Tuomas Höylä

Evaluation of academic funding system by simulation

Master's thesis, information technology

30. September 2014

University of Jyväskylä

Faculty of Information Technology

Author: Tuomas Höylä

Contact details: tuomas.a.hoyla@student.jyu.fi

Directors: Timo Tiihonen and Christoph Bartneck

Name of thesis: Evaluation of academic funding system by simulation

Title in Finnish: Akateemisen rahoitusjärjestelmän evaluointi simuloinnin avulla

Thesis: Masters's thesis

Alternative subject: Mobile systems

Number of pages: 95+1

Abstract: Thesis examines the usage of social simulation as a decision-making tool. The system under examination is the funding of university research and its effect to scientific output. From the system a model is created, which is implemented into a simulation to imitate the functionality of the original system. The results of the simulation are presented by varying the parameters of grant allocation, and finally it is evaluated how much conclusion can be drawn from the results.

Abstract in Finnish: Pro-gradu-työ tutkii sosiaalisen simulaation käyttöä päätöksenteon apuna. Tutkittavana systeeminä toimii yliopistontutkimuksen rahoitus ja sen vaikutus tieteelliseen tuottoon. Systeemistä luodaan malli, joka implementoidaan simulaatioksi jäljittelemään alkuperäisen systeemin toimintaa. Simulaation tulokset esitetään varioimalla rahoitukseen liittyviä parametrejä, ja lopuksi arvioidaan kuinka paljon johtopäätöksiä voidaan tehdä simulaation tuloksista.

Keywords: Simulation, Social simulation, Academia

Figures

Tables

Index

1 Introduction

In a world of increasingly scarce resources, researchers compete for funding, people, infrastructure, promotions and publications. Acceptance rates of funding proposals are low, and conferences proclaim a low acceptance rate as an indicator for the quality of the event.

To understand better the problem around funding proposals a few examples can be given:

- *"One-third (of grant makers) received fewer than 50 proposals; 38 percent of them funded at least half of the proposals.*
- *6 percent received more than 1,000 proposals each; 11 percent of them funded at least half of the proposals.*
- *Overall, 35 percent funded 50 percent or more of the grant requests they received.*
- *Corporate foundations receive a higher volume of proposals, compared to independent and community foundations.*
- *Foundations that reported giving of less than $1 million funded a larger share of their grant requests than foundations with giving of $10 million or more."*
  (Foundation Center 2005)

"2012, according to National Institute of health, the average rate was, depending on the type of funding, 19.3%, 16.8%, or 16.3% (NIH 2013), or *"acceptance rate within the financial year 2012-2013 was 34%"* (EPSRC 2013). Receiving tenure is not an easy task either; since the 70's to 2010, the proportion of tenured staff has decreased from 50% to 21% (Carson, Bartneck and Voges 2013). Similar behavior can be seen all over OECD-countries: the amount of positions and tenures are decreasing (Afonso 2013).

For example, approximately three quarters of funding proposals are turned down (Geard and Noble 2010). Currently in academia, there is a rising concern about the high time-usage for funding proposals. This does not concern the time of the applicant, but also the people and/or parties processing the application. On the other hand, the reason for funding process is to allocate the resources to researchers who are believed to give the best results with said resources and to prevent possibly unsuitable researchers from receiving those.

Conversely, researchers don't have enough resources to produce research, because the time spent writing grant applications. In 2012, Australian researchers spent on estimate around 550 years preparing 3727 proposals, which 21% were funded. This does not include time spent in administration or reviewing the proposals. This would mean roughly 66 million Australian dollars, whereas the total amount that was available was 459.2 dollars (Herbert, Coveney, et al. 2013), (NHMRC 2013). This lack of time reduces the time available for actual research while suffering from constant rejections, which affects the quality of research. Meanwhile, researchers are measured against high expectations for their promotions, which affect the researcher's productivity and motivation to apply for future funding. In addition, the review systems in place to judge the proposals or publications create an enormous overhead. Senior researchers often spend more time writing funding proposals and reviewing papers than actually doing research. This thesis tries to understand what influences competitiveness has on the productivity of researchers and the quality of their work.

To better understand this issue, as well as to possibly provide solutions, this thesis posits a model which analyzes the effect of these acceptance rates on productivity and alternatives how the grant system can allocate funding to research.

A heuristic model for simulation purposes of the system is derived, to measure scientific output. This model is implemented in form of a stochastic simulation, and the model is validated by comparing the results from the model data with data gathered from real world. The final step is varying different factors of the model in order to find optimal resource distribution strategy. The model discussed herein would provide a decision and evaluating tool for grant-allocation scheme.

Structure of the thesis

A typical simulation modeling cycle begins from the definition of the problem and system identification, where the boundaries of the systems are defined, leading to conceptual model. After identification the model needs to be designed, and data collected. Now the implementation of operational model can be done, which needs to be validated. After this follows the experimentations whose results are analyzed with statistical methods (Page and Kreutzer 2005), (Tiihonen 2014).The thesis follows a similar structure.

The next chapter identifies the key characteristics of the system on an abstract level, such as what kind of elements are there in the real-life model. In this chapter it is defined what information will be brought to conceptual model and what is excluded.

The needed concepts related to simulation are introduced in third chapter. The simulation types are introduced and how to avoid typical potholes in them. The mathematical tools of the topic are reviewed, how to design an experiment, and finally how to make analytical results out of the experiment.

The fourth and fifth chapters define a simulation model by means of connecting the entities and methods based upon their interaction and gathering data from related literature. These answers to questions such as which factors we want to measure (choosing the dependent and independent factors), which factors we want to vary (the dependent factors), how the information from outside sources can be adapted to created model, and how the model correlates with the previous data. The fourth chapter presents all the data of the system and compounds that are used in model. The fifth chapter ties compounds together with mechanisms.

The sixth chapter is dedicated for the simulation experiments and their results. In this chapter the concrete functions, variables, number and length of experiments, and scenarios will get their actual values. Before any experiments the system is calibrated. A handful of test runs are made for the system by means of choosing different dependent factors and how they reflect to literature shown in real-world model. In a way, the virtual model is observed from different angles to make it imitate more the actual observed model.

In chapter seven the experiments are done, with means of factorial design. In this chapter the factors of the simulated system will receive their exact values.

The last chapter, eight presents a response surface model for the dependent and independent factors we chose for experiment, presenting optimal levels for the dependent factors, results. After this follows the discussion part; are the results plausible enough and what could make them more accurate.

2 Characteristics of academia

To determine what kind of role grant allocation system actually plays in the production of research it is needed to examine what happens between funding of research and the produced research. By so it is possible to understand how resources allocated for research could provide higher quality and quantity of research. The risen concerns needs to be examined in order to determine the actual system and determine the key characteristics what could be implemented into the model. Elements which lie between these two ends are presented in this chapter. Starting from the end, quality of research can be measured in many ways, but always research needs to be made. Research is made by researchers often working in academia whose accomplishments are measured in various ways. This chapter explains, for the purpose of our modeling, what research is, who is doing it, and how it is done.

Research is *"Creative work undertaken on a systematic basis in order to increase the stock of knowledge, including knowledge of man, culture and society, and the use of this stock of knowledge to devise new applications"* (OECD 2002). Research is used to create new theorems, solve problems, or to prove and examine previous theories and hypotheses. In this thesis we don't make a difference between the types of research, but simply state research as an entity which needs inputs to provide outputs. Inputs can be considered as resources and the people using these resources, researchers. Research produces latent artifacts, which are commonly called as scientific output.

Measuring the success of academic research organization is typically done through amount and quality of research. It can also be done by examining other indicators, such as patents, monetary value, founded companies, and amount of graduates. History has a particularly rigorous way of revealing the value of different scientific theories and efforts. Good science leads to novel ideas and changes the way we interpret physical phenomena and the world around us. Good science influences the direction of science itself, and the development of new technologies and social policies. Poor science leads to dead ends, either because it fails to advance understanding in useful ways or because it contains important errors (Kreiman and Maunsell 2011). This thesis concentrates on one area, which is the scientific impact of academic organization.

Many measures of scientific output have been devised or discussed (Hirsch 2005). Because most scientific output takes the form of publication in peer-reviewed journals, these measures focus on articles and citations.

2.1 Publications

Scientific literature is normally divided in to four groups: Primary, secondary, tertiary and grey literature. Primary literature refers to a research which has been published in peer-reviewed scientific journals or conferences. The difference of secondary literature is that the authors can be summarizing the work of other authors. Tertiary literature has a different audience, meaning it is aimed at people who have an interest about the research, not being professionals of that field. Grey literature does not mean that the information is not valid or scientifically correct, but the means of distribution is limited (Schembri 2007) and is not necessarily published via journals. In general, primary and secondary literatures are the measured publications in academia. Scientific literature can include following kinds of publications:

- Scientific articles published in scientific journals
- Patents
- Books
- Edited Volumes
- Presentations at academic conferences
- Publications on the World Wide Web
- Technical reports, pamphlets and working papers
- Blogs and science forums

The research that has been done is explained in scientific publication, and its impact is measured by how well the particular research has been received among the society of science. Publications are evaluated how often they are cited among other publications, making it a symbiotic, self-feeding cycle. In this way science evaluates itself again, in theory. Reality reveals quite different examples. (Economist 2013) The publications are more likely to be cited if they are gone through peer-review process hence considered as a primary and secondary literature.

In academia peer-review is often used to determine academic paper's (or other type of publication) suitability for publication. The most typical way to get paper published in journals is to get it approved on peer-review process. This is not the only way, since there are other open boards where anyone can submit their publications (such as arXiv.org). This work focuses on primary and secondary papers which are published. This can be publications such as articles, monographs, summarizing work and reviews.

2.2 Researcher

So who is doing research? There is no law on research who is eligible to do it, and how it should be done. Around two-thirds of research and development in scientific and technical fields is carried out by industries, and 20% and 10% respectively by universities and government (OECD 2002). However, currently, among private companies, the academia (universities etc.) plays a key role in doing research and it is very common from industry to order research from universities since the knowhow is often there. Also the governments and private sector tend to play big role when funding universities (Koelman and Venniker 2011). Government-funded research can either be carried out by the government itself, or through grants to academic research units outside government. In this study we focus on academia.

As mentioned in the research problem section, researchers are expecting resources from the same scarce pool of resources. Researchers often face setbacks with funding and the results and the quality of their produced research. This can cause frustration, and decrease in motivation. It is also mentioned that academics often have to make the decision whether they decide to apply for funding or not. When applying, it can be a time consuming process. The modeled characteristics of researchers are factors such as position, time usage, motivation and other personal traits which are connected to quality of produced scientific output.

It is very common for researcher in academia that they apply for private funding. When reached a certain position, the researchers are not completely dependent on funding they have to apply, but they have a contract which is either fixed or permanent and have a base resource income. This is often dependent on position, having tenure, and country. The funding has often been separated in two, where the research institutes and research projects are funded separately (Koelman and Venniker 2011).

2.3 Research organization

Since in the thesis the focus is in academia, other producers of the scientific output are excluded from the study. The scientific output is produced by academics (or researchers) who work in academia. Since academia is the community of students and scholars engaged in education and research, there are many different positions which can be considered as working in academia but not necessarily among research, such as a student, a teacher, or general staff of the university.

Researchers work in different positions. The types of positions differ from country to country, but after graduating from university typical approach are doing a Ph.D. first, and after this doctor can approach towards teaching and research. Typically after this there will be a post-doctoral period. The final phase can be considered as professor rank. This professor rank is normally divided into multiple different positions. Academics can be on different positions with different contracts and funding, and the might or might not participate to producing scientific output.

2.4 Funding

The original examined system, funding process of a research community, is a complex process with multiple stakeholders. The resources are coming from multiple sources, and this depends on which university or research organization currently surveyed. Private corporations are an important source for academic research funding. The funding occurs when government or companies releases their funding calls, companies or such have a need for researchers, among others. The researchers apply for grants by writing applications, and grant providers evaluate applications and provide the resources for researchers they expect to carry out the job the best.

However, the reality of grant process is more profound. Usually applicant needs to compete with other similar researchers while there exists different kind of grants for different ladders of researcher's career. The researcher needs to sell their ideas to reviewers, by balancing between clarity and depth. Often this means using collaborators and consultants to strengthen the application. If the researcher does not have the required professionalism often the collaborators (professors) need to ensure the grant funder that researcher has the sufficient assistance for research project. (Yang 2005). Continuously decreasing success rates for grants, in 2012 and 2013 rates fell from 20.5% to 16.9% and heaviness of the grant system have been causing a lot of dissatisfaction which has been leading to trying to find alternatives for current grant system. One suggestion is use a formula (Barnett, Herbert and Graves 2013) where department would receive the grants as a function of number of degrees, publications, and support from other agencies and industry.

When researcher applies for funding, it is not obvious that one will receive it. There exists only a limited amount of resources in the same pool of resources, and multiple researchers apply for the same resources. In order to receive funding researcher has to write a funding proposal and meet the requirements assigned by the funder. To write a funding proposal takes a lot of time, and it is not obvious even then that the resources will be received since there are other researchers competing from the same resources. So at the end of the day it is funder's decision. The reason of existence of the system is that the allocated resources go to right researchers and who are capable for providing greatest amount of scientific output (Koelman and Venniker 2011).

