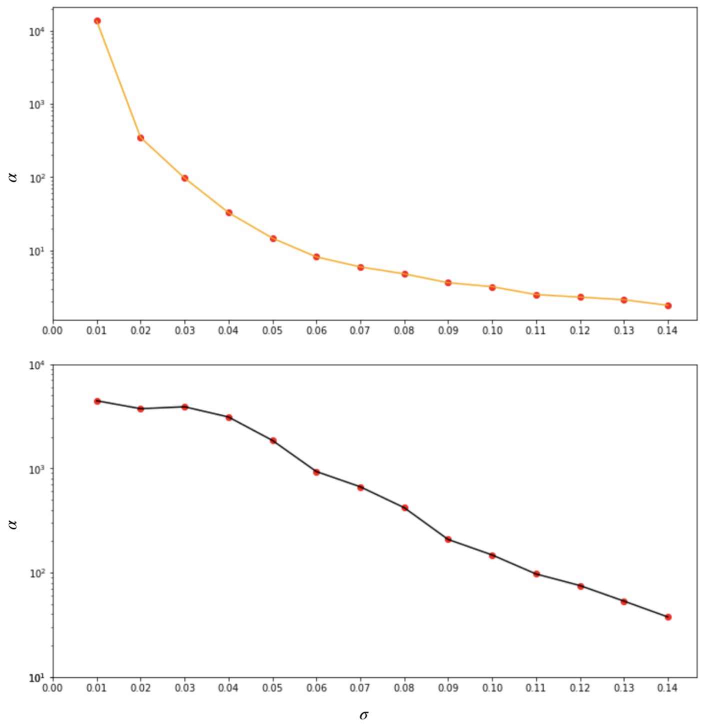


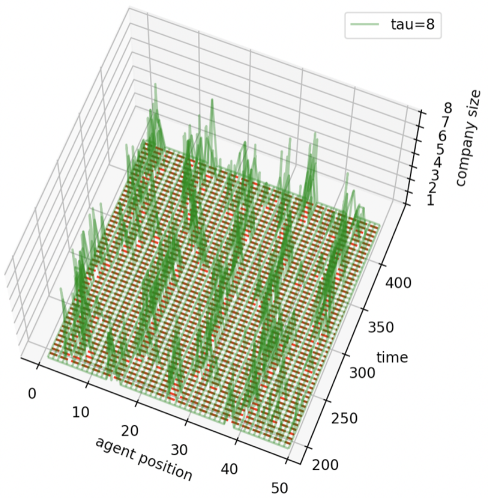
Figure 5.8: ***Model 1:*** α ***as a function of*** σ: *Plots of the correlation time interval as a function of* σ *for systems with* $\tau = 0.15 \times 200 = 30$ *(upper plot) and* $\tau = 0.6 \times 200 = 120$ *(lower plot) The remaining parameters are kept fixed:* $L = 200$, $T = 10,000$, $\gamma = -0.75$

In the upper plot, we see that the value of α exceeds the value of T for $\sigma = 0.01$. We therefore discard this value of α. In the lower plot we discard the data points for $\sigma = 0.01$, $\sigma = 0.02$, $\sigma = 0.03$ and $\sigma = 0.04$, as they do not agree with what we have previously seen for the curves of α.

Similar to figure 5.6, the value of α decreases as a function of σ in both plots. By comparing the slope of the curves, we see that α decreases slower for larger values of τ. We also notice that the values of α are generally higher for larger τ. This means that the correlation time interval is longer for larger τ, thus the system takes longer to become uncorrelated.

This is consistent with that we see in figure 5.9, which shows 3D plots of systems with two different values of τ. The correlation time interval is again marked with red lines, where $\alpha \sim 5$ for $\tau = 0.15 \times L$, and $\alpha \sim 420$ for $\tau = 0.6 \times L$. However, the correlation for the system with $\tau = 0.6 \times L$ does not decrease by $1/e$ within the time frame of the simulation for the 3D plot.

We again see that the correlation increases with the order in the dynamics.



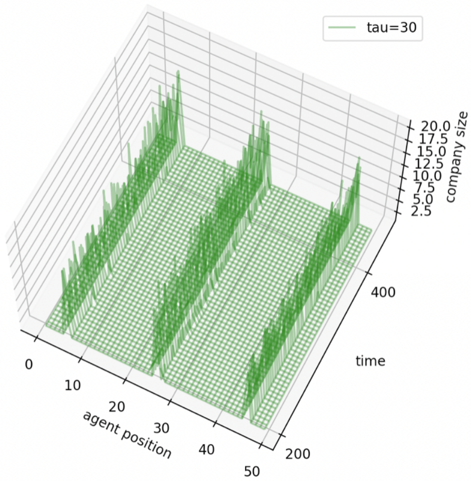


Figure 5.9: ***Model 1: Dynamics for different values of*** τ*: 3D plots of the dynamics for systems with* $\tau = 0.15 \times 50 \sim 8$ *and* $\tau = 0.6 \times 50 = 30$*. The remaining parameters are kept fixed:* $L = 50$*,* $\gamma = -0.75$*,* $\sigma = 0.08$*. The red dashed lines mark the value of the time interval* α*, where the correlation has decreased with* $1/e$

To emphasize the claim of the influence of τ on the correlation, where a higher τ leads to a higher α, we plot in figure 5.10 α as a function of τ. For this plot, we discard the value of α for $\tau = 0.05 \times L = 10$ since it is a clear outlier. The remaining data points follow a straight line, which means that α increases exponentially as a function of τ. We indeed see that higher τ leads to higher α.

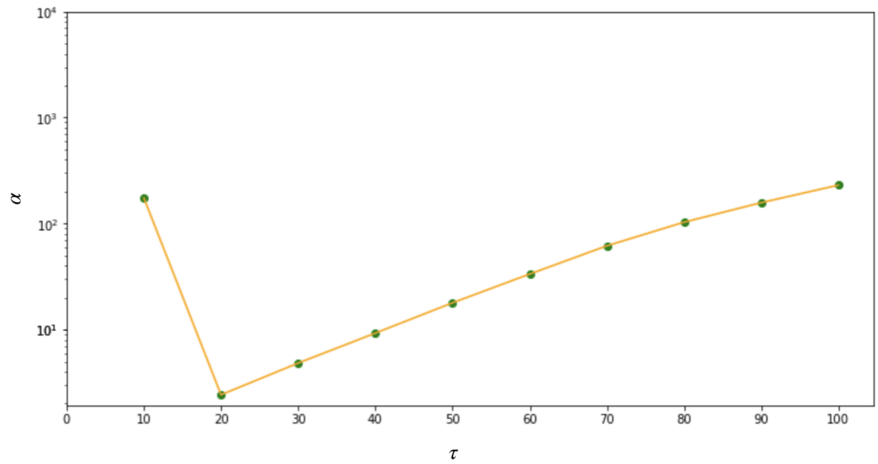


Figure 5.10: ***Model 1:*** α ***as a function of*** τ*: Plots of the correlation time interval as a function of* τ *for systems with* $L = 200$*,* $T = 10,000$*,* $\gamma = -0.75$ *and* $\sigma = 0.08$

Lastly, we will show how γ affects the value of α. In 5.11, a plot is shown for α as a function of σ for systems with $\gamma = -0.25$ (instead of $\gamma = -0.75$). We notice that the curve for the low σ "breaks" (there is a lack of flow), and we therefore discard the data points for $\sigma = 0.01$, $\sigma = 0.02$, $\sigma = 0.03$. By comparing this figure with figure 5.4, we see that the correlation for $\gamma = -0.25$ decreases faster than for $\gamma = -0.75$. In figure 5.12, 3D plots of systems with $\gamma = -0.25$ and $\gamma = -0.75$ are shown. The correlation time interval for $\gamma = -0.25$ is $\alpha \sim 3$, and $\alpha \sim 30$ for $\gamma = -0.75$.
The observation is the same; the dynamics for the system with $\gamma = -0.25$ is higher, thus the system is faster to become uncorrelated.

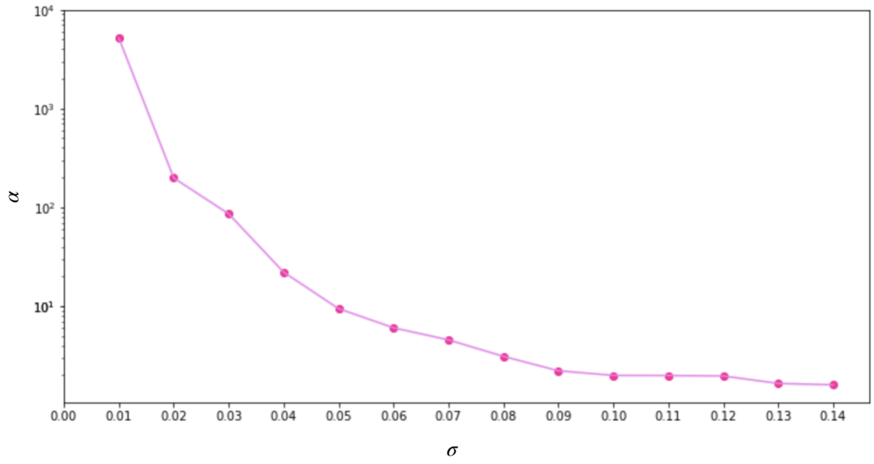


Figure 5.11: ***Model 1: α as a function of σ:*** *Plots of the correlation time interval as a function of σ for systems with $L = 200$, $T = 10,000$, $\gamma = -0.25$ and $\tau = 0.3 \times 200 = 80$*

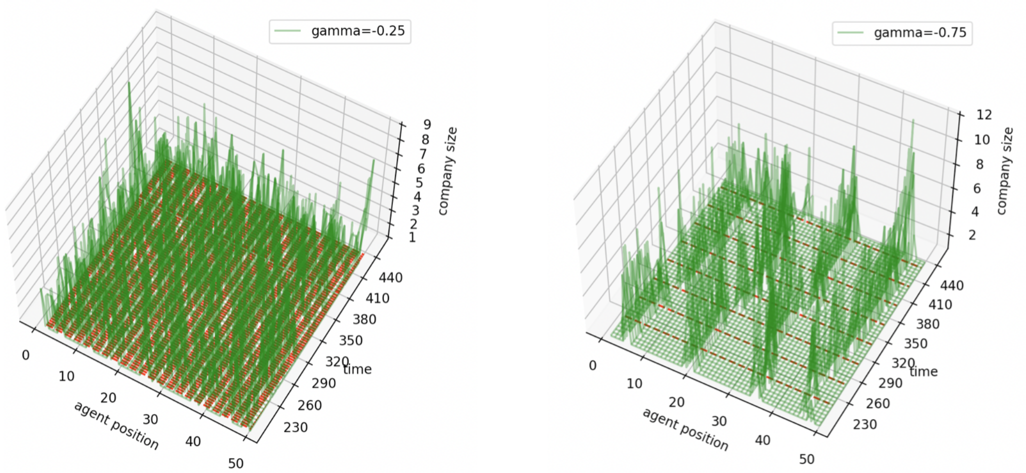


Figure 5.12: ***Model 1: Dynamics for different values of γ:*** *3D plots of the dynamics for systems with $\gamma = -0.25$ and $\gamma = -0.75$. The remaining parameters are kept fixed: $L = 50$, $\tau = 0.3 \times 50 = 15$ and $\sigma = 0.08$. The red dashed lines mark the value of the time interval α, where the correlation has decreased with $1/e$*

Graphs of α were shown as a function of σ in figure 5.8 and as a function of τ in figure 5.10. For the α/σ dependency, the lower graph in figure 5.8 showed that the value of α decreases approximately exponentially as a function of σ. For the α/τ dependency, figure 5.10 showed that the value of α increased approximately exponentially as a function of τ. This result encourage us to try to set up a function that could describe the relationship between the three parameters:

$$\alpha(\tau, \sigma) = e^{\tau/\tau_0} e^{-\sigma/\sigma_0} \tag{5.1}$$

Where σ_0 and τ_0 might be functions of γ, but we have no ideas about the forms of these functions.

In total, we observed a relation between the correlation and the degree of order in the dynamics of a system. We saw that a higher order in the dynamics led to a higher correlation time, and vice versa. The order of the dynamics is high when the changes in company positions and sizes are low, whereas the order of the dynamics is low when the changes are high. Changes in company positions and sizes are related to the lifetime of the companies, since any system with rapid (slow) changes will have shorter (longer) company lifetimes. The lifetime of the companies is discussed in the following section.

5.1.4 Company Lifetimes

The distribution of the lifetimes for companies with size $S > 1$ is shown in figure 5.13 for different values of the transportation cost. The lifetimes are power-law distributed, which is the reason why the histograms are logarithmic binned (see section 4.4.2.2.2). A plot of the mean lifetime as a function of σ is shown in the same figure.
The five histograms show the distributions of the lifetimes for systems with $\sigma = 0.06$, $\sigma = 0.08$, $\sigma = 0.10$, $\sigma = 0.12$ and $\sigma = 0.14$. Common for the histograms is that the density for the shorter lifetimes are much larger than for the longer lifetimes. This indicates that a majority of the companies only manage to maintain size $S > 1$ for a few time steps, before returning back to being an one-man business with size $S = 1$. These companies will not become very large, and will most likely oscillate between size $S = 1$ and $S = 3 \pm 1$.

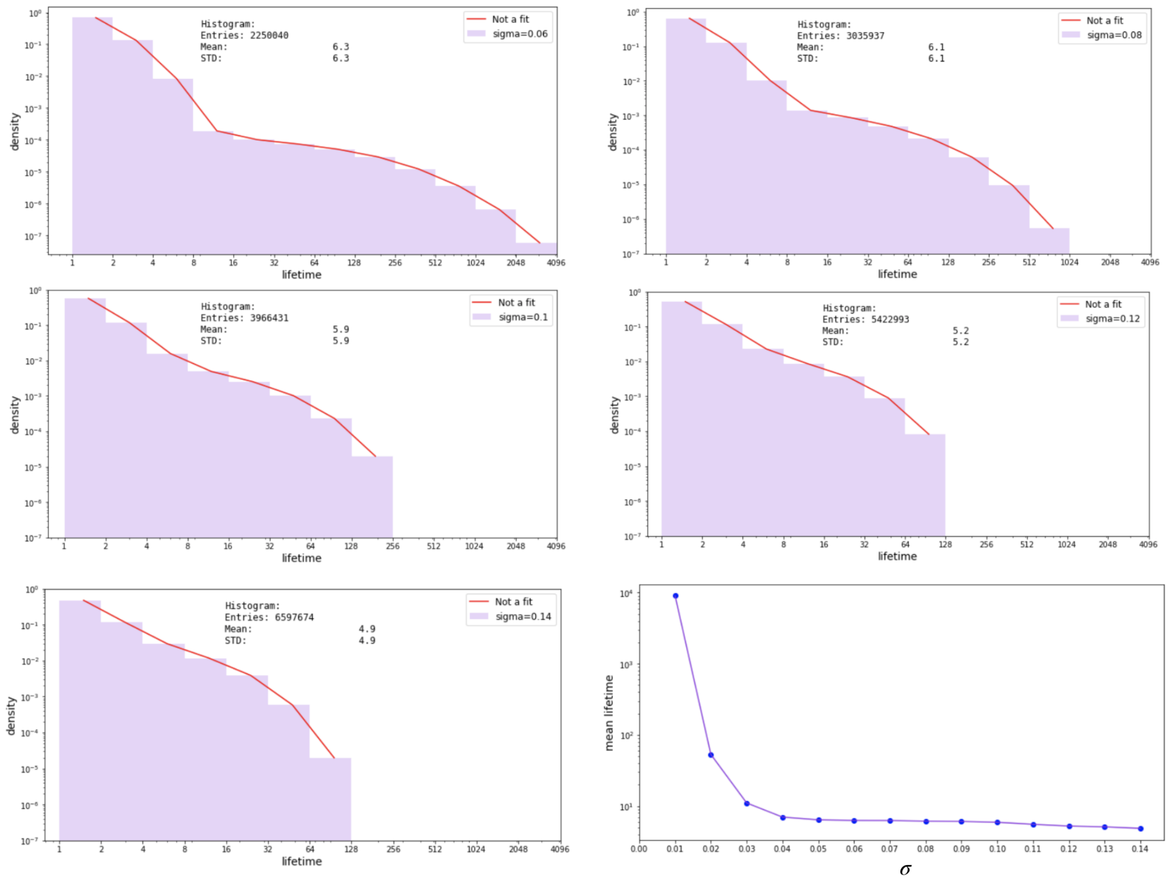


Figure 5.13: ***Model 1: Lifetime of companies with*** $S > 1$: *Histograms of the lifetime of companies with* $S > 1$ *for different values of the transportation cost* $\sigma = 0.06$, $\sigma = 0.08$, $\sigma = 0.10$, $\sigma = 0.12$ *and* $\sigma = 0.14$. *The other parameters are fixed:* $L = 200$, $T = 10,000$, $\gamma_A = -0.75$ *and* $\tau = 0.4 \times 200 = 80$. *The lower right corner shows a plot of the mean lifetime of companies with* $S > 1$ *as a function of* σ

In figure 5.14, we ignore the small unstable companies and instead only look at the lifetimes for companies with size $S \geq 4$. For these histograms, we see that the lifetimes are more smoothly distributed, although there is still a tendency for many companies to have short lifetimes. The mean of the lifetimes is greater, when the small companies are ignored. The reason for this is that larger companies obviously have more flexibility to cope with fluctuations in the company sizes; A large company can survive to be cut in half, whereas a small company will struggle.

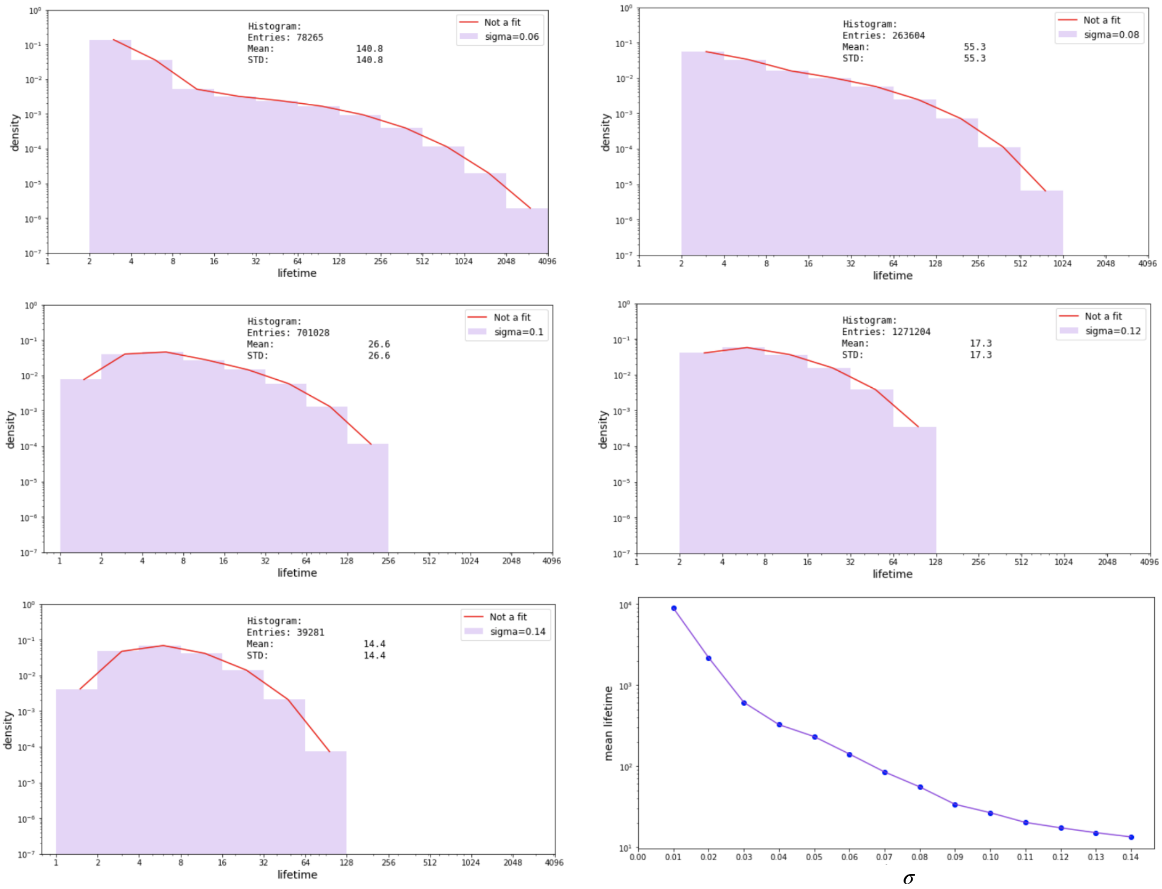


Figure 5.14: ***Model 1: Lifetime of companies with*** $S \geq 4$: *Histograms of the lifetime of companies with* $S \geq 4$ *for different values of the transportation cost* $\sigma = 0.06$, $\sigma = 0.08$, $\sigma = 0.10$, $\sigma = 0.12$ *and* $\sigma = 0.14$. *The other parameters are fixed:* $L = 200$, $T = 10,000$, $\gamma_A = -0.75$ *and* $\tau = 0.4 \times 200 = 80$. *The lower right corner shows a plot of the mean lifetime of companies with* $S \geq 4$ *as a function of* σ

The mean lifetime of the companies decreases with an increase of the transportation costs. This is in agreement with figure 5.6, where α decreases with an increase in σ. This shows that there is indeed a relation between the correlation (dynamics) of a system and the lifetimes of the companies; For short lifetimes, the dynamics of the system is high and therefore the correlation of the system is small. For long lifetimes, the dynamics of the system is small and the correlation of the system is high. This implies that the average lifetime of companies with high σ, low τ and small negative values of γ must be short, whereas the average lifetime of companies with low σ, high τ and larger negative values of γ must be long (see figure 5.6, figure 5.8 and figure 5.11).

The above observations are also illustrated in figure 5.15, showing three phase diagrams of the logarithmic mean lifetime of companies with $S \geq 4$ for systems with different values of τ and σ. The left plot shows the phase diagram for $\gamma = -0.25$, middle plot for $\gamma = -0.5$ and right for $\gamma = -0.75$.

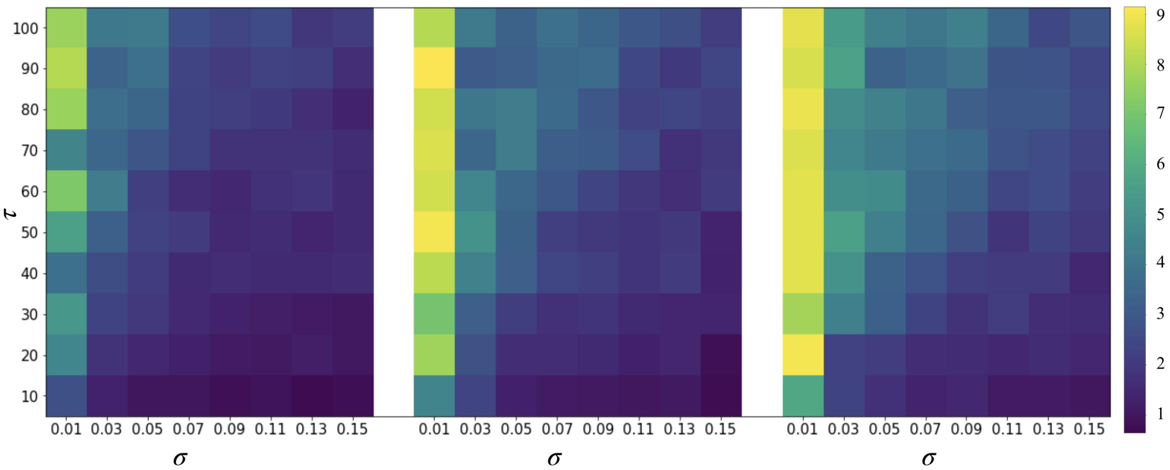


Figure 5.15: ***Model 1: Phase Diagram**: Phase diagrams of the logarithmic mean lifetimes of companies with $S \geq 4$ for systems with different values of τ and σ. Each system has size $L = 200$ and $T = 10,000$. The left plot shows the phase diagram for $\gamma = -0.25$, middle plot for $\gamma = -0.5$ and right plot for $\gamma = -0.75$*

5.1.5 Distances, Sizes and Number of Companies

The distributions of the distances between companies with $S > 1$ and the company sizes for companies with $S > 1$ are shown in figure 5.16 for one system.

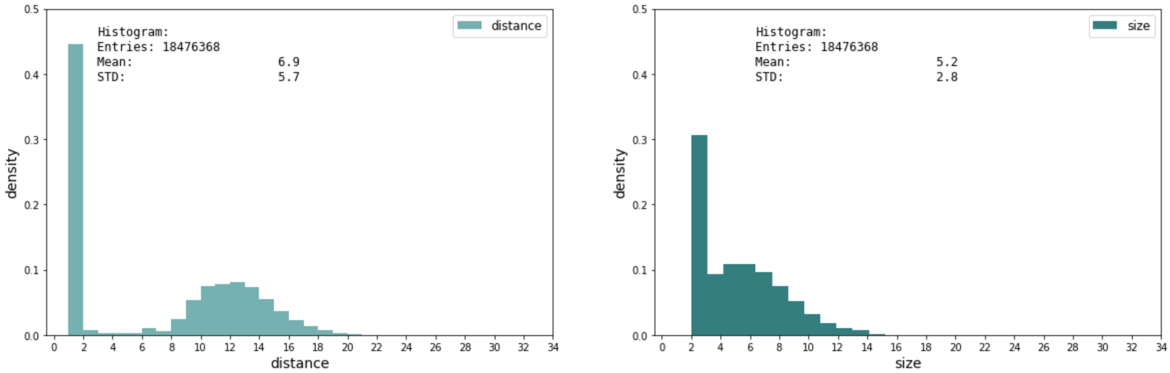


Figure 5.16: ***Model 1: Distances and Sizes of Companies with*** $S > 1$: *Histograms of the distances between companies with* $S > 1$ *and company sizes for companies with* $S > 1$ *for the system* $L = 200$, $T = 10,000$, $\gamma = -0.75$, $\tau = 0.4 \times 200 = 80$ *and* $\sigma = 0.08$

For the distances, we see that $\sim 45\%$ of the entries have a distance equal 1, and the remaining $\sim 55\%$ are distributed between 2 and 31. For the sizes, $\sim 25\%$ of the entries have size equal 2, and the remaining have sizes between 3 and 24. This could indicate that there is a correlation between the number of companies having $S = 2$ and the density of distances equal 1.

In figure 5.17, we only included companies with $S > 2$. In the distribution of the company sizes, obviously, the peak for $S = 2$ vanishes. This leads to a higher mean size for the system. The mean distance changes from $7 \pm 6/\sqrt{100}$ to $12 \pm 4/\sqrt{100}$, and the percentage of the data with distance equal 1 is reduced to $\sim 0.8\%$. This result indicates that there is a correlation between the company sizes and the distances between the companies.

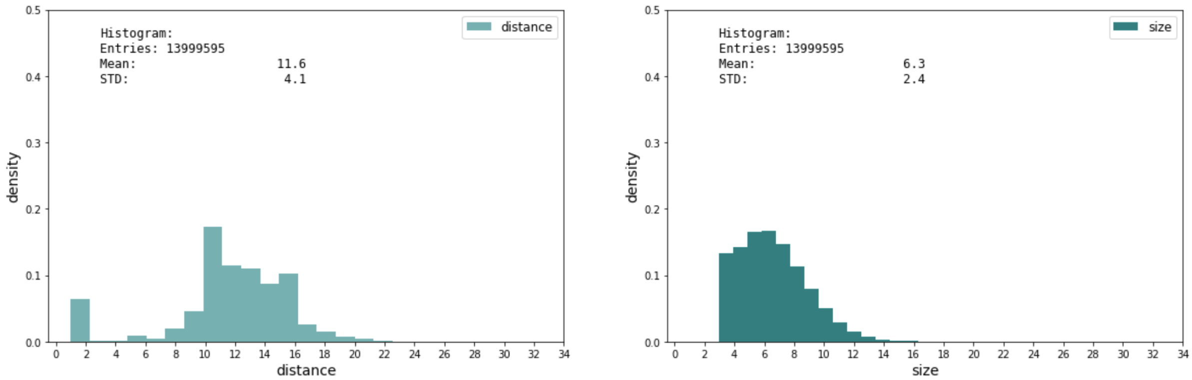


Figure 5.17: ***Model 1: Distances and Sizes of Companies with*** $S > 2$: *Histograms of the distances between companies with* $S > 2$ *and company sizes for companies with* $S > 2$ *for the system* $L = 200$, $T = 10,000$, $\gamma = -0.75$, $\tau = 0.4 \times 200 = 80$ *and* $\sigma = 0.08$

Figure 5.18 shows three plots; The upper plot for the mean distance between companies with $S > 1$, the middle plot for the mean size of companies with $S > 1$ and the lower plot for the number of companies with $S > 1$. In the plots, we see that the distances and sizes decrease as a function of σ, whereas the number of companies increases as a function of σ.

The value of σ can be interpret as the degree of globalization in the world, where a low σ corresponds to a low transportation cost and thus a high degree of globalization. For high σ (low globalization), the number of companies is high as each fragment of the world (country) needs local suppliers. These local companies only sell to local customers, which leads to small company sizes. As there are more companies, the distance between them is short. For low σ, the countries will share the suppliers, causing a decrees in the number of companies. As the companies produce and sell the same product type, companies place themselves at a longer distance from other companies. This will ensure them a solid and large customer base leading to stable and large companies.

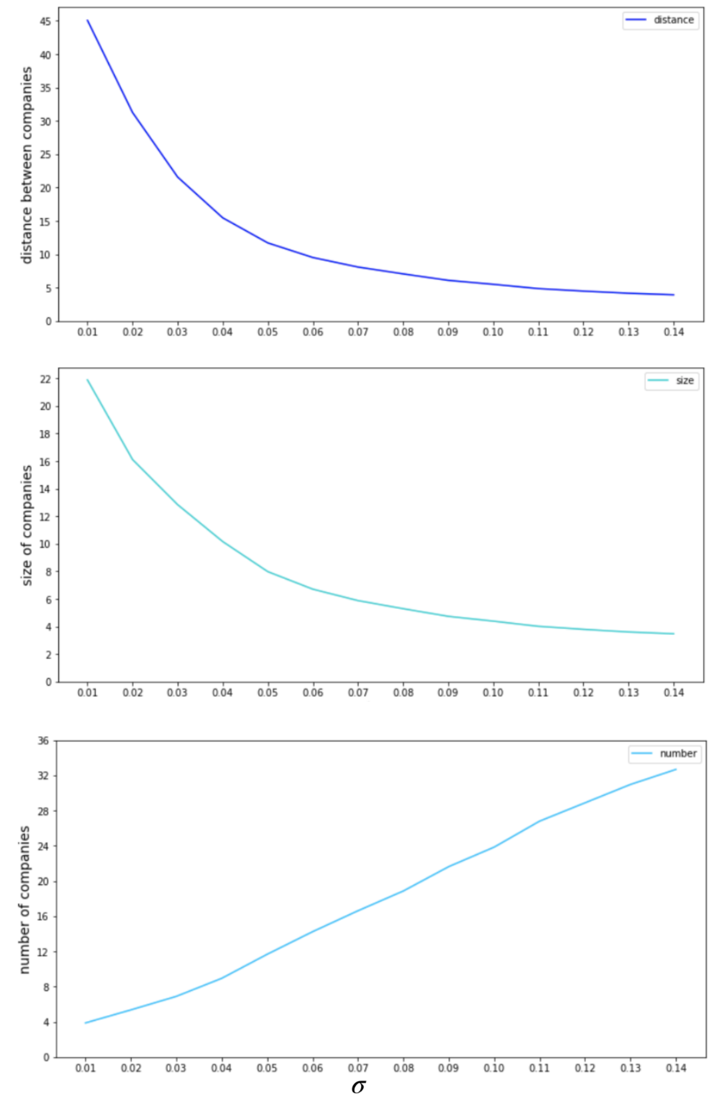


Figure 5.18: ***Model 1: Mean Distance, Mean Size and Mean Number of Companies*** *Top plot: The mean distance between companies with* $S > 1$ *as a function of* σ*. Middle plot: The mean size of companies with* $S > 1$ *as a function of* σ*. Lower plot: The number of companies with* $S > 1$ *as a function of* σ*. For the system* $L = 200$*,* $T = 10,000$*,* $\gamma_A = -0.75$ *and* $\tau = 0.4 \times 200 = 80$

5.2 Model 2

Model 2 consists of two company types Company A (A) and Company B (B). Model 2 is divided into two sub-models Model 2.a and Model 2.b, where the environment is static in Model 2.a and dynamic in Model 2.b. The difference in Model 2.a and Model 2.b is small and not significant. The comparison of the two sub-models are shown in appendix A.

In the result section for Model 2, we review a system for Model 2.a containing companies from the agricultural and manufacturing sector ($\gamma_{agriculture} = 0$ and $\gamma_{manufacturing} = -0.75$).

All systems of interest for Model 2.a and Model 2.b have reached a steady state behavior at time step $t_0 \leq 200$ (see appendix B.4), hence we start saving data at $t_{data} = 200$.

5.2.1 Model 2.a

5.2.1.1 Dynamics

The dynamics of Model 2.a is shown in figure 5.19 for a system with $\gamma_A = -0.75$ and $\gamma_B = 0$. The two company types can be interpreted as the manufacturing sector and the agricultural sector, respectively. The companies of type A are plotted with the color green and the companies of type B with the color blue.