2.5 Research problem

The current resource allocation system should be evaluated to determine whether it is possible to find alternative, more effective way to distribute the resources for academia. This is the underlying motivation to validate simulation as a decision-making tool. In other words, how the funding problem can be examined by the means of the simulation and how strong conclusion can be drawn from the simulated model.

The proposal-evaluation-grant system is operated by the same resource pool that could be allocated to research instead. The reduction of the resources from grant application system would lead to less administration but a decrease efficiency in resource allocation, meaning that the resources are less probable to go to right persons. On the other hand, when too many resources are assigned for surveillance, the less resources there are for the actual research. Also, the resources that are allocated to research could be allocated more effectively between the researchers within the population. The key for this current study is the limited resources for doing research, since this is the case also in real life. This will be zero-sum game, so the amount of resources will always be the same compared to population.

These two clauses, grant processing time and distribution within population, (attributes) are mirrored against scientific output; by making changes in resource allocation scientific output is monitored for optimal output. This leads to result in to which direction the research funding process distribution should be improved.

The study uses social simulation to model the research environment. In order to create a simulation, the system needs to be evaluated and translated into more formal, simplified form (and eventually into a computer-understandable form).

Now the generated model itself can be validated; how trustworthy are the results, what are the weakest parts of the model, and from which variables the model is the most dependent on. Achieved results are validated against the trustworthiness of the model, and conclusions are given about using social simulation as a tool in this type of problem-solving. It is practically impossible to conduct experiments on allocation strategies in real life. Simulation can be a tool for experimenting and observing the consequences that otherwise would take a significant amount of time.

3 Model design

In this chapter the focus is on identification and to answer what kind of system is modeled in this case (present the entities, submodels, and the literature sources to fit the instances together), and what kind of data there is to be observed. In other words, this chapter presents *what* is supposed to be simulated.

3.1 Background for system analysis

A system is defined as a group of objects that are joined together in some regular interaction toward the accomplishment of some purpose. System can be considered as an observed subset of reality. System is imitated by a model, and using the model a simulation is created to represent the system (Banks, Carson, et al. 2010).

Model is a representation of actual system. Model should be complex enough to answer the questions raised (in this case research questions), but not too complex. Model does never fully explain the actual system, but is a simplification of a real world system (Banks, Carson, et al. 2010).

In simulation model, various concepts can be identified that are necessary for functional simulation. An **element** (or entity, object) represents an object that requires explicit definition. Entities are system components in the system. They communicate with other entities, request resources and so on. **Attributes** are variables storing an entity's state, and **parameters** are the values assigned for variable. Attributes of interest in one investigation may not be of interest in another investigation (Page and Kreutzer 2005). Depending on a view, variables have different names.

For example, researcher is an entity, its attributes are skill, years in academia, frustration, and parameter is when years in academia is set to "25". A behavioral relationship connects elements and attributes together as a system structure.

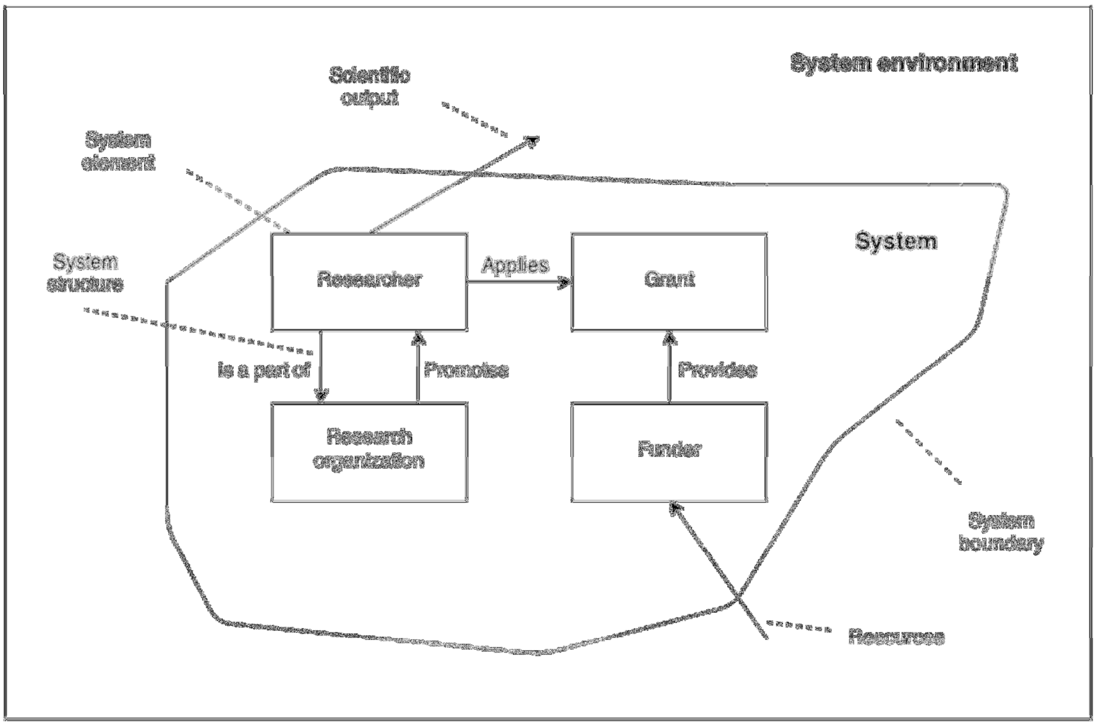

Figure 1: Terminology of the system

Simulated models consist from three parts: input, output and functional model. Models can be categorized into different ways:

- Concrete vs. Abstract model (miniature vs. computer model)
- Deterministic vs. stochastic model (known vs. randomized data)
- Analytical vs. numerical model (formula vs. approximate)
- Continuous vs. discrete (infinite vs. finite number of states and changes) (Tiihonen 2014)

Models are also categorized by their purpose. Explanatory model aims to show why and how the system works as it does, predicative tries to predict the probability of an outcome. Design model is based on analysis and architectural requirements of the system. Optimization model is used to find the best possible choice out of a set of alternatives (Page and Kreutzer 2005), which is used in this thesis.

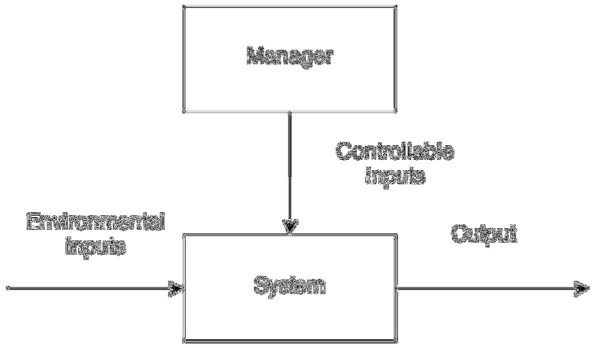

Figure 2: User view of the modeled system

3.2 Entities and attributes

Heuristic model is created according to all the entities mentioned in chapter two and attributes of them are presented. The exact values for attributes are left out, but the dependencies which are mentioned among the way. In order to measure the scientific output against resource allocation and effectiveness of the evaluation process, the mechanisms between them needs to be identified and implemented. In the beginning, the simulating system can be referred as “Black box” (Figure 3), where input and output are known but the inner logistics, for now, stay unknown.

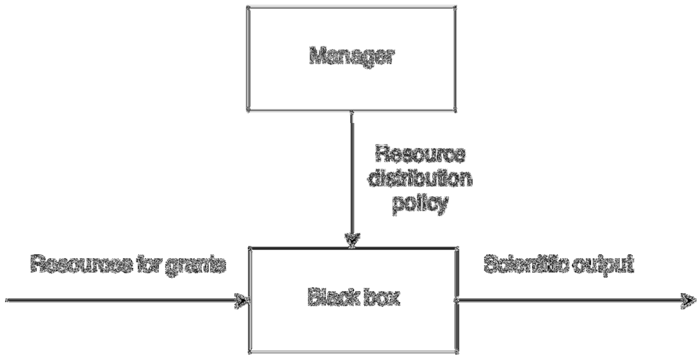

Figure 3: input (grants) and output (Scientific output) of the simulated system

The unknown black box determines how the grant funds turn eventually into measurable scientific output. The comparison is done by making changes into resource distribution, and the scientific outputs of the data points are compared. The important elements (or entities) that need to be presents between input and output (or parameter and response) are publications, publications providing researcher population and the resources that keep the population running. These elements need various mechanisms to imitate the system.

- Since scientific output is to be measured, the researcher entity provides measurable artifacts. Hence the research and publication mechanisms need to be present.
- Research population consists of unique researchers, decision making agents.
- Positions of the researchers within organization are not vital, but it makes the systems structure and dynamics accurate since it is connected frustration, and eventually to leaving decision.
- A mechanism where researchers sacrifice resources for receiving more (or any) resources. Researchers have a certain expectation for the resources they might receive which can lead to increase or decrease on frustration. This is generally a grant writing process, followed by its consequences.
- To make a difference in quality of publication, a publication process where the publications are treated differently depending by who and how they were written. In examined original system this is often done by peer-review process, which is left out from this study.

The model can be transferred into a simulated system, and the results of the simulation are presented in response surface model, where the response is the scientific output and explanatory variables are distribution of resources and noise between the grant-providing party and resources allocated for research.

3.3 System identification

In order to determine what lies between the affecting parameters and the response, the examining of the model begins from how the scientific output is determined from research, and what fuels the research and its environment.

3.3.1 Scientific output production

In order to measure impact of scientific output, the artifacts of research needs to be determined. Measuring a scientific output includes a broad range of approaches, such as total number of citations and journal impact factors (Garfield 2006). Also scientific output can be page ranks, article download statistics, and comments using social (Mandavilli 2011). The disadvantage of just measuring the number of papers it does not measure the importance of the paper. With total number of citations the disadvantage is that the scientific output can be inflated by a small number of really good papers (Hirsch 2005). But then again, does this make researcher worse if the quality between publications varies since it is up to citers to choose which source to use. Other ways to measure scientific output can be such as citations per paper (problem is that researcher cannot produce any bad papers to keep the productivity high), number of significant papers (favors and disfavors researchers, depending on the case) and number of citations to each of the q most-cited papers (lacks a single indicator of the quality (Hirsch 2005). J. E. Hirsch proposes a h-index to characterize the scientific output of the researcher by calculating citations and papers, and also time used for those. Study (Franceschet 2009) presented 13 different bibliometric indicators to measure scholar performance. These indicators can be divided into three groups: Paper-, citation-, and hybrid based metrics and some of mentioned above are included. For the sake of simplicity of the implementation, the quality of the research is measured only by the amount of citations. This is not the best way to measure it, but it loosely takes heed to amount of publications so the quantitative and qualitative measurements are present (Kreiman and Maunsell 2011).

Impact measures, such as citation counts, allow one to find out where and how often an article is cited. This provides an estimation of the importance of an article. Citation statistics are widely used for the allocation of funds, promotion, tenure decisions and determining research influence. Typically it takes two to five years to produce meaningful measurable citations counts but from journal impact factors it is possible to draw conclusions in shorter period of time (Boyack and Börner 2003), (Franceschet 2009).

The scientific output is a function of work and skill. Now that the scientific output is created by papers and their citations next step is to present how the papers and citations are formed. Publication has a method and attributes to create citations, which is affected by the personal traits of a researcher and the researchers citing the paper. Paper can be considered as one entity wrapped inside a publishing mechanism.

For a mechanisms to count citations of a single paper as a function of time, (Wang, Song and Barabási 2013), presents a formula for probability how the citations are formed over the time. The formula is presented more accurately on latter chapter but for now it is only necessary to know that it is dependent on personal attributes of the researcher. However, the formula doesn't pay attention to non-personal attributes such as how other researchers receive the paper, how it is distributed and reviewed.

3.3.2 Career of academic

The staffs of the universities are not equal, but the position within organization is dependent on seniority and merits during the career. According to University of Canterbury HR, academic career starts from Lecturer position (But the US term Assistant professor from here on).

Every rank is divided into steps, and one step means one year. However, before moving to next rank the steps can be skipped by applying for promotion (Canterbury University HR 2013).

Ranks in US	Ranks in NZ	Ranks in Finland	Av. years in position
PhD. Student	PhD. Student	PhD. Student	*4*
Post-doctoral researcher	Post-doctoral researcher	Post-doctoral researcher	*4*
Assistant professor	Lecturer	University researcher	*6*
Associate professor	Senior Lecturer	Adjunct Professor	*8*
Professor	Associate Professor	Professor	*4*
“named” Chair	Professor	Professor	*7*

Table 1: Ranks compared between countries

First step is to come up with a solution how the population is distributed between positions within organization. This really depends on the country that is to be examined. The titles can be different as well as the structure. Some academic organizations have four different positions, some can have almost ten. Some positions are meant for researchers, others can be for teachers only. In the system created, a simplified model of US model ranks is used. The chair-rank is taken out and combined with professor. Also, the levels before post-doctoral rank are taken out. Undergraduate students leave are left out because their small contribution to scientific output. PhD students are left out (for simplicity, even though their contribution can be seen through grant-received researchers) as well mainly because the used data does not show accurate figures from this position, and because this is thought as a part of training for actual research positions. This leaves us with 4 different positions with a doctoral degree: Post-doc, assistant professor, associate professor and full professor.