In the 3D plot, we see two large companies of type A surrounded by small companies of type B. The difference in size and number of companies between the two types arises as the companies have different economies of scale. With constant return to scale, the companies of type B are not at all able to reduce their production cost; Regardless of the company size, the cost of production of one unit product will always be equal to one unit payment, i.e. $S^0 = 1$. For this reason, the total cost (equation 4.1) for a company of type B is entirely dependent on the distance to the customer. This requires that the companies are distributed throughout the environment, where they can supply products to their nearest customers. With increasing return to scale, the companies of type A are able to reduce their production costs with an increase in company size. This means that for companies of type A, the production cost is mainly dependent on the company size and less on the distance to the customers. This is why we see large but few companies of type A. By increasing the value of γ (to a lower negative value) and thus reducing the effect of economies of scale, we can increase the number of companies and decrease the company sizes.

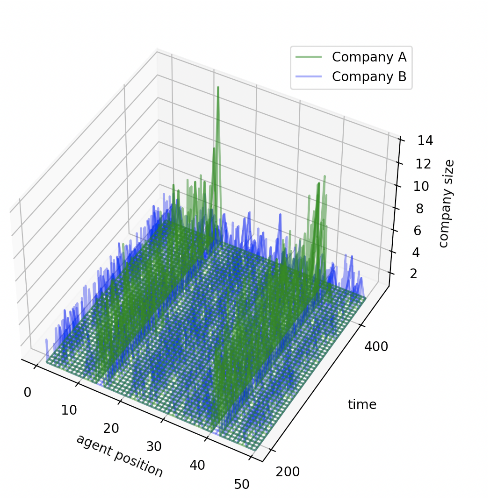


Figure 5.19: ***Dynamics of Model 2.a****: 3D plot showing the dynamics of Model 2.a for $L = 50$, $T = 450$, $\gamma_A = -0.75$, $\gamma_B = 0$, $\tau = 0.4 \times 50 = 20$ and $\sigma = 0.04$. The companies of type Company A are plotted with the color green, and the companies of type Company B with the color blue*

5.2.1.2 Distances, Sizes and Number of Companies

The distribution of the number of companies with $S > 1$ is shown in figure 5.20 for type A in the left histogram and for type B in the right histogram.

In the histograms, we indeed see that the number of companies is distributed along the small values for type A and large values for type B. With the mean values $7.5 \pm 0.9/\sqrt{100}$ for type A and $28 \pm 3/\sqrt{100}$ for type B, the number of companies is 4 times larger for type B than for type A. This difference can be interpreted as the companies of type A are 4 times more efficient than the companies of type B. Thus, economies of scale increase the efficiency of companies.

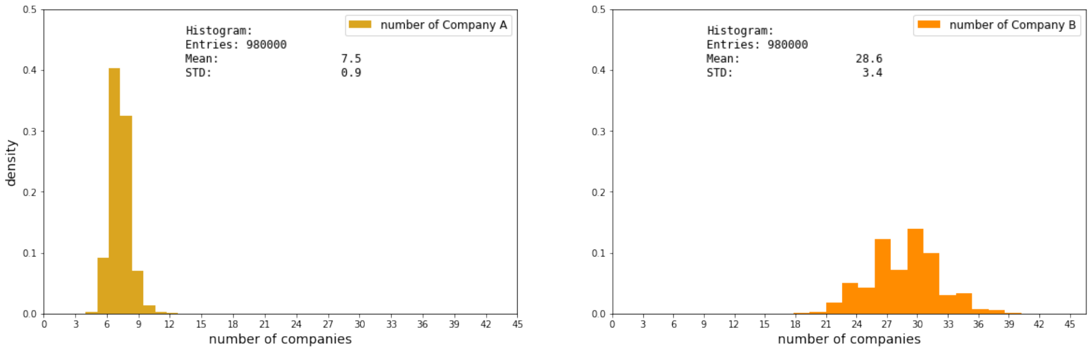


Figure 5.20: ***Model 2.a: Number of Companies**: Histograms of the number of companies with $S > 1$ for type Company A (left) and Company B (right). The model parameters are $L = 200$, $T = 10,000$, $\gamma_A = -0.75$, $\gamma_B = 0$, $\tau = 0.4 \times 200 = 80$ and $\sigma = 0.04$*

The distribution of company sizes for $S > 1$ is shown in figure 5.21 for type A in the left histogram and for type B in the right histogram. In the histogram for type B, we see that 70% of the data have size equal 2 and therefore the mean size is $2.4 \pm 0.7/\sqrt{100}$. For type A, the mean size is $6 \pm 3/\sqrt{100}$. This difference in size shows that companies of type A benefit more from larger company sizes than companies of type B, as an increase in size leads to a decrease in production cost for companies of type A.

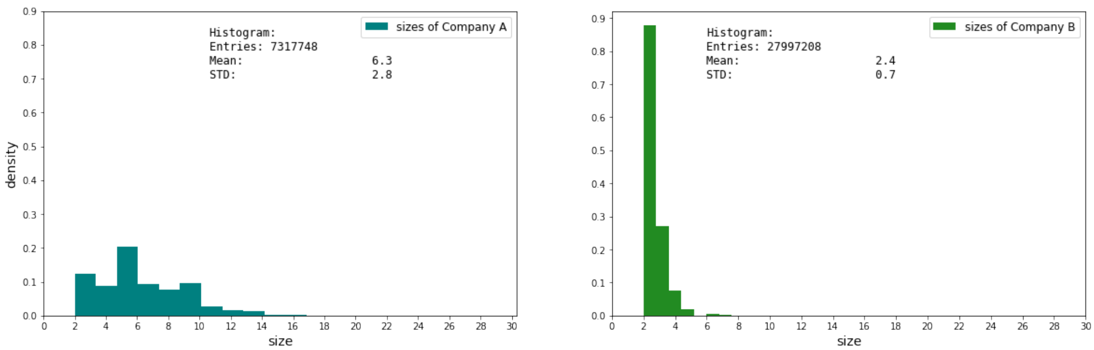


Figure 5.21: ***Model 2.a: Company Sizes**: Histograms of company sizes for $S > 1$ for type Company A (left) and Company B (right). The model parameters are $L = 200$, $T = 10,000$, $\gamma_A = -0.75$, $\gamma_B = 0$, $\tau = 0.4 \times 200 = 80$ and $\sigma = 0.04$*

The distribution of the distances between companies with $S > 1$ is shown in figure 5.22. The left histogram shows the distances between a company of type A and its nearest company of type B (d_{AB}). The right histogram shows the distances between a company of type B and its nearest company of type A (d_{BA}). The mean of d_{AB} is $3 \pm 3/\sqrt{100}$, and $7 \pm 5/\sqrt{100}$ for d_{BA}. This result occurs as the number of type B companies is high, making it more likely for a company of type A to be close to a company of type B.

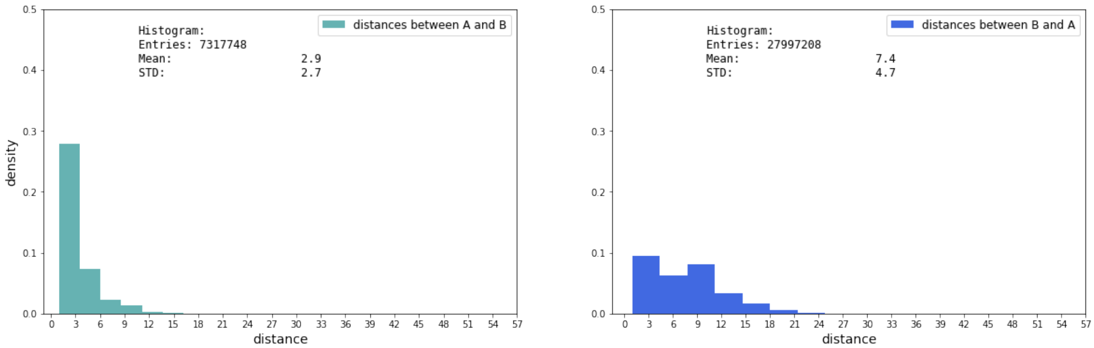


Figure 5.22: ***Model 2.a: Distance between Companies**: Histograms for the distribution of the distances between companies with $S > 1$. The left histogram shows the distribution of the distances between a company of type Company A and its nearest company of type Company B. The right histogram shows the distribution of the distances between a company of type Company B and its nearest company of type Company A. The model parameters are $L = 200$, $T = 10,000$, $\gamma_A = -0.75$, $\gamma_B = 0$, $\tau = 0.4 \times 200 = 80$ and $\sigma = 0.04$*

In this section, we have seen that different values of the economies of scale exponent γ change the behaviour of the companies. With increasing return to scale, the companies of type A benefit from gathering their production into few large single point production sites. With constant return to scale, the companies of type B do not benefit at all from economies of scale, and their total costs are therefore only dependent on the transportation costs. Since the cost of type B products depend only on the distance to the buyer, and since the demand of these products occur at every position, there will always exist a need for a local supplier. This resulted in companies of type B emerging frequently and evenly distributed in the simulated environment. The sizes of these companies are small, as the companies do not benefit from large company sizes but instead of many companies.

5.3 Model 3

In this section, we will present results for Model 3. Similar to Model 2.b, Model 3 consists of two company types (A and B), where only companies with opposite type can interact. The environment of the model is dynamic.

In Model 3, there is a positive feedback between the two companies types. The positive feedback is designed to choose the purchasing company in proportion to the size of the company, and therefore larger companies are forced to make more purchases. Companies of opposite type will therefore have an advantage in being close to larger companies as it would lead to more frequent sales.

For this model, we will mainly focus on the largest company of each type in each time step. Thus, the moving average is found with time series of the sum of the largest company for each type. The graphs are shown in appendix B.5 and B.6, where it appears that all systems have reached a steady state behavior at $t_0 \leq 2000$. Therefore, we start saving data at $t_{data} = 2000$.

5.3.1 Dynamics

The dynamics of Model 3 is shown in figure 5.23 for two systems with two different values of τ. Panel (a) shows the dynamics for $\tau = 0.4 \times L$, and panel (b) for $\tau = 0.3 \times L$. As previously discussed, the value of τ has an impact on the company sizes; a larger τ leads to larger company sizes. Therefore, the companies in panel (a) are larger than the companies in panel (b).

For each time step, the largest company of type A is marked with a red dot and the largest company of type B with a black dot. From the distances between the red and black dots, it appears that large companies prefer to be close. This is indeed due to the positive feedback, where close companies boost each other (see the illustration in figure 5.24). Less larger companies will not have same impact on companies of opposite type, therefore the distances between the companies seems longer for the system shown in panel (b) than for panel (a). To conform this, distributions of the company sizes and distances are shown in figure 5.25 for the two systems in figure 5.23

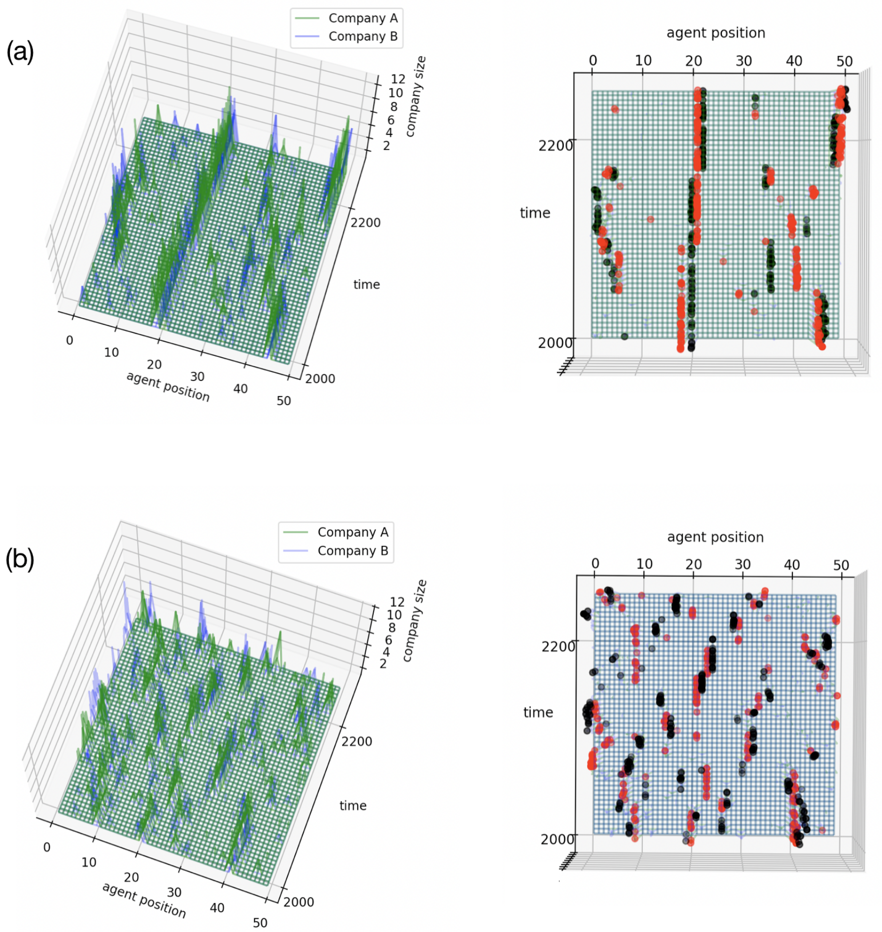


Figure 5.23: ***Model 3: Dynamics****: 3D plots showing the dynamics of Model 3 for $L = 50$, $T = 2250$, $\gamma_A = \gamma_B = -0.5$ and $\sigma = 0.05$. Panel (a): $\tau = 0.4 \times 50 = 20$. Panel (b): $\tau = 0.3 \times 50 = 15$. For each time step, a red dot marks the largest company of type Company A and a black dot marks the largest company of type Company B*

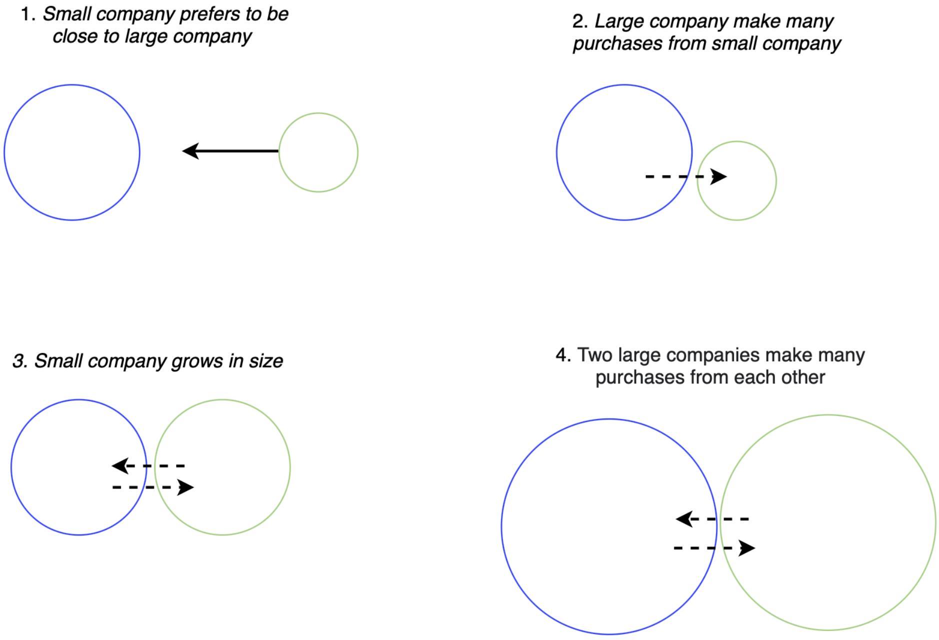


Figure 5.24: ***Positive Feedback Between Companies:*** *The illustration shows a large company of type B (blue) and a small company of type A (green). 1: As larger companies are forced to make more purchases, small companies of opposite type will have an advantage in being close to larger companies. 2: The large company makes many purchases from the small company as the small company is closest and therefore can offer the lowest lowest total cost. 3: The many purchases increases the size of the small company. 4: The two companies now make many purchases from each other*

5.3.2 Distances and Sizes of Companies

Figure 5.25 shows four histograms, where the two upper histograms show the distributions for the system in panel (a) and the two lower for the system in panel (b). The two left histograms show the distribution of the sizes of the largest company of type A (type B is similar as $\gamma_A = \gamma_B$). The two right histograms show the distribution of the distance between the largest company of type A and its nearest company of type B with $S \geq 3$.

In the histograms, we see that the mean size of the largest company is $9 \pm 2/\sqrt{100}$ for the system in panel (a), and $7 \pm 1/\sqrt{100}$ for the system in panel (b). This is as expected, a higher τ leads to larger companies. The mean distance is $4 \pm 5/\sqrt{100}$ for the system in panel (a), and $7 \pm 8/\sqrt{100}$ for the system in panel (b). The difference in size and distance indicates, that companies of some type prefer to be close to larger companies of opposite type. The larger a company is, the closer do companies of opposite type prefer to be.

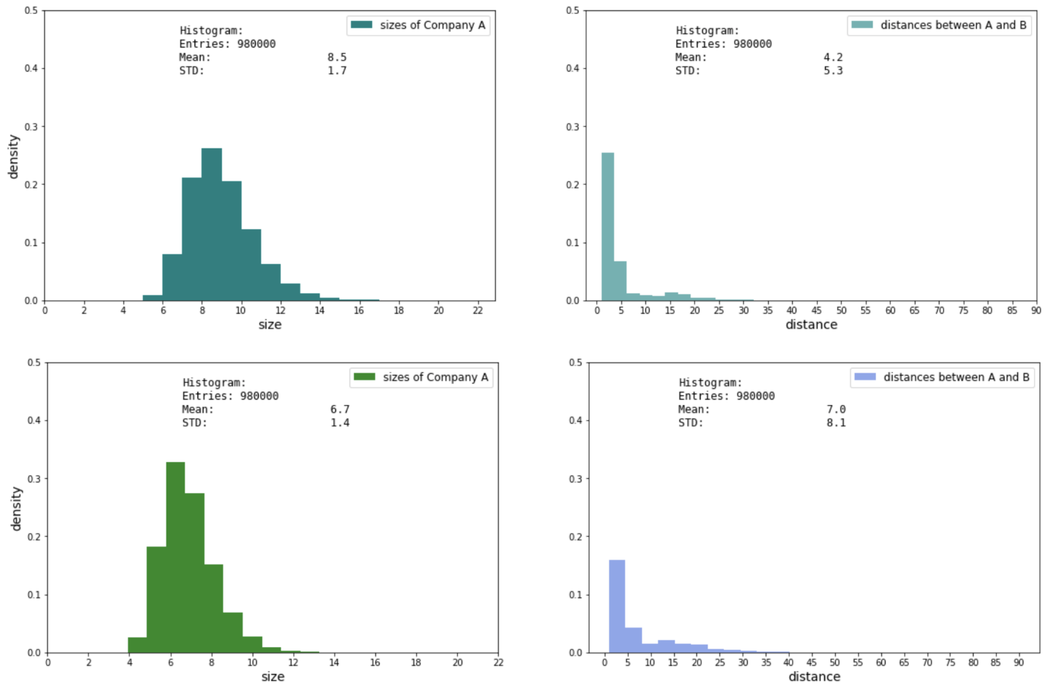


Figure 5.25: ***Model 3: Size and Distance between Companies***: *Histograms for the two systems shown in figure 5.23. The two upper histograms for the system in panel (a) and the two lower for the system in panel (b). The left histograms show the distributions of the sizes of the largest company of type Company A. The right histogram shows the distributions of the distance between the largest company of type Company A and its nearest company of type Company B with* $S \geq 3$

5.3.3 Phase Diagrams

Figure 5.26 shows three phase diagrams. For each phase diagram, one pixel of the diagram corresponds to a system with the given τ and σ value. For all systems $\gamma_A = -0.5$ and $\gamma_B = -0.5$. Phase diagrams are shown for the logarithmic mean lifetime of companies of type A, the mean size of the largest company of type A and the mean distance between the largest company of type A and its nearest company of type B with $S \geq 3$.

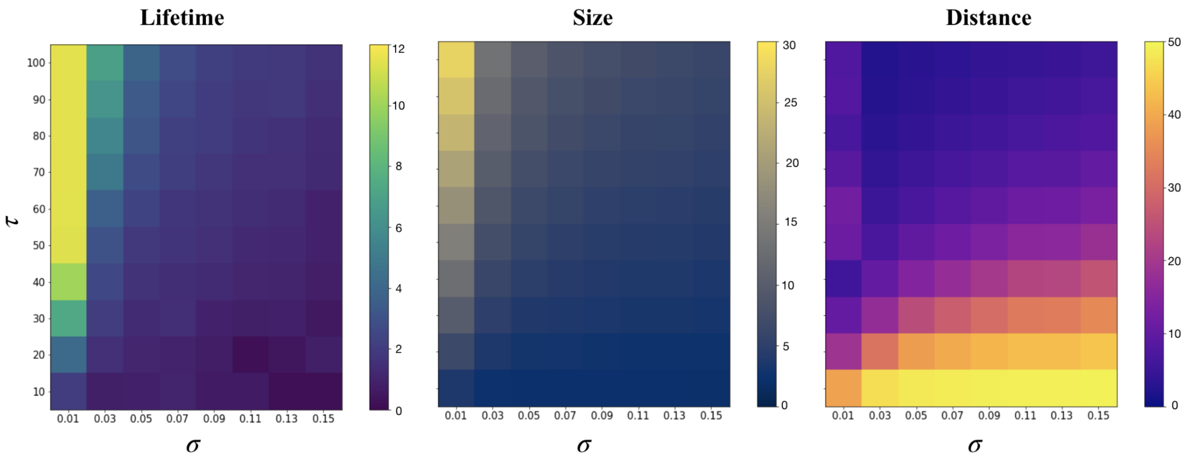


Figure 5.26: ***Model 3: Phase Diagrams****: Phase diagrams for systems with* $L = 200$*,* $T = 100,000$ *and* $\gamma_A = \gamma_B = -0.5$*. Left phase diagram: Logarithmic mean lifetimes of companies of type Company A. Middle phase diagram: Mean size of the largest companies of type Company A. Right phase diagram: Mean distance between the largest company of type Company A and its nearest company of type Company B with* $S \geq 3$

As we have been seen before, small values of τ and large values σ increase the dynamics of a system and thus reduce the length of the lifetimes of the companies. The length of the lifetime has an impact on the size of the company, where we see longer lifetimes lead to larger company sizes. The interesting aspect of Model 3 is found by comparing the phase diagrams for the sizes and the distances. Here, we clearly see a correlation between the size of the largest company and how close a company of opposite type prefers to be; The companies do indeed prefer to be close to larger companies.

Figure 5.27 shows phase diagrams for systems with $\gamma_A = -0.75$ and $\gamma_B = -0.25$. As $\gamma_A \neq \gamma_B$, phase diagrams for each company type is needed. The upper phase diagrams show the behavior of companies of type A and the lower for companies of type B. The result is similar to figure 5.26; longer lifetimes lead to larger company sizes. The distance to the nearest company with opposite type is short when the size of the company is large.

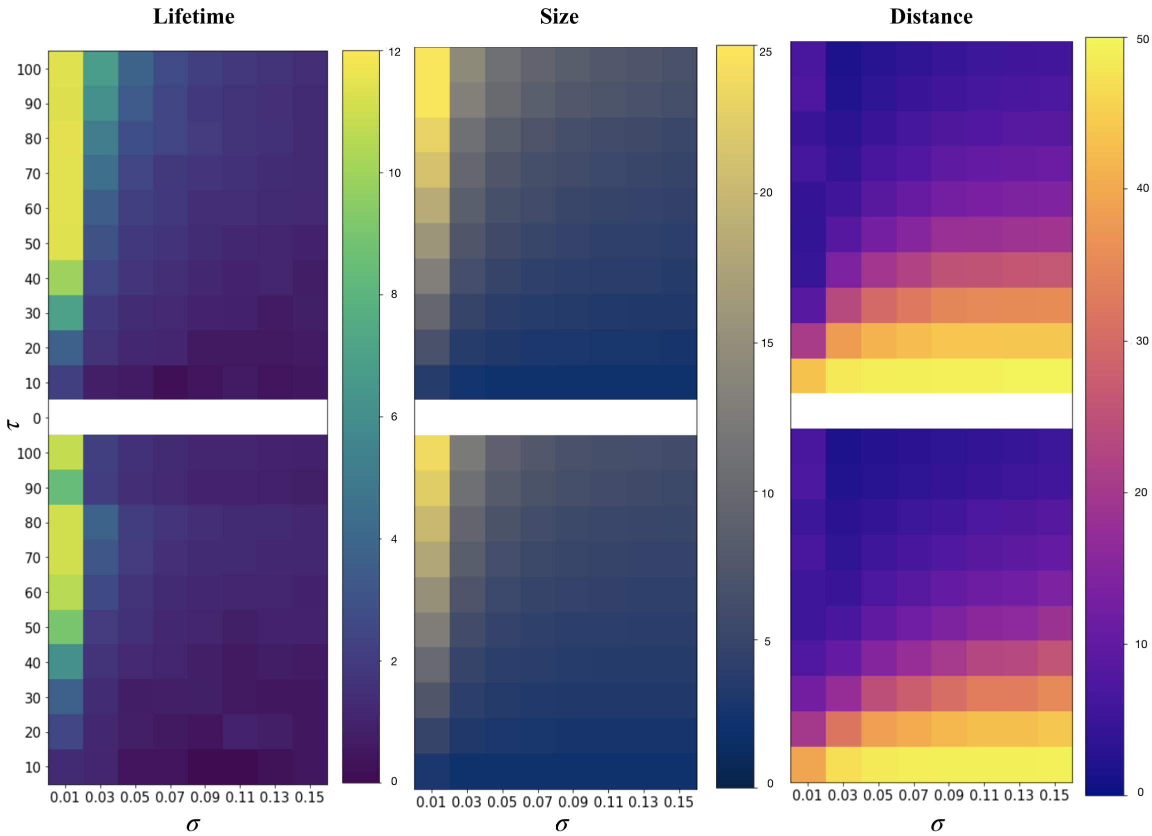


Figure 5.27: ***Model 3: Phase Diagrams:*** *Phase diagrams for systems with $L = 200$, $T = 100,000$, $\gamma_A = -0.75$ and $\gamma_B = -0.25$.* ***The upper phase diagrams:*** *Left: Logarithmic mean lifetime of companies of type Company A. Middle: Mean size of the largest company of type Company A. Right: Mean distance between the largest company of type Company A and its nearest company of type Company B with $S \geq 3$.* ***The lower phase diagrams:*** *Left: Logarithmic mean lifetime of companies of type Company B. Middle: Mean size of the largest company of type Company B. Right: Mean distance between the largest company of type Company B and its nearest company of type Company A with $S \geq 3$*

5.4 Model 4

In this section, we will present the results for Model 4. Similar to Model 3, Model 4 consist of two company types (A and B), where only companies with opposite type can interact. The environment is dynamic and there exist a positive feedback between companies of opposite type. In Model 4 there also exist a negative feedback between the companies. The negative feedback occurs in the calculation of the total cost for each of the selling companies. In the calculation of the total cost for a company, the production cost S^{γ} is multiplied with a factor $\big(log_{10}(S_{factor}) + 1\big)$, where S_{factor} is the sum of company sizes of companies in a distance equal to the size of the selling company.
The moving average is again found with a time series of the sum of the largest company for each company type. The graphs are shown in appendix B.7, where it appears that all systems reaches a steady state behavior at time step $t_0 \leq 4000$. Thus we start saving data at $t_{data} = 4000$. To be able to compared phase diagrams of Model 3 and Model 4, the total number of simulation time steps for Model 4 is set to be equal to the total number of simulation time steps for Model 3, $T = 100,000$.

5.4.1 Dynamics

The dynamics of Model 4 is shown in figure 5.28 for three systems with three different values of τ. The systems shown in panel (b) and (c) have similar parameter values as the systems in panel (a) and (b) in figure 5.23. Comparing these, we see that the negative feedback increases the dynamics of the systems for Model 4. The reason is that companies are torn between the positive and negative feedback; By moving closer to a large company, the sale of the company increases because of the positive feedback. However, the total cost of the company also increases as its production cost increase due to the negative feedback.
In panel (a), we increase the value of τ to stabilize the system and to obtain a more structured plot of the system dynamics. Comparing the system in panel (a) with the systems in figure 5.23, we see that the largest companies are further apart in the system for Model 4.

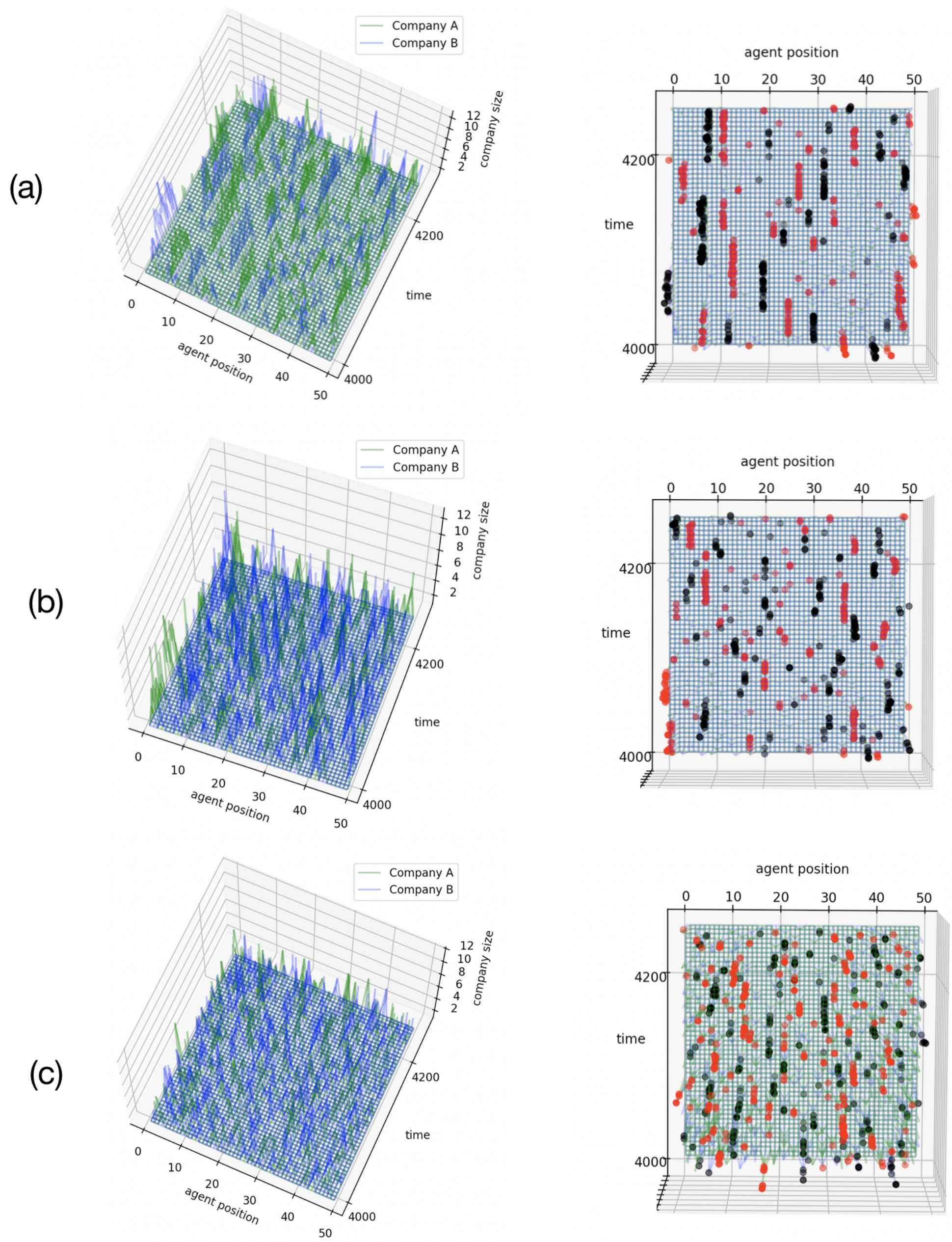


Figure 5.28: ***Model 4: Dynamics:*** *3D plots showing the dynamics of Model 4 for* $L = 50$, $T = 2250$, $\gamma_A = \gamma_B = -0.5$ *and* $\sigma = 0.05$. *Panel (a):* $\tau = 0.5 \times 50 = 25$. *Panel (b):* $\tau = 0.4 \times 50 = 20$. *Panel (c):* $\tau = 0.3 \times 50 = 15$. *For each time step, a red dot marks the largest company of type Company A and a black dot marks the largest company of type Company B*

5.4.2 Distances and Sizes of Companies

To understand the behavior of Model 4, we show three histograms for the systems shown in panel (a) and panel (b) in figure 5.28. The histograms for the system in panel (a) are shown in the first row of figure 5.29, and the histograms for the system in panel (b) is shown in the second row. The left histograms show the distribution of the size for the largest company of type A, the middle histograms show the distribution of the distance between the largest company of type A and its nearest company of type B with $S \geq 3$, where the right histograms show the distribution of the size for the nearest company of type B.