Data (Canterbury University HR 2013) does not include the information about post-docs so the estimated values are used. The assumptions are that post-docs mostly do research and apply for grants.

In Australia, unlike the United States, but similar to Britain and New Zealand, the academic hierarchy has distributed as follows. In 1983,

- *9.3% of university teaching and research staff were professors,*
- *12.3% were associate professors/readers,*
- *32.1% were senior lecturers,*
- *22.5% were lecturers,*
- *6.0% were principal tutors, senior tutors/demonstrators.*

The remaining 17.8% were tutors/demonstrators on contract who cannot compete in the promotion stakes (Moses 1986).

In New Zealand, (Canterbury University HR 2013) the distribution was in 2013 as follows:

Position	Amount	Percentage
Full professor	*28*	*22 %*
Senior lecturer	*54*	*43 %*
Lecturers	*22*	*17 %*
post-doc	*22*	*17 %*
Total	***126***	*100 %*

Table 2: Distribution between positions

The features that both sources have in common are that the shape of the organization is like Grecian urn, where senior lecturers are the largest group. There is also evidence (Jones 2013) that in computer science field the amount of post-docs is vastly larger. The assumption is that those positions prosper well where the grants are directed.

In order to get promoted, academic needs to fulfill certain traits before he/she is eligible for a higher position. Researcher is eligible to apply for promotion after a certain amount of years and a proven track record. In the 1990's, Ph.D. was gained at age of 31 and the first professorship was gained at the age of 43.9, on average. Part-time positions have been slightly increasing after year 2000, now reaching over 40% of the positions. Full-time tenured positions have been reducing at the same rate, plateauing to 17%. Full-time tenure-track positions (Which means, they will receive a tenure once it is possible) have been also decreasing, approaching to 15%, while Full-time non-tenure-track positions have been plateauing around 15% for years (Curtis and Thornton 2013).

It needs to be also examined whether the size of the organization has a relative correlation to scientific output. Study (University Alliance 2011) states about the dependency between organization size and its performance as follows:

"There is no continuous relationship between research unit size and performance in RAE2008. Among smaller research units there may be a significant positive correlation between size and performance but above a certain threshold no further improvement is evident. It is also apparent that there are small and median-sized units which perform as well as, and in some cases better than, the largest units. There is no evidence that funding on the basis of scale would improve overall performance."

Finally, what is the cause for researcher to leave the research organization, or, the whole academia? One study (Johnsrud and Rosser 2002) found that there are many reasons, such as salary, integration, communication and centralization. Opportunity is also one reason that affects psychological factors such as job satisfaction, morale and commitment. Either gender or ethnicity does not explain changes in morale. Study (Barnes, Agago and Coombs 1998) presents similar factors as study mentioned above, but comes down to two important predictors; frustration due to time commitments and lack of a sense of community. Especially researchers who haven't gained tenure are quite likely to leave research organization, if they have received offers elsewhere. From all the factors previously mentioned this study condensates the behavior to one factor, frustration which is dependent on a handful of predictors, Salary and recognition gained through promotions and citations. Retirement age varies greatly, and many professors tend to continue their career years after even they have the possibility for retirement.

To join the academia there are various reasons which are left out from this study. The trend is that there exist fewer possibilities to have a career in academia than there are students willing to have a career within academia.

3.3.3 Attributes of a researcher

There have been various studies which try to define the factors that affect to the researcher's ability to provide research results. Gender, age, education, rank of the university, talent, luck, effort, tenure, rank, and seniority in rank are also mentioned attributes (Kelchtermans and Veugelers 2005). If further examination is wanted there are plenty of personal features, such as courage, intelligence, writing skill, charisma, social contacts and determination and so on. Ingrid Moses has made research about rewarding academic staff and motivation. She mentions that Maslow's hierarchy of needs also applies here, safety (having a tenure etc.) increases motivation. Also, sense of achievement such as recognition for teaching, promotions, to supervise, give papers and applied grant increase motivation. And vice versa, lack of these recognitions is causing demotivation (Day 2011). There is also introduced one possible factor, endurance. This indicates how much negative feedback researcher can tolerate without it affecting to their motivation.

The study “Modeling academic research funding as a resource allocation problem” (Geard and Noble 2010) presents a model where they define a few attributes such as research skill R which is distributed uniformly between 0 and 1. There is also attribute G which increases researchers scientific output if grant awarded. Researchers also have an attribute called effort. If they apply for a grant, this application time is taken away from actual research time. A few assumptions are also made: The quality of researcher’s application correlates with his/hers research skill. However, researcher can compensate lower research skill with using more time for application, and there is a diminishing returns on time invested. Also, the funding decision is not accurate, hence there is a noise factor allocated.

It is also very common for university staff to teach. This leads to natural question whether there is a correlation between research and teaching, if the teaching provides additional boost to research, which is not the case (Hattie and Marsh 1996). Of course, teaching time reduces the time for research time and vice-versa.

There is a small correlation between gender and publications, in general females tend to publish less at the beginning of their career compared to males but this reverses later on (Kelchtermans and Veugelers 2005). Also, there is a quadratic relationship between age and number of publications. Luck can be considered as random, not ordinary behavior which is favorable (or for bad luck, unfavorable) for entity. Hence the simulation needs contain to stochastical elements. These random values producing elements can be considered as “luck” and chance.

Tenure-centered attributes are also clearly involved into scientific output: People with higher position tend to have higher scientific output. This is also connected to time since study shows (Link, Swann and Bozeman 2008) that the fewer researchers have time for actual research when they advance in their career. However, they’ll get their name into more publications because of supervising role.

For the model it is wise to recognize all the attributes, but to choose from only a handful of well-defined attributes for the model. In the literature common mentioned factors for researchers are the following:

- personal properties (talent, skill, intelligence)
- career-centered attributes (tenure, seniority)
- effort (motivation, determination)
- time (available in total, used for applying, used for research, used for teaching)

Also it should be said that writing the application is actually doing research and is helping later with research, so it is not just time wasted, assuming that the application has been approved.

Motivation has also other consequences but just lowering the productivity. The study (Barnes, Agago and Coombs 1998) presents one criterion variable, Intent to leave academia, which is dependent on four predictor variables, the stressors (reward and recognition, time commitment, influence and student interaction), and from two moderator variables (Interest in discipline and sense of community). They showed that there is a positive correlation between stress and intention to leave academia. From the four stressors, time commitment had the strongest correlation for intention to leave the academia.

These mentioned four types of attributes that affect to researchers performance; personal properties, seniority, motivation and time need a submodel how they are formed. Personal properties don't really need a separate submodel because they're considered as constant variables from the beginning of the career. Of course, skills are gained during the career but the capability to learn and produce is fixed at least for the simplicity. Also for motivation, researchers reach differently to setbacks and successes. Hence a heuristic model needs where some actions raise motivation and others lower it. A general model for motivational factors is presented later in this chapter.

3.3.4 Grant allocation

To get the idea in what kind of environment academics often work it is needed to dig deeper the fundamentals of the frustration factors, such as what are the non-personal characteristics that help to survive in the competitive environment.

In England, Economic & Social Research Council (ESRC) has faced an incline of applications of 33% between years 2006 to 2010 (from 2,000 to 2,800) which is a good thing for science. Unfortunately, the budget of ESRC has not kept pace with the rapid increase in applications. ESRC receives more applications every year while the budget does not keep in the same pace. This means that within the same time period the success rates of grants have declined from 26 per cent to 12 per cent, on average (ESRC 2010).

The Large Grants Scheme of the Australian Research Council (ARC) is generally regarded as a most prestigious source of research funding for Australian academics. They state the success rates for grants hovers around 22 per cent (Bazeley 2003). Among ARC, The median age for grant applicants was 40-49 years. Younger applicants (< 40 years) who applied solo or first named researcher tended to be less successful than those who were older, but younger researchers in overall were not disadvantaged. However, grant applications with a young researcher as second or third named investigator were at least as successful. Male researchers receive more grants in general, but it is more because lack of seniority within female researchers than the gender itself (Bazeley 2003). This indicates that there is no significant difference either in age or gender.

Only 7 out of 174 successful ARC applicants did not have a PhD or greater qualification. It is fair to assume in the model that only research with PhD qualification or greater is examined.

Receiving a grant is dependent on the University that researcher is representing, as well as it is dependent on academic status; having a research-only position had a strong relationship to success in winning in funding. Research fellows and readers are more likely also to be successful on average than teaching and research academics at an equivalent level (Bazeley 2003).

Combining full-time researchers and university professors in labs tend to preserve incentives. Size of the labs, the individual promotions, and the role of non-permanent researchers and of non-researchers are also underlined (Carayol and Matt 2004).

There exists a correlation between previous received grants and grants yet to be received; Researchers who received grants before are way more likely to receive (Only 4.1% of funded researchers/teams had not had any funding during the previous three years for any projects) grants than those who did not receive any. When it comes to publications, successful ARC applicants have been solo or first authors for more books, articles and chapters than unsuccessful ones (5.31 vs. 2.73 relevant publications within 5 years, respectively (Bazeley 2003). It is also mentioned that validation of applications is considered as one task of being a professor.

The time usage for grant writing is a tradeoff between teaching, research, service and other general tasks within academia. This all is dependent on the position, and whether the personnel have tenure. The differences in time allocations between researchers are shown in table 3 (Link, Swann and Bozeman 2008). The table 3 indicates to averages, naturally the grants are for different periods of time.

Mean fraction of time	Total	Teaching	Research	Grant writing	Service
Assistant professor	100 %	32 %	40 %	12 %	17 %
Associate professor	100 %	35 %	35 %	8 %	23 %
Full professor	100 %	30 %	35 %	7 %	28 %

All assistant professors are untenured, rest hold tenure

Sample size = 1365, population size 33,813

Table 3: Time allocation of researchers

By using these presented values it is possible to come up with a submodel where the time used for research and grant writing are connected to scientific output.

3.3.5 Validation of research

Before the publications (or whatever measuring parameters are used) can be considered as scientific output, they typically need to go through a validating process. This chapter focuses on peer-review process, since that is the main channel within academia. It is notable that the validation of research and validation of the grant allocation are using similar mechanisms. In the model the validation is not implemented but treats everyone equally.

The basic idea of peer-review process is that researcher's publication goes through a reviewing process, where one (or more) experts of the field go through the paper and evaluate its quality. If the quality requirements are not met, the publication is returned with comments so researcher can make the decision to improve the publication and to try again or not. The actual review process works so that the publication goes through different steps where it is validated. Figure 4 below shows Elsevier's example.

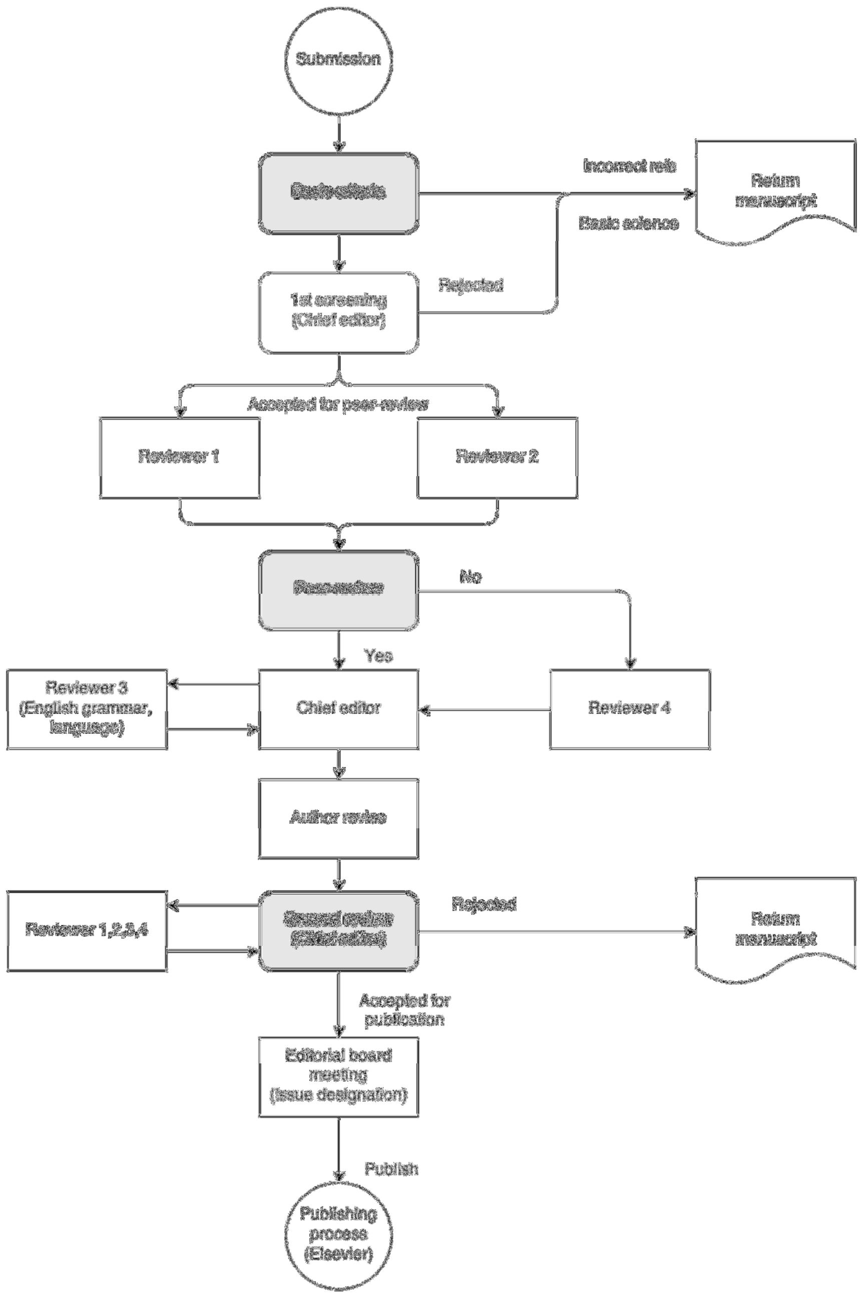

Figure 4: Example of Peer-review process flowchart (Elsevier)

There are a few common methods for reviewing, such as single and double blind review, and open review. In the first mentioned the author does not know the identity of the reviewer, and in double blind review both are unknown to each other. Open review means that the identities of reviewer and author are known to each other.

Peer reviews are often the method which affects to papers, promotions, grants and recognition within academia. There has been a lot of debate of the consistency of the peer-review. Two different peers can have completely different view for paper. Fictive paper which contained errors made on purpose was sent for reviews. “78% were returned. 59% of those were rejections, 33.5% revisions and 7.5% acceptances. Even so, even the rejective ones didn’t find all the errors on the papers. The conclusions of the paper were not supported by the results, and 68% of the reviewers failed to see this (Baxt, et al. 1998). Study (Peters and Ceci 1982) revealed that already published research articles, with slight changes, were resubmitted to the same journals, and only three out of 38 detected resubmission.

Not only unreliable, but there are also costs in Peer review system: *“The average cost per manuscript for salary and fees only, excluding overheads – infrastructure, systems etc., was 250 US dollars per paper (…) other source states that expenses for publication in two OA journals peg costs at around 300 US dollars. (…) The more editing and selective work is done the more expensive the process becomes* (Andersson 2013).

The cost of peer-review per paper estimated for British Medical Journal was around 100 British pounds, whereas paper that got published straight was around 1 000 pounds. It has been also proposed that authors would pay for peer review. There is also evidence that there exists bias against woman and less known institutions (Smith 2006).

3.4 General functionality of the entities

Now that the system and its characteristics (with some numerical values for latter experimental use) have been introduced, general functionality of the model can be generated. The process is iterative, so it can be considered as simultaneous, interactive loops (figure 5).

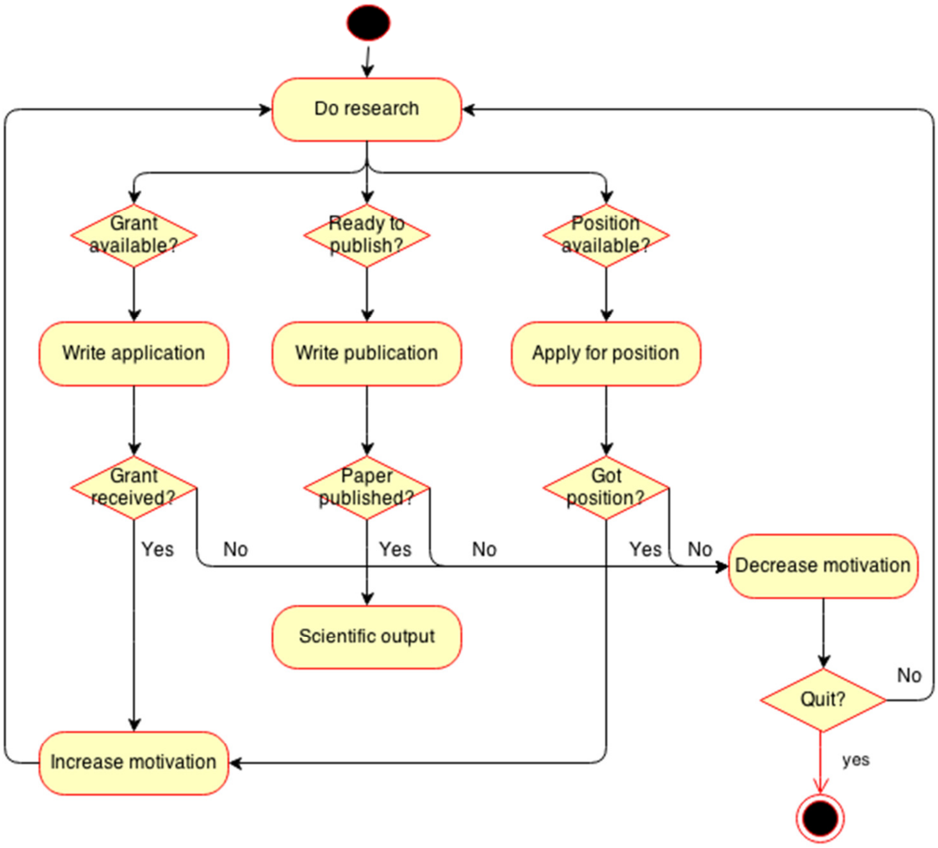

Figure 5: A general activity diagram of a researcher

In reality, research happens all the time while grants, promotions and publications occur within undefined periods of time. To move between positions, researchers need to fulfill a certain demands in order to get promoted, and when staying long enough in the same position frustration also grows more rapidly. Promotions should follow the guidelines of organization structure, so they can only happen if there is position available on the next ladder.

The occurrence of citations does not happen overnight but over the time so the impact of publications depends on the period of time. Figure 5 in previous chapter explains the process of publications. A stripped version is used in the model.

The production of publications is increased from motivation increasing events, such as promotions, received grants, accepted publications etc. These events are self-feeding system since promoted researchers are more likely to be more active researchers on publishing.

The scientific output of a researcher is measured, and if there are positions available and enough evidence for promotion this happens. However, there is always a small possibility for resignation every year. Figure 6 below explains how the promotion process is running in this case. When certain changes occur in the system, the event might trigger changes in the organization structure. When occurred, more might follow.

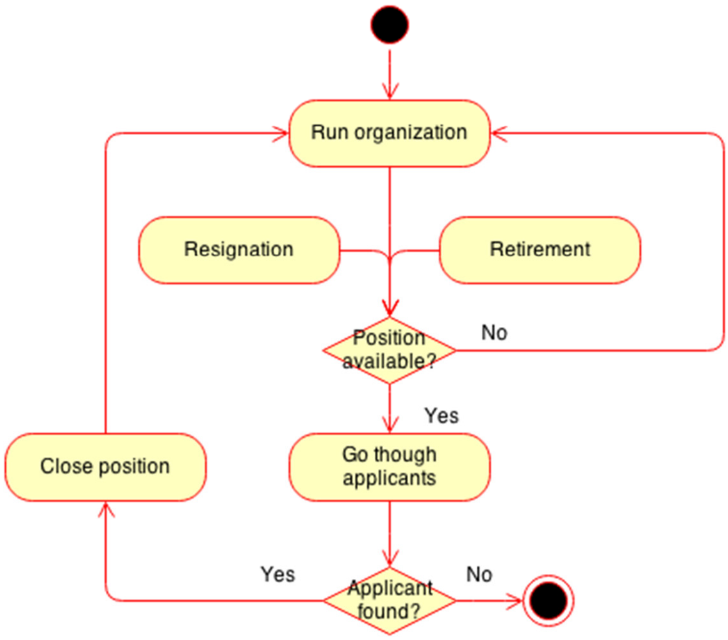

Figure 6: Activity diagram from research organization perspective

Academics perform research, and there are different kinds of grants available. If an academic decides to apply for grant, he consumes time for application writing. This application is either accepted or rejected, depending on if the need and demand meets. Negative grant decision lowers morale which leads to less publications, and, eventually, to resignation. Grants can be for different periods of time, and during that time researcher can focus on research and/or apply for upcoming grants.

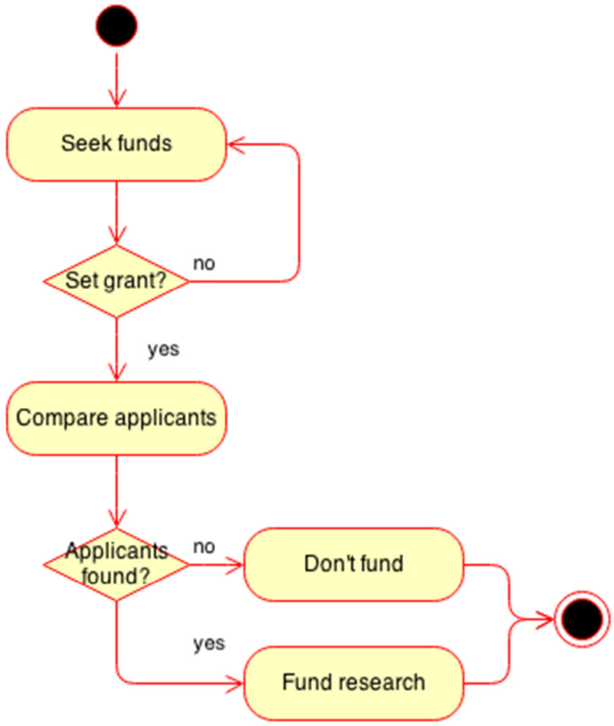

Figure 7: Activity diagram from grant provider's perspective

3.4.1 Researcher

Researcher entity contains many quantitative factors which personalize its behavior. Some attributes are defined in the creation of the entity, and some of them vary over the time.

Research skill

Researcher has an attribute skill. Research skill determines how good papers are published, and how likely grants are received. This is assigned in a same manner as intelligence quotient, IQ, which is normally distributed. However, since it is assumed that the schooling system is working on at least in some level, the assumption is that the lowest percentiles are not counted to this population.

Applying skill

A study (Day 2011) states that high research skill does not necessarily indicate good performance in publishing. This factor determines how well researcher applies to grants. It is completely independent on research skill, and they together affect how likely it is for researcher to receive a grant. It is assigned in same manner as research skill.

Frustration

Frustration is affected by many other methods, including promotions, grant decisions, promotions etc. Generally, frustrated researchers produce fewer papers (Day 2011). Eventually researcher leaves the organization when trigger value has been reached. In reality would not mean the researcher leaves academia, but just the currently observed organization. However in the simulation researcher, as entity, disappears from the observed system.

As the research artifacts of one organization are measured, a closed system is examined. This means frustrated researchers can leave to other research organizations too, but this is not observed in simulation.

Years in academia

Every step in simulation is considered a “year”, and this indicates how many steps researcher has been around. This value is used in retiring, and promotions can’t be achieved until a certain amount of years in research.

Productivity

Productivity tries to imitate situation where the researcher responds to the latest year’s event of resources allocation, and also to event history of previous years of resources allocation and promotions. Productivity is dependent on two factors, called commitment and monetary productivity, which are summed up hence creating productivity. Commitment is dependent on the current frustration and resources productivity depends on received resources on last grant round. Resources productivity compares resources needed to be effective (explained below) and received resources. If the expected amount of resources is not reached, productivity is lowered, and if exceeded, productivity grows.

Resources needed to be effective

To control the frustration the researchers needs to have an expectation how their usage of time for grants is made up for them. This parameter indicates how much resources researcher expects to receive. This model is hypothetical, and it is assumed that researchers on lower positions are more dependent on grants than the researchers who have already tenure. The logic behind this is that in the lower positions researchers are more dependent on outside funding than on the top.

3.4.2 Grant

The purpose of funding entity is to provide resources for researcher entity. In order to avoid frustration and increase productivity, researcher needs to receive funding. The resource distribution follows a pattern where a resources pool (attribute amount of funding) is defined. From this pool, some of the resources are distributed between researchers as a base funding. The resources that are not set as base funding are the ones researchers have to apply for. From these resources a part is defined for grant system administration. All of these three attributes, amount of resources, base funding and level of administration are defined by user.

Amount of funding

When running the simulation, user can define the level how much funding will be allocated to the system. Of course the scarceness depends on the size of the population. The basic idea is that there is not enough funding for the whole population, as in real life. Funding can be considered as a fuel for researcher to keep the productivity high. The resources needed to be effective attribute defines how much more time the researcher will receive with positive funding decision.

3.4.3 Paper

The paper-entity provides citations. Because citations of the paper occur as a function of time, that is one parameter for the entity. The quality of the paper is also dependent on who wrote (skill) it and how much time has been used for production. Citations are a single number which is measured in response surface against the two independent factors, resource distribution and resources to evaluation system. The next chapter subchapter presents the formal model how the citations are formed depending on the attributes of other entities.