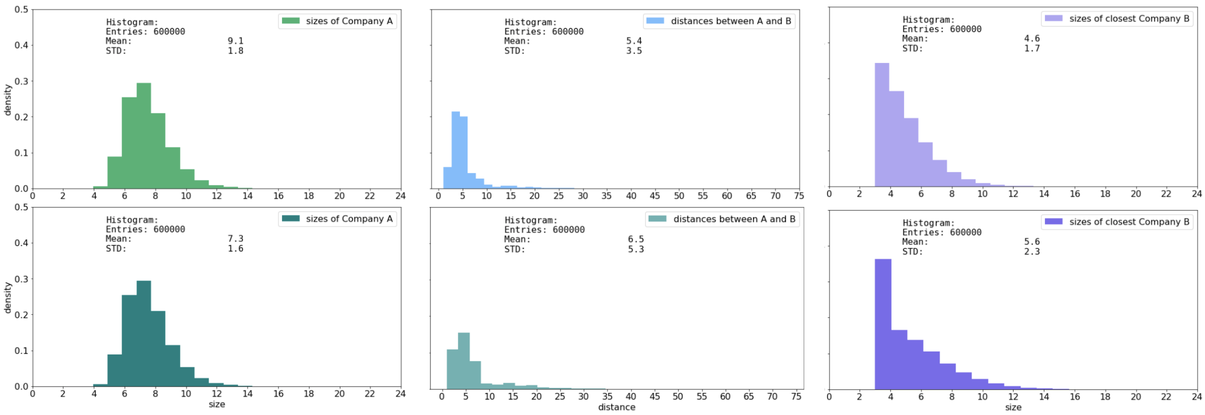


Figure 5.29: ***Model 4: Size and Distance between Companies****: Histograms for the systems shown in panel (a) and panel (b) in figure 5.28. The three upper histograms for the system in panel (a) and the three lower for the system in panel (b). The left histograms show the distributions of the size for the largest company of type Company A. The middle histograms show the distributions of the distance between the largest company of type Company A and its nearest company of type Company B with $S \geq 3$, where the right histograms shoe the distribution of the size for the nearest company of type B*

The histograms shown in the second row of figure 5.29 and in the first row of figure 5.25 are for systems with same values of the model parameters but for different models; The former is for Model 4 and the latter for Model 3. By comparing these histograms, we see that the company sizes are larger for Model 3 than for Model 4. We also see that the distances between the companies is shorter for Model 3 than for Model 4. Thus, the negative feedback reduces the company sizes and increases the distances between the companies (see the illustration in figure 5.30).

By comparing the first row in figure 5.29 with the second row, we see that a higher τ increases the company sizes and decreases the distances. This is similar to what we saw in Model 3, where a higher τ also increased the company sizes and decreased the distances. However, the closest companies of type B manage to grow larger in size for the system with the lower value of τ. Notice how both sizes of closest company B are barely smaller than the distances to the largest company This will all be discussed further in chapter 6.

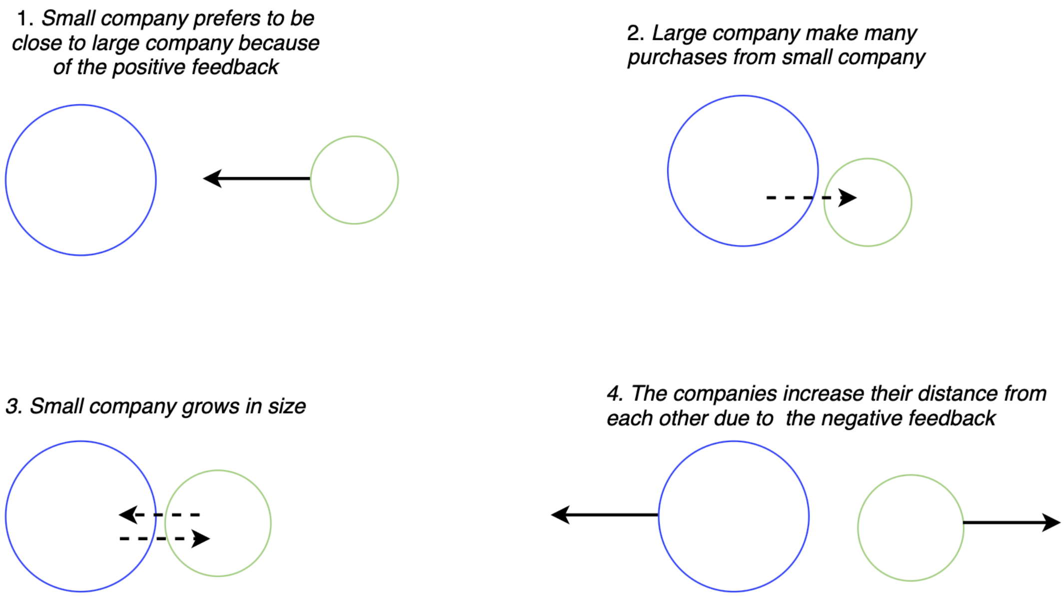


Figure 5.30: ***Positive and Negative Feedback Between Companies:*** *The illustration shows a large company of type B (blue) and a small company of type A (green). 1: As larger companies are forced to make more purchases, small companies of opposite type will have an advantage in being close to larger companies (positive feedback). 2: The large company makes many purchases from the small company as the small company is closest and therefore can offer the lowest total cost. 3: The many purchases increases the size of the small company. 4: The increased company sizes will force the companies to increase their distance to each other (negative feedback).*

5.4.3 Phase Diagrams

Figure 5.31 shows phase diagrams for systems of Model 4. As in figure 5.26, the left diagram shows the logarithmic mean lifetimes of companies of type A, the middle diagram shows the mean size of the largest company of type A and the right diagram shows the mean distance between the largest company of type A and its nearest company of type B with $S \geq 3$.

By comparing the phase diagrams of Model 4 with the phase diagrams of Model 3, we see that the negative feedback decreases the company lifetimes and the company sizes, and increases the distances between the largest company of one type and the nearest company of opposite type with $S \geq 3$.

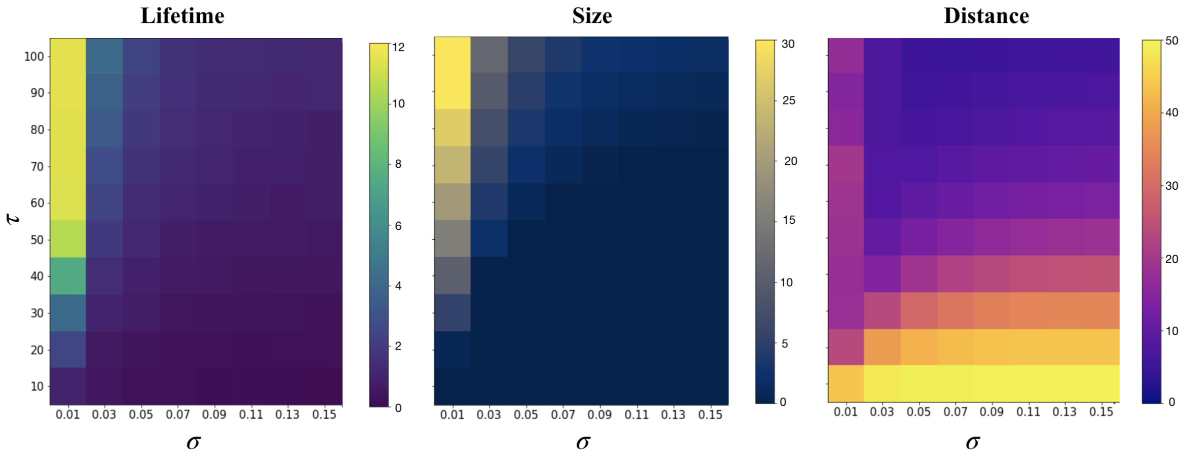


Figure 5.31: ***Model 4: Phase Diagrams**: Phase diagrams for systems with $L = 200$, $T = 100,000$ and $\gamma_A = \gamma_B = -0.5$. Left phase diagram: Logarithmic mean lifetimes of companies of type Company A. Middle phase diagram: Mean size of the largest companies of type Company A. Right phase diagram: Mean distance between the largest company of type Company A and its nearest company of type Company B with $S \geq 3$*

Chapter 6

Discussion

The overall goal with this thesis was to develop and analyze agent based models, that are able to describe the interplay between companies in a globalised market. In the development of the models, we have chosen to incorporate some parameters and assumptions and omit others. For the parameters, we chose to define the companies with γ and τ to incorporate economies of scale and the time it requires to rebuild production apparatus. The parameter σ was used to determine the transportation cost and thus the degree of globalization in our models. Furthermore, we assumed that in a market containing two different company types, producing two different product types, companies of the same type would not buy products from each other i.e. a company producing Product A would not buy Product A from another company. We also assumed that larger companies make more purchases, as their production is higher and therefore their need for intermediate goods is higher (positive feedback). Lastly, we made the assumption that areas with large companies or a large number of companies will get an increased production cost, as the demand for labor is high. The high demand for labor, will force an increase in the wage rate to ensure the necessary labor supply, and thus the increase in production cost (negative feedback).

The critical issue with agent based modelling is that the necessary level of details is unclear; Which elements have to be part of the model and which can be left out? The chosen parameters all need to play a significant role for the outcome, whereas the irrelevant parameters play an insignificant role. Omitting less significant parameters of course affects the accuracy of the model, but the simpler model increases the understanding of the relationship between the fewer chosen parameters. Moreover the complexity of the model increases rapidly for each added parameter, making both

computational power and analysis vastly more complicated. The fewer details a model has, the fewer assumptions have to be justified. Also, fewer parameters have to be calibrated, and the implementation and handling of the model becomes easier. However, if the model is too simplistic, the significance of the results might be compromised [7].
Our models are clearly too simple to fully describe the interplay between companies in the current world, and therefore we have problems finding relevant real life data to make a comparison with our data. However, we still believe that our models manage to describe some of the trends that can be observed in the current world. The reason for this believe is based on our choice of parameters and assumptions for the models, which have been chosen because of their resemblance to real world behaviour. If our models are correctly formulated it is reasonable to assume that data outputs compares well to tendencies in the real world.

6.1 Discussion of Time Correlation and Dependent Data

In order to perform statistical analysis on our data, we treated dependent data as independent. To understand the impact of this decision, we defined a function for the time correlation as a function of time delay (see equation 4.21).

> The time correlation was plotted and fitted with an exponential function (see equation 4.22) for different systems, and values of α were estimated. The value of α gave the time interval at which the correlation of a system has decayed with $1/e$. The results showed that for systems with low negative γ, low τ and high σ, the time intervals α were short compared to those for systems with high negative γ, high τ and low σ. From 3D plots, showing the dynamics of systems, we understood that the value of α reflected the degree of dynamics of the system. For systems with high dynamics i.e. systems with more frequent changes in company size and position, the correlation decreases faster, and thus the system had a shorter α. For systems with low dynamics, the correlation decreases slower, and thus the system had a longer α.

The main purpose of the measurements of the time correlation was to understand the possible consequence of treating dependent data as independent on our statistical results. Therefore, we will look at an example where the value of α is long. This could be the system with $\gamma = -0.75$, $\tau = 0.3 \times L = 60$ and $\sigma = 0.03$, where $\alpha \sim 742$. To ensure that the correlation has decreased with at least $1/e$, we increase the value to $\alpha = 1000$. Thus, we only use data for every 1000'th time step. The full data set is plotted in the left histogram of figure 6.1, and the subset in the right histogram.

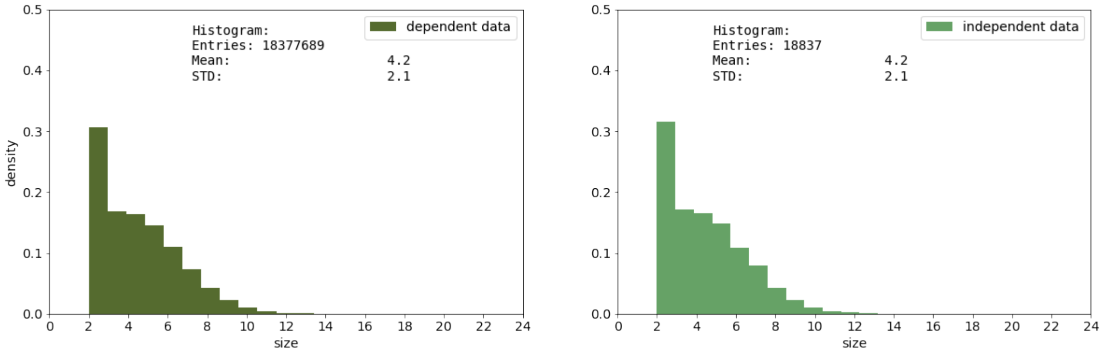


Figure 6.1: ***Dependent and Independent Data**: The two histograms both show the distribution of company sizes for the system $L = 200$, $T = 10,000$, $\gamma = -0.75$, $\tau = 0.3 \times 200 = 60$ and $\sigma = 0.03$. For this system, the correlation time interval is $\alpha \sim 742$. We however increase the value to be $\alpha = 1000$ to ensure that the correlation has decreased with at least $1/e$. The full data set is plotted in the left histogram and the subset (data point for every 1000'th time step) in the left histogram. As the distribution, mean and standard deviation are the completely the same, we do not see any problem in treating dependent data as independent*

By comparing the two histograms, we see absolutely no difference; The distribution, mean and standard deviation are completely the same. As we assert $1/e$ to be a relatively high number, we assume that when the correlation has decreased with $1/e$ the data point is completely independent. This means that we do not see any problems with assuming dependent data as independent. However, we would still like to emphasize that the quality of our statistics is less than it appears - even though the left histogram claims that we have $18,377,689$ data points, we actually only have $18,837$ data points.

6.2 First Research Question

From the graphs and considerations in the result section, we will now try to answer the research questions posed in section 3.2. To answer the first research question, we will, as described in section 4.2.6, primarily use the results regarding Model 1.

1. What influence do the following parameters have on companies and thus the market the companies exist in?
   - *Economies of scale* (γ)
   - *Time to rebuild production apparatus* (τ)
   - *Transportation cost* (σ)

6.2.1 Discussion of τ

Previously, we have referred to the parameter τ as analogous to the time a company would use to rebuild their production apparatus. An alternative way to interpret this parameter is as an analogy to the time it takes a newly opened company to assess if it can get a sufficiently big market share to exist, with its chosen product and location, in competition with its surroundings. So in this way τ is comparable to a time window of potential sales for a company. Consider for instance that you have a newly opened bike shop in a geographical area where two other well established bike shops are already located. After a month, you realize that all the passing customers choose to visit the competitors. This means that even though τ is high (you have had a lot of potential buyers), you do not have any sales because your competitors are too established and therefore can offer a lower cost on bikes. When the month is over (i.e. 1 time step has passed in our mode), you realize that you will have to close your company due to lack of sales. At this time you have two options, either to reopen at the same location with another product, or to keep your product and relocate instead.

> Our findings in the result section were that when τ is low, the dynamics of the system are high and companies do not grow as large in size as when τ is high. The systems natural growth limit is determined by the value of τ during each time step, since the model is defined to reset company sizes to be equal to 1 plus the accumulated sales of

the previous time step. Therefore no company size can ever exceed the value of τ, which explains why companies are larger for a system with higher τ. So since companies never become very large when τ is low, they will never get full benefits of the economies of scale parameter γ, and thus the position of the company plays a more significant role for the total cost. Hence, new companies with marginal better positions have a greater chance to emerge, shown as the higher dynamics in the 3D plot in figure 5.1 panel (a). Conversely, when τ is high we see the opposite pattern, where companies grow larger and benefit more from the economies of scale. This makes the total cost less dependant on the geographical location, making the competition high to newcomers and explains the more static dynamics of the system in figure 5.1 panel (b).

The two remaining parameters γ and σ both have a more obvious analogy to real life. The economies of scale exponent relates to the unit cost's dependency on the scale of the production; The higher the production, the lower the cost per unit product and vice versa. The transportation cost obviously corresponds to real life transportation cost's dependency of distance because of the time, fuel and resources needed.

6.2.2 Discussion of σ

The transportation cost is assumed to have a linear dependence to distance. In real life, one would assume an initial transportation cost and a cost per unit distance, where the rate of the increase of the latter typically decreases with distance. This indicates that an exponential dependence on distance could have been more realistic, but a linear dependence was chosen for simplicity.

From the result section, we observed that an increase in transportation cost σ led to more local and small companies emerging. This increased the number of companies and thus reduced the average distance between companies (see figure 5.18). The opposite was seen for a decrease in transportation cost, where few large companies existed.

This behavior is intuitively expected and corresponds very well to observations of real life - if shipment is expensive, then one prefers to buy locally. In modern day and age, with well developed infrastructure throughout almost all industrial countries, transportation costs are lowered to a minimum. This has led to a vast increase in globalization, and goods are bought and transported from all over the world. As a consequence, this has led to fewer and larger companies concentrating the production of goods.

6.2.3 Discussion of γ

The economies of scale is incorporated in our models via the term S^{γ}. Company size S is always a natural number $S \geq 1$ and γ has been chosen to be $-1 \leq \gamma \leq 0$. With these parameter settings, the value of the production cost is always $S^{\gamma} \leq 1$. For a fixed negative γ, an increase in S will always lower the production cost per unit product, resembling the real life economies of scale feedback. A decrease in γ (to a higher negative value) would increase the effects of the economies of scale on the production cost. Whereas the overall behaviour of this function satisfies our need to simulate real life economies of scale, the specific value of γ for different sectors and products in real life is hardly possible to determine. What can be said, however, is that a sector such as agriculture would have a low benefit of the economies of scale, i.e. a γ close to 0, whereas more technologically advanced production sectors benefit more from economies of scale i.e. a higher negative γ.

> From our results, we observed that a higher negative γ increased the company lifetimes and thus lowered the dynamics of a system (see figure 5.15). Since the benefits of economies of scale increases after *the bigger the better* principle, competition has an even less chance to emerge, leading to larger and fewer companies.

In total, the analysis of the effects on the market from different parameter settings shows that all the parameters individually have intrinsic capabilities to adjust market mechanisms such as company size, geographical location and the number of companies. This suggests that a well considered balance between these parameters is necessary to obtain a specific desired market structure, and a change in just one parameter can completely alter the market dynamics.

6.3 Second Research Question

All of the above addresses markets with only one company type. To better compare our models to real life, we will increase the number of company types and discuss the interplay between them. With the results from Model 2, we will try to answer the second research question.

2. What impact do different returns to scale have on co-existing company types?

Different settings of γ can be interpreted as different real life returns to scale sectors. A sector with constant return to scale, i.e. $\gamma = 0$, compares well to preindustrial productions with no technology and without machine assistance. It can therefore seem unreasonable to assume that a modern sector produces with $\gamma = 0$. However, it seems reasonable to assume that the agricultural sector produces with almost constant return to scale as for instance a harvest with a yield of x kg potatoes requires an area of size y and a harvest with a yield of $2x$ kg potatoes requires an area of size $2y$.
In the manufacturing sector, the initial cost to establish a production facility is very high thus the production of e.g. $100,000$ units is made with a much lower cost per unit than the production of $1,000$ units. Therefore, the manufacturing sector produces with a much higher return to scale than the agricultural sector. For this reason, the case with $\gamma_A = -0.75$ and $\gamma_B = 0$ shown in the result section for Model 2 can be compared with return to scale values for the manufacturing sector and the agricultural sector, respectively.

> In the result section for Model 2, we considered a case with two company types, the manufacturing sector and the agricultural sector. In the interplay between them, we saw two distinctly different distribution patterns of the two company types. The companies with constant returns to scale, obviously, did not benefit from economies of scale, and therefore the total cost was only dependent on the transportation cost. The consequence of this led to small company sizes, as there were no benefits for the companies to grow in size. Since the cost of agricultural products depends only on the distance, and since the demand for these product can occur at every location, there always exists a need for a local supplier. This resulted in agricultural companies to emerge frequently and evenly distributed in the environment. For companies with increasing returns to scale, the total

cost mainly depended on the production cost and partly depended on the transportation cost. This was observed in fewer, but significantly larger companies.

In our model, we consider the products offered by the same company type to be of equal quality. Thus, the only consideration relevant to the customer in the agricultural sector is the final price of the product - determined solely by the distance to the company. Therefore, for a company to access the market the competition from neighbouring companies is not a problem. In real life, companies that do not mass produce a single product and therefore have no benefits of a large single manufacturing location, are often seen evenly distributed geographically and thus easily accessible to local customers. The opposite is true for manufacturing companies that with the advantage of economies of scale, benefit from large single point production sites. To further decrease production costs, these companies are often located in developing countries, where the price of labor and of intermediate goods are markedly lower. Under these conditions, such companies are able to minimize the production cost to such a level that the transportation cost is so insignificant, that they are able to supply goods to all over the world.

6.4 Third Research Question

In order to answer the third research question, we will discuss Model 3 and Model 4. We will consider companies producing two different products which can either be produced with same or different economies of scale.

3. How do different feedback mechanisms affect the distance between interacting companies producing two different products?

6.4.1 Discussion of the Positive Feedback

Up until Model 3, demand has been equally distributed in the environment, in order to better observe behaviour changes of the remaining variables. Therefore, the demand so far has been more comparable to single customers with random distribution. In Model 3, however, we introduced a demand distribution more similar to that of real life. The model's change in demand behaviour was set up in such a way, that the size of demand was proportional to company sizes. In this way, demand in Model 3 is differentiated between one-man businesses, small and medium companies and large companies. The model still consists of two company types A and B, where companies only can interact with companies of opposite type. The interaction rule and demand behaviour corresponds to an industry where a company have to acquire intermediate goods from another company with a different production than its own, and with a demand size proportional to the size of the demanding company. The companies of the two types should therefore *no longer* be interpreted as the agricultural sector and the manufacturing sector, as in the discussion of Model 2, but as two manufacturing sectors where companies of type A use intermediate goods, manufactured by companies of type B, to produce intermediate goods for the companies of type B.

> In the results from Model 3, phase diagrams were displayed, depicting the mean lifetime of all companies, mean size of the largest companies and mean distance between the largest company and its nearest company of opposite type with $S \geq 3$ (see figure 5.26 and figure 5.27). In the phase diagrams, we saw that when parameter settings produced larger companies, the lifetime increased as well, however the mean distance between the companies decreased and vice versa. Observing the phase diagrams vertically, we can see that whenever τ increases, the company sizes increased and thus distances decreased. If we instead view the phase diagrams horizontally, we see that an increase in σ, decreased the mean size of companies and increased distance between the companies.

The behaviour of distance relative to company sizes seem to be opposite of previous observations in figure 5.18 for Model 1, where an increase in σ led to a decrease in mean distance and subsequently a decrease in company size. However, the two different observations might not be quite comparable since the mean distance measured in Model 1 is the shortest distance between any two companies with $S > 1$. Whereas in Model 3, the mean distance is a measure of the mean distance between the largest company and its nearest company of opposite type with $S \geq 3$. So the reason that we observe the opposite size/distance-relationship in Model 3, compared to Model 1, could be explained by the fact that companies never grow very large when σ is high. This means that many companies of smaller sizes $S < 3$ exists and thus the distance to the nearest company of $S \geq 3$ increases, giving us the large distance output. For a decrease in σ, the companies grow larger and the distance decreases.

In section 2.2, we presented the Krugman-Venables Model (1995), in which they considered two economies both producing agricultural and manufacturing goods. Within these economies, the agricultural sector produces non-traded goods and the manufacturing sector produces traded goods. The products of the manufacturing sector are divided into two categories; final goods and intermediate goods, the latter being sold between manufacturing companies. In our Model 3, the trade between the two company types is comparable to the trade within the manufacturing sector of the Krugman-Venables Model. As we ignore the agricultural sector in our model, the following considerations also ignore the agricultural sector.
The conclusion of the Krugman-Venables Model were that a decrease in transportation costs would lead to an increase in company size and ultimately agglomeration of the manufacturing companies. This is similar to our results, where we observed that a decrease of σ also led to an increase in company size and to a reduction of distance between the producer and the supplier of intermediate goods. Intuitively, this behavior of our model might be surprising as an increase in σ leads to an increase in distance and a decrease in σ leads to a decrease in distance - Why should companies move closer with a decrease in transportation cost instead of with an increase in transportation cost? The reason behind this behavior is not due to the transportation cost, but the side effects of it. We have previously seen that a low σ increases the company sizes and a high σ decreases them (see figure 5.18). With the larger company sizes, the largest company is forced to make more purchases from the company of the opposite type. To obtain the lowest cost, the purchasing company buy as locally as possible, thereby boosting the nearest neighbour of opposite company type. Thus, the results of Model 3 should be interpreted in real life as: Larger companies with higher demand for intermediate goods attract local suppliers of their demanded goods.

6.4.2 Discussion of the Negative Feedback

In Model 4 we introduced a negative feedback between the companies, where the cost of the selling company is affected by all companies and the sizes of these, within some range dependent on the size of the selling company. In all other aspects, Model 4 is similar to Model 3.
For Model 4, we assumed that the size of a company is proportional to the number of employees. This means that in areas with large and/or many companies, the number of employees is high. In order to accommodate the high demand of labor, companies are forced to increase labor wages and thus the production cost per unit increases. It is an assumption of the model, that labor is always available to the right price, making the companies production size theoretically unlimited. This

should be interpreted in the meaning that the considered companies always can attract labor, with an increase in labor wage, from other non-simulated sectors existing in the environment.

In the results of Model 4, we presented phase diagrams similar to those of Model 3. We saw that the lifetimes and sizes of the companies in Model 4 are smaller than for Model 3, which explains the more dynamic behaviour of the systems of Model 4 (see figure 5.28). In all other aspects, the behaviour of the systems observed from the phase diagrams is similar to those shown in Model 3; Larger sizes lead to longer lifetimes and decreasing distances, and vice versa (see figure 5.31). For further discussion of the behaviour of the phase diagrams, see the above discussion of Model 3.

In Model 4, the negative feedback forces large companies to buy from companies further away. Consider a large Company (A) that wants to buy from a Company (B), then the distance to (B) has to be bigger than the size of (B); otherwise the cost per unit product from (B) contains the negative feedback from containing Company (A) in (B)'s S_{factor} (see equation 4.4). If company (A) wants to buy from a (B) that *does* contain the negative feedback, the cost per unit product will be much higher. This does not mean, however, that such a sale could not take place, if for instance transportation costs are very high. This behaviour is the reason why we see that the distance between the largest company and its nearest company of opposite type with $S \geq 3$ has increased compared to Model 3 (compare figure 5.26 and figure 5.31).
Another influence from the negative feedback on the companies is the constant adjustment of size dependent on the number and sizes of companies within the neighbourhood, and especially the distance to the largest company. This intrinsic self-adjustment of size explains why the company sizes are smaller in Model 4 than in Model 3.

The consideration made above regarding the distance/size-relationship of Model 4 can be visually observed in figure 5.29, where histograms are shown for different values of τ. The mean size of the closest company of type B is barely smaller than the distance to the largest company of type A, suggesting exactly the preference of the companies to be in a distance just greater than the size of the company. Comparing the three upper histograms for the system with large τ with the three lower histograms for small τ, we see that for large τ the companies of type A are larger (as seen in all previous models) increasing the effect of the positive feedback. With increased positive feedback, the distance to the nearest company of type B with $S \geq 3$ is reduced. As the size of B must be barely smaller than the distance to the largest company, the size of B is also reduced. For small τ, the distance is greater which could be explained by the decreased positive feedback due to smaller company sizes.

From the discussion of Model 3 and Model 4, we can now extract the main motivations for how the companies choose locations. Model 3 suggest that trading companies exhibit a mutual interest of being within the same geographical proximity, thereby both reducing transportation costs and, through mutual trade, boost their respective growth. Model 4 however shows a more diffuse picture, that indicates that for markets in which negative feedback exist, for instance a labor supply/increased wages relationship, companies with shared trade would benefit from choosing locations with an optimal distance between them, considering both the degree of negative feedback and the benefits from the positive feedback. Model 4 also suggested that the size of companies is affected by the distance to their primary trading partner, where companies could benefit from being further away from each other, if the degree of negative feedback is sufficiently large.

Chapter 7

Conclusion

We have presented four agent based models that aim to describe companies in a globalized market. The model parameters are γ, τ and σ, which respectively incorporate economies of scale, the time a company would use to rebuild their production apparatus and transportation cost, the latter comparable to the degree of globalization in the world. Positive and negative feedback mechanism were also added to the models, to observe behaviour changes of the companies. The positive feedback changed the demand behaviour of companies from single customers with random distributions to being proportional to company sizes. The negative feedback considered labor supply deficits in areas with high company concentrations. Lack of labor supply increased wages and thus production cost. The model setups were inspired by previous works of Kim Sneppen and Stefan Bornholdt (2018), and of Paul Krugman and Anthony J. Venables (1995).
Based on simulation outputs containing the company sizes, number, lifetimes and distances, we used statistical analysis to observe model response to changes in the parameter settings.

Answer to the first research question:
With different settings of the model parameters, we understood the impact of the parameters on the companies and the market in which they exist:

- The value of τ was mainly set to determine the dynamics of the system, thereby preventing frozen state of the system and reducing noise in the output. The degree of dynamics of the system were raised with a decrease in τ and lowered with an increase in τ.
- When increasing transportation cost σ, we observed an increase in the number of companies,

a decrease in mean size, mean lifetimes and mean distance - and vice versa for a decrease in σ. This is in agreement with real world observations; With an increase in transportation cost, more local suppliers would emerge, securing low costs to the consumer. With a decrease in transportation cost, fewer but larger companies can supply goods to larger customer bases.

- With an increase in economies of scale, i.e. a higher negative γ, the production cost of the companies were reduced. This reduced production cost, secure more sales for larger companies increasing their growth. This made it difficult for newcomers to enter the market. These market conditions led to fewer, larger and more stable companies.

Answer to the second research question:
By assigning different returns to scale to co-existing and interacting company types (sectors), we understood the impact of economies of scale on companies:

- With increasing return to scale i.e. negative γ, the companies benefited from large single point production sites. This led to fewer but larger companies.
- With constant return to scale i.e. $\gamma = 0$, the total cost of companies were determined solely by distance to the customer. The companies therefore had no benefit of large single point production sites, but instead to evenly distribute in the environment and thus easily accessible to local customers.

Answer to the third research question:
By adding positive and negative feedback, we understood that the distance between the companies can be affected.

- An increased positive feedback showed that trading companies exhibit a greater mutual interest of being within the same geographical proximity, thereby both reducing transportation costs and, through mutual trade, boost their respective growth.
- When both positive and negative feedback exist, our model showed that with an increase in the negative feedback, companies with shared trade would benefit from choosing locations in such a distance that the influence of the negative feedback was minimized, suggesting an existence of an optimal distance between the companies keeping production expenses to a minimum.

Perspectivation:

In the following, we will consider how to extrapolate features of our conclusions to real world behaviour of manufacturing companies.

Today's generally low global transportation costs reduce the demand for local supply and the companies are therefore no longer limited to locate themselves next to the site of demand. This enables companies to concentrate their production in larger and fewer production sites.

With an increase in production size, manufacturing companies can benefit from economies of scale, which enable them to reduce their cost per unit product. From advanced technology and automatisation, companies will further optimize the effects of economies of scale.

In our model, the negative feedback from high wages forced companies to move beyond the negative feedback zone. Similar phenomenons motivates Western companies in the real world to relocate their production to developing countries, exactly to avoid "the negative feedback" of the Western countries, by high labor wages and strict regulations among other things.

The trade of intermediate goods between companies yielded a positive feedback in our model, decreasing the distance between them. With the vast emigration of manufacturing companies to developing countries, a similar relocation of companies producing raw material or intermediate goods could be expected.

The overall global trend of relocating manufacturing companies to developing countries has increased trade deficits in many Western countries. In practice a consequence of this is that capital is transferred from Western countries to developing countries. One could speculate that continuous trade deficits could eventually lead to a shift in welfare standards between these two global regions. The final outcome of this development is hard to predict, but it does not seem unreasonable to assume that at some point in the future, trade restrictions or traffic barriers could be introduced to stabilize economy growth, and as consequence of this we might see company production once again return to the West.

Bibliography

[1] Ellen-Wien Augustijn. *Components of an Agent Based Model (ABM)*. URL: https://www.futurelearn.com/info/courses/geohealth/0/steps/19290. (accessed: 05.19.2021).

[2] S. H. Kim C. Alexopoulos. "Output Data Analysis for Simulations". In: *Winter Simulation Conference* (2002).

[3] R. C. H. Cheng C. S. M. Currie. *A Practical Introduction to Analysis of Simulation Output Data*. University of Southampton: Mathematical Sciences, 2016.

[4] H. Dembinski. *Introduction to iminuit*. URL: https://indico.cern.ch/event/833895/contributions/3577808/attachments/1927550/3191336/iminuit_intro.html. (accessed: 06.23.2021).

[5] J. T. Dunlop. "The Demand and Supply Funnctions for Labor". In: *The American Economic Review Vol. 38, No. 2* (1948).

[6] BC CAMPUS OPEN EDUCATION. *Demand and Supply at Work in Labor Markets*. URL: opentextbc.ca/principlesofeconomics/chapter/4-1-demand-and-supply-at-work-in-labor-markets/. (accessed: 10.10.2021).

[7] A. L. C. Bazzan F. Klügl. "Agent-Based Modeling and Simulation". In: *AI Magazine* (2012).

[8] R. Giri. *Increasing Returns to Scale as a Determinant of Trade*. Instituto Tecnologico Autonomo de Mexico: Centro de Investigacion Economica.

[9] S. Bornholdt K. Sneppen. "Globalization in a nutshell". In: *American Physical Society* (2018).

[10] L. R. Klein. "Theories of Effective Demand and Employment". In: *Journal of Political Economy, Vol. 55, No.2* (1947).

[11] T. Levitt. "The Globalization of Markets". In: *Harvard Business Review* (1983).

[12] World Trade Organization. *Special and Differential Treatment Provisions.* URL: https://www.wto.org/english/tratop_e/devel_e/dev_special_differential_provisions_e.htm. (accessed: 10.19.2021).

[13] G. Gancia P. Epifani. "Global Imbalances Revisited: The Transfer Problem and Transport Cost in Monopolistic Competition". In: *University of Nottingham, CREI, Barcelona GSE and CEPR* (2017).

[14] A. J. Venables P. Krugman. "Globalization and the Inequality of Nations". In: *The Quarterly Journal of Economics, Vol. 110, No. 4* (1995).

[15] A. Pal PKS. Prakash. *Practical Time Series Analysis.* Birmingham: Packt Publishing, Limited, 2017.

[16] python.org. *multiprocessing — Process-based parallelism.* URL: https://docs.python.org/3/library/multiprocessing.html. (accessed: 06.25.2021).

[17] K. Sneppen. *Complex Physics.* University of Copenhagen: Compendium to the course: Complex Physics, 2019.

[18] A. M. Sørenssen. *Percolation theory using Python.* University of Oslo: Department of Physics', 2020.

[19] United Nations Conference on Trade and Development. *Fact sheet 7: Gross Domestic Product.* URL: https://unctad.org/system/files/official-document/tdstat44_FS07_en.pdf. (accessed: 10.20.2021).

[20] H. R. Varian. *Microeconomic Analysis, 3. edition.* Uniyed States of America: Norton, 1992.

[21] www.saylordotorg.github.io. *Trade Imbalances and Jobs.* URL: https://saylordotorg.github.io/text_international-economics-theory-and-policy/s17-02-trade-imbalances-and-jobs.html. (accessed: 10.19.2021).

Appendices

Appendix A

Additional Results

Model 2.a Versus Model 2.b

The dynamics of Model 2.a and Model 2.b for $\gamma_A = -0.75$ and $\gamma_B = -0.25$ are shown in panel (a) and panel (b) of figure A.1. We see that the systems contain a larger number of companies of type B than of type A. We also see that the company sizes are smaller for type B than for type A. This difference in size and number of companies is due to the different values of γ.

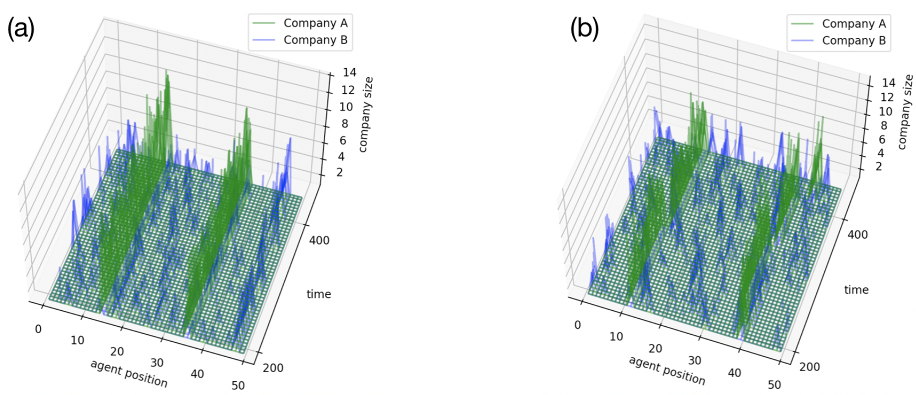


Figure A.1: ***Dynamics of Model 2.a and Model 2.b****: 3D plot showing the dynamics of Model 2.a (panel (a)) and Model 2.b (panel (b)) for $L = 50$, $T = 450$, $\gamma_A = -0.75$, $\gamma_B = 0$, $\tau = 0.4 \times 50 = 20$ and $\sigma = 0.04$. The companies of type Company A are plotted with the color green, and the companies of type Company B with the color blue.*

The dynamics of the two sub-models appear to be similar, and they are indeed. In the figures A.2, A.3, A.4, histograms of the number of companies, the company sizes and the distances are shown for the systems in figure A.1. Without further discussion, we see that the differences in the histograms are small and not significant.

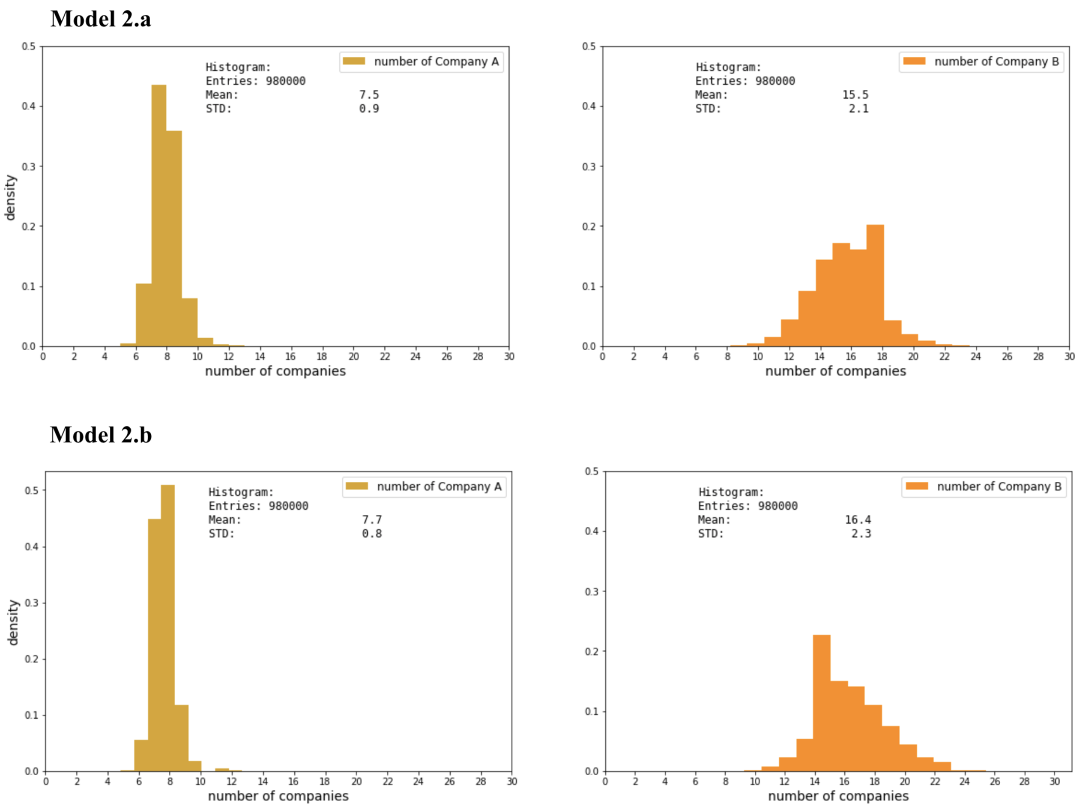


Figure A.2: ***Model 2.a and Model 2.b: Number of Companies****: Histograms of the number of companies with* $S > 1$ *for type Company A (left) and Company (b). The model parameters are* $L = 200$, $T = 10,000$, $\gamma_A = -0.75$, $\gamma_B = -0.25$, $\tau = 0.4 \times 200 = 80$ *and* $\sigma = 0.04$

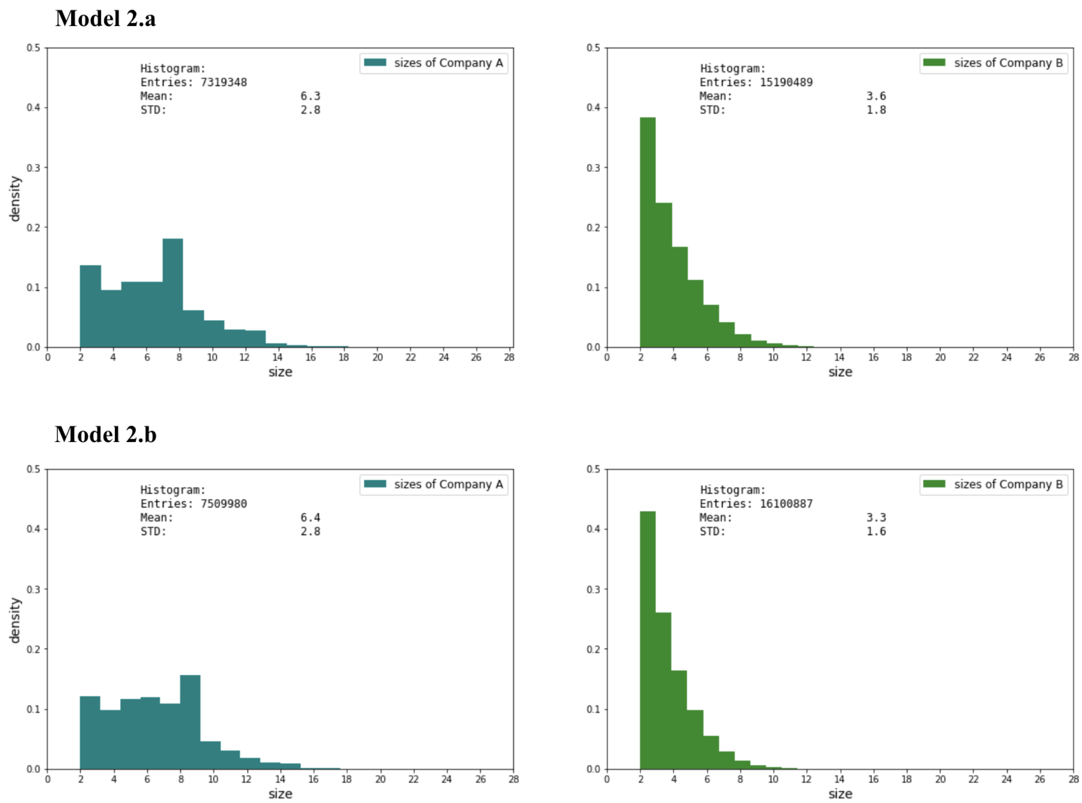


Figure A.3: ***Model 2.a and Model 2.b: Company Sizes***: *Histograms of company sizes for $S > 1$ for type Company A (left) and Company B (right). The model parameters are $L = 200$, $T = 10,000$, $\gamma_A = -0.75$, $\gamma_B = -0.25$, $\tau = 0.4 \times 200 = 80$ and $\sigma = 0.04$*

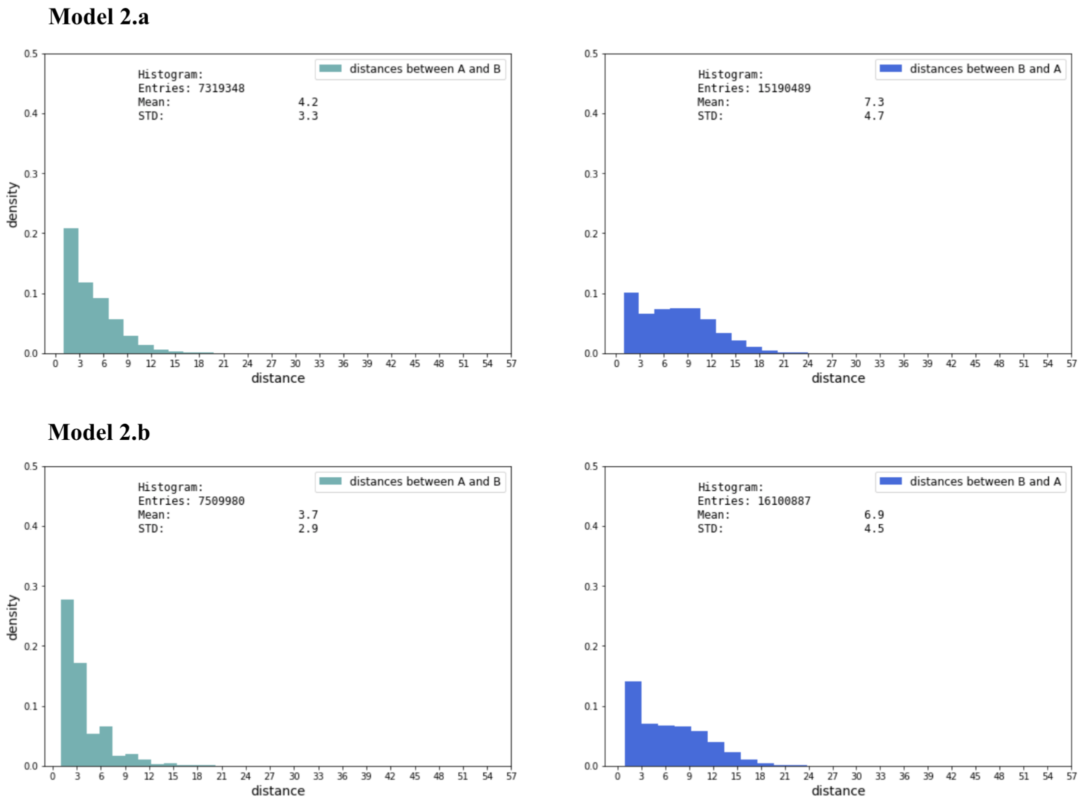


Figure A.4: ***Model 2.a and Model 2.b: Distance between Companies***: *Histograms for the distribution of the distances between companies with $S > 1$. The left histogram shows the distribution of the distance between a company of type Company A and its nearest distance to a company of type Company B. The right histogram shows the distribution of the distance between a company of type Company B and its nearest distance to a company of type Company A. The model parameters are $L = 200$, $T = 10,000$, $\gamma_A = -0.75$, $\gamma_B = -0.25$, $\tau = 0.4 \times 200 = 80$ and $\sigma = 0.04$*

A.1 Model 2.b Versus Model 3

In the next section we will present the results for Model 3. Similar to Model 2.b, the model consists of two company types (A and B) and a dynamic environment. In Model 3 there is a positive feedback between the two company types. The positive feedback has a impact on the distances between the company types. In figure A.5, we show the distribution of the distances between the largest company of type A and its nearest company of type B with $S \geq 3$ (the distribution of the distances between the largest company of type B and its nearest company of type A with $S \geq 3$ is similar as $\gamma_A = \gamma_B = -0.5$).

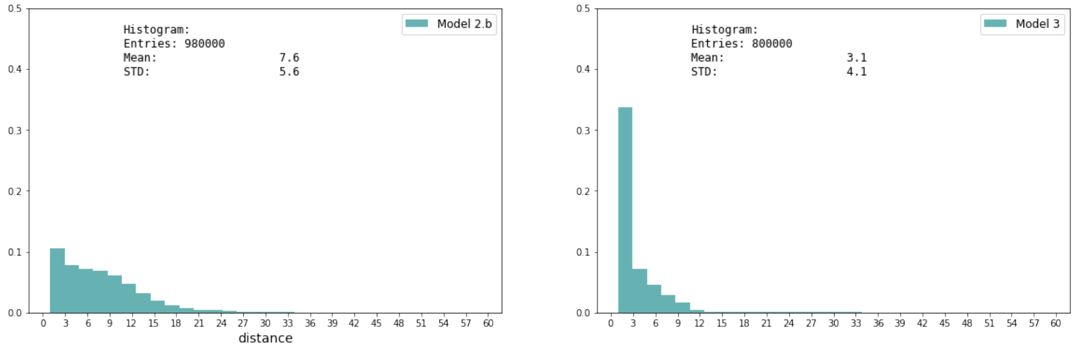


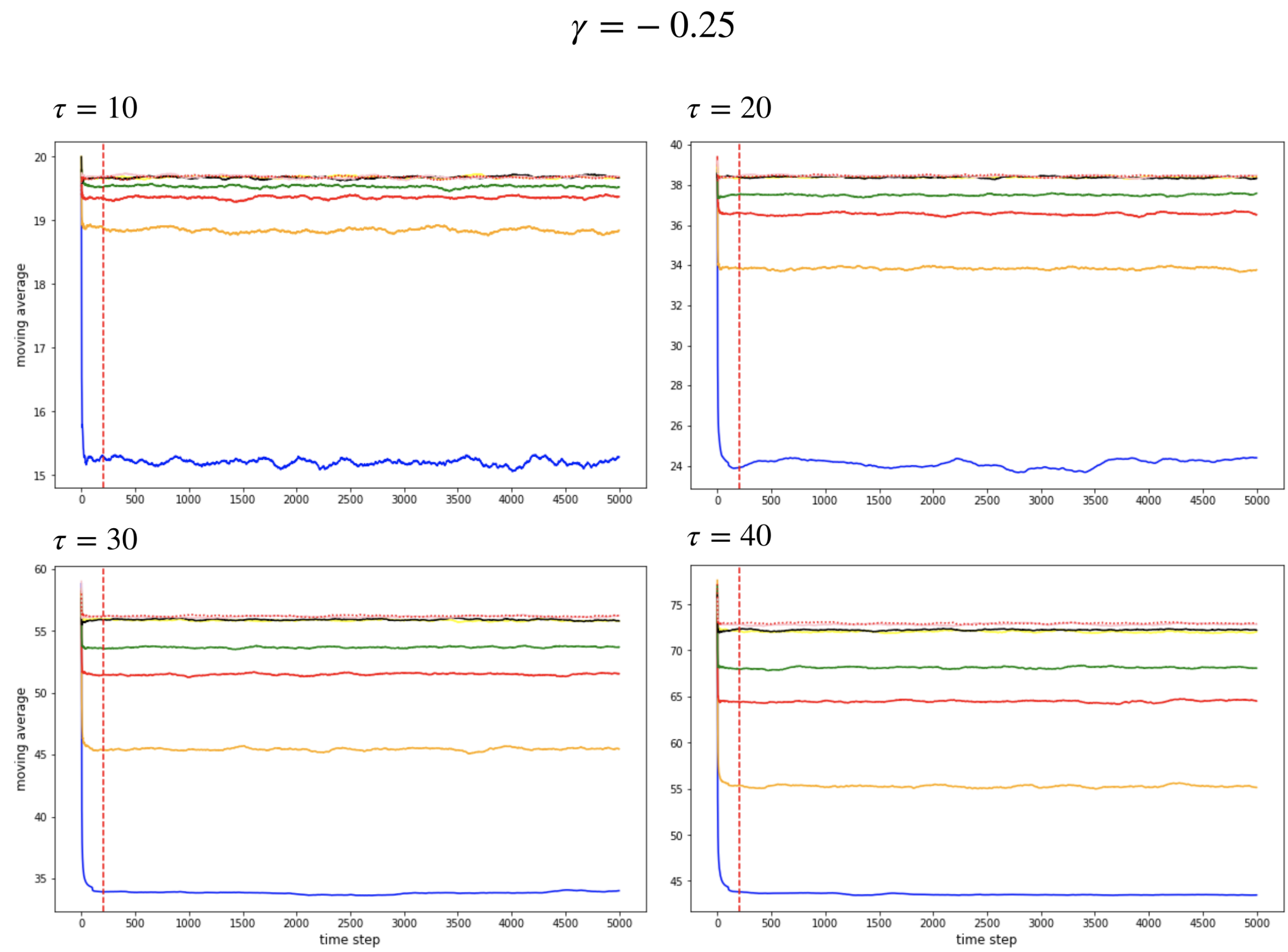
Figure A.5: ***Model 2.a and Model 2.b: Distance between Companies:*** *Histograms of the distribution of the distances between the largest company of type A and its nearest company of type B with* $S \geq 3$*. The right histogram shows the distribution for Model 2.b, and the left for Model 3. The model parameters are* $L = 200$, $T = 10,000$, $\gamma_A = \gamma_B = -0.5$, $\tau = 0.4 \times 200 = 80$ *and* $\sigma = 0.03$

The mean distance for Model 2.b is $8 \pm 6/\sqrt{100}$, and $3 \pm 4/\sqrt{100}$ for Model 3. This difference in the means, with Model 3 having the shortest mean distance, is an indication of the positive feedback between the two company types. In the next section, we will continue this discussion.

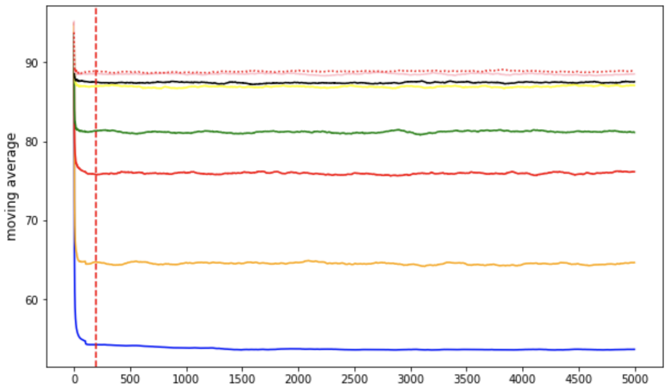
Appendix B

Steady State

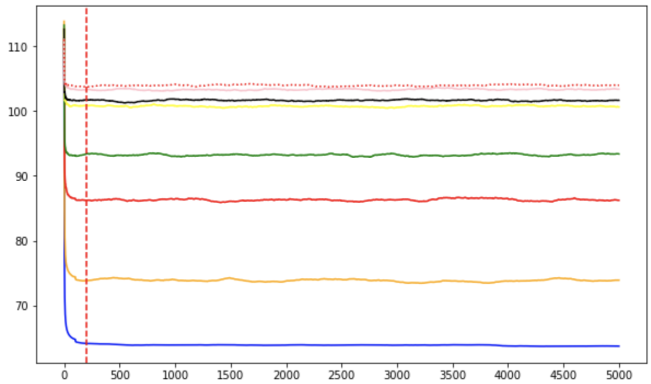
Model 1


γ = − 0.25
τ = 10
τ = 20
τ = 30
τ = 40
moving average
moving average
time step
time step

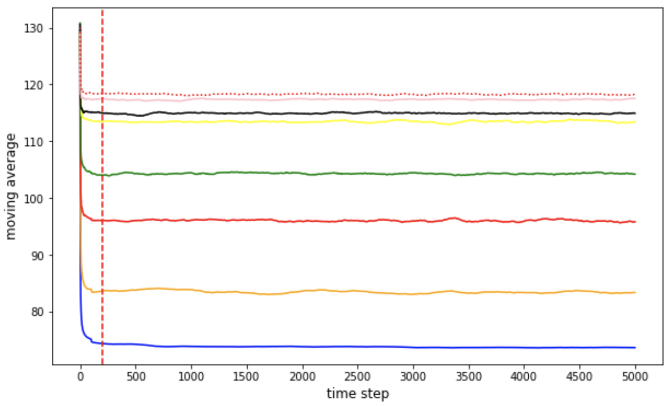
$\tau = 50$



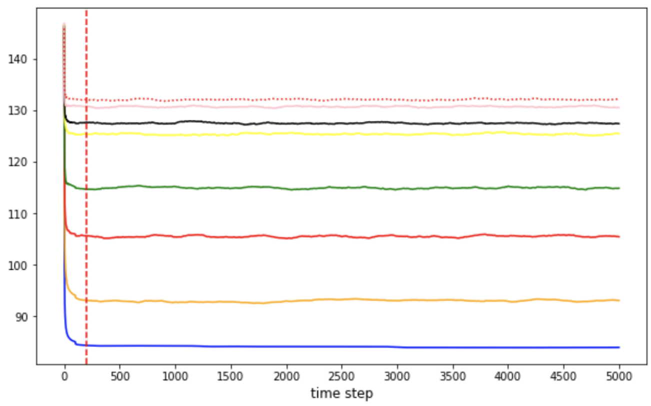
$\tau = 60$



$\tau = 70$



$\tau = 80$



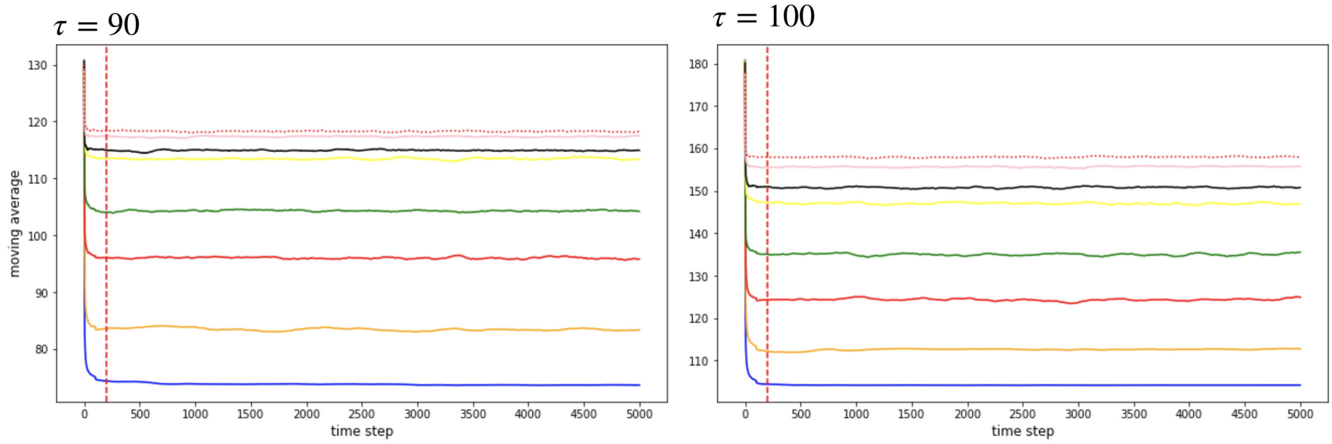
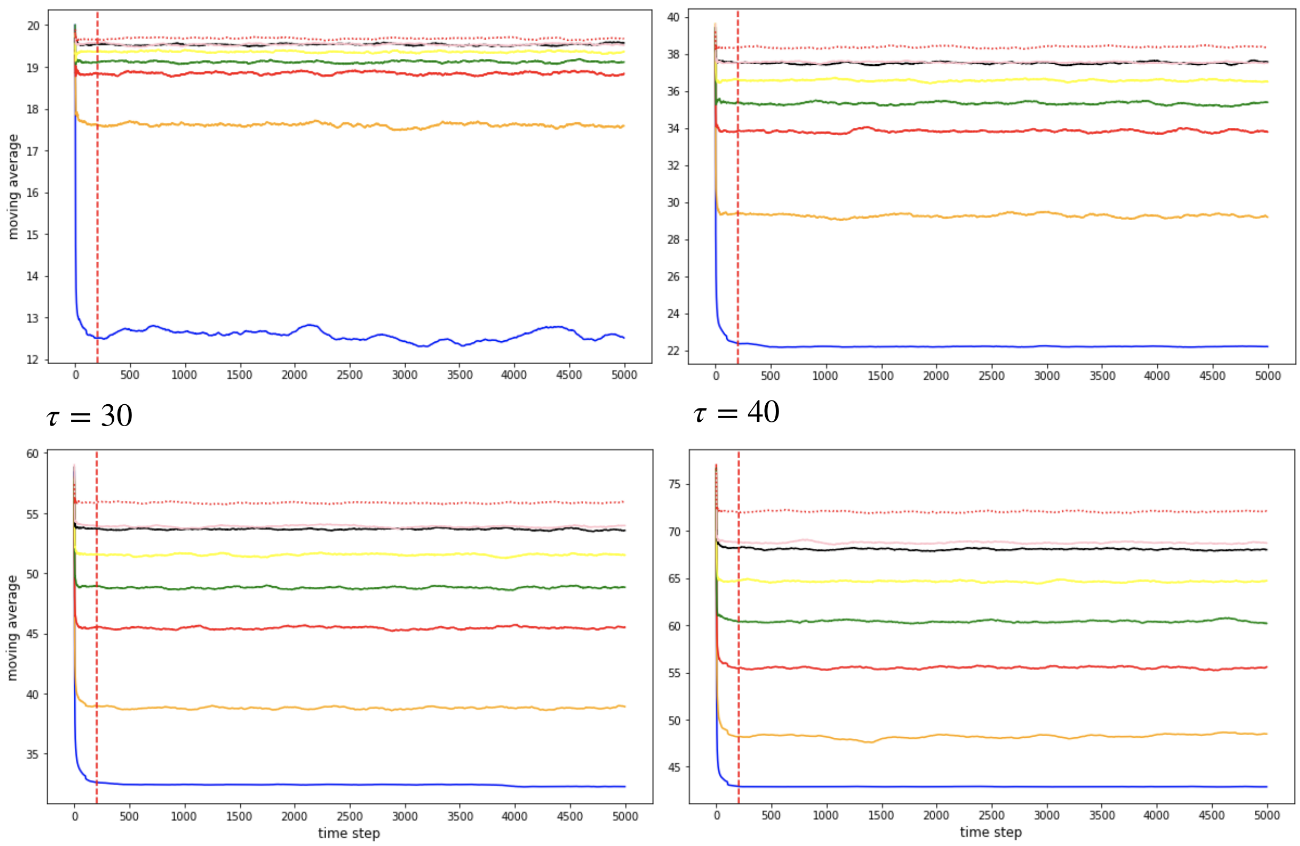
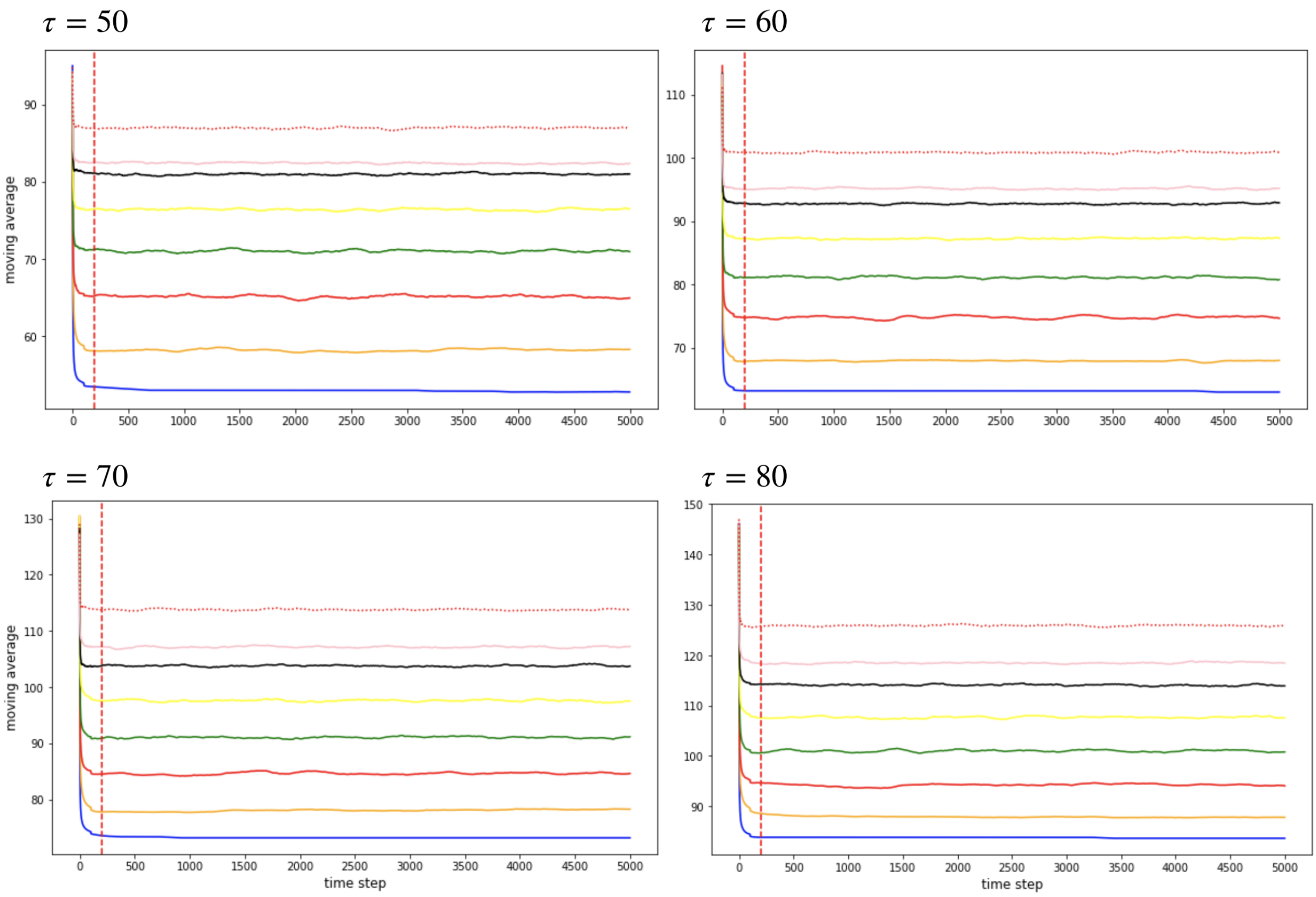


Figure B.1: ***Model 1: Steady State***: *Plots of the moving average with $w = 100$ as a function of time steps for different values of Δt (color representation is shown in figure 5.2). All system have size $L = 200$ and $T = 10,000$, and $\gamma = -0.25$*