3.5 Formal presentation of the model

Because model contains numerical procedures it has a closed form solution. This means it is possible to write down a single equation for the whole system. For simplicity, instead of presenting the whole system as one equation to explain the functionality of the simulation model, it is presented using equations and formulas where the quantity of attribute can be presented as a function of other attributes. However, ***attachment 1*** presents how the entities and attributes are connected without presenting the mechanisms between them. The basic idea is that every year the citations of the population are measured and summed together. As the citation output of a research population as a whole is from a certain period of time is wanted to be measured, the need is to define which are the factors that

- Affect to citation output
- Are to be adjusted

First, a formal overall equation and later the mechanisms that affect to that equation can be presented. The amount of citations is presented as a function of time, meaning citations per year. All citations that the whole population produced during the time period *y*, C^O are calculated from citations of all researchers.

$$C^O(y) = \sum_{P \in P^O(y)} C(P, y) \tag{1}$$

Where $C(p, y)$ contains all citations C of the organization population p in the year *y*. The publications of the organization $P^o(y)$ in year *y* is defined as follows:

$$P^O(y) = P^o(y-1) \bigcup_{r \in R^o(y)} P^r(y) \tag{2}$$

where *r* is output of publications on year *y* of researchers R^o and y is the year. Citations per researcher are formed as follows:

$$C(r,t) = C(r,t-1) + \sum_{P \in P(r)} C(P,y) \tag{3}$$

After this, there can be recognized two paths: Other one is how citations of the publication are formed and the other one is how the amount of publications is formed. First, it is present how the citations of one publication are formed and in latter chapter it is explained the behavior for how the publications of researcher are formed.

Citations *C* of one publication *P* in year *t*,

The formation of citations is a function of the quality of paper and point of time *t*, since the citations are not formed in one instance but over the time instead.

$$C(P,t) = C(P,t-1) + f(t - t_P, q(P)) \tag{4}$$

where t_p is the year published, and $q(p)$ indicates the quality of the publication *p*.

Quality of publication, *q*

The quality is formed as a function of researchers skill S

$$q(P) = \hat{f}(S(r,P)) \tag{5}$$

Amount of publications, p_r

Amount of new publications, during year t, is a function of time available for research τ_r and researchers productivity θ_t. This, of course, is measured at one point of time, t.

$$P_r(t) = f(b(\theta), \tau_r(t)) \tag{6}$$

Productivity

Productivity θ consist from two parts, resources productivity (depending on funding decision) and commitment (which is stochastic function of promotion decision).

$$\theta(t) = f(\theta_b(t), \theta_m(t)) \tag{7}$$

Where θ_{b} is the commitment, which is dependent on the frustration $\theta_{\mathrm{b}}(t) = f(f_r, t)$ and $\theta_{\mathrm{m}}(t) = f(m_{w,}, m_r, t)$ is a function of expected resources $m_{w,}$ and received resources m_r

Joining / leaving from organization, *O*

The population size O at the point of time t can be represented as:

$$O(t) = O(t) \backslash L(t) \cup J(t) \tag{8}$$

$L(t)$ is the leaving population and $J(t)$ is the joining population at the point of time t. Researcher leaves from organization if one of these conditions is true; frustration f_r grows too large (exceeds 1) or time spent in academia a grows too high which is considered as retirement (over retirement age, λ). More formally:

$$r \in L(t), if \begin{cases} f_r \geq 1 \\ or \\ r(a,t) \geq \lambda \end{cases} \tag{9}$$

On contrary, joining to organization is set so that it is a percentage e from the population size O. So $J(t) = O(t, e)$.

Promotions

Whether researcher gets promoted, it is dependent on one's career success (citations, c_r), years in academia a, researchers current position p and positions available on next ladder p_n. Furthermore, this is dependent on amount of researchers on next ladder $p(r, t-1) + 1$.

$$p(r,t) = f\big(p(r,t-1) + 1, p(r,p_n)a(r,t), c(r,t)\big) \tag{10}$$

The positions that are available are not fixed, but the amount of positions scales depending on the size the population as a whole. This way the promotions are dependent on researcher's personal traits, not just from static amount of seats since new positions open occasionally.

Frustration, f_r

In this case, frustration is measured for following reasons:

- It effects the productivity of a researcher
- It makes one to seek other options outside research organization

Assumption is that frustration grows from:

- Every negative funding decision (partially, or fully)
- Every promotion which researcher thinks that one deserves but don't receive

Frustration carries memory from its previous values, hence frustration can be written as:

$$f_r(r,t) = \sum_t f\big(f_r(r,t,a,p) + (r,t,m_w,m_j)\big) \tag{11}$$

where r = researcher, a = years in academia and p = position. This is dependent also from how many resources were wanted m_w and received m_j at the point of time t.

Quality of application, *A*

The quality of application is dependent on researcher's personal traits and how much time one has to prepare the application.

$$A = f(r, \tau_a, f_r, s_r, s_a) \tag{12}$$

where τ_a is time used for applying, f_r is frustration, s_r is research skill and s_a is applying skill. This value determines how good changes the researcher *r* has for winning a grant. *N* is the distribution and evaluation caused overhead.
Selection outcome of the grant applications are reviewed with limited resources, hence the perceived quality of applications contains evaluation error $N = 1 - \sqrt{h}$, where h is overhead. This error is presented in Figure 9.

$$A_s = A + N \tag{13}$$

Time available for research, τ_r

Time available for research is dependent on the τ_{tot}, total time available after teaching and similar tasks, and how much time from this research time were used for applying τ_a and how much resources were received through grants τ_g, and how much resources were received through base funding τ_b.

$$\tau_r(r, y) = \tau_{tot} + \tau_g(r, y) - \tau_a(r, y) + \tau_b + e \tag{14}$$

Time used for applying, τ_a is dependent on the level of frustration f_r and from the time researcher uses for applying on that position, on average. The time used for applying varies uniformly, $e \in unif[-0.116, 0.116]$.

$$\tau_a(r) = f(P_r, f_r, r) \tag{15}$$

Resources for research, τ_g

Resources for research consist of uniformly distributed resources among researcher population O, where the resources are distributed depending on the quality of researchers application. As mentioned before, the resources received through grants τ_g are summed to user-defined base resources τ_b which is dependent only from the percentage user wants to share equally between researchers.

The amount of received grant funding τ_g is dependent on quality of the application A_s, when it is compared to quality of other researcher's applications.

$$\tau_g = \begin{cases} 0, if\ A_s(r) \leq A_{others} \\ \tau_q, if\ A_s(r) \geq A_{others} \end{cases} \tag{16}$$

4 Mechanisms of the model

As presented, there are limited resources for researcher population where everyone cannot be satisfied. The factors that are artificially modified by user to see changes in the results, as mentioned in the research question, are:

- Whether the resources are allocated for grant process administration, making it more accurate but causing a decrease in resources that are available for research
- How the resources that are allocated to research instead of administration are distributed between the researchers

Now that all the used attributes are presented, the interaction of attributes can be designed. This chapter defines all the mechanisms used in the model between the entities, defining the formulas of presented functions Some subchapters describes the model functionality formally, and some chapters explains it through algorithms or even verbally.

4.1 Resource allocation between researchers

The resource allocation to grant-review-process is a tradeoff between resources and accuracy of the system. The mechanism presented defines how the user of the simulation allocated the *funding model*, x. The more resources are allocated for researchers (less overhead), the less accurate the evaluation process will be and vice versa. On the other hand, the resources that user decides to share to researchers, can be shared to population in different manners.

The population can be thought as a queue of researchers, and it is up to simulation user to define how much resources each one gets. The fundamental idea is that there exist 100% of resources (whatever the resources actually are) and this 100% of the resources available can be assigned between administration and researcher however the user wants. The funder does not react to results of the funding and carries no memory about previous actions, meaning the model is static and there is no interaction besides the grant giving between the funder and researcher. Funder sets the resources depending on parameters mentioned in previous chapter. The real-world system is more complex, with social connections between, but for simulation purposes a stripped model is used.

Figure 8 below illustrates example where on x-axis the population of 10 researchers. The researchers are sorted in a manner where on the right are the most skilled researchers with the best applications. Figure 8 presents a situation where 25% of the resources are distributed equally, and the remaining 75% is given to applying researchers.

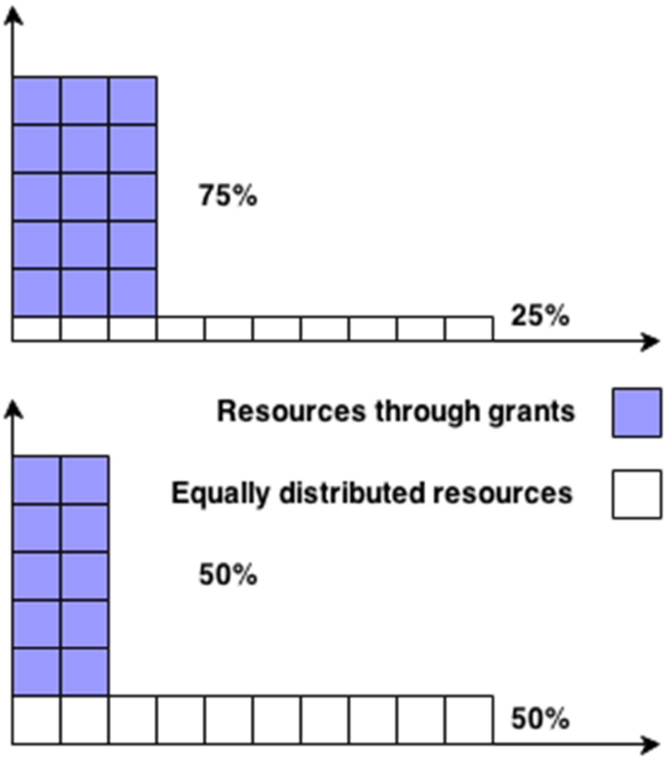

Figure 8: example of distribution of resources between the researchers

This funding model is one (other one is error in overhead) of the two independent factors how the input for the model is varied in order to see changes in scientific output.

4.2 Amount of publications

Amount of produced papers of a researcher at reached point of time t, (t = *0…t*) follows formula:

$$P(r,y) \sim Exp\left(\frac{|y| * \theta * \tau}{\omega}\right) \tag{17}$$

Where θ was, as mentioned before, productivity and τ research time available, and ω is constant defined through calibration.

4.3 Productivity

Commitment is created with formula $\theta_{\mathrm{b}} = 1 - f_r * \omega$, where f_r is current frustration and ω is pre-defined constant. Resources productivity θ_{m} however is dependent on how much resources was expected, m_w and how much was received, m_r, more formally:

$$\theta = \theta_{\mathrm{b}} + \theta_{\mathrm{m}} = (1 - f_r) + \max\left(1, \frac{m_r}{m_w}\right) \tag{18}$$

4.4 Frustration

Frustration consists also from resources and promotions that were not received, as follows:

$$f_r(r) = f_m + f_p \tag{19}$$

Monetary frustration f_m is a function of expected resources m_{e}, resources received m_{r}.

$$f_m = \left(1 - \frac{m_r}{m_w}\right) \tag{20}$$

However, the promotional frustration is dependent on how much time researcher spends on each position. If too much time has been spent, frustration starts to grow:

$$f_p = \sum_{p_o}^{L} f_p(a * p_o) \tag{21}$$

where p_o is the average year limit on position when the frustration starts to grow and a is the years in academia, and L is the leaving decision.

4.5 Promotions

How many researchers is there for each position, is dependent on percentage how many researchers are allowed for each position $\{p_{a1}, p_{a2}, p_{a3}, p_{a4}\}$ which are {Post-doc, assistant professor, associate prof, full professor}, respectively. If a researcher fulfills a set of demands, promotion is possible. These expectations that must hold are:

- Researcher can only get promoted one step at the time $P(r,t) \leq P(r,t-1) + 1$
- There is space on next position researcher is intending to get promoted $\hat{\imath}|P(\hat{\imath},t) = P(r,t-1) + 1 \leq c$
- Has been working a years in academia (a is set depending on position $n = 1,2,3,4$)
- Enough recognition through citations, meaning the following equation must hold: $C(r) \geq \left(\frac{\omega}{R*C}\right)$ where ω is constant, c_i citations of a researcher, C all citations of researchers together and R is the population size. Researchers are sorted by their publications, and the most published is the first one in the queue.

When promoted, the frustration lowers. Promotional frustration goes naturally to zero and monetary frustration halves since the assumption is that promotion temporarily satisfies the researcher's monetary needs. A book (Lyubomirsky 2014) argues that happiness is more a mindset, and not the circumstances. While this sounds more plausible theory, modelling the frustration for this kind of purpose is rather simplified hence only circumstances are considered as mood-changing factors.

4.6 Overhead

How much resources there are to distribute between researchers, depends on how much error in evaluation is tolerated. Higher overhead (resources for evaluation) makes the grant system more accurate allocating more resources to application evaluation, leaving fewer resources to share between researchers. These resources are directly converted for research time.

The grant-system time usage (Link, Swann and Bozeman 2008) is presented in table 4. The research states also that service time includes supervision of the grant applications. The assumption is that a third is used for this activity (except on post-doc position).

Mean fraction of time	Teaching	Research	Grant writing	Service	Total
Full professor	*30,00 %*	*35,15 %*	*7,30 %*	*27,55 %*	*100 %*
Associate	*36,40 %*	*32,90 %*	*7,60 %*	*23,10 %*	*100 %*
Assistant	*31,50 %*	*40,10 %*	*11,60 %*	*16,80 %*	*100 %*
Post-doc	*10,00 %*	*60,00 %*	*20,00 %*	*10,00 %*	*100 %*

Table 4: time usage of a researcher, with post-doc added

On average, 42.04% of the time is used for research. If all the time used for grants would be used for research, this would add research time 19.75% (average of grant writing and a third of service time). This means when no resources are allocated for overhead, research time is 65.79% (and added 3.31% administrative cost, see chapter 4.7.1). On contrary, when all the resources are allocated for overhead, the resources for research are 0. The net resources for research are reduced linearly as seen on figure 9. Error in evaluation however, is just an assumption error growth is a square of overhead, as seen also on figure 9. However the assumption is that error cannot reach zero, while getting closer. Error in evaluation is a standard deviation of normally distributed accuracy of quality of grant application (average being 0.5).