$\gamma = -0.5$

$\tau = 10$ | $\tau = 20$ | $\tau = 30$ | $\tau = 40$





τ = 50
τ = 60
τ = 70
τ = 80
moving average
time step

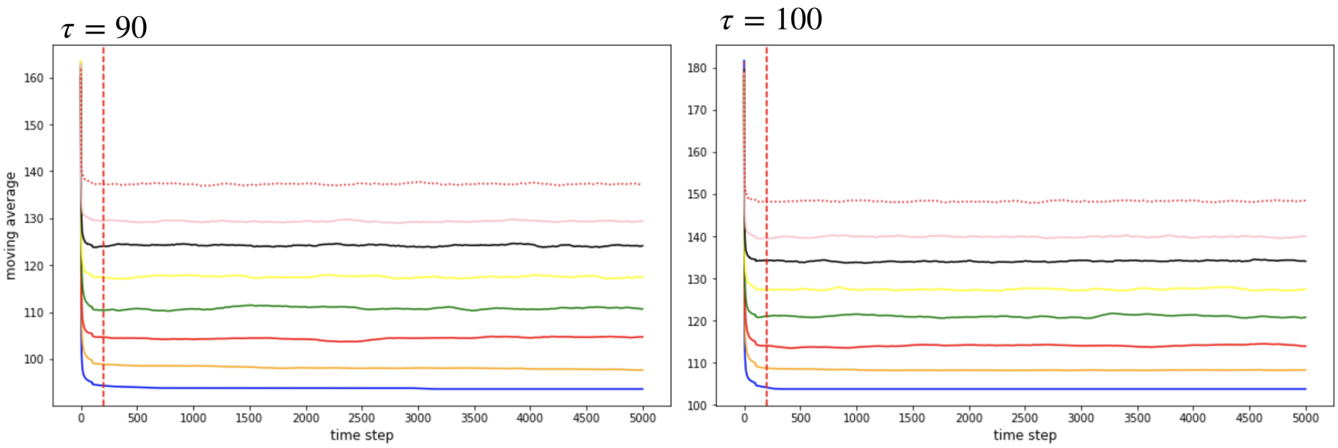
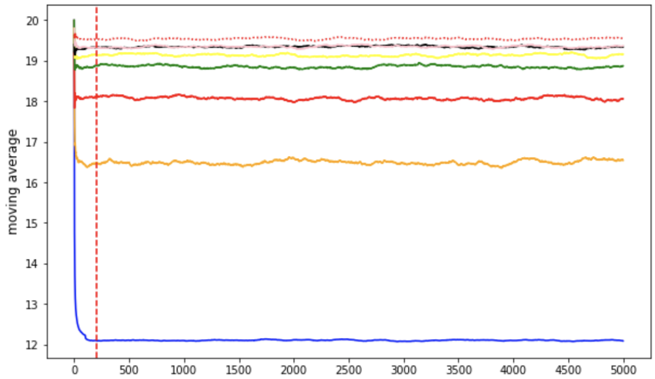


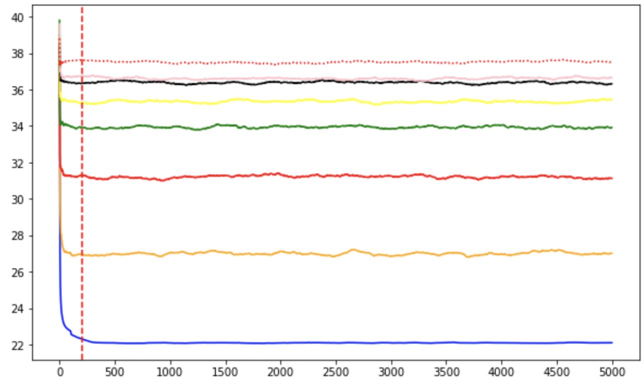
Figure B.2: ***Model 1: Steady State:*** *Plots of the moving average with $w = 100$ as a function of time steps for different values of Δt (color representation is shown in figure 5.2). All system have size $L = 200$ and $T = 10,000$, and $\gamma = -0.5$*

$\gamma = -0.75$

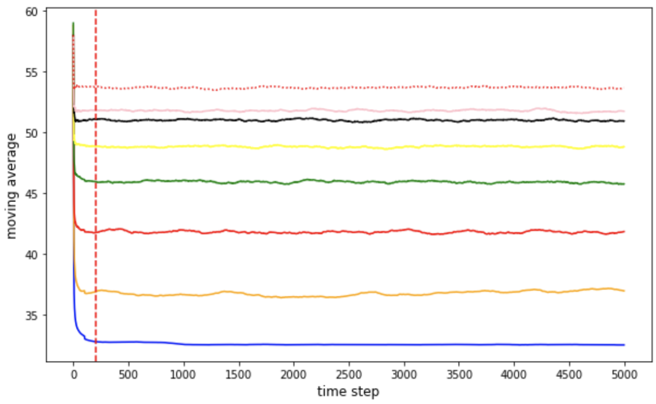
$\tau = 10$



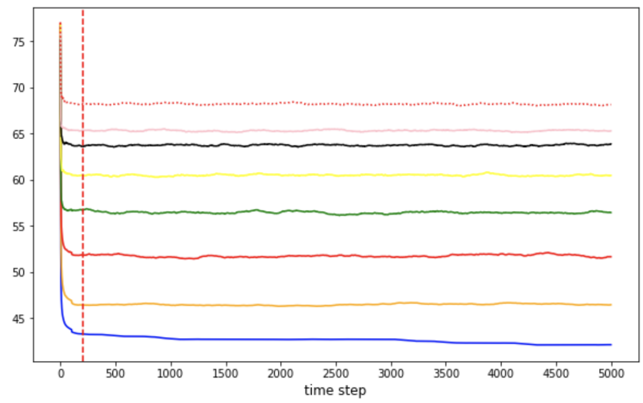
$\tau = 20$



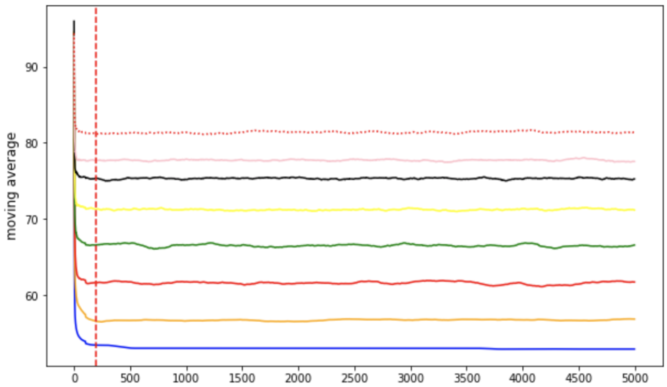
$\tau = 30$



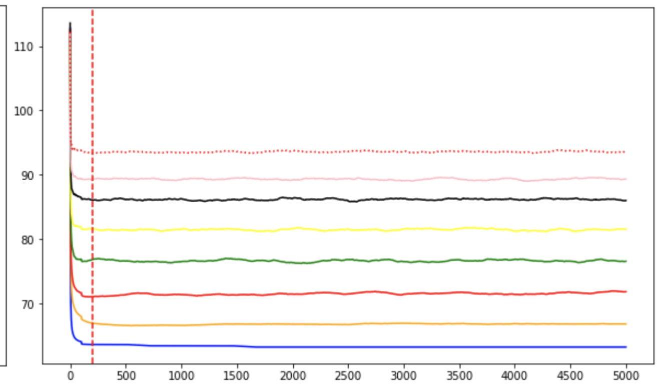
$\tau = 40$



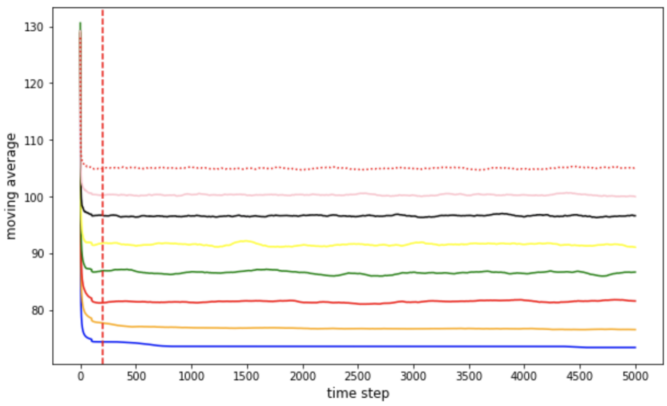
$\tau = 50$



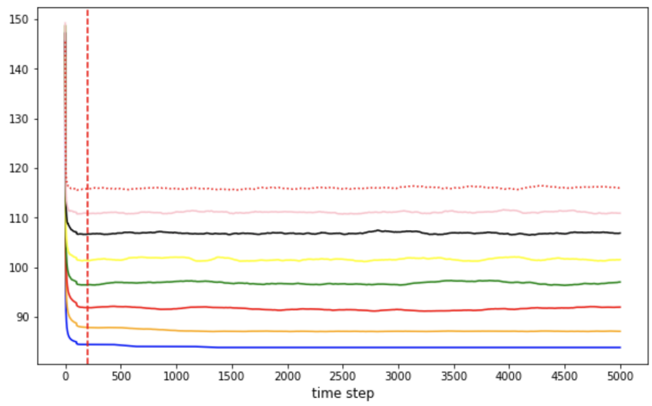
$\tau = 60$



$\tau = 70$



$\tau = 80$



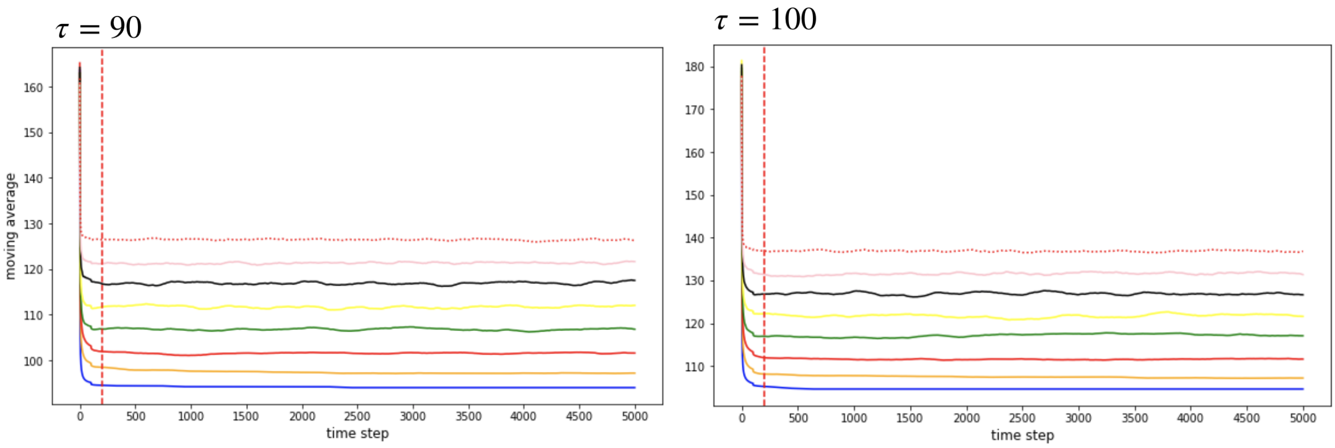


Figure B.3: ***Model 1: Steady State***: *Plots of the moving average with $w = 100$ as a function of time steps for different values of Δt (color representation is shown in figure 5.2). All system have size $L = 200$ and $T = 10,000$, and $\gamma = -0.75$*

Model 2

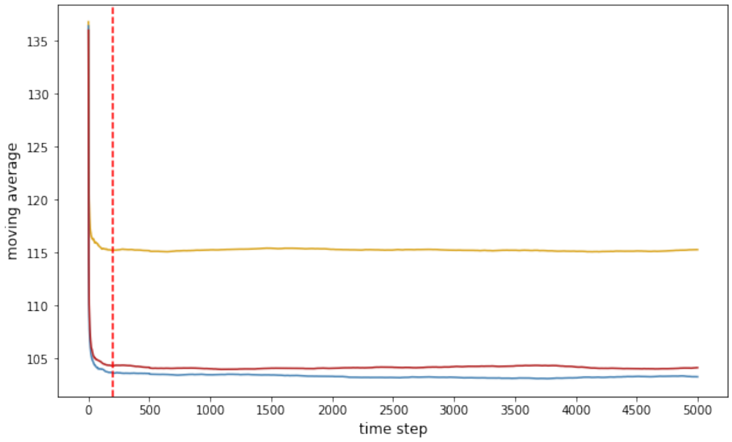
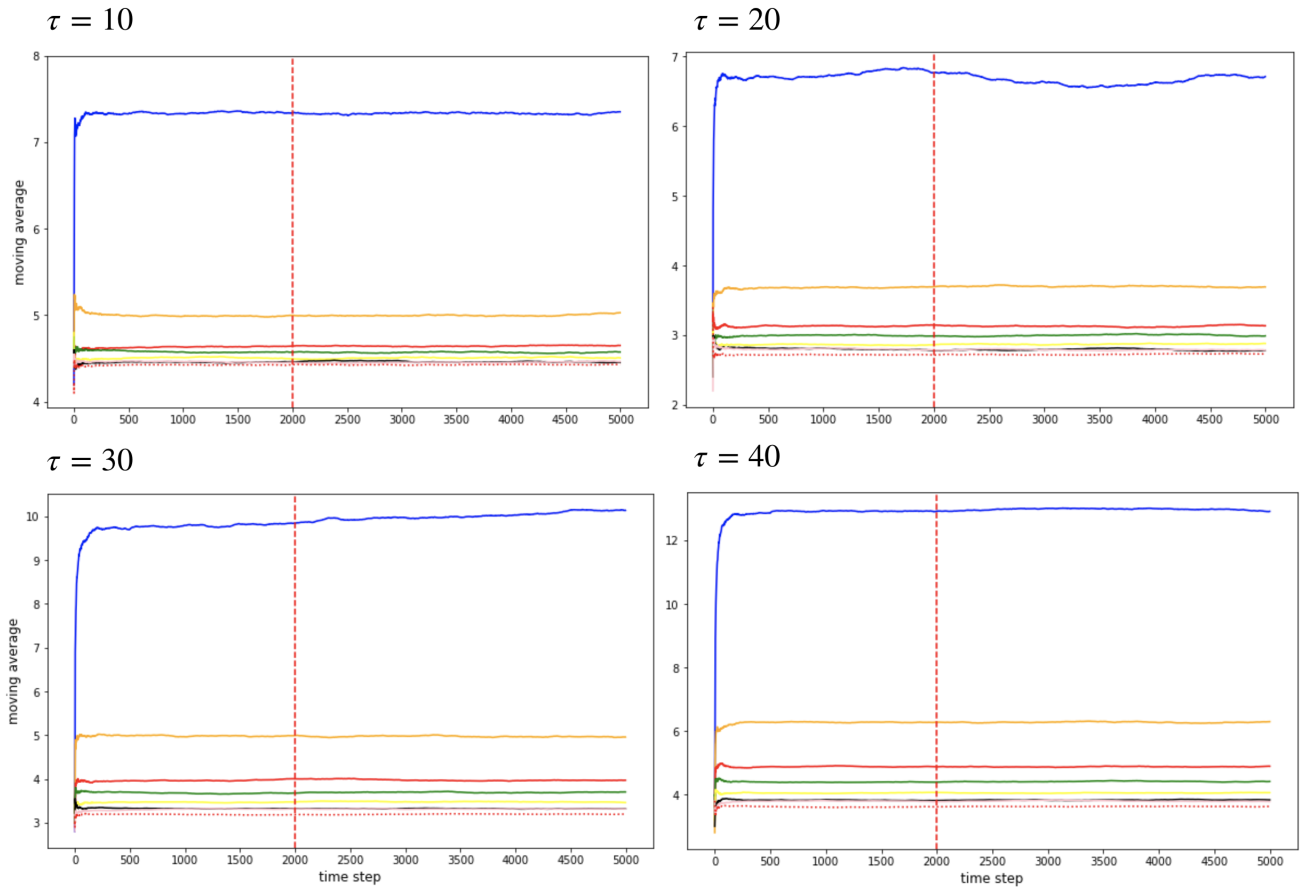
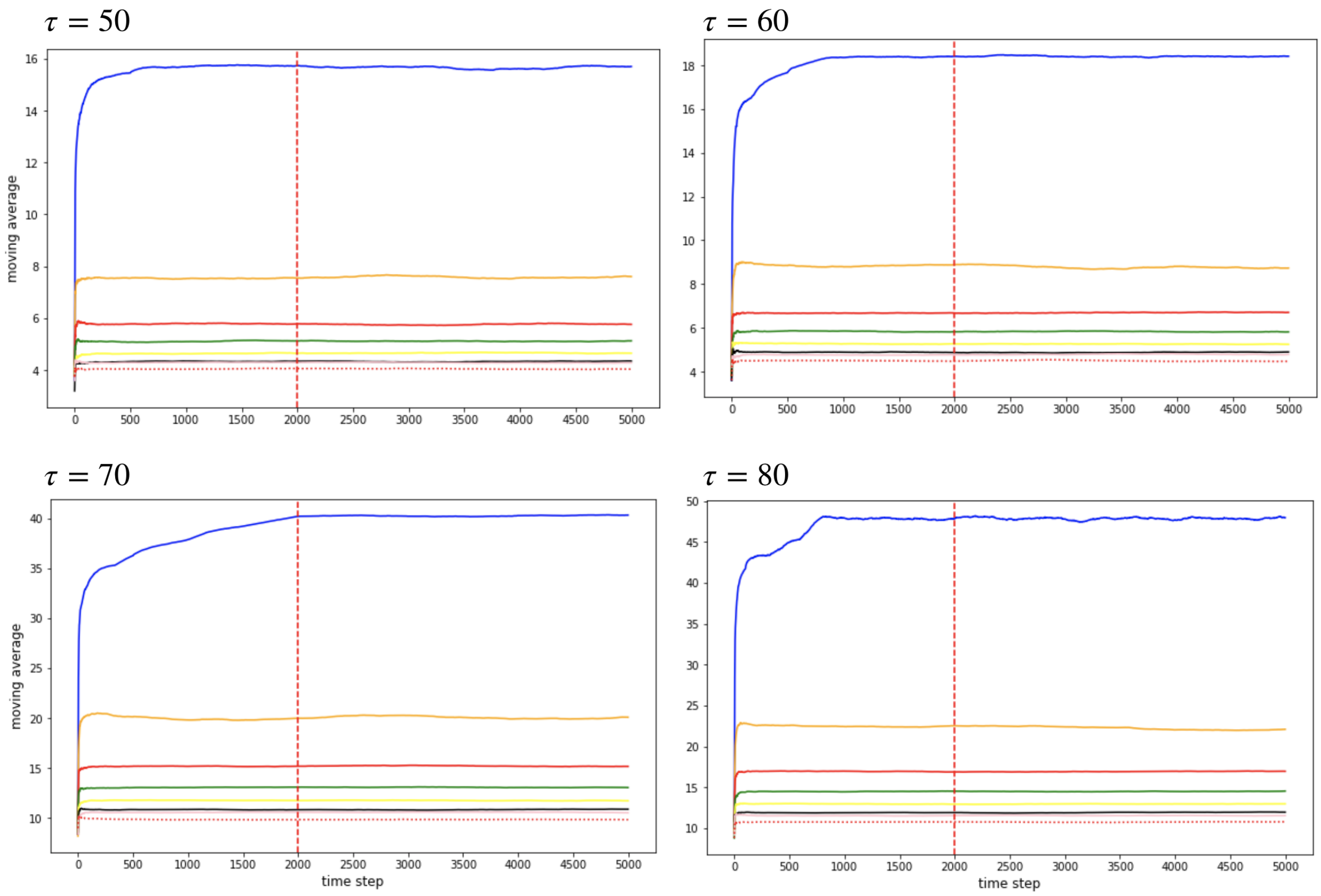


Figure B.4: ***Model 2.a and 2.b: Steady State****: Plots of the moving average with w=150 as a function of time steps. All systems have size $L = 200$ and $T = 10,000$.* ***Model 2.a*** *Yellow curve: $\gamma_A = -0.75$, $\gamma_B = 0$, $\tau = 0.4 \times 200 = 80$ and $\sigma = 0.04$. Blue curve: $\gamma_A = -0.75$, $\gamma_B = -0.25$, $\tau = 0.4 \times 200 = 80$ and $\sigma = 0.04$.* ***Model 2.b*** *Red curve: $\gamma_A = -0.75$, $\gamma_B = -0.25$, $\tau = 0.4 \times 200 = 80$ and $\sigma = 0.04$. Green curve: $\gamma_A = \gamma_B = -0.5$, $\tau = 0.3 \times 200 = 80$ and $\sigma = 0.04$*

Model 3

$$\gamma_A = \gamma_B = -0.5$$





τ = 50
τ = 60
τ = 70
τ = 80
moving average
moving average
time step
time step

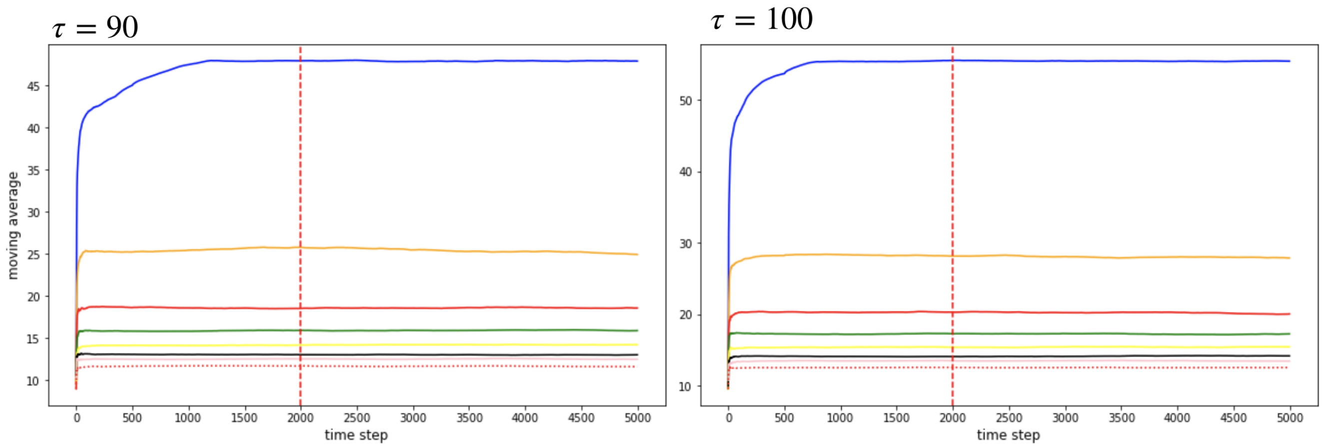
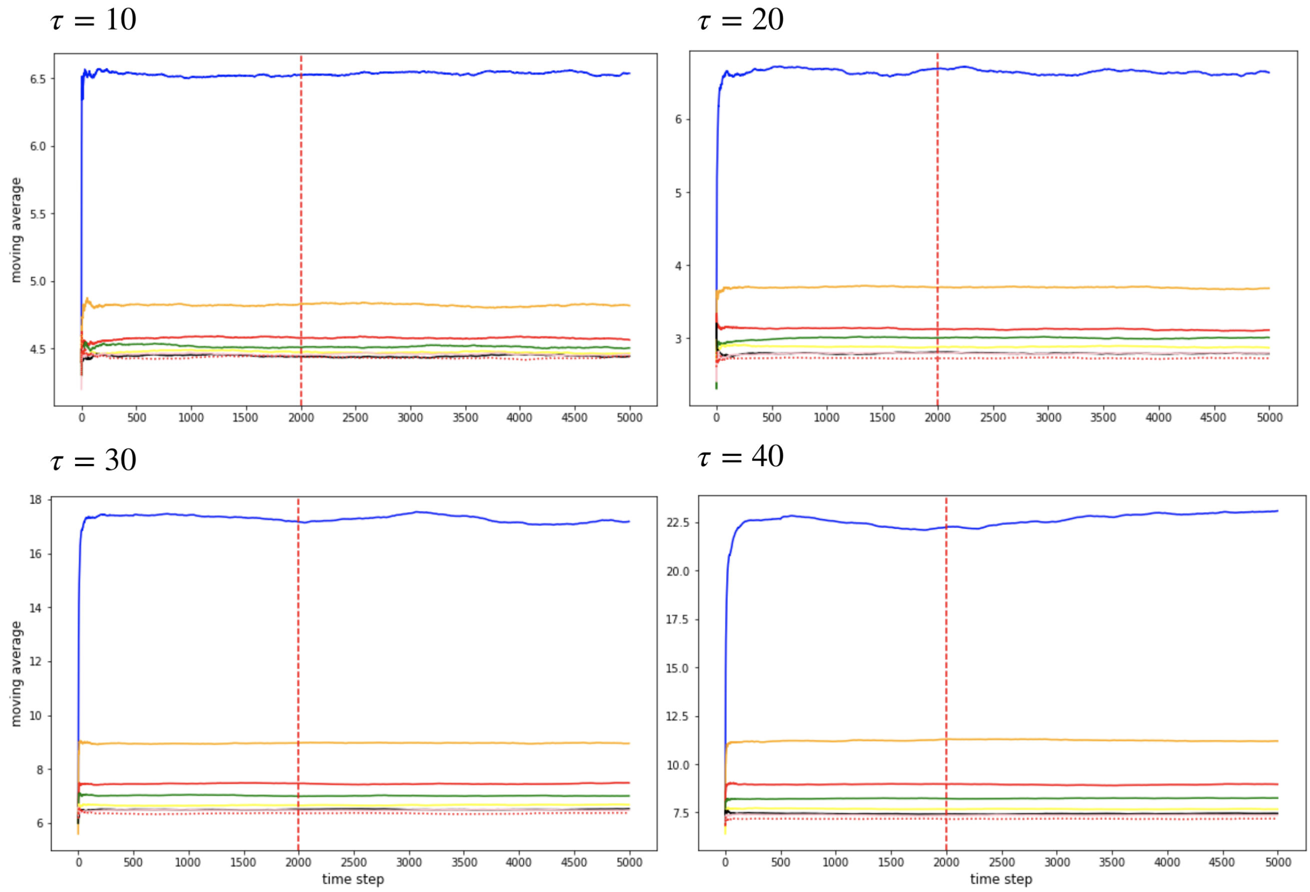
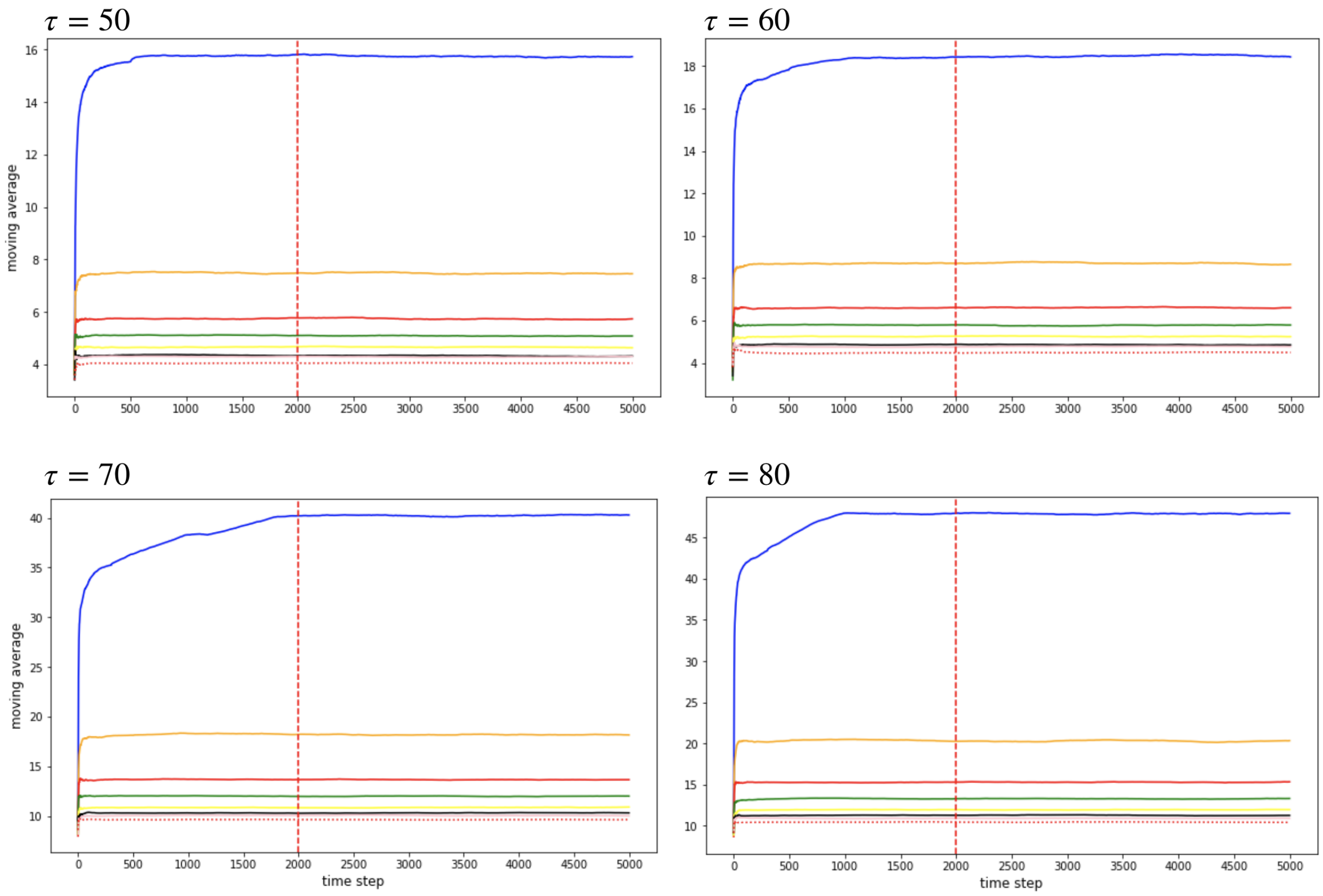


Figure B.5: ***Model 3: Steady State***: *Plots of the moving average with $w = 500$ as a function of time steps for different values of Δt (color representation is shown in figure 5.2). All system have size $L = 200$ and $T = 10,000$, and $\gamma_A = \gamma_B = -0.5$*

$\gamma_A = -0.75 \quad \gamma_A = -0.25$




τ = 50
τ = 60
τ = 70
τ = 80
moving average
moving average
time step
time step

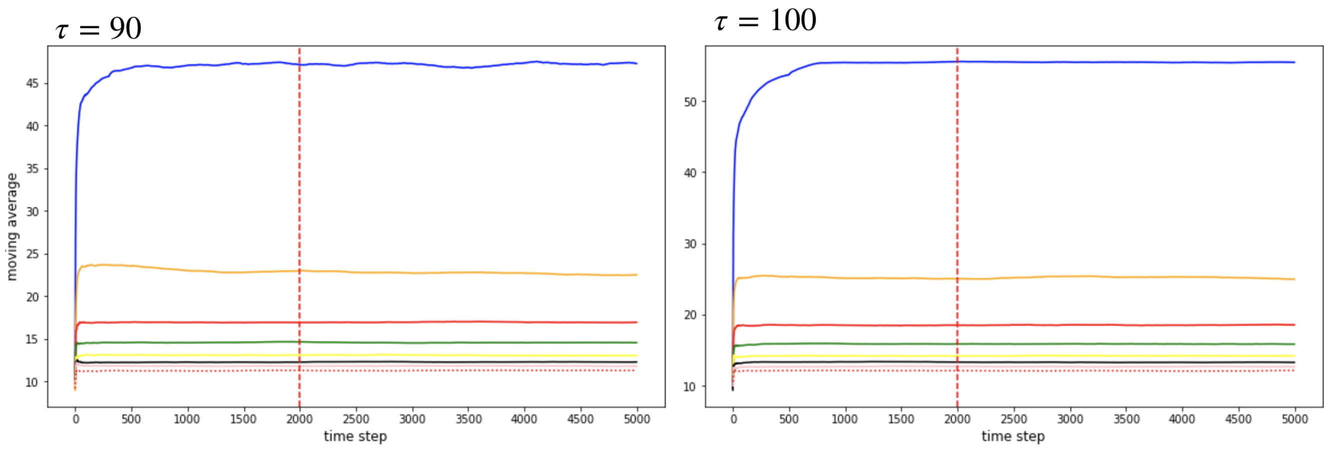
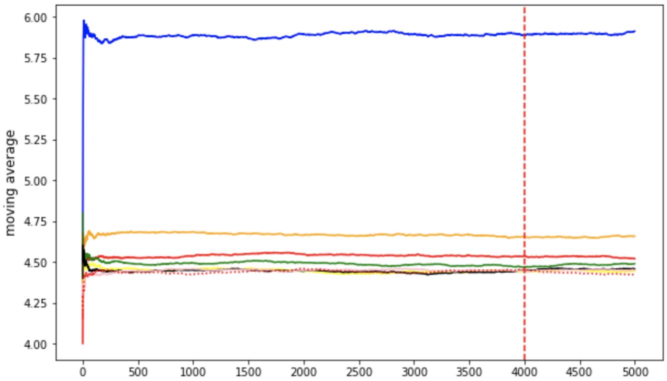


Figure B.6: ***Model 3: Steady State***: *Plots of the moving average with* $w = 500$ *as a function of time steps for different values of* Δt *(color representation is shown in figure 5.2). All system have size* $L = 200$ *and* $T = 10,000$*, and* $\gamma_A - 0.75$*,* $\gamma_B = -0.25$

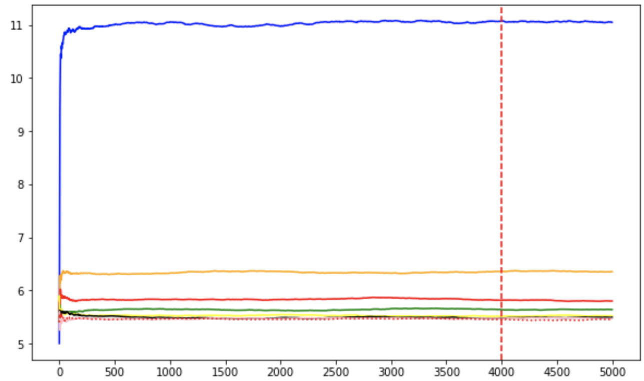
Model 4

$$\gamma_A = \gamma_B = -0.5$$

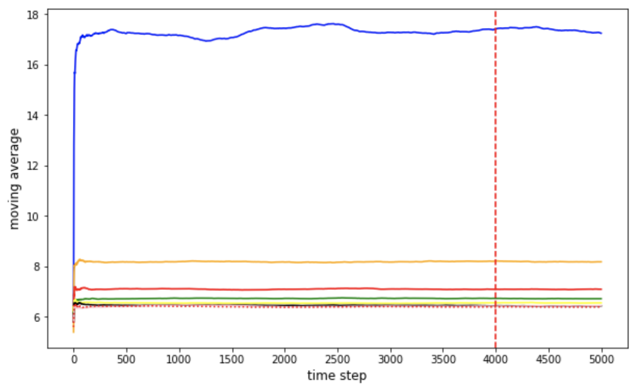
$\tau = 10$



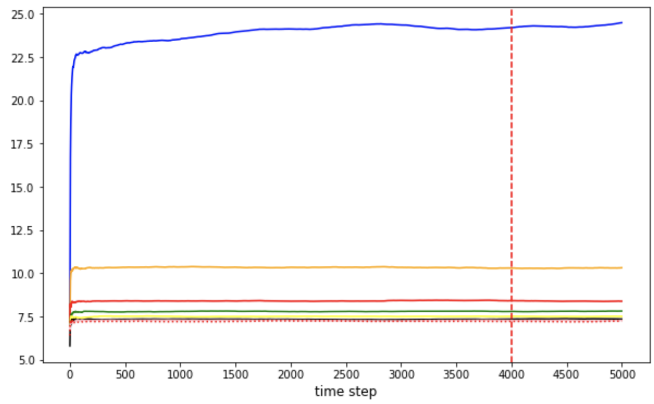
$\tau = 20$

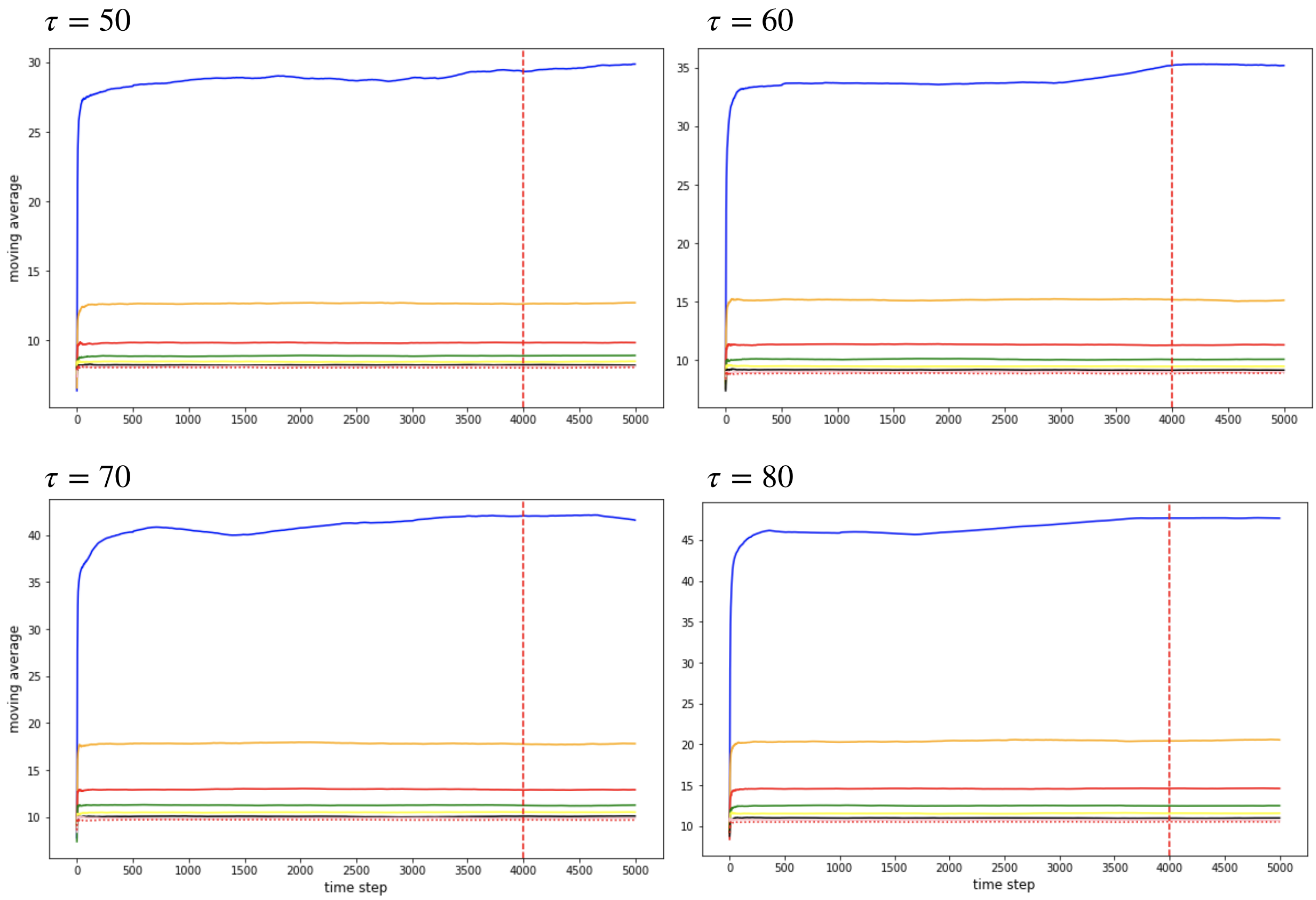


$\tau = 30$



$\tau = 40$




τ = 50
τ = 60
τ = 70
τ = 80
moving average
time step

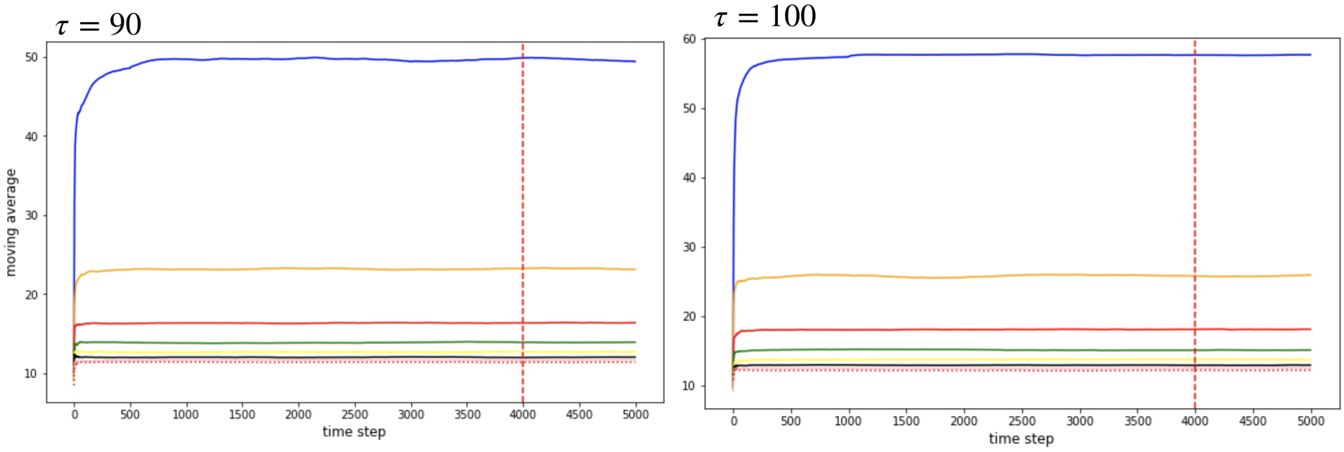


Figure B.7: ***Model 4: Steady State***: *Plots of the moving average with $w = 500$ as a function of time steps for different values of Δt (color representation is shown in figure 5.2). All system have size $L = 200$ and $T = 10,000$, and $\gamma_A = \gamma_B = -0.5$*

Appendix C

Python Code

Model 1

```python
def model1(L, gamma, tau, sigma, T, t_data):
    '''Input: L = the number of agens, gamma= economies of scale exponent, tau=number of interactions,
    T=total simulation time steps, t_data=time step at which data is saved.
    Possible outputs: S=array with all sizes, Size=list with all sizes S>1, sum_size=sum of all sizes S>1,
    number= number of companies with S>1, distance= distance between companies with S>1'''

    #List/array to save data
    lifetime=[] #company lifetime
    Size=[] #company sizes
    sum_size=[] #sum of company sizes
    number=[] #number of companies
    distance=[] #distance between companies
    S=np.zeros((T, L)) #array with all size

    current_S=np.ones(L) #keep track of the company sizes
    life=np.zeros(L) #list with lifetime of companies

    for t in range(T):
        size=np.zeros(L) #list with the number of rewards for time step t

        for _ in tau:
            x=random.choice(range(L)) #The position of the buying agent
            cost=[] #total cost of selling agents
            index=[] #position of selling agents

            for i in range(L):
                if x!=i:
                    distance = np.min([abs(x-i), abs((i-x)-L)]) #distance from buyer to seller
                    c = current_S**gamma+distance*sigma #total cost
                    cost.append(c)
                    index.append(i)
            if not(cost):
                break
            else:
                min_cost = np.min(cost) #finding the smallest cost
                Y = np.where(cost==min_cost)
                Y = int(random.choice(Y[0])) #if more than one company offering the same low cost
                Y = int(index[Y]) #the position of the selling company
                size[Y] = size[Y] + 1 #company is rewarded

        current_S=np.ones(L)
        current_S=current_S + size #opdating company sizes
```

```python
# ------------------------------------------------------------------------------------
        # DATA
        for i in range(L):
            if current_S[i]==1:
                S[t,i]=0
            if current_S[i]>1:
                S[t,i]=current_S[i]

        if t<t_data:
            for i in range(L):
                _, life = lifetime_model1(i, current_S, life, t, T) #counting life

        if t>=t_data:
            # Lifetime
            for i in range(L):
                l, life = lifetime_model1(i, current_S, life, t, T) #counting life and saving lifetime

                if l>0:
                    lifetime.append(l)

            #Distance, size, number of companies
            n, s, sumS, d = distance_size_model1(L, current_S)

            Size.append(s) # company sizes
            sum_size.append(sumS)  # sum of company sizes
            number.append(n)  # number of companies
            distance.append(d)  # distance between companies

    return S, Size, sum_size, number, distance, lifetime
```

Figure C.1: *Code: Model 1. Functions found in figure C.2 and C.6*

```python
def lifetime_model1(i, current_S, life, t, T):
    '''lifetime of companies'''
    l=0

    if current_S[i]>1:
        life[i] = life[i] + 1 #one count is added to the lifetime

    if life[i] >0 and t==T-1: #last time step
        l=life[i]

    if life[i]>0 and current_S[i]==1: #size drops to S=1
        l=life[i]
        life[i]=0

    return l, life
```

Figure C.2: *Code: Model1, finding company lifetime*

```python
def distance_size_model1(L, current_S):
    '''Output: Number of companies, company sizes, sum of company sizes, distance between companies '''
    #Size and number of companies
    sum_size=0
    size=[] #size of each company with S>1
    position=[]
    for i in range(L):
        if current_S[i] >1:
            sum_size=sum_size+current_S[i] #sum of company sizes
            size.append(current_S[i]) #size of each company with S>1
            position.append(i)
    number=len(position) #number of companies with S>1

    #distanc
    distance=[]
    for i in range(number):
        D=[]
        for j in range(number):
            if i!=j:
                d=np.min([abs(position[i]-position[j]), abs(L-abs(position[i]-position[j]))])
                D.append(d)
        distance.append(np.min(D)) #shortest distance

    return number, sum_size, size, distance
```

Figure C.3: *Code: Model 1, finding company size, number of companies and distance between companies*

Two Types of Companies (Model 2, Model 3, Model 4)

```
def model2(L, gamma_A, gamma_B, tau, sigma, T, t_data):

    #List/array to save data:
    lifeA=[] #lifetime of companies with type A
    lifeB=[] #lifetime of companies with type BB

    sizeA=[] #company sizes
    sizeB=[]
    sum_size=[] #sum of all company size

    numberA=[] #number of companies
    numberB=[]

    distanceAA=[]
    distanceBB=[]
    distanceAB=[]
    distanceBA=[]
```

```python
#Defining the companies
gamma=[gamma_A, gamma_B]
current_company=np.zeros(L) #containing information about the company type of the position
for i in range(L): #choosing the company type of the position
    current_company[i]= random.randint(0,1)

current_S=np.ones(L) #initial company sizes
life=np.zeros(L)

for t in range(T):
    size=np.zeros(L)

    for _ in range(tau):
        cost=[]
        index=[]
        x = random.choice(range(L)) #choosing the purchasing company

        for i in range(L):
            if not(int(current_company[i])==int(current_company[i])): #interaction between opposite type
                distance=np.min([abs(x-i), abs(abs(x-i)-L)]) #periodic boundary conditions
                c=current_S[i]**gamma[int(current_company[i])] + distance*sigma
                cost.append(c)
                index.append(i)

        if not cost: # if empty
            break
        else:
            min_cost=np.min(cost)
            Y=np.where(cost==min_cost)
            Y=int(random.choice(Y[0]))
            Y=int(index[Y])
            size[Y]=size[Y]+1 #reward to the selling company

    current_S=np.ones(L)
    current_S= current_S+size
```

```
    if t<t_data:
        for i in range(L):
            _, _, life=lifetime_model2(i, current_S, current_company, life, t, T)


    if t>= t_data:
        #lifetime:
        for i in range(L):
            lA, lB, life=lifetime_model2(i, current_S, current_company, life, t, T)

            if lA >0:
                lifeA.append(lA)
            if lB >0:
                lifeB.append(lB)

        #number of companies, company sizes and distances:
        nA, nB, ss, sA, sB, dAA, dBB, dAB, dBA = distance_size_model2(L, current_company, current_S)


        numberA.append(nA)
        numberB.append(nB)

        sum_size.append(ss)
        sizeA.append(sA)
        sizeB.append(sB)


        distanceAA.append(dAA)
        distanceBB.append(dBB)
        distanceAB.append(dAB)
        distanceBA.append(dBA)

return sizeA, sizeB, distanceAA, distanceBB, distanceAB, distanceBA
```

Figure C.4: *Code: Model 2, Model 3, Model 4 . Functions found in figure C.5 and C.6*

```python
def lifetime_model2(i, current_S, current_company, life, t, T):
    '''Lifetime for models with two company types'''
    lA=0
    lB=0

    if current_S[i]>1:
        life[i]=life[i]+1
    if life[i]>0 and t==T-1: #last time step.. save
        if int(current_company[i])==0: #typeA
            lA=life[i]
        if int(current_company[i])==1: #typeB
            lB=life[i]

    if life[i]>0 and current_S[i]==1:
        if int(current_company[i]) == 0:  # typeA
            lA = life[i]
            life[i]=0
        if int(current_company[i]) == 1:  # typeB
            lB = life[i]
            life[i]=0

    return lA, lB, life
```

Figure C.5: ***Code: Model 2, Model 3, Model 4, finding company lifetime***

```python
def distance_size_model2(L, current_company, current_S):
    '''For models with two company types.
    Outputs: Number of companies, company sizes, sum of companies, distance between companies'''
    indexA=[]
    indexB=[]
    sizeA=[]
    sizeB=[]
    sum_size=0

    dAA=[]
    dBB=[]
    dAB=[]
    dBA=[]
    #Size and number of companies:
    for i in range(L):
        if current_S[i]>1 and int(current_company[i])==0:
            sizeA.append(current_S[i])
            sum_size=sum_size+current_S[i]
            indexA.append(i)

        if current_S[i]>1 and int(current_company[i])==1:
            sizeB.append(current_S[i])
            sum_size=sum_size+current_S[i]
            indexB.append(i)

    #number of companies:
    numberA=len(indexA)
    numberB=len(indexB)
    number= numberA + numberB
```

```python
    #_____Distance from A to A____#distance between a company of type A to a company of type A
    if numberA==0 or not(indexA):
        dAA.append(0)
    else:
        for i in range(len(indexA)):
            daa=[]
            for j in range(len(indexA)):
                if i!=j:
                    d=np.min([abs(indexA[i]-indexA[j]),abs(L-abs(indexA[i]-indexA[j]))])
                    daa.append(d)
            dAA.append(np.min(daa))

    #_____Distance from B to B___________________
    if numberB==0 or not(indexB):
        dBB.append(0)
    else:
        for i in range(len(indexB)):
            dbb=[]
            for j in range(len(indexB)):
                if i!=j:
                    d=np.min([abs(indexB[i]-indexB[j]),abs(L-abs(indexB[i]-indexB[j]))])
                    dbb.append(d)
            dBB.append(np.min(dbb))
```

```python
# _____Distance from A to B___________
indexAB= indexA + indexB
indexAB = sorted(indexAB)

for i in range(len(indexAB)):
    dab=[]
    for j in range(len(indexAB)):
        if i!=j:
            if indexAB[i] in indexA and indexAB[j] in indexB:
                d = np.min([abs(indexAB[i] - indexAB[j]), abs(L - abs(indexAB[i] - indexAB[j]))])
                dab.append(d)
    if not(not(dab)):
        dAB.append(np.min(dab))

# _____Distance from B to A___________

for i in range(len(indexAB)):
    dba = []
    for j in range(len(indexAB)):
        if i != j:
            if indexAB[i] in indexB and indexAB[j] in indexA:
                d = np.min([abs(indexAB[i] - indexAB[j]), abs(L - abs(indexAB[i] - indexAB[j]))])
                dba.append(d)
    if not (not (dba)):
        dBA.append(np.min(dba))

if not(dAA):
    dAA.append(0)
if not(dBB):
    dBB.append(0)
if not(dAB):
    dAB.append(0)
if not(dBA):
    dBA.append(0)

return numberA, numberB, sum_size, sizeA, sizeB, dAA, dBB, dAB, dBA
```

Figure C.6: *Code: Model 2, Model 3, Model 4, finding company size, number of companies and distance between companies*

Dynamic Environment

```python
# Dynamic enviornment:
for i in range(L):
    if int(current_S[i]) == 1: # if S=1 (size equal 1, no sales)
        current_company[i] = random.randint(0,1) #new company type
```

Figure C.7: ***Code: Dynamic environment***

Positive Feedback

```python
for _ in range(tau):
    cost=[]
    index=[]
    X=random.randint(0,1) #randomly choosing between the type of the purchesing company

    xx=[]
    Sxx=[]
    for i in range(L):
        if int(current_company[i])==x: #type of the company
            xx.append(i)
            Sxx.append(current_S[i])
    x = random.choices(xx, weights=Sxx, k=1) #the probability is weighted with company size
    x=int(x[0])

    #Below: calculation of the total cost
```

Figure C.8: ***Code: Positive Feedback***

Negative Feedback

```
for i in range(L):
    S_factor=0
    if not(int(current_company[i])==int(current_company[x])):
        l=np.arange(L)
        J=np.roll(l, int(len(l)-i+current_S[i]))
        J=J[0:int(2*current_S[i]+1)] #only sizes for companies within the distance current_S[i]

        for j in J:
            if j!=i:
                S_factor=S_factor+current_S[j]

#Below: calculation of the total cost
```

Figure C.9: *Code: Negative Feedback*

Multi Processing

```
import pickle
from multiprocessing import Pool, cpu_count

def run(i):
    print('Running number', i)
    return function(L, gamma, sigma, tau, T, t_data)

if __name__=='__main__':
    pool = Pool(cpu_count())
    data = list(pool.map(run, list(range(N))))
    with open('filename.pkl', 'wb') as f:
        pickle.dump(data, f)
```

Figure C.10: *Code: Multiprocessing*

Welch's Method

```
def mean_sum(file, index):
    #Open file:
    with open(file, 'rb') as f:
        data=pickle.load(f)

    #Data: Sum of sizes:
    sum_size = [data[i][index] for i in range(len(data))]

    #Find the mean for each time step:
    S=[]
    for i in range(len(sum_size[0])):
        s=[]
        for j in range(len(sum_size)):
            a=sum_size[j]
            s.append(a[i])
        S.append(np.mean(s))

    return S
```

```
#Moving average:
def moving_avg(S, w):
    length=len(S)
    m_avg=[]

    for t in range(length):
        if t%1000==0:
            print(t)
        if t<=w:
            if t==0:
                y=np.sum(S[0:1])
                m_avg.append(y)
            else:
                y=np.sum(S[0:(2*t)])/(2*t)
                m_avg.append(y)

        if t>w and t<=length-w:
            y=np.sum(S[(t-w+0-1):(t-w+2*w)])/(2*w+1)
            m_avg.append(y)
    return m_avg
```

Figure C.11: ***Code: Welch's Method***

Mean of Variable

```
def mean_func(data, index):
    '''Mean of variable'''
    mean_list=[]
    z_list=[data[i][index] for i in range(len(data))] #z is a list containing all time series

    for i in range(len(z_list)):
        z=z_list[i] #time series nr i
        mean_list.append(np.mean(z)) #the mean of the time series

    z_mean=np.mean(mean_list)

    return z_mean
```

Figure C.12: *Code: Mean of Variable*

Data for Histograms

```
def open_file(file_name):
    with open(file_name, 'rb') as f:
        data=pickle.load(f)

    return data
```

```
#Data used for histograms:
def histogram_data(data, index, data_form):
    '''Data of system is given as a input. Variable of interest is placed on index.
    If data is saved as list of lists then data_form=True, else False. A list containing
    all data of variable is given as output. '''

    data_list=[]
    data_output=[data[i][index] for i in range(len(data))]

    for i in range(len(data_output)):
        d=data_output[i]

        if data_form==False: #If data not list of lists
            for j in range(len(d)):
                data_list.append(d[j])

        if data_form==True: #If data list of lists
            for k in range(len(d)):
                for l in d[k]:
                    data_list.append(l)

    return data_list
```

Figure C.13: *Code: Data for Histogram*

Logarithmic Binning

```
#logarithmic binning
def log_bin(max_life, a=2):
    '''the value of the maximum lifetime and basis a=2 is given as inputs.
    A list containing logarithmic binning is returned'''
    n=0
    while 2**n<max_life:
        n=n+1

    logbin=[]
    for i in range(n):
        logbin.append(a**i)

    return logbin
```

Figure C.14: *Code: Logarithmic Binning*

Time Correlation

```
#Time Correlation
def correlation(S):
    '''The time correlation function takes a array with sizes as inpunt,
    and returnes a list containg the value of time correlation for each time lag '''
    C=[]
    d1=np.arange(0, 600, 1)
    dt2=np.arrange(600, len(S), 200)
    DT=np.concatenate((dt1, dt2)) #time lag

    for dt in DT:
        num=0 #numerator
        den=0 #denominator

        for t in range(len(S[:,0])):
            for x in range(len(S[0,:])):
                if t + dt <= len(S[:,0]) -1:
                    num = num + S[t,x]*S[t+dt, x]
                    den = den + (S[t,x])**2
        C.append(num/den)

    return C
```

Figure C.15: *Code: Time Correlation*

Exponential Fitting

```
from iminuit import Minuit
```

```
t=dt
y=data
ey=np.mean(y) #The curve is exponential
```

```
#Fitting function:
def fit(t, a, b, alpha):
    return a+b*np.exp(-t/alpha)
```

```
def chi2_owncalc(a, b, alpha):
    y_fit = fit(t, a, b, alpha)
    chi2 = np.sum(((y - y_fit) / ey)**2)
    return chi2
```

```
minuit = Minuit(chi2_owncalc, pedantic=False, a=, b=, alpha=)

# Perform the actual fit:
minuit.migrad();
minuit_output = [minuit.get_fmin(), minuit.get_param_states()]

# Here we extract the fitting parameters and their errors
a_fit = minuit.values['a']
b_fit = minuit.values['b']
alpha_fit = minuit.values['alpha']

sigma_a_fit = minuit.errors['a']
sigma_b_fit = minuit.errors['b']
sigma_alpha_fit = minuit.errors['alpha']

Nvar = 3                    # Number of variables (a, b, alpha)
Ndof_fit = len(y) - Nvar    # Number of degrees of freedom = Number of data points - Number of variables

# In Minuit, you can just ask the fit function for the Chi2:
Chi2_fit = minuit.fval                          # The chi2 value
Prob_fit = stats.chi2.sf(Chi2_fit, Ndof_fit)
```

Figure C.16: ***Code: Exponential Fit***

Appendix D

The Sneppen-Bornholdt Model (2018)

Globalization in a nutshell

Kim Sneppen[1] and Stefan Bornholdt[2]
[1]*Niels Bohr Institute, Blegdamsvej 17, DK-2100 Copenhagen, Denmark*
[2]*Institute for Theoretical Physics, University of Bremen, 28359 Bremen, Germany*

(Received 22 January 2018; published 29 October 2018)

The geographical distribution of production is getting an increased amount of attention in economics. Distributed trade opens for production where it is cheapest, which in turn is reinforced by *economies of scale*. Using a simple agent-based model for the geographical interplay between transportation costs, economies of scale, as well as information costs, we address here the transition between local and distributed economies. The model naturally recapitulates that decreased transportation and information costs favor large companies. This suggests history dependence in the sense that new companies typically reemerge in the vicinity of old ones. Further, it suggests that company stability depends on transport costs, and that the transition from a local economy to a global economy is naturally driven by reduced transportation costs and an increased information horizon.

DOI: 10.1103/PhysRevE.98.042314

I. INTRODUCTION

Globalization is the prominent characteristic of today's manufacturing and trade on an intercontinental scale. This has not always been the case. Global trade emerged on a large scale as transportation costs decreased over the course of more than a century, fueled by technology and innovations such as the construction of railroads [1] and the invention of containers [2]. This brought about competition between different geographical regions, with further optimization of manufacturing of simple products for global markets [3], allowing manufacturers to take advantage of *economies of scale* up to the global market. On a regional scale, however, this philosophy was not fully in effect, as large margins persisted in retail, boosting consumer prices much higher than the theoretical minimum, at least from the perspective of the manufacturer.

This has changed over the past two decades because the invention, and broad adoption, of the Internet made prices transparent to consumers beyond their regional scope and on a truly global scale. Local retailers often lost the resulting competition with other retailers located far away, and a few particularly competitive retailers suddenly found themselves turning into quickly growing Internet retailers. This "winner-take-all" type of dynamics, as a consequence of the Internet-driven removal of information barriers for consumers, replaced a large fraction of local or regional retail enterprises by national or international suppliers, operating at smaller margins and thus potentially selling higher numbers of a given product [4]. Thereby, the Internet extends the efficiency of *economies of scale* down to the smallest scale of a single consumer in the third (or Internet) phase of globalization. This adds a new facet of the effects of the Internet to the classical discussions of the pros and cons of agglomeration [5] and globalization [6–8], where a quantitative understanding has only recently begun to emerge [9].

Here we contribute a simple agent-based model to illustrate the interplay between transportation costs and information barriers in a model of manufacturing and trade. As even low-dimensional economic models with a small number of degrees of freedom turn out to be highly complicated when considered along with agents operating in space [10], we decided to adopt the econophysics approach of modeling markets with many agents, chosen to be identical and interacting according to standardized rules [11]. We explicitly do not average over agents, as is common in economic toy models, because we aim to characterize collective effects rather than average properties. We keep a large number of degrees of freedom by considering a large number of identical agents while still focusing on the reductionist approach of limiting the number of model parameters to the utmost minimum. This allows us to go beyond the low dimensionality of standard economic toy models and to study phenomena that those models cannot represent.

Classical globalization affects nations and trading patterns between nations [7,12] as well as local societies. This calls for mathematical modeling at least of its large-scale dynamics, yet such models are scarce. Krugman [7] introduced a two-region model of trading between the developing and the developed model, pinpointing the possibility for a *winner-takes-all* outcome for one of the regions.

Instead of two regions, here we take the opposite approach and study an agent-based model with a large number of agents on a one- or two-dimensional lattice. Each agent attempts to buy a product where it is offered at the lowest price, and producers are assigned an *economies of scale* advantage based on their previous popularity. The spatially extended model allows us to discuss a number of issues associated with the interplay between distance, information horizon, transportation costs, and the startup threshold for a new company. The model is defined as follows.

II. MODEL

Define a lattice with L agents at position $x = 1, 2, 3, \ldots, L$. In our model, agents for simplicity act both as buyers and sellers. At each update, one random agent at position x rewards (makes one unit payment to) the agent y in the system that can supply him with the good at lowest cost. This selected reward specifies a position y that as a consequence accumulates a gain. The cost comes from production, transport, and possibly associated tariffs, which all are simply added together. Thus, for an agent at position x, the cost of a product from position y is set to

$$c = s(y)^\gamma + |x - y|\sigma. \quad (1)$$

A reward of one unit is subsequently assigned to that position y, which provides the lowest cost for x:

$$s(y) \rightarrow s(y) + 1 \quad \text{for } y \text{ with minimal } c. \quad (2)$$

In Eq. (1), s is the company size at position y, and s^γ with $\gamma \leqslant 1$ is the production cost per unit for that company. A $\gamma \sim 0$ means that production cost does not diminish much with company size, whereas a lower γ would reflect an increasing effect for an *economies of scale*. $\gamma < -1$ is not realistic, as it would mean that the total production cost of many product units is cheaper than for just one unit. But $\gamma \sim -1$ may be realistic for software or movie industries, whereas traditional factory production may have negative γ closer to 0.

σ quantifies a transportation cost that, for simplicity, is assumed to be linear in distance between a consumer at x and a producer at y. Notice that this proportional dependence is markedly different from the exponential "iceberg cost" assumed in the economic literature [13]. In fact, one may even expect modern shipping costs to increase more slowly than proportional to distance, however we keep the linear dependence for simplicity here. In any way, it will act as a more or less rigid barrier for long-distance trade. In an alternative interpretation of our model, what we define as transportation cost here may alternatively be seen as an increased cost of information search when—before the Internet—exploring a larger neighborhood was necessary for finding the best price. Last but not least, a tariff parameter β quantifying an eventual customs barrier between position x and position y may be added to Eq. (1).

The model is executed in time steps, where each time unit consists of τ trading updates as defined above. After these τ updates, new company sizes $s(y)$ are assigned to be equal to one plus the accumulated orders at site y during these updates. There is no direct memory of the earlier size of the company, but they tend to remain localized because of the sensitivity of orders to production capacity at the previous production period. Alternatively, the model may be formulated by updating all $s(y)$, $y = 1, \dots, L$ at each update by $s(y) \rightarrow s(y) = 1 + e^{-1/\tau} s(y)$ if y is chosen as a producer, and $s(y) \rightarrow s(y) = e^{-1/\tau} s(y)$ if y is not chosen as a producer, respectively.

The model has three parameters: γ, σ, and τ. In addition, tariffs may be added for externally imposed customs, and the model may also be extended to include prefactors in front of s^γ in order to take into account the variation in labor costs that is considered in model descendants of [7]. The system size L is irrelevant, as long as it is much larger than the domain scale set by the other parameters. γ and σ quantify incremental production cost and transportation cost for a unit of product, whereas τ is proportional to the time it takes to rebuild the production apparatus for the considered product type. The model can be directly generalized to the more realistic two-dimensional case. Below, we present simulations in one and two dimensions, and we also explore variations where one only allows consumers to explore producers within some maximum information horizon h.

III. RESULTS

Figure 1 explores the dynamics of the one-dimensional model with periodic boundary conditions using an intermediate level of *economies of scale* exponent $\gamma = -0.5$. The first three panels illustrate the emergence of production centers (denoted "companies" in the following), reflecting the positive feedback between consumers and the *economies of scale*. Panel (a) illustrates that a given manufacturer may collapse while others emerge. Notice further that the emergence of new companies often occurs close to the positions of the previously collapsed ones. This partly pinned inheritance reflects the memory associated with the geography of surrounding companies that survive the collapse of a particular company. In other words, when a company disappears, it leaves open a wide business niche because of the cost associated with the distance for local customers to deal with companies farther away in the larger neighborhood.

Comparing panels (a) and (b) in Fig. 1, one notices that a low τ destabilizes companies. A low τ corresponds to the case where it only costs a few product units to build a new production facility. Therefore, a higher startup cost will tend to stabilize existing production centers.

Comparing Figs. 1(c) with 1(a), one can see that lower distribution costs may stabilize even a sector with low τ. Figure 1(d) introduces an information horizon, where agents are only allowed to explore prices of the nearest $h = 10$ neighbor agents in search for the lowest price (modeling an offline economy). This panel uses the same other parameters as panel (c) and illustrates that a low information horizon has an effect comparable to higher transportation cost [compare with panel (a)].

Figure 2 displays a sequence of snapshots of a two-dimensional model with colors that mark the agent's association with different production centers. Each production center is marked with a black disk with its radius proportional to its size (number of products during the last τ transactions). Note that companies disappear, their customers being taken over either by a neighboring agent, then turning into a production center, or merging into the readjusted range of neighboring producers.

Figure 3 explores the density of companies, defined by the number of production centers with size $s > 1$ per unit length (or area in two dimensions). For relatively small transportation cost, the productions centers become big while the noise in allocating customers in the time interval τ becomes small. In that limit, an "equilibrium" production center should supply customers up to a distance x where the gain by *economies of scale* $s^\gamma \propto x^{D\gamma}$ balances the transport cost:

$$x^{D\gamma} \sim \sigma \cdot x \Rightarrow x \sim (1/\sigma)^{1/(1-\gamma D)}. \quad (3)$$

Here D is the spatial dimension. Figure 3 shows that this scaling fits the obtained behavior for small production costs σ.