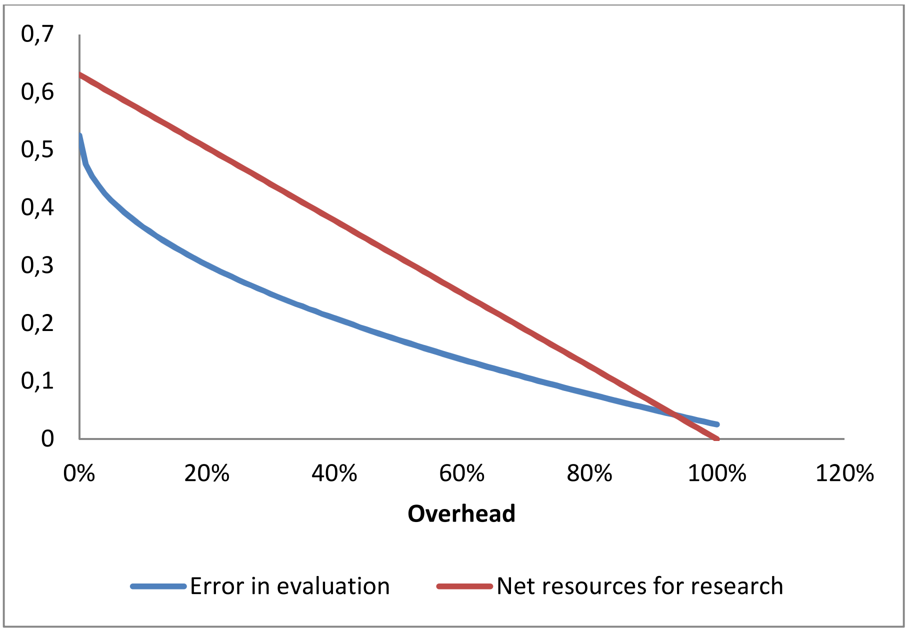

Figure 9: How the resources cause error in grant application ratings

4.6.1 Administrative cost

To make the system time usage more accurate, there is also an administrative cost from having the grant proposal handling process. Administrative staff is hired for handling the funding proposals and otherwise these resources that are allocated for administrative staff, could be also used to research. The basic idea is that the amount of resources (time) that are allocated for administration is also lowered when the time for grant writing is lowered and when resources assigned to grant process is zero percent, also this administrative cost can be fully assigned for research. Naturally, grants that are available through private sector are not in this calculation. So, it is needed to modify the information of previous graph into the direction where the grant time is systematically higher, and other time consumption is lower. To determine how much higher it should be as a reference, New Zealand's research quality evaluation system is used as an example (Tertiary Education Comission 2012). Under the review system there are 6 757 researchers. The average hourly cost of researcher is 51 and yearly hours per one researcher are 1950. PBRF runs every 6 years with total cost of 50 750 200 dollars. So, the cost for each researcher is 1251.80 dollars, or 24.5 hours. Rough calculations can be made:

PBRF money available every year	$262 000 000,00
PBRF total cost for 6 years	*$52 000 000,00*
PBRF total cost for one year	*$8 666 666,67*
Overhead % of PBRF	***3,31 %***

Table 5: Overhead caused by PBRF

From here it can be seen that there is 3.31% overhead of the resources that would otherwise go to research. This means that researchers would have 3.31% more resources in use without PBRF which would go straight to research without application process. This includes also the time spent on reviewing the proposals. In England, similar system, called RAE, where administrative cost of this is 0.8% (Koelman and Venniker 2011). This overhead is corresponding to research time; the less grant time, the less overhead. Eventually it can be assumed that when there is no grant system, all the overhead resources can be assigned for research time.

4.7 Quality of grant application

The quality of the grant application that a researcher produces is dependent how much time has been used for preparation, research & applying skill and the average time.

$$Q = \sqrt{T_r * s_r} * s_a \tag{22}$$

Where s_r is the research skill, s_a applying skill and T_r is the time used for applying.

The evaluated quality of the application depends on the actual quality Q and error e caused by overhead h.

$$Q_t = Q + e(h) \tag{23}$$

Figure 10 presents an example how the quality of application behaves when both skills are set to 1, without any interference.

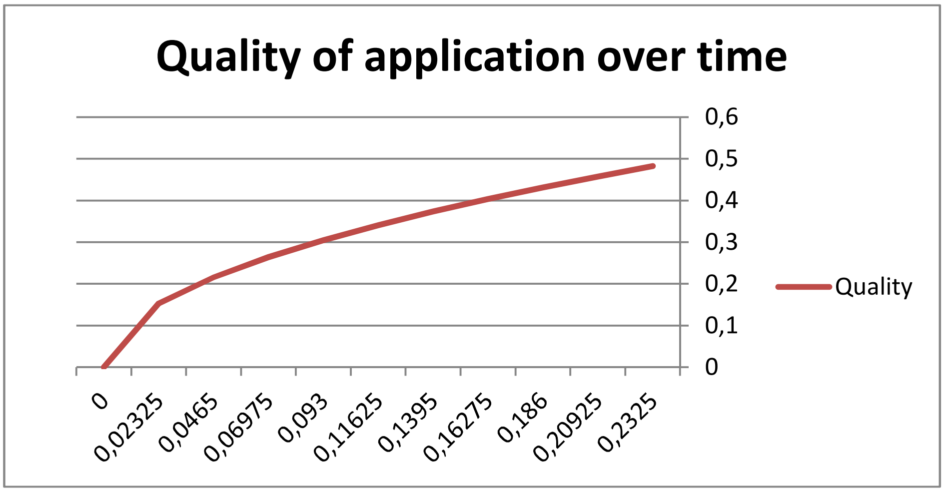

Figure 10: Example quality of application depending on the time

4.8 Citations

For a mechanism to count citations of a single paper p_j as a function of time the following formula (24) (Wang, Song and Barabási 2013), is presented for probability that paper *i* is cited at time *t* after publication as

$$\prod\nolimits_i (t) \sim \eta_i\, c_i^t P_i(t) \tag{24}$$

By using this formula, a separate submodel is created and adjusted, where η_i indicates fitness of the paper c_i^t indicates received citations so far, and $P_i(t)$ aging as a function of time with following formula:

$$P_i(t) = \frac{1}{\sqrt{2\pi}\sigma_i t} \exp\left(-\frac{(\ln t - \mu_i)^2}{2\sigma_i^2}\right) \tag{25}$$

The more simple form for the formula is:

$$c_i^t = m\left(e^{\frac{\beta\eta_i}{A}\Phi\left(\frac{\ln t - \mu_i}{\sigma_i}\right)} - 1\right) \tag{26}$$

Where $\Phi(x)$ is the cumulative normal distribution, m, β and A are global parameters and relative fitness is presented as $\lambda_i \equiv \frac{\beta\eta_i}{A}$. μ_i indicates *immediacy,* governing the time for a paper reach its citation peak and σ_i is *longevity*, capturing the decay rate. Also, time *t* comes from creation year subtracted from current year.

In this case, fitness of the publication λ_i, is fixed to researchers research skill s_r, which means they get the same value. Immediacy and longevity are fixed to value 1.0. Value t means the current point of advancing time, since the citations are not formed in one instance of time but over the time.

Since the citation production of the model is highly dependent on the parameter research skill, the value and distribution of research skill needs to be connected to fitness parameter so that the model (Formula 26) provides similar results as observed in literature (presented in chapter 5.3.6).

5 Implementation

Now that the model has been defined, it can be implemented into a computer simulation. This chapter presents used techniques and paradigms, and why they were chosen. In other words this chapter shows *how* the simulation is done. Simulation, as a program can be described using UML (Unified Modeling Language), which is a common way to visualize the system to be created. For simulation to be actually functional, it needs to be calibrated.

Typical simulation theory covers about different types of simulation paradigms and common concepts, which is different from one instance of simulation. There are a few model classifications which are presented briefly, but in this case the implementation follows time-driven discrete model, meaning that the created system advances by pre-defined time units Δt. The reason for choosing this type of approach is that data in literature often refers to activities as yearly basis rather measured events of time. Second reason for this kind of implementation is its simplicity; event-driven approach runs in multiple threads meaning concurrency, whereas time-driven implementation is far more simple approach after the occurrence order of events has been recognized.

The background that is needed begin with general knowledge of simulations, randomness, presentation how state and time can be seen in simulation paradigms, especially in dynamic discrete-event simulations. Static models are left out, as well as finite state-machines and Monte-Carlo methods are explained only briefly.

5.1 Background for simulation

Simulation is a representation of an existing, or non-existing system. Simulation can be used to describe and analyze the behavior of a real-world system, using a representation of the actual system. This can only be as large as the actual examined system itself (Banks, Introduction to simulation 1999). In Computer simulation the model is modeled on a computer so that its functionality can be studied. By making changes in variables, assumptions can be made about the system behavior (Banks, Carson, et al. 2010).

Simulation as a tool is usable when the original system is too big, expensive, dangerous, complex or impractical to examine in real-life, or to test a product design. Simulation is not reasonable tool when the real-life experimentation (or other alternative) would be cheaper, easier and safer. Also vital is to know the functionality of the original system fully, and all the data has to be available. Finally, the key is to have an actual reason to build a simulation (Becker and Parker 2011).

A particular finite state machine is defined by a list of its states, and the triggering condition for each transition. Finite-state machine is in only one state at the time. The state at any given time is called the current state. It can change from one state to another when initiated by a triggering event or condition; this is called a *transition* (Arbib 1969). Hence the system is not infinite but discrete. This movement through states within a machine is sometimes called a Markov chain.

5.1.1 Simulating a system

Simulation can either follow discrete, continuous or static model. In discrete model nothing relevant happens outside events, meaning no computations are needed between events. Continuous model follows time slavishly, and there are events in simulation when state does not change, excluding the timer. Static model holds no time aspect and is only a representation of created system of that moment of time. In discrete time model, state of elements can only change in specific moments of time, called events. State presents a condition of the simulated system at the given instant of time *t*. The events can be time-driven (at moment of time *t* occurs an event) or time can be event-driven (at the moment of occurring event Δt time has passed) (Page and Kreutzer 2005).

Typical discrete model contains **list processing**, where resource-waiting entities are in a queue of some sort (waiting for **activity**). The resource allocation could be done so that researcher-agents are in a queue and accessing resource-providing entity one by one, either in states receiving grants, waiting, blocked, failed or starving. Instead in the implementation the resources (or lack of resources) are assigned all to all of them at the same instance of time. Hence list processing exists in the implementation, but not as a queue. The duration of the activity is constant, with no **delays** which are caused by a certain combination of system conditions (Banks 1999).

5.1.2 Discrete-event simulation

Discrete event simulations can be divided into two types, terminating and steady-state. In terminating type measures to be observed are measured at fixed instants of time with finite simulation length. In steady-state measures to be observed depend on instants (or intervals) of time whose starting points are taken to be $t \to \infty$ (Page and Kreutzer 2005).

However, discrete event simulation can be divided into various paradigms. In activity-oriented paradigm jobs arrive at random times, and time units are broken into tiny increments. At each time point, the program scans any occurred events. The downside of activity-oriented paradigm is its slow execution, since there can be many time units without events. The event-oriented paradigm tackles this problem by going from event to event. This paradigm is not possible in models in which the states are continuous by nature. However the event-oriented approach is easy to implement since no thread-programming is needed and flexible since one event can, for example trigger two other events. It is fairly efficient on computation speed (Matloff 2008).

Process-oriented paradigm is currently dominating paradigm among the simulation community. The events are still there, but the simulation runs in various threads which are communicating which each other. For example, one thread would produce customers and other thread would execute the operation for these (Matloff 2008).

Discrete-event simulations can be either continuous (they run as long as user wants) or they can terminate at the certain point of time. The paradigm can be sequential (events occur in specified order) or parallel (such as process-oriented).

Generalization of Discrete-Event simulation process works as follows:

Start - Initialize Ending Condition to False, clock and system state variables

While End condition(s) are not met do following:

Advance the time to the next event of time

Do next event and remove from events list

Update statistics

End - Generate a statistical report

5.1.3 Markov chains

Sometimes, it is important to have not only the number of states but also their order. The next state is generally determined by the last few states. If it is assumed that the next state request depends only on the last state of the system, then the system follows a **Markov model**. For now, it is sufficient to understand that such models can be described by a *transition matrix*, which gives the probabilities of the next state given the current state. Transition matrices are used not only for resource transitions but also for application transitions (Jain 1991) .

More formally, Markov chain can be considered as a successful movement through a set of states $s_1, s_2, \dots s_n$, from s_i to state s_j with probability p_{ij}. These probabilities are visible in transition matrix (Kemeny, et al. 1959).