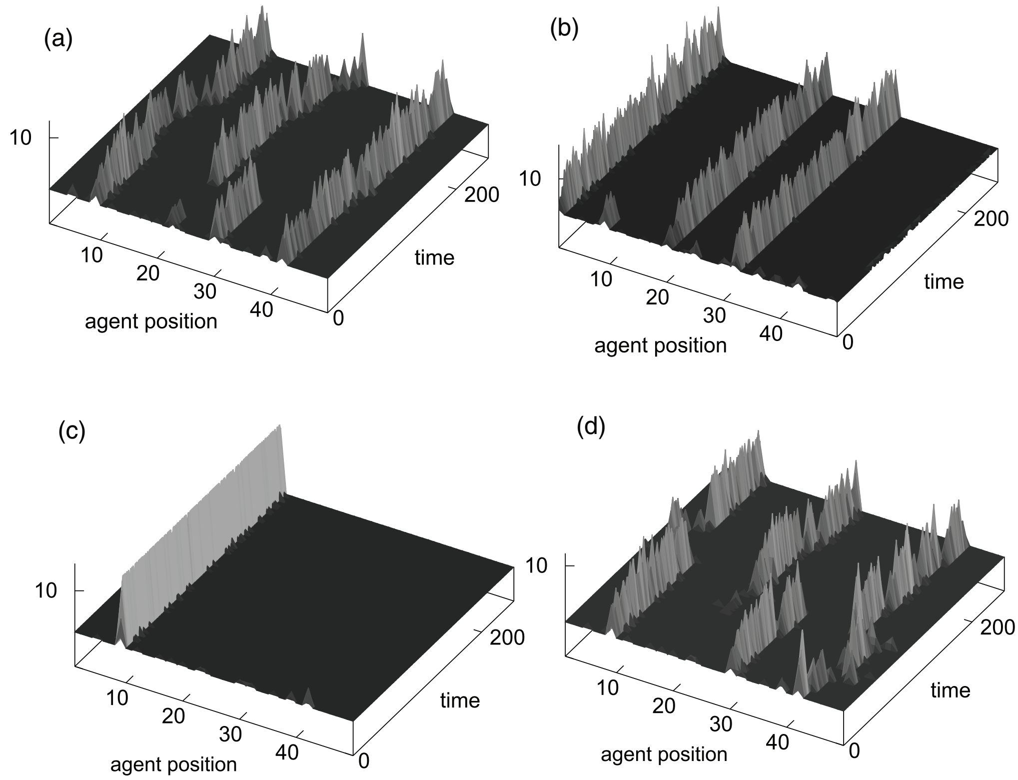


FIG. 1. Company location and size x as a function of time for $L = 50$ agents organized on a line and a production cost per unit that scales as $c \propto s^{-0.5}$. (a) High noise ($\tau = 15$ transactions per unit time), high transport cost $\sigma = 0.05$, and information horizon $h = \infty$. (b) Low noise $\tau = 25$, $\sigma = 0.05$, and $h = \infty$. (c) $\tau = 15$, low transport cost ($\sigma = 0.01$), and $h = \infty$. (d) As in (c) but with a finite information horizon $h = 10$.

Noticeably, the outcome of the model is relatively insensitive to increased τ [blue dots in panel (a)].

We further explored an economy with zero transportation cost, but only with access to knowledge about prices within a limited information horizon h. Figure 4 shows that diversity then becomes independent of γ, provided that there is an *economy of scale*, i.e., $\gamma < 0$. Importantly, this situation then allows for a simple discussion of the positive aspect of globalization through the Internet, namely that the minimum product price a consumer experiences decreases with a larger information horizon. This decrease in consumer price results from the long-term feedback between the customer agents and the sizes of the manufacturers.

IV. DISCUSSION

Starting with the seminal papers by Krugman [7,8], globalization models have focused on trade labor and the interplay between different products with different *economies of scale*, complemented by geography and associated transportation costs. Our approach is further simplified and ignores the important interplay between labor cost and where products are produced. We do this to concentrate on the central theme of transport cost versus production geography for a production with an *economies of scale* advantage of mass production being cost-effective. Our proposed model goes beyond other globalization models in introducing actual space, and not just two compartments. In particular, our model suggests that the

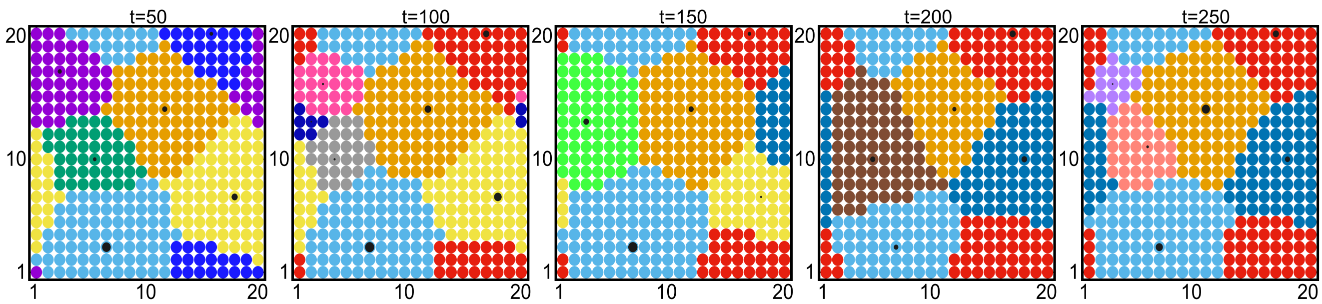


FIG. 2. Five subsequent snapshots of a simulation with 400 agents placed in a two-dimensional geometry with standard (*economies of scale*) exponent $\gamma = -0.5$.

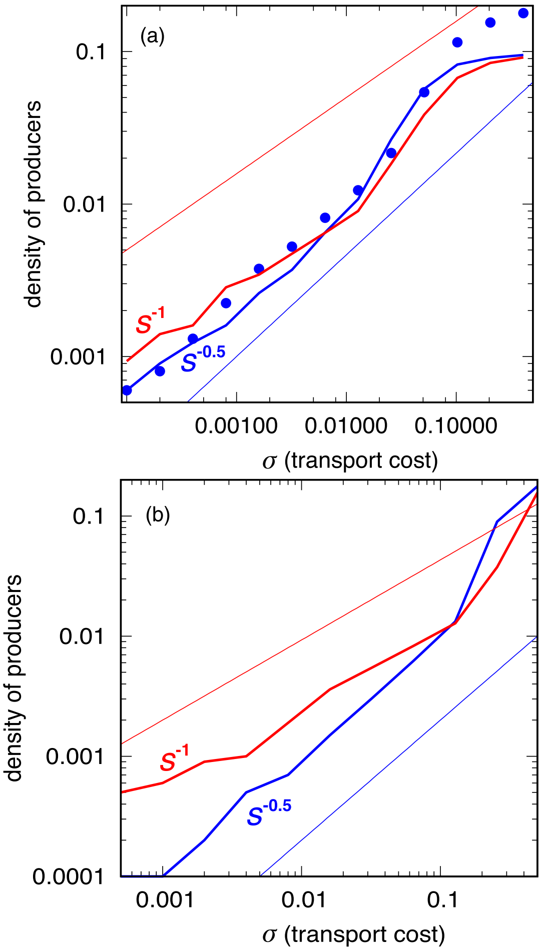


FIG. 3. Density of production centers as a function of production cost. (a) One-dimensional system with $L = 10\,000$ agents and (b) two-dimensional system of 100×100 agents. Solid curves are done with $\tau = 0.1$ times the number of agents in the system, meaning that each agent has 10% chances to make a purchase between each reassignment of production centers. Simulations are done for $\gamma = -0.5$ (the solid curve starting lower on the left side of the plot) and $\gamma = -1$ (the solid curve starting higher on the left side of the plot). The thin straight lines show the expected scaling for the $\tau \to \infty$ case: density $\propto \sigma^{D/(1-\gamma D)}$. The dots in panel (a) refer to the case of $\gamma = -0.5$ and $\tau = 0.2 \times L$.

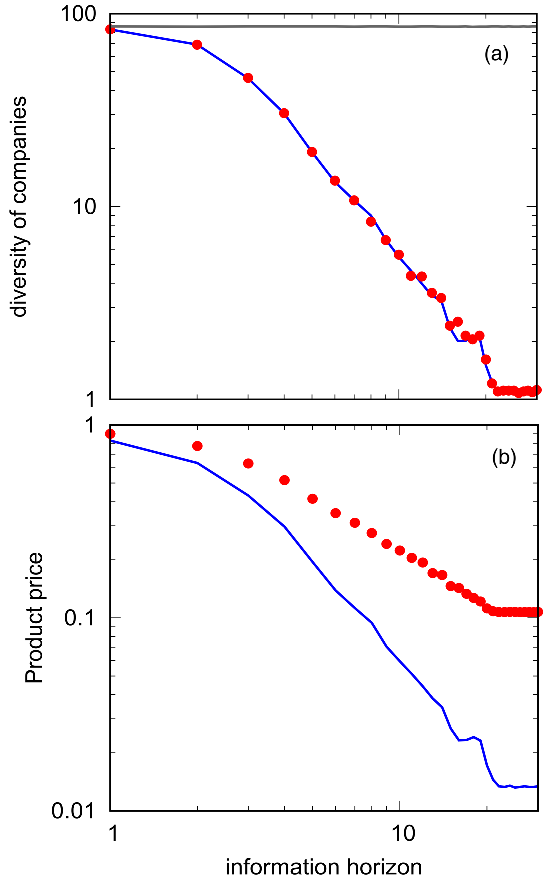


FIG. 4. Two-dimensional simulation with zero transportation cost while varying the information horizon instead (for a 30×30 system with $\tau = 90$, i.e., the same relative size of τ as in Fig. 3). The different curves refer to different economies of scale sectors, with dots marking $\gamma = -0.5$ and the lines corresponding to $\gamma = -1$. (a) Diversity decreases proportional to $1/h^2$ as the system self-organizes into regions of areas $\propto h^2$ that each are served by one company. (b) Consumer prices decrease as the information horizon increases.

emergence of production centers depends primarily on the *economies of scale* exponent γ and the transport cost σ.

Overall, the size of a production center or its associated customer base is governed by the balance between the positive feedback of an *economy of scale* and the negative feedback set by transport. In this analogy, the patterns in Figs. 1 and 2 are reminiscent of those found in reaction diffusion systems [14] with local positive feedback and spatially extended inhibition.

Consumer price versus local or global production also have political ramifications. From a national scale to a more global economy, one recurrently considers protective tariff barriers. Figure 5(a) illustrates how a gradual decrease in transport cost favors fewer and still larger companies. Figure 5(b) subsequently illustrates how this centralization is stopped by customs trade barriers (marked with red lines). Notice that not all regions are of equal size, and as a result the producer in the center of the large region [1, 20] becomes particularly large, and thereby provides the consumers in this region with lower prices than in the smaller regions. In our model, these consumer benefits did not result from more efficient competition, but rather from their larger country favoring a more efficient production of scale.

Interestingly, if noise is larger (smaller τ, the time-scale separation between fast trade and slow construction of production capacities, thus considering a fast and therefore "noisy" change in production infrastructure), then the relative gain of the larger region is increased because each of the smaller regions will have several very small companies with only small economies of scale benefits. Figure 5(c) thereby tells us about an added industrial advantage of a larger economy compared to a smaller one. If one imagines a sudden lift of trade barriers, the odds are that the companies in the larger region would prevail against the smaller ones from the smaller

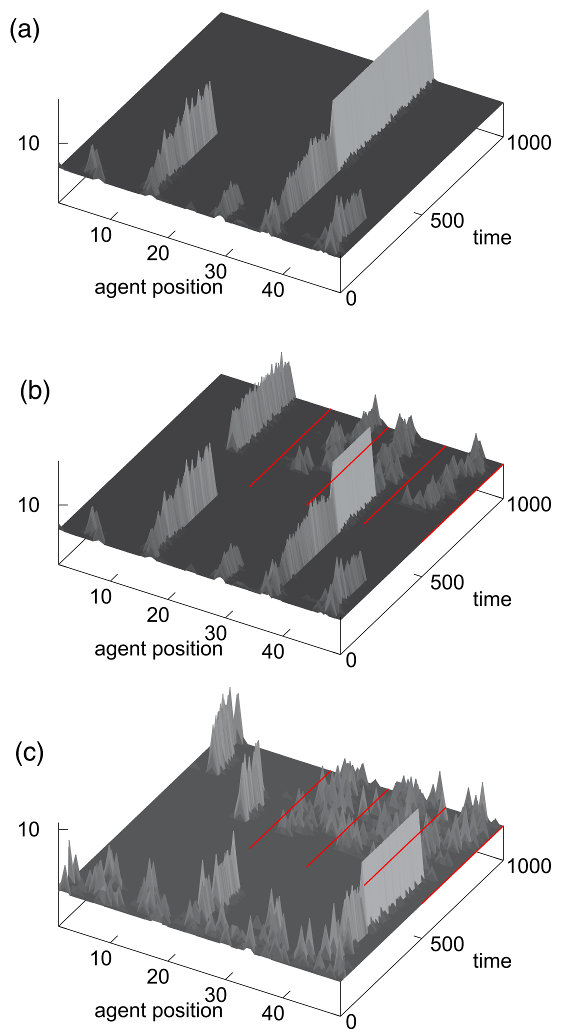


FIG. 5. (a) Company size and location as transportation is gradually lowered as $0.2\exp(-0.005t)$ where t counts time ($N = 50$ agents and $\tau = 25$). One observes a coarsening of the economy until a single monopoly takes over. (b) At time $t = 500$, increasingly hard customs barriers are introduced in the form of an added tariff cost of $0.005(t - 500)$ whenever a product has to traverse a position marked with a red (gray) line. We only assign costs of the barrier crossed last (i.e., they are not additive). (c) As above, but with $\tau = 10$, corresponding to increased noise of the updates.

region. History benefits the current winner, provided that the currently dominating region can enforce the removal of barriers.

Thus our model represents additional mechanisms of globalization over classical models, whose major results are reproduced in our simulations as well: In [7,8] the focus was on decreasing transportation costs ($\sigma \to 0$) and not on increasing an economy of scale (γ decreasing from 0 to -1). Reference [8] predicted that when transport cost falls below a critical threshold, then industry is concentrated in one region. This is recapitulated here by using $\gamma < 0$.

A major simplification of our model is that the consuming agents only assign resources to the producers, whereas the consumers do not lose anything by their consumption. Thereby, the model ignores feedback associated with generating richer consumers, as well as higher wages in the neighborhood of big production facilities. However, our aim was to define the possibly simplest agent-based model that demonstrates aspects of the underlying consequences of the "winner takes all" dynamics of the Internet economy that challenges local economic and political structures across the planet.

A next step of the model could be to include the effects of wages, but also to introduce several types of products, each developing production centers in analogy with the above market rules. The interaction between the products should be facilitated by selecting a customer at position x of product type i to be proportional to his production ability of all other products, $\sum_{j\neq i} s_j(x)$. Such an extended model would recapitulate the overall finding of the simpler model, with positive reinforcement within strong economies where the production of some products favors other production facilities in their neighborhood.

In conclusion, while the prevailing belief is that specialization and associated *economies of scale* are a benefit for all, our model highlights their tendency to generate spatial structures with long distances between eventual workplaces. Locally, this means that globalization comes at a cost for villages and towns, with implicit gains for larger population centers. Multiple production centers and *economies of scale* appear mutually exclusive, much as competitive exclusion in well-mixed ecosystems only allows for the coexistence of relatively few biological species [15].

ACKNOWLEDGMENTS

This research has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007 2013)/ERC Grant Agreement No. 740704.

[1] D. Donaldson, Railroads of the Raj: Estimating the impact of transportation infrastructure, Am. Econ. Rev. **108**, 899 (2018).
[2] M. Levinson, *The Box: How the Shipping Container Made the World Smaller and the World Economy Bigger* (Princeton University Press, Princeton, NJ, 2016).
[3] T. Levitt, The globalization of markets, in *The Marketing Imagination* (The Free Press, New York, 1983), pp. 20–49.
[4] E. Brynjolfsson and M. D. Smith, Frictionless commerce? A comparison of Internet and conventional retailers, Manage. Sci. **46**, 563 (2000).
[5] M. Fujita, P. R. Krugman, and A. J. Venables, *The Spatial Economy: Cities, Regions, and International Trade* (MIT Press, Cambridge, MA, 2001).
[6] F. Cairncross, *The Death of Distance: How the Communications Revolution is Changing Our Lives* (Harvard Business School Press, Cambridge, MA, 1997).

[7] P. Krugman and A. J. Venables, Globalization and the inequality of nations, Quart. J. Econ. **110**, 857 (1995).
[8] P. Krugman, Increasing returns and economic geography, J. Polit. Econ. **XCIX**, 483 (1995).
[9] K. L. Kraemer, J. Gibbs, and J. Dedrick, Impacts of globalization on E-commerce use and firm performance: A cross-country investigation, Inf. Soc. **21**, 323 (2005).
[10] V. Albino, N. Carbonara, and I. Giannoccaro, Innovation in industrial districts: An agent-based simulation model, Int. J. Production Econ. **104**, 30 (2006).
[11] E. Samanidou, E. Zschischang, D. Stauffer, and T. Lux, Agent-based models of financial markets, Rep. Prog. Phys. **70**, 409 (2007).
[12] M. Karpiarz, P. Fronczakand, and A. Fronczak, International Trade Network: Fractal Properties and Globalization Puzzle, Phys. Rev. Lett. **113**, 248701 (2014).
[13] P. McCann, Transport costs and new economic geography, J. Econ. Geogr. **5**, 305 (2005).
[14] A. M. Turing, The chemical basis of morphogenesis, Philos. Trans. R. Soc. B **237**, 37 (1952).
[15] G. Hardin, The competitive exclusion principle, Science **131**, 1292 (1960).

Appendix E

The Krugman-Venables Model (1995)

THE
QUARTERLY JOURNAL
OF ECONOMICS

Vol. CX November 1995 Issue 4

GLOBALIZATION AND THE INEQUALITY OF NATIONS*

PAUL KRUGMAN AND ANTHONY J. VENABLES

A monopolistically competitive manufacturing sector produces goods used for final consumption and as intermediates. Intermediate usage creates cost and demand linkages between firms and a tendency for manufacturing agglomeration. How does globalization affect the location of manufacturing and gains from trade? At high transport costs all countries have some manufacturing, but when transport costs fall below a critical value, a core-periphery spontaneously forms, and nations that find themselves in the periphery suffer a decline in real income. At still lower transport costs there is convergence of real incomes, in which peripheral nations gain and core nations may lose.

In recent years there has been growing concern among many observers in the advanced nations over the impact of globalization on their ability to sustain high living standards. As growth has surged in developing countries such as China, these observers fear that Third World growth—led by an expansion of manufactures exports—will come at Western expense. The most extreme expression of this fear was Ross Perot's warning that the North American Free Trade Agreement would lead to a "great sucking sound" as American jobs moved to Mexico. Yet more respectable voices raise similar concerns. Indeed, the White Paper of the Commission of the European Communities [1993], in effect asserted that the rise of Third World manufacturing nations has already had serious adverse impacts. It claimed that the single most important reason for the secular upward trend in European unemployment rates was the rise of countries that "compete, even on our own markets, at

*We are grateful for the support of the Centre for Economic Performance (London School of Economics) financed by the U. K. Economic and Social Research Council.

The Quarterly Journal of Economics, November 1995

cost levels that we simply cannot match"—Eurospeak for low-wage competition from the Third World.

To anyone who remembers the debate over the New International Economic Order during the 1970s, all of this sounds rather surprising. Many of the participants in that debate asserted that globalization, rather than benefiting all nations, tended to produce gains for some at the expense of others, but the general view was that integration of world markets produced "uneven development," a rise in the living standards of rich nations at the expense of the poor, rather than the other way around. The claim that a global marketplace tends to widen inequality among nations was used to justify demands for aid and commodity price support schemes. More radical theorists argued that the South could develop only if it "delinked" its economies from the too well-established North.

What accounts for this reversal in the conventional wisdom? In large part, of course, it simply reflects events: the most dramatic feature of the development landscape circa 1974 was the failure of development efforts to narrow the North-South gap, while the most striking feature twenty years later is the contrast between the rapid growth of East Asian economies and the economic troubles of the advanced nations. It is possible, indeed, to dismiss both the old concern about uneven development and the new concern about immiserization of the North as intellectual fads rather than serious analytical propositions. As an empirical matter, one might well argue that divergent growth performance generally reflects internal factors, not the inevitable consequence of national roles in the international economic system.[1]

Nonetheless, in this paper we propose to take seriously concerns about effects of globalization on real national incomes. To do so, we will develop a model in which there are no inherent differences among national economies, yet in which an international division of labor can nonetheless spontaneously arise, and in which some nations may fare better under this division than others. That is, we offer a model in which the world economy may organize itself into a core-periphery pattern. In the context of this model, we can then ask how does increased globalization, a closer integration of world markets, affect the real incomes of core and periphery nations. Does globalization, as free-trade enthusiasts

1. For skeptical assessments of the supposed impact of developing country exports on advanced nations, see Krugman and Lawrence [1994] and Krugman [1994].

might assert, always benefit all nations? Does it hurt the periphery, as so many thought during the 1970s? Or does it hurt the core, as many now believe?

Our somewhat surprising answer is that both concerns about uneven development and worries about maintaining First World living standards in the face of Third World competition have some justification. In particular, they appear, to correspond to different stages in the process of globalization. Suppose that transportation and communication costs fall gradually over time. Then our model predicts an early stage of growing world inequality: when transport costs fall below a critical value, a core-periphery pattern spontaneously forms, and nations that find themselves in the periphery suffer a decline in real income. As transport costs continue to fall, however, there eventually comes a second stage of convergence in real incomes, in which the peripheral nations definitely gain and the core nations may well lose.

It turns out, then, that a relatively simple model predicts a U-shaped pattern of global economic change, of divergence followed by convergence. We are aware that any explanation of such large-scale and long-term economic trends in terms of a single cause must be offered with tongue firmly in cheek (our working title for this paper was "History of the World, Part I"). Moreover, it is highly likely that other factors, such as changing technology of production, have played a more important role than falling transportation costs have in driving changes in regional advantage. Nonetheless, we believe that the model is suggestive of some of the forces at work in the real world economy.

It is also interesting that the surprising conclusions of this paper arise from relatively small changes in the assumptions of fairly standard models in international trade, and that these changes are, generally speaking, in the direction of greater realism. Indeed, our model may be seen as a sort of hybrid of two well-known models of trade under monopolistic competition. One is the model of costly trade in differentiated final goods introduced in Krugman [1980]. To this we realistically add trade in intermediate goods, drawing ingredients from the model of costless trade in differentiated intermediate goods introduced by Ethier [1982]. Despite being constructed in this way from off-the-shelf components, our model exhibits behavior different from that of either antecedent: the interaction between transport costs and trade in intermediates creates country-specific external economies, which may lead to agglomeration of industrial activity. These externalities are similar to those that arise from the interaction between transport costs and labor mobility in recent models of economic geography (e.g., Krugman [1991]). However, our model differs from these in important ways. The mechanism creating the externalities is linkages between firms (through the input-output structure), rather than linkages between firms and worker/customers (as in Krugman ([1991]). Since we do not assume labor mobility, the model is applicable to international as well as to interregional economics. Immobility of labor also changes results in important ways. Simple geography models like Krugman [1991] respond in a monotone way to declining transport costs: when these costs fall below a critical level, industry concentrates in one region. Here, because labor is immobile (and thus wage differentials between regions emerge), continuing reductions in transportation costs eventually lead to a reindustrialization of the low-wage region. We believe that this is not just an artifact of the model: it represents a real distinction between interregional and international economics because labor is in fact far less mobile between than within nations.

The remainder of this paper is in six parts. Section I offers an informal exposition of the model's logic. Section II sets out the formal model, while Section III shows how equilibrium is determined, and how this equilibrium changes as the world economy becomes increasingly integrated. Section IV then shows how national welfare changes as globalization proceeds. Section V explores the effects of trade policy, and finally, Section VI offers some conclusions and suggestions for further research.

I. The Basic Story

We imagine a world consisting of two regions, North and South. Each region can produce two kinds of goods: "agricultural" goods that are produced with constant returns to scale, and "manufactured" goods that are subject to increasing returns. The manufacturing sector produces both final goods sold to consumers and intermediate goods used as inputs in production of other manufactures. All countries are equally proficient in both sectors: neither region has any inherent comparative advantage in manufacturing.

We suppose that initially transportation costs between the two regions are very high. Clearly, in this case each region will be essentially self-sufficient, and each region will produce both manufactured and agricultural goods.

Now imagine gradually reducing transportation costs. There will now be the possibility of trade between the regions. If (as we will assume to be the case) there are many differentiated manufactured products, some two-way trade in manufactures will arise. So long as transport costs are high enough, however, there will be no specialization at the aggregative level.

At some point, however, a circular process arises that leads to regional differentiation. Suppose that one region for some reason has a larger manufacturing sector than the other. This region offers a large market for intermediate goods, and thus makes the region, other things equal, a more attractive place to locate production of such goods. (This effect corresponds to the traditional development concept of "backward linkages.") But if one region produces a greater variety of intermediate goods than the other, better access to these goods will, again other things equal, mean lower costs of production of final goods (an effect corresponding to the concept of "forward linkage"), leading to a further shift of manufacturing to that region, and so on. When transportation costs fall below some critical point, then the world economy will spontaneously organize itself into an industrialized core and a deindustrialized periphery.

If the manufacturing sector is large enough, this differentiation of roles will be associated with a divergence in real wages as well. The self-reinforcing advantage created by backward and forward linkages will drive up demand for labor in the industrializing region, while the decline of industry in the other region will lead to falling labor demand. Thus, real wages will typically rise in the region that becomes the core and fall in that which becomes the periphery. Global economic integration leads to uneven development.

But now suppose that transportation costs continue to fall. As they do so, the importance of being close to markets and suppliers—and thus the importance of forward and backward linkages—will decline as well. Meanwhile, the peripheral region will offer potential producers the advantage of a lower wage rate. At some point the decline in transportation costs will be sufficient that the lower wage rate in the periphery more than offsets the disadvantage of being remote from markets and suppliers. At this point manufacturing will have an incentive to move out from the core to the periphery once again, forcing a convergence of wage rates.

This intuitive story suggests that a single cause—the long-term decline in transportation costs, leading to growing integration of world markets—can produce first a division of the world into rich and poor regions, and then a convergence in incomes and economic structure between those regions.

To study the insights of this intuitive story, however, we must turn next to building a formal model.

II. A Formal Model

We assume the existence of two economies, North and South, which are identical in endowments, preferences, and technology. We describe the Northern economy, simply noting that analogous conditions hold in South.

North is endowed with L units of labor, with wage rate w. It contains two sectors, agriculture and manufacturing. The representative consumer in each country receives only labor income, and has Cobb-Douglas preferences between agriculture and manufacturing. These preferences can be represented by an expenditure function $Q_A^{(1-\gamma)}Q_M^{\gamma}V$ in which V is utility, Q_A is the price of agriculture, Q_M is the price index for manufactures, and γ is the share of manufactures in consumer's expenditure. The budget constraint takes the form,

$$wL = Q_A^{(1-\gamma)}Q_M^{\gamma}V. \tag{1}$$

The manufacturing sector produces a number of varieties of differentiated products, which are aggregated by a CES subutility function into a composite good. The price index of this manufacturing composite is Q_M, and takes the form,

$$Q_M = [np^{1-\sigma} + n^*(p^*t)^{1-\sigma}]^{1/(1-\sigma)}, \tag{2}$$

where n is the number of varieties produced in North. In equilibrium these are all sold at the same price p. Similarly, n^* is the number produced in South and sold at price p^*. Southern products sold in North incur iceberg transport costs at a rate t; i.e., a proportion $1/t$ of the good arrives implying a consumer price p^*t. $\sigma > 1$ is the elasticity of demand for a single variety.

Turning to the supply side, we assume that agriculture is perfectly competitive, and uses only labor with constant returns to scale. We let agriculture be the numeraire ($Q_A = 1$), and assume that it can be costlessly traded. Choosing units such that one unit of labor produces one unit of output gives the equilibrium condition:

(3) $$w \geq 1.$$

The wage rate equals one if the economy produces agriculture, and exceeds it only if agricultural production is zero.

Firms in manufacturing use labor and a composite manufacturing intermediate good to produce output. We make the major simplifying assumption that the composite intermediate good is the same as the composite consumption good. Thus, the price index of the intermediate is Q_M, as defined in (2) above. Labor and the intermediate are combined with a Cobb-Douglas technology with intermediate share μ. Each firm produces output for domestic sale (y) and export (x), with production using α units of the input as a fixed cost and β per unit output thereafter. Each firm's total cost function is therefore

(4) $$TC = w^{1-\mu}Q_M^{\mu}[\alpha + \beta(y + x)].$$

Given this description of preferences and technology, we can now characterize equilibrium as follows. First, define the total value of expenditure on manufactured goods in the Northern economy as E. Then we have

(5) $$E = \gamma wL + \mu(x + y)pn.$$

The first term on the right-hand side is consumers' expenditure on manufactures, and the second intermediate demand, where we have used the fact that proportion μ of costs (and since there are no profits, of revenue) is spent on intermediates.

Next, note that firms mark up price over marginal cost by a factor $\sigma/(\sigma - 1)$, so that prices are set according to the condition,

(6) $$p(1 - 1/\sigma) = w^{1-\mu}Q_M^{\mu}\beta.$$

Now note that Northern and Southern demand for a single variety take the form,

(7) $$y = p^{-\sigma}Q_M^{\sigma-1}E, \qquad x = p^{-\sigma}t^{1-\sigma}(Q_M^*)^{\sigma-1}E^*.$$

With free entry and exit of firms, there is a zero profit condition that, as usual in this type of model, establishes a unique size of firm,

(8) $$y + x = (\sigma - 1)\alpha/\beta.$$

We choose units of measurement such that the right-hand side of this equation is equal to unity, and use (7) in (8) to express the zero profit condition as

$$(9) \qquad 1 = p^{-\sigma}[Q_M^{\sigma-1}E + t^{1-\sigma}(Q_M^*)^{\sigma-1}E^*].$$

Equilibrium is now characterized by equations (2), (3), (5), (6), and (9) (and analogous equations for the other region) which can be used to find equilibrium values of variables Q_M, w, p, n, and E.

Before discussing the solution of the model, it is important to understand the way in which n, the number of firms in manufacturing, affects firms' profitability. It does this through three channels. The first is the standard one. An increase in n reduces the price index Q_M, (equation (2)), thus shifting the demand curve for each firm down (equation (7)) and reducing firms' profitability (equation (9)). The second and third channels operate only if μ is positive; i.e., manufacturing uses manufacturing as an input. The reduction in Q_M associated with an increase in n now reduces total and marginal costs ((4) and (6)), thus raising firms' profits. This is a *cost,* or *forward linkage* between firms. An increase in n also increases total expenditure on manufactured products, E (equation (5)), thus raising demand and profits of each firm (equations (7) and (9)). This is the *demand,* or *backward linkage* between firms. It is the presence of these linkages that generates the effects we describe in this paper.

III. Output and Employment

In order to see how the model works, we first see what determines the allocation of manufacturing between the two countries, and the allocation of labor in each country between activities. Analytical study of the equilibrium is algebraically complex, so our main tool for exposition of the properties of the model is numerical simulation. Analytical results are derived in the Appendix.

This is a general equilibrium model, and as in any general equilibrium model of trade each industry must in effect compete on two fronts. On one side, it must compete for markets with foreign firms in the same industry. On the other side, it must compete with the other domestic industry for inputs. It is possible to represent the determination of equilibrium in terms of at least two diagrams, each of which focuses attention on one of these competitive fronts.

One such diagram is illustrated in Figure I. On the axes of this figure are the number of manufacturing firms n and n^* in North

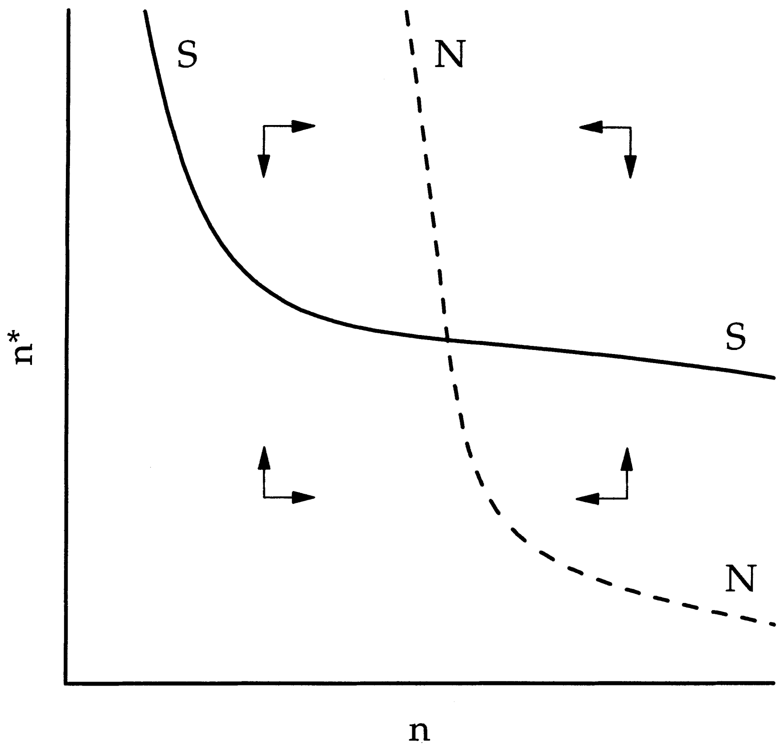


FIGURE I
Zero Profit Loci in North and South

and South, respectively. The schedules NN and SS indicate loci along which firms in North and South earn zero profits. On the assumption that firms enter if profits are positive, and exit if they are negative, the dynamics are indicated by the arrows. For the parameters used to draw this figure, there is a unique stable equilibrium that is symmetric, with each country having the same number of firms. (Values of parameters underlying the figures are given in the Appendix.)