5.1.4 Randomness

A random number is a number which exact value is unknown, but it has an expected value. In simulations, generally pseudo-random numbers are used. The difference between random and pseudo-random numbers is that between *n* random numbers there is no connection (different seeds), where pseudo-random numbers only seem to be random. After repeating pseudo-random numbers infinite amount of times, a pattern will occur (numbers start to repeat themselves). Depending from the usage, most of the times this pseudo-random behavior is enough. (Becker and Parker 2011).

For example, a roll of a dice has no known value, but the expected value is an integer within range [1,6]. The probability for each value at the dice roll is ~16.66%, this type of probability distribution is called *uniform distribution*.

When the probability for some value to occur is higher than others, this is called *non-uniform distribution*. The most common this type of distribution is called *normal distribution* (Formula 27) or *Bell curve* because it's shape. This type of distribution is often seen in the nature.

$$f(x,\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{27}$$

Where μ is the mean (or expected value), x is the value that's probability of occurrence is wanted and σ is the standard deviation, which indicates the the variation from the average. Often in simulation purposes samples from the distribution are taken. In standard normal distribution ($\mu = 0.0$, $\sigma = 1.0$), the probability of x having a value within one standard deviation [-1,1] is 68.2%, standard deviation of 2 95.4% and so forth. These two type of distributions are used in simulation, plus one called *log-normal distribution* (Wolfram 2014). The shape is similar to normal distribution, but with skewed tail to one side, reduced from the other side.

5.1.5 Monte Carlo methods

When a certain Finite-state machine has been run, it will create a single path between the states (a Markov Chain). However, when the movement between the states varies, being stochastic, by *n* different runs different results are received, hence leading to different results. The purpose of Monte Carlo methods is to obtain a numerical value from a stochastic (random) system. The key is a repetition of the experiment: By repeating the experiment for the system *n* times, it is possible to receive a distribution of the *n* results. Naturally, larger number of repeats shows larger accuracy for the distribution. From this set of results is possible to calculate mean and other characteristics for an unknown probabilistic entity.

5.1.6 Microsimulation

Microsimulation models are computer models that operate at the level of the individual behavioral entity, such as a person, family, or firm (Bourguignon and Spadaro 2006). Dynamic microsimulations estimates changes over the life-cycle of the examined system. These models are common, for example, estimating pensions (Kärkkäinen 2013)

Microsimulation systems used in economics are nonlinear, complex, stochastic and period-oriented system which cannot be solved analytically but have to be solved numerically. They are applied in theory development and in decision support. Typically they are used in determining consequences of changes in taxation, investment decisions etc.

In microsimulation a representative sample of statistical units is transferred from period t to period $t + 1$. Microsimulations are often divided into static, dynamic and longitudinal microsimulations. The main difference between static and dynamic is that in static simulation the behavior of agents remains unchanged, whereas in dynamic simulation each agent (researcher, in this case) change their behavior depending on systems state. In longitudinal microsimulation not a representative sample is transferred from period t to period $t + 1$, but synthetic sample of micro units, that is a sample of a certain age cohort (Troitzsch, et al. 1996).

The common structure of microsimulations for redistribution analysis comprises three elements (Bourguignon and Spadaro 2006):

- a micro-dataset, containing the economic and socio-demographic <u>characteristics</u> of a sample of individuals or households
- the <u>rules</u> of the policies to be simulated – that is, the budget constraint facing each agent
- a theoretical <u>model</u> of the behavioral response of agent

The implemented model follows dynamic principles with simple built-in artificial intelligence. The researchers have certain attributes where they change their behavior, for example the frustration changes the applying time and motivation has an effect to paper production.

5.2 Implemented simulation

This chapter explains how the model is designed. The methods (mechanisms) and attributes are separated to classes they belong to.

It is fairly easy to define that used system is abstract since concrete model is not suitable to lie out. To run the system there are some stochastic elements and it can be considered numerical because it is possible, but not reasonable to write down one closed formula for behavior of the model. It is possible to present a finite-state machine for our system, where all the states of the system are presented. The implementation is a discrete-event simulation which advances in pre-defined steps of time. Simulation contains characteristics of microsimulation as well. To be more accurate, there are characteristics of dynamic microsimulation, as explained in previous chapter.

Figure 11 explains the class structure of the simulation. Class Simulation is called with assigned parameters, which calls the rest of the instances.

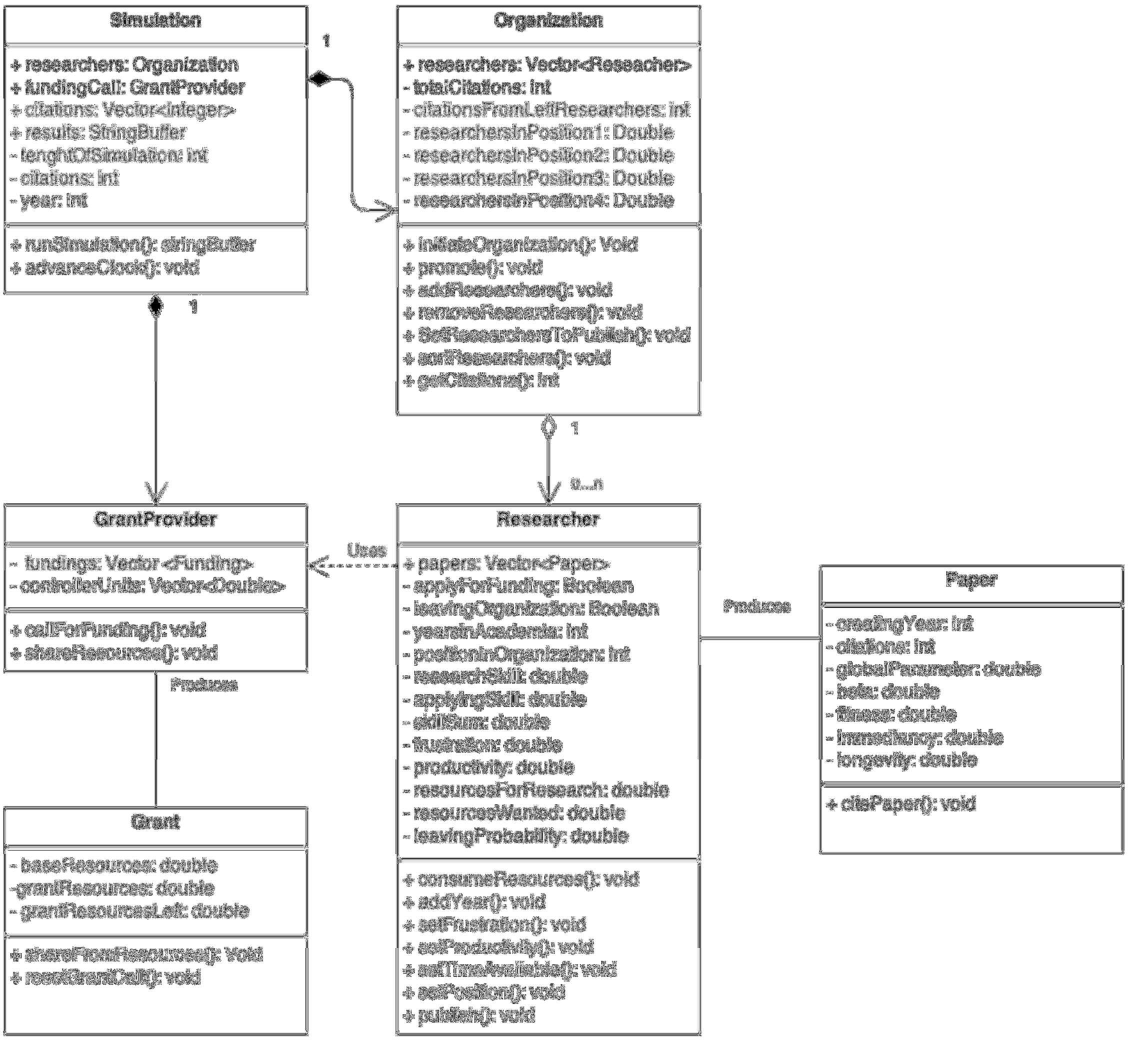

Figure 11: The class diagram of the simulation

Simulation classes function is to run the simulation with given parameters, advance the clock and finally return the results for the user.

GrantProvider-class provides the resources for the researchers and making the decision who is eligible for funding, as well as carries the information about user-assigned grants available (container for funding-classes).

Grant-class contains all the information about how much resources is available, and what kind of researcher has the possibility to attain those. Class funding has assistive class funding provider, which function is to combine assigned funding to research population with user-defined parameters. Class contains the checks whether researcher applies for funding, and tells researcher how much time they spend on average for applying.

In general, class GrantProvider contains mechanism for grants, while funding the attributes. This class receives researchers and provides the grants depending on the quality of the grant application.

Organization and researcher works in a same manner. Class Organization has a container where population is assigned to. Class has the information about needed levels of researchers attributes for promotion. Class Organization tells researchers to send grant applications and to react received grants, and researcher-class reacts in a defined way. Also the class has methods for sorting the researchers and a container for removed researcher's papers and citations.

Class researcher contains all the attributes of researcher, but also a container for publications of this researcher. All the information about researcher's current state is stored into its attributes. Figure 12 illustrates the interaction between objects.

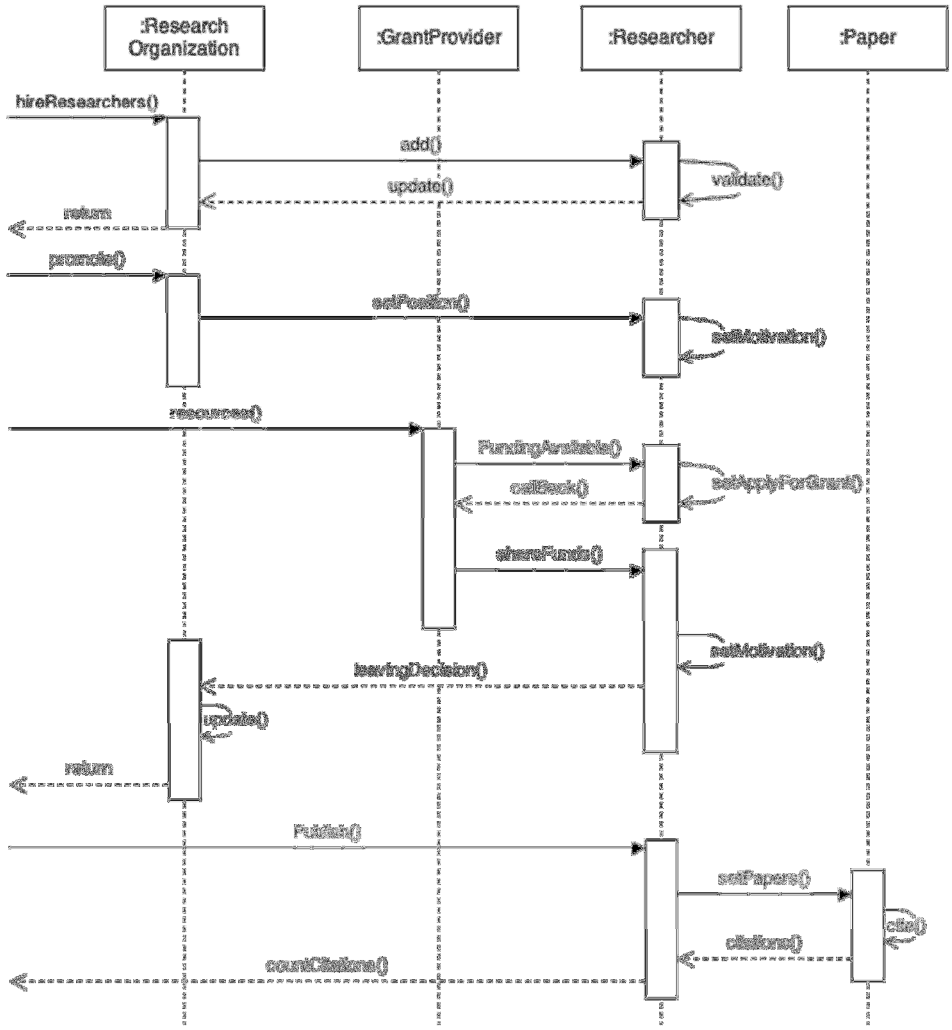

Figure 12: Interaction diagram of the system main loop

5.3 Calibration of the system

The simulated model needs to be examined from different angles to see whether it correlates with real world behavior. Instead of full simulation runs, the model should be verificated in "modules". (Ross 1997). In this case, modules are such as citation occurrence and population behavior. The real world questions can be such as "How many papers researcher produces every year", "what is the maximum amount of citations that a paper can receive," "how likely is it for a researcher to resign" or "How much resources average researcher receives each year". By using this kind of observations simulated system can be steered into direction that makes it more a presentation of the examined real world system.

Some experimental design needs to be performed here, such as sample sizes, population and observed timeframe, since the calibration needs to be done with functional system.

The difficulty of calibration is that everything is tied together meaning a change in population behavior can affect, for example, to publication behavior so calibration needs to be done multiple times from different corners of the system.

5.3.1 Examined timeframe of the simulation

There is no a pre-defined way to tell for how long the simulation should run. Usually it is a trade-off between accuracy and computing budget.