An alternative representation of the same case, which emphasizes the competition for factors, is the variant of the "scissors" diagram of two-sector general equilibrium theory (especially the specific-factors model) shown in Figure II. In that figure the length of the horizontal axis is L, the total labor force. Northern employment in manufacturing, L_M, and in agriculture, L_A, are measured

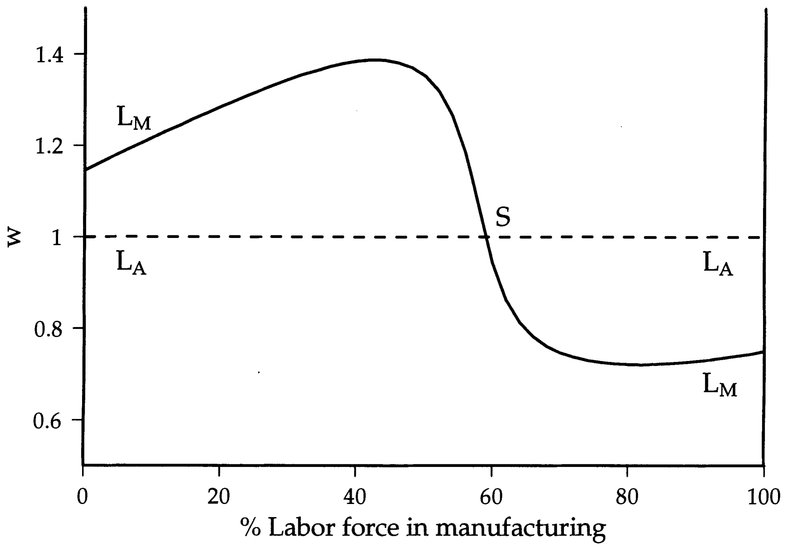


FIGURE II
Labor Demand: High Trade Costs

from the left- and right-hand ends, respectively. The vertical axis is the wage w.

The broken line L_AL_A is the demand function for agricultural labor, it represents equation (3), and our simple structure ensures that it is horizontal at height unity. The solid line L_ML_M is demand for labor in manufacturing. It gives the maximum wage that Northern firms can pay and break even as a function of Northern manufacturing employment, L_M, given that Southern manufacturing is in equilibrium with $w^* = 1$. (To put it another way, one may think of deriving this schedule by sliding down SS in Figure I, and calculating the maximum wage consistent with nonnegative profits in Northern manufacturing at each point.) The schedule is computed as follows. Northern employment in manufacturing is related to the value of output by the equation,

$$wL_M = (1 - \mu)np(y + x). \tag{10}$$

That is, a proportion $(1 - \mu)$ of firms' revenue is devoted to the wage bill. We assume that agriculture is active in the other country, so $w^* = 1$, and then use equations (2), (5), (6), and (9) (and

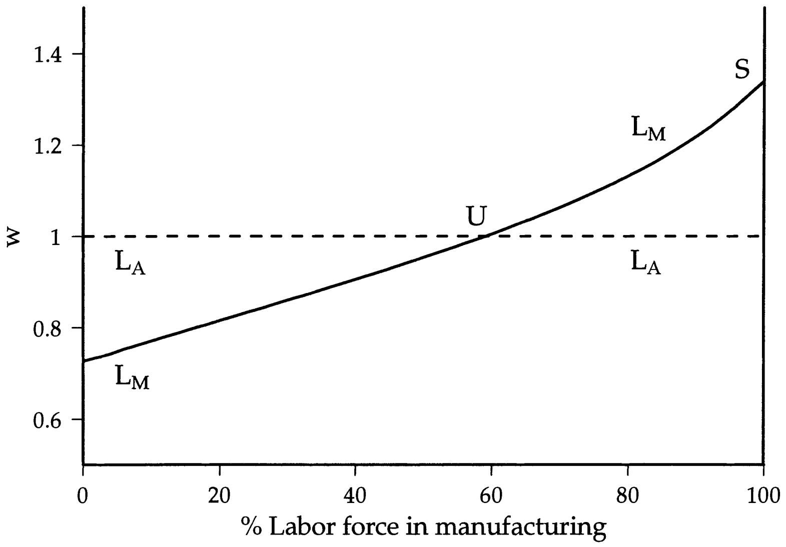


FIGURE III
Labor Demand: Low Trade Costs

their foreign analogs) to trace out manufacturing equilibrium as a function of w. Using this with equation (10) gives the illustrated relationship between w and L_M.

Equilibrium is at the intersection of the two curves, point S. In the case illustrated in Figures I and II, the equilibrium is symmetric with each economy having a wage equal to unity and producing both agricultural and manufacturing output. The proportions L_M/L and L_A/L are equal to γ and $1 - \gamma$, respectively. There is no net trade (although there is intraindustry trade in manufactures), so employment shares are determined by shares of the two sectors in final consumption.

Figures I and II are constructed with a high level of trade costs ($t = 3$). Figure III is the analogous diagram with a much lower level of trade costs ($t = 1.5$). The striking point is that the slope of the manufacturing labor demand curve is now positive, so that the equilibrium at U is unstable, and there is another equilibrium at point S. At this point North specializes in manufacturing and has a wage above the value marginal product of labor in agriculture. All agricultural output is produced in South, which may also produce manufactures.[2] The figure is produced with the assumption that South becomes the agricultural exporter, but country labels may of course be reversed. There are therefore three equilibria: U, S, and a further stable equilibrium with agriculture operating in North and manufacturing concentrated in South. In our discussion this case will be ignored.

The reason for the reversal of slope of the manufacturing labor demand schedule is the presence of linkages between manufacturing firms. Imagine relocating a firm from South to North. This raises demand for Northern firms' output, via the demand linkage, since at positive trade costs firms' demand for intermediates falls disproportionately on firms at the same location. It also reduces Northern firms' costs, via the cost linkage, as another variety of intermediate does not have to bear trade costs. Both these linkages create forces for agglomeration of manufacturing in a single location. At high trade costs (Figures I and II) these forces are dominated by the need to be near final consumer demand. At lower trade costs they are powerful enough to make the symmetric equilibrium unstable and cause manufacturing agglomeration.

If the share of manufactures in final consumption (γ) is less than or equal to 1/2, then all manufacturing agglomerates in a single country, and the equilibrium has $w = w^* = 1$. In this case world manufacturing demand is small enough to be met from a single location. But if $\gamma > 1/2$ (as illustrated), then the equilibrium must involve $w > w^*$. One country specializes in manufacturing, any further demand for manufacturing is met by the other, and the international wage differential offsets the locational disadvantage suffered by Southern firms distant from their markets and suppliers.

Figure IV illustrates the structure of equilibria at an intermediate level of trade costs ($t = 2$). Four equilibria are illustrated (a fifth in which South has no agriculture is not shown). At this intermediate level of trade barriers, linkages are not powerful enough to destabilize the symmetric equilibrium. But if North has all its labor employed in manufacturing, then linkages are sufficient to ensure that this is an equilibrium.

Figures I–IV suggest that as the level of trade costs is reduced, there are two points at which the qualitative character of the set of equilibria changes. At high levels of trade costs, the unique, stable

2. North is specialized in manufactures; but does South specialize in agriculture, or does it produce some manufactures as well? This is a somewhat difficult question to analyze; we discuss it further in Appendix 3.

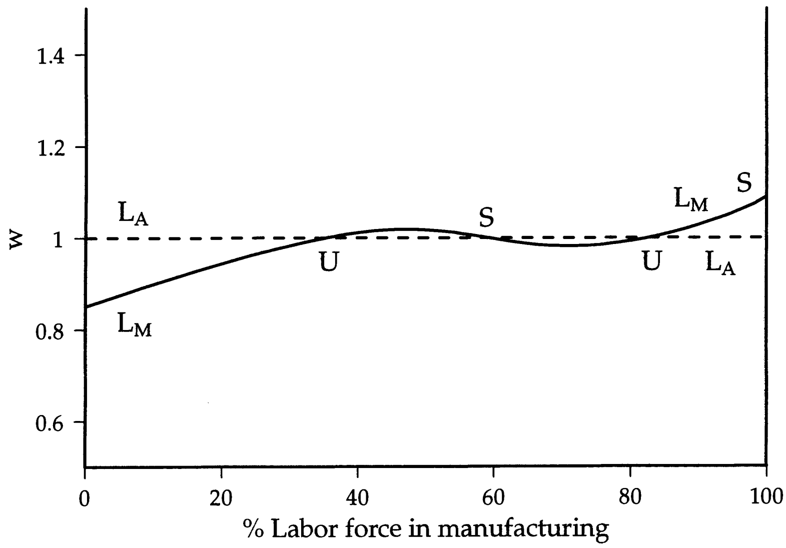


FIGURE IV
Labor Demand: Intermediate Trade Costs

equilibrium is one in which manufacturing is equally divided between the countries. At some point additional, asymmetric equilibria emerge. Finally, when transport costs fall to a critical level, the symmetric equilibrium becomes unstable. If we think of a historical sequence in which trade costs gradually fall over time, it is this latter level at which symmetry is broken and the core-periphery pattern emerges. In the Appendix we show that the critical level of t is defined by

$$t^{\sigma-1} = \left(\frac{1+\mu}{1-\mu}\right)\left(\frac{\sigma(1+\mu)-1}{\sigma(1-\mu)-1}\right). \tag{11}$$

What determines this critical level? From inspection of (11) we first notice that asymmetry only arises if there is a significant role of manufactured goods as intermediates. If μ were close to zero, there would be few forward and backward linkages. Indeed, we can see from (11) that $\mu = 0$ would imply a critical t of unity; i.e., any $t > 1$ (any positive trade costs) would imply symmetry between the economies.

At the opposite extreme, if $\sigma(1 - \mu) < 1$, the expression becomes negative. The interpretation of this is that a core-

TABLE I
CRITICAL VALUES OF t

σ / μ	3	5	7
0.3	2.21	1.42	1.25
0.5	4.58	1.90	1.50
0.7	/	3.04	1.96

periphery pattern will emerge no matter how high trade costs are. This will occur either if economies of scale are very large—which will be true in equilibrium if σ is small—or if the share of intermediates in costs, and thus the importance of backward and forward linkages, is very high. Both factors tend to make the de facto external economies in manufacturing larger.

For values of μ in the interval $[0,(\sigma - 1)/\sigma]$, it is certainly the case that there is a critical value of t at some number greater than unity. It is messy to derive the effects of changing parameters on the critical level of transport costs, but easy to calculate a table from equation (11). Table I gives the critical value of t for a range of values of σ and μ. The critical value is higher—and hence the region of multiple equilibria greater—the lower is σ, and the higher is μ. In other words, the greater are firms' price cost markups, and the greater is the share of intermediates in production, the more powerful are the forces for agglomeration. (At $\mu = 0.7$ and $\sigma = 3$, $\sigma(1 - \mu) < 1$, so the symmetric equilibrium is unstable at all levels of trade costs. Even under autarky, adding a further firm reduces price less than it reduces costs.)

IV. TRADE AND WELFARE

What are the implications of this structure of equilibria for real income and welfare? Figure V illustrates the dependence of real wages in each country on trade costs. The curves in the figure are wages divided by the consumer price index in each country (i.e., are utility, V, as given in equation (1)). The solid line (V) gives real wages in North, and the dashed line (V^*) real wages in South. Only stable equilibria are illustrated.

The figure illustrates the three stages of development analyzed in the preceding section. At high levels of trade costs North and South are symmetric, each operating agriculture and therefore

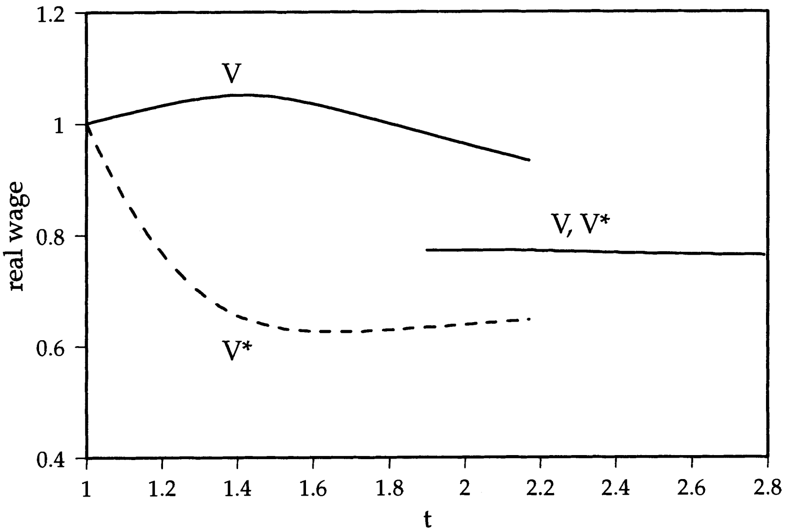


FIGURE V
Trade Costs and Real Wages

having the same wage relative to agriculture ($w = 1$) and relative to the consumer price index (V).

Reducing t creates an interval in which there are multiple equilibria (as in Figure IV). Beyond some point the symmetric equilibrium becomes unstable, and the world economy develops an asymmetric structure. At this point real wages rise in North while falling in South, that is, a process of uneven development occurs. Why does this divergence occur? In South the wage in terms of agriculture stays at unity, but real wages fall because a high proportion of manufactures now have to be imported, thus incurring transport costs. In North real wages rise for two reasons. Manufacturing labor demand causes an increase in the wage relative to agriculture (if $\gamma > 1/2$). And a smaller proportion of manufactures are imported and subject to trade costs, thus reducing the consumer price index and raising V further.

The third stage is one of factor price equalization. As trade costs become small enough, the wage differential that holds firms indifferent between locating in core and periphery narrows. Both the relocation of firms to South and the decline in the Northern wage in terms of agriculture reduce the Southern consumer price index, raising real wages. The movement of Northern real wages is

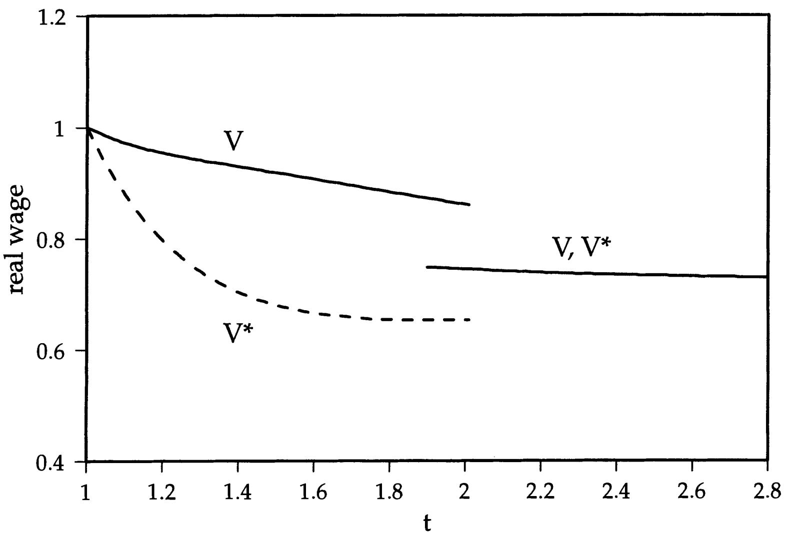


FIGURE VI
Real Wages: High γ

more ambiguous. Relocation of firms reduces wages in terms of agriculture. The consumer price index may, however, move in either direction. On the one side, an increasing proportion of manufacturing is being imported and is thus subject to trade costs. On the other side, these trade costs themselves are being reduced, directly tending to reduce the price index. Thus, real wages can move either way. In the case illustrated in Figure V, real wages fall as trade costs are brought down to very low levels.

Figure V was constructed for the same parameter values as Figures I–IV. Figures VI and VII indicate the effects of changing two parameters, the share of manufactures in demand and the share of intermediates in manufacturing. In Figure VI the share of manufactures in demand is increased. This increases the amount of manufacturing activity in South, and thereby reduces real wage differences. In this case the ‘‘globalization'' phase does not involve falling real wages in North.

Figure VII illustrates the case when the share of intermediates in manufacturing is raised. Agglomeration forces are now stronger, creating a wider real wage differential. Whereas in previous figures there is manufacturing activity in both North and South, there is

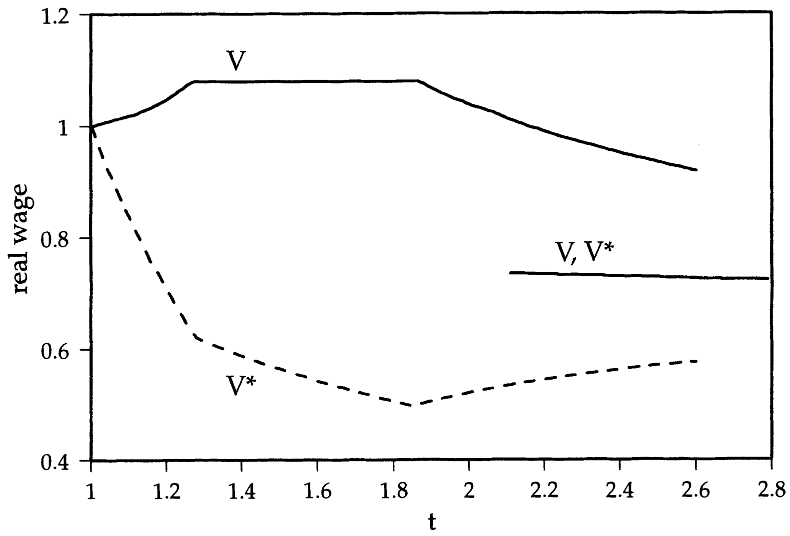


FIGURE VII
Real Wages: High μ

now a range of transport costs—the interval $t \in [1.28, 1.85]$—over which all manufacturing is concentrated in North. Northern wages are determined by the condition that the value of manufacturing output equals the value of expenditure. This generates a large North-South wage gap, but not large enough for Southern manufacturing to be profitable. Conditions under which this concentration of manufacturing occurs are given in Appendix 3. Notice that Northern real wages are constant in this range, as trade costs are assumed to affect only manufactures.

These results are, of course, based on numerical examples, that is, on particular parameter values. Nonetheless, the general picture—in particular, the sequence of phases with initial separation into core and periphery followed by a return to factor price equalization—is general given this model. As long as there are some linkages ($\mu > 0$) but these are not too strong ($\sigma(1 - \mu) > 1$), there is always a critical level of t below which the equal-wage equilibrium is unstable. And it is always the case that as $t \rightarrow 1$ the wage differential between countries must also disappear. Thus, while the details depend on parameters, the general picture of a U-shaped response of relative wages to transport costs does not.

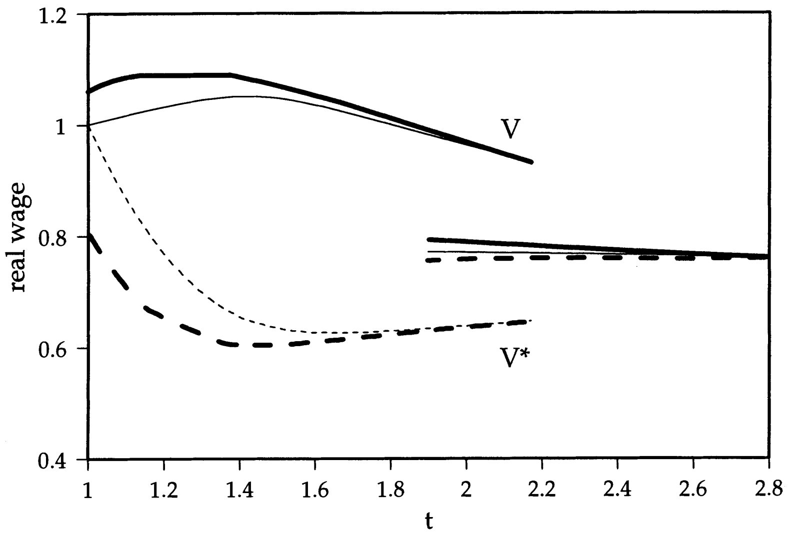


FIGURE VIII
Real Wages with a Northern Tariff

V. NOTES ON TRADE POLICY

The final phase of the process of globalization described by our model, in which the spread of industry to the South reduces relative and perhaps absolute Northern wages, obviously corresponds to the fears of many commentators on the world economy. Among these, some, such as billionaire-turned-pundit Sir James Goldsmith, whose recent book [1994] has been a European best-seller, advocate protectionist policies to prevent global competition from depressing wages.

A Northern tariff affects location of industry in two ways. First, by worsening Southern producers' access to the large Northern market, it tends to draw firms to the North. Against this, Northern firms now pay more for intermediate goods imported from South. The net effect is to attract firms to North from South, widening wage differentials.[3] This is illustrated in Figure VIII,

3. To see this, consider a small Northern tariff $d\tau$ at a point close to $t = 1$ (at which manufacturing is divided equally between North and South). The tariff raises Southern firms' cost on half their sales by $d\tau$. It raises Northern firms costs on all their sales by $\mu d\tau/2$ (since half of intermediates are imported). The former effect is larger than the latter, so Southern firms are hit harder by the tariff than Northern ones.

which compares Northern and Southern real wages as t declines toward unity under two different scenarios: free trade, and a Northern tariff of 33 percent on manufactures imports. Parameters and free trade real wages (the lighter lines) are as in Figure V. The heavy lines give real wages with the Northern tariff. Northern wages are higher than under free trade. Thus, the claims of some free traders that protectionism is necessarily a self-defeating policy are not borne out. Additionally, North receives tariff revenue, which is not included in the figure.

Two crucial cautions should, however, be made about these results. First, in supposing that North as a whole imposes a tariff against South, we have in effect gone beyond regarding North and South as *regions* and treated them as political units or at least customs unions. A general outbreak of protectionism, in which high-wage nations restricted imports from each other as well as low-wage nations, would clearly produce a very different outcome. By raising the prices of intermediates traded intra-North, Northern industry would suffer. To put it differently, the trade policy experiment described by Figure VIII is one in which trade policy has in effect been taken over by disciples of Goldsmith, who wants free trade in manufactures among high-wage nations while preventing imports from low-wage competitors.

Furthermore, it is important to point out that the model does not at all bear out the claims of some modern protectionists that a regime which allows trade only between countries with similar wage rates is somehow in the interests of labor everywhere. On the contrary, in the scenario described by Figure VIII Northern workers are protected from wage decline only by suppressing incipient Southern industrialization, and thereby also keeping Southern real wages low.

VI. Summary and Suggestions for Further Research

The conventional wisdom of economic analysis is that while greater global integration may hurt particular interest groups, it will normally raise the overall real income of just about every nation. There are exceptions to this rule even in the most conventional model: barriers to trade, natural as well as artificial, may sometimes act as de facto optimal tariffs, and their removal may therefore leave some countries worse off. Nonetheless, standard trade models do seem to suggest a presumption that integration is an all-around good thing.

Critics of this conventional wisdom have long argued that, on the contrary, greater integration usually produces national winners and losers. Traditionally, heterodox critics have argued that integration fosters inequality, that an integrated world economy divides into a rich core and a poor periphery, and that the wealth of the center comes at the periphery's expense. Only recently has the contrary argument, that globalization benefits the periphery at the core's expense, gained ground.

What we have shown in this paper is that a simple model in which regional differentiation is driven by the interaction between scale economies and transport costs makes sense of both old and new arguments. The world economy must achieve a certain critical level of integration before the forces that cause differentiation into core and periphery can take hold. When that differentiation occurs, the rise in core income is partly at peripheral expense. As integration proceeds further, however, the advantages of the core are eroded, and the resulting rise in peripheral income may be partly at the core's expense.

There are obviously many ways in which this analysis could be extended. We would, however, emphasize three directions in particular.

First, it would be desirable to get more geography into this model. As it stands, we postulate the existence of two exogenously defined regions, which then take on endogenously derived roles. In practice, the core has gradually spread into what was the periphery, with such areas as the southern United States, much of southern Europe, Japan, and now some of East Asia effectively making a transition from agricultural suppliers to manufacturing exporters. So we would like to extend the analysis to a multiregion setting, perhaps even one with continuous space.

Second, our model excludes capital mobility: indeed, it has no capital. Yet much of the political debate over integration focuses on the alleged impacts of capital movement rather than (or along with) trade flows. Thus, a natural step would be to add capital movements.

Finally, it is obviously important to discipline this analysis with some real numbers. We have offered a stark, one-factor explanation of vast global trends. This is in itself worrying. Worse yet, it is an explanation that will appeal to the prejudices of many people. Thus, it is crucial to do at least rough empirical work to see whether the kind of story described here is at all likely to be a large part of the explanation of real global economic trends. We suspect that as a practical matter growing integration with the South is at best a minor factor in the economic woes of the North, but the important point is that one should be careful about assuming that something that is possible in principle actually happens in practice.

In spite of these cautions, we regard it as a useful exercise to construct a minimalist model of the kind described here, and find it remarkable that so simple a structure can give rise to such a sweeping picture of divergence and convergence in the global economy.

APPENDIX 1: THE ALGEBRA OF SYMMETRY-BREAKING

Define $\tau \equiv t^{1-\sigma}$ and the ratios of Northern to Southern values of endogenous variables as follows:

$$\text{(A1)} \qquad \tilde{Q}_M \equiv \frac{Q_M}{Q_M^*}, \qquad \tilde{p} \equiv \frac{p}{p^*}, \qquad \tilde{E} \equiv \frac{E}{E^*}, \qquad \tilde{w} \equiv \frac{w}{w^*}.$$

Using equation (10), we can express the ratios of Northern and Southern expressions (2), (5), (6), and (9), as

$$\text{(A2)} \qquad \tilde{Q}_M^{1-\sigma} = \frac{L_M\tilde{w}\tilde{p}^{-\sigma} + \tau L_M^*}{\tau L_M\tilde{w}\tilde{p}^{-\sigma} + L_M^*}$$

$$\text{(A3)} \qquad \tilde{E} = \tilde{w}\left[\frac{\gamma(1-\mu)L + \mu L_M}{\gamma(1-\mu)L + \mu L_M^*}\right]$$

$$\text{(A4)} \qquad \tilde{p} = \tilde{w}^{1-\mu}\tilde{Q}_M^{\mu}$$

$$\text{(A5)} \qquad \tilde{p}^{\sigma} = \frac{\tilde{Q}_M^{\sigma-1}\tilde{E} + \tau}{\tau\tilde{Q}_M^{\sigma-1}\tilde{E} + 1}.$$

Eliminating $\tilde{Q}_M$ and $\tilde{E}$ gives

$$\text{(A6)} \qquad \tilde{p}^{(1-\sigma)/\mu}\tilde{w}^{(\mu-1)(1-\sigma)/\mu} = \frac{L_M\tilde{w}\tilde{p}^{-\sigma} + \tau L_M^*}{\tau L_M\tilde{w}\tilde{p}^{-\sigma} + L_M^*}$$

and

$$\text{(A7)} \qquad \tilde{p}^{(\sigma-1)/\mu}\tilde{w}^{(\mu-1)(\sigma-1)/\mu}\,\tilde{w}\left[\frac{\gamma(1-\mu)L + \mu L_M}{\gamma(1-\mu)L + \mu L_M^*}\right] = \frac{\tau - \tilde{p}^{\sigma}}{\tau\tilde{p}^{\sigma} - 1}.$$

These equations express $\tilde{w}$ and $\tilde{p}$ as a function of L_M and L_M^*. By inspection, if $L_M = L_M^*$, there is a solution to these equations with

$\tilde{p} = 1$, and $\tilde{w} = 1$. Consider a small change dL_M with associated change $- dL_M^*$ in the neighborhood of these values. Totally differentiating and applying Cremer's rule yield

$$\text{(A8)} \quad \frac{d\tilde{w}}{dL_M} = \left(\frac{\tau - 1}{\tau\gamma L}\right) \frac{(\mu - 1)[\sigma(\mu - 1) + 1] - \tau(\mu + 1)[\sigma(\mu + 1) - 1]}{2\sigma(\sigma - 1)(1 - \mu) + (\tau - 1)[\sigma(\mu + 1) - 1]}.$$

This derivative is the slope of the $L_M L_M$ curve at the symmetric equilibrium. Given that $\tau \in (0,1)$, a sufficient condition for the denominator of this expression to be positive is $\sigma(1 - \mu) > 1$. The numerator is positive or negative according to whether t is greater or less than the value implied by equation (11) of the text.

Appendix 2: Parameter Values

The simulations of Figures I–V set $L = L^*$, $\gamma = 0.6$, $\mu = 0.5$, $\sigma = 5$. In Figure I, $t = 3$, and in Figures II–IV, $t = 3$, $t = 1.5$, $t = 2$, respectively. In Figure VI γ is increased to 0.7. In Figure VII μ is increased to 0.55, and γ is returned to 0.6.

Appendix 3: The Pattern of Specialization

As long as the share of manufacturers in final demand exceeds one-half, any asymmetric equilibrium must involve specialization by one core country in manufactures. But does the other, periphery country also produce some manufactures? It is possible to shed some light on this question algebraically.

Suppose that all manufacturing is in North and therefore $w^* = 1$. We then have

$$\text{(A9)} \quad E + E^* = \frac{wL}{1 - \mu}, \quad E^* = \gamma L^*, \quad E = \gamma wL + \mu\left(\frac{wL}{1 - \mu}\right).$$

The first of these says that total expenditure equals the value of output (which is the wage bill divided by the labor share). The second and third give manufacturing expenditure in each location, with Northern expenditure including intermediate demand. Setting $L = L^* = 1$ and solving,

$$\text{(A10)} \quad w = \frac{\gamma}{1 - \gamma}, \qquad E^* = \gamma, \qquad E = \frac{\gamma}{1 - \gamma}\left(\gamma + \frac{\mu}{1 - \mu}\right).$$

A necessary condition for this to be an equilibrium is that it is not profitable for any firm to start producing in South. This can be

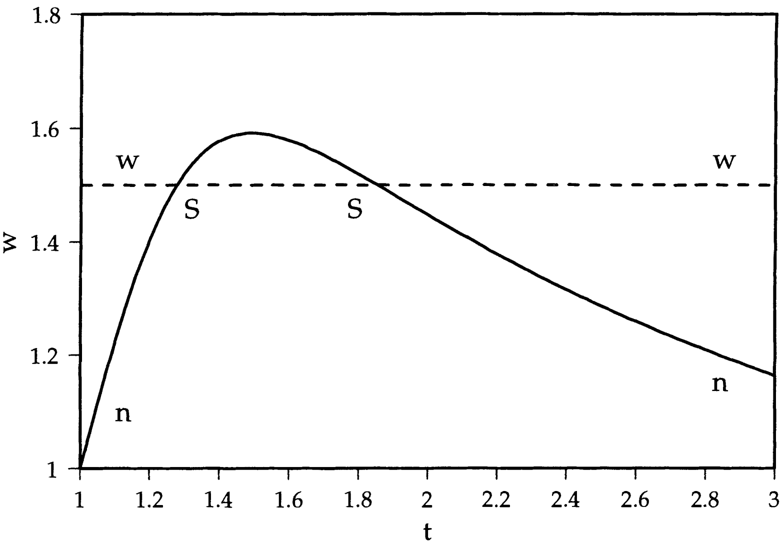


FIGURE IX
Southern Specialization

written as a condition imposing an upper bound on the Northern wage w. Using (9), together with (2) and (6) and noting that when $n^* = 0$, then $Q_M^* = tQ_M$, the condition is

$$w^{\sigma(\mu-1)} \leq t^{-\mu\sigma}\left[\frac{E^*t^{\sigma-1} + Et^{1-\sigma}}{E + E^*}\right]. \tag{A11}$$

This relationship is illustrated as curve nn in Figure IX (for the same values of parameters as used in Figure VII, and for constant values of E and E^* from (A10). Above the line nn the Northern wage is high enough that it would be profitable for a firm to establish in South, and below it this is unprofitable. This maximum wage gap peaks at intermediate values of t, and is larger the larger is Northern expenditure E, relative to Southern E^*. The curve nn is given by industrial location considerations. The actual wage is as given by (A10) and illustrated by the dashed line. Equilibrium is fully specialized in the interval SS, in which the equilibrium wage in North is low enough that no firm wants to set up in South.

STANFORD UNIVERSITY
LONDON SCHOOL OF ECONOMICS

References

Commission of the European Communities, "Growth, Competitiveness and Employment: The Challenges and Ways Forward to the 21st Century," White Paper, *Bulletin of the European Communities,* 6/93 (1993).

Ethier, Wilfred, "National and International Returns to Scale in the Modern Theory of International Trade," *American Economic Review,* LXXII (1982), 389–405.

Goldsmith, James, *The Trap* (London, UK: Macmillan, 1994).

Krugman, Paul, "Scale Economies, Product Differentiation and the Pattern of Trade," *American Economic Review,* LXX (1980), 950–59.

——, "Increasing Returns and Economic Geography," *Journal of Political Economy,* XCIX (1991), 483–99.

——, "Is the Third World a Threat?" *Harvard Business Review,* (July/August 1994), 113–21.

Krugman, Paul, and Robert Lawrence, "Trade, Jobs, and Wages," *Scientific American,* CCLXX (1994), 22–27.