Besides the length of the timeframe of the simulation, the system needs to reach its *equilibrium* state in order to be examined. This means the samples of the running simulation system need to be taken from the point where the initial values of the simulation cannot be seen anymore. To determine when the system has left its transient state, and reached sufficient equilibrium, can be examined by eye, using statistical tests or taking subsequent samples and comparing them (Tiihonen 2014), (Chen 1995). Series of population size behavior taken in time window between 0 and 60 can be examined manually by eye (Figure 13).

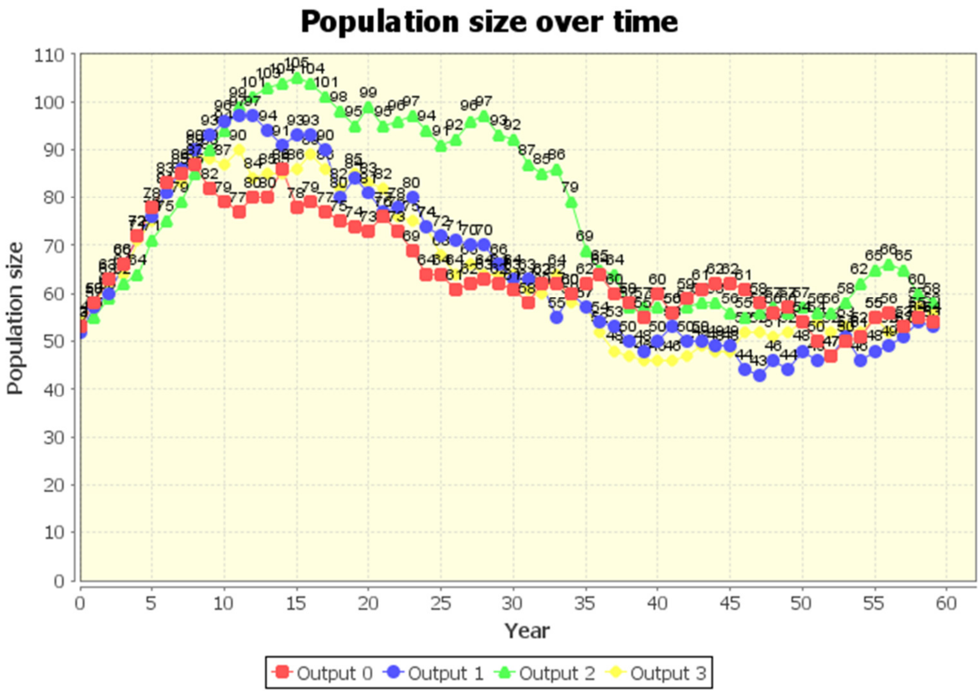

Figure 13: Four sample examples of system in transient state

In figure 13 a clear ramp-up period can be seen by eye from year 0 to 15. Decay towards steady-state can be seen from 15 to 40, and after that it is hard to tell by eye whether autocorrelation exists between samples. The manual examination is not sufficient, so instead evaluating results by eye, similarities between samples autocorrelation is calculated as a function of mean and variance, using the following formula:

$$\hat{\rho}_k \equiv \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{c\hat{o}v(R_{it}, R_{i,t-k})}{v\hat{a}r(R_{it})} \tag{28}$$

The observations are said to be independent if autocorrelation is zero. Instead of four sample series above, 10 sample series are taken from 40 to 190 time units. Using equation (28), autocorrelation of the samples can be calculated (Table 6). Autocorrelations are just presented for whole sample length, and just averages for reduced timeframes.

Samples of 40-240	1	2	3	4	5	6	7	8	9	10	
Variance	50,87	48,45	46,08	82,16	59,90	32,62	78,73	69,59	54,09	38,29	
Standard deviation	7,13	6,96	6,79	9,06	7,74	5,71	8,87	8,34	7,35	6,19	
Confidence interval	*1,01*	*0,99*	*0,97*	*1,29*	*1,10*	*0,81*	*1,26*	*1,19*	*1,05*	*0,88*	
Covarance 1 & n		-17,96	4,70	0,38	8,17	9,56	13,12	-16,74	-19,17	6,51	
Acorrelation 1 & n		**-0,36**	**0,10**	**0,01**	**0,15**	**0,23**	**0,21**	**-0,28**	**-0,37**	**0,15**	
Covariance 2 & n	-17,96		-0,20	20,11	-5,13	2,77	2,08	12,34	16,00	-10,79	
Acorrelation 2 & n	**-0,36**		**0,00**	**0,32**	**-0,10**	**0,07**	**0,03**	**0,21**	**0,31**	**-0,25**	
Covariance 3 & n	4,70	-0,20		9,30	-2,44	-7,66	16,42	9,58	-16,31	3,48	
Acorrelation 3 & n	**0,10**	**0,00**		**0,15**	**-0,05**	**-0,20**	**0,27**	**0,17**	**-0,33**	**0,08**	
Covariance 4 & n	0,38	20,11	9,30		-11,20	-9,55	2,75	19,67	-8,05	-18,82	
Acorrelation 4 & n	**0,01**	**0,32**	**0,15**		**-0,16**	**-0,18**	**0,03**	**0,26**	**-0,12**	**-0,34**	
Covariance 5 & n	8,17	-5,13	-2,44	-11,20		14,66	8,21	-10,54	-12,42	-6,49	
Acorrelation 5 & n	**0,15**	**-0,10**	**-0,05**	**-0,16**		**0,33**	**0,12**	**-0,16**	**-0,22**	**-0,14**	
Covariance 6 & n	9,56	2,77	-7,66	-9,55	14,66		0,71	-12,84	12,76	0,82	
Acorrelation 6 & n	**0,23**	**0,07**	**-0,20**	**-0,18**	**0,33**		**0,01**	**-0,27**	**0,30**	**0,02**	
Covariance 7 & n	13,12	2,08	16,42	2,75	8,21	0,71		-26,30	-4,94	-10,88	
Acorrelation 7 & n	**0,21**	**0,03**	**0,27**	**0,03**	**0,12**	**0,01**		**-0,36**	**-0,08**	**-0,20**	
Covariance 8 & n	-16,74	12,34	9,58	19,67	-10,54	-12,84	-26,30		-15,77	-4,10	
Acorrelation 8 & n	**-0,28**	**0,21**	**0,17**	**0,26**	**-0,16**	**-0,27**	**-0,36**		**-0,26**	**-0,08**	
Covariance 9 & n	-19,17	16,00	-16,31	-8,05	-12,42	12,76	-4,94	-15,77		8,41	
Acorrelation 9 & n	**-0,37**	**0,31**	**-0,33**	**-0,12**	**-0,22**	**0,30**	**-0,08**	**-0,26**		**0,18**	
Covariance 10 & n	6,51	-10,79	3,48	-18,82	-6,49	0,82	-10,88	-4,10	8,41		
Acorrelation 10 & n	**0,15**	**-0,25**	**0,08**	**-0,34**	**-0,14**	**0,02**	**-0,20**	**-0,08**	**0,18**		**Average**
Av. Acorr. 40-240	-0,02	0,03	0,02	0,00	-0,02	0,04	0,01	-0,08	-0,06	-0,06	***-0,017***
Av. Acorr. 40-90	-0,21	0,00	-0,05	-0,06	0,01	0,09	-0,10	0,00	0,02	-0,19	**-0,050**
Av. Acorr. 90-140	-0,04	0,02	-0,01	0,04	0,01	0,10	-0,15	0,03	-0,03	0,01	**-0,003**
Av. Acorr. 140-190	-0,05	0,10	-0,09	-0,03	0,14	0,12	0,09	-0,30	-0,08	-0,14	**-0,024**

Table 6: Autocorrelations of population samples

The average of calculated averages for the whole timeframe is -0.017 and the averages do not get any more accurate after 140 samples. This means the samples further does not provide more information than previous ones hence this sample length is sufficient.

Now that the length has been found sufficient, the size of the timeframe needs to be determined. This is done by choosing amount of citations as observed variable. By comparing two sample series, accuracy for timeframes of different sizes can be seen (Table 7).

Timeframe	*40-70*	*40-90*	*90-120*	*90-140*	*Difference in accuracy*	
Lenght	*30*	*50*	*30*	*50*	*30*	*50*
Av. of results	**32890**	**64121**	**33715**	**63848**	**2,45 %**	**0,43 %**
St. Dev	3040	5373	5673	4834		
Conf. Interval	1332	2355	2486	2119		

Table 7: Citations of different sample series

When taking averages of two subsequent sample series (average of 20 samples) of citations, timeframe of 30 time units gives accuracy of 2.47%. By making the timeframe even wider, to 50, the accuracy is 0.43%. Hence the timeframe with length of 50 is used for more accurate results.

The last phase is to determine whether for every sample point separate sample runs is more effective or taking one long sample and take subsamples from a single run (Figure 14). This single long run method is called batch means. If the distant segments have low autocorrelation, the usage of the separate samples is plausible (Page and Kreutzer 2005).

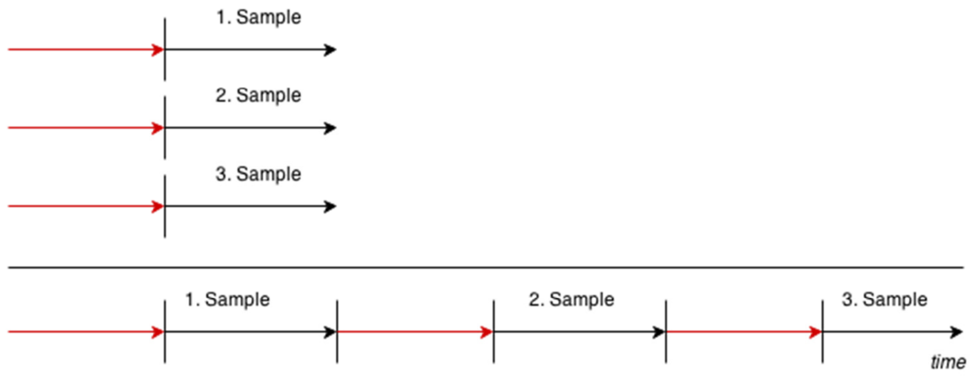

Figure 14: Sample points from multiple runs versus a single run

A ramp-up time of 90 time units is always present, but the time between subsequent samples is dependent on how autocorrelated the samples are after a certain period of time. A single run of 2000 time units is made, and the gap distance between subsequent samples is made (Table 8).

Distance	1	10	20	30	40	50
Average	0,97	0,61	0,35	0,24	0,02	-0,13
St. Dev	0,0076	0,0790	0,2201	0,1974	0,2958	0,2822

Table 8: Average of autocorrelation between samples

Average is approaching to zero at 40, meaning that this is sufficient gap between samples, when no significant autocorrelation is present. Hence it is more effective to take one long sample, discard the first 90 years because of the transient state, and take snapshots of 50 years between every 90 years until 31 samples have been gathered. This means run-length is

$$90 + 50 * 31 + 40 * 31 = \mathbf{2880}$$

5.3.2 Determining sample size

This chapter explains how and why a certain number of samples have been chosen to calculate a single sample point, and how many runs is enough. If standard deviation is known, a rough estimate can be given by previous simulation results and calculating average for them. Eighty runs are taken, and their average is 1495 work years, with standard deviation of 220. The average is always a subset of an infinite set. Wanted confidence interval is 95%, hence the area of a normal distribution is within 1.96 standard deviations of the mean.

Now, to calculate the sample size, following formula is used:

$$n = \left(\frac{Z_{\frac{\alpha}{2}}(\sigma)}{E}\right)^2 \tag{29}$$

Where

n = Sample size

$Z_{\frac{\alpha}{2}}$ = Confidence interval

E = Margin of error

σ = Standard deviation

More general equation for calculating 95% confidence interval goes as:

$$\bar{x} \pm 1.96\left(\frac{\sigma}{\sqrt{n}}\right) \tag{30}$$

where σ is the standard deviation and *n* is the sample size.

Now *n* can be calculated as follows, margin of error is set to 5% (Which is 74.80 from the average 1495)

$$n = \left(\frac{1.96 \times 220}{74.80}\right)^2 = 33.25 \approx 34$$

5.3.3 Determining population size

The initial values of the population size (how many researchers should there be in the beginning) is not relevant, but the system reaches equilibrium over the time through the researchers joining and leaving parameters. The system will pass the transient state faster when the initial population size is closer to the average of the transient state, which is around 50, depending on the funding.

5.3.4 Population behavior calibration

To see how the promotions, leaving and joining to and from organization functions the population behavior needs to be analyzed. It is expected that on average the population structure would remain the same over the time. However, since the researchers are set to go through the whole career path, the expectancy for structure is that on lower positions fill up first and less and less researchers are entitled for higher positions.

From the study (Shaw and Vaughan 2008) directional years in position spent can be seen in table 9.

Rank	Minimum	Maximum	Mean	Median	SD
Assistant professor	4	32	11.1	10	6.7
Associate professor	4	30	14.5	14.5	8.0
Full professor	12	42	24.9	24	6.6

Table 9: Population time spent in position, years

The initial setting is to set the population into positions as follows: {Post-doc = 15%, Assistant professor = 15%, Associate professor = 37%, full professor = 33%} (Canterbury University HR 2013). 31 samples of population are taken over 140 time units. The following table of means can be presented:

Position	Post-doc	Assistant prof.	Associate prof	Full professor	Sum
(Initial values)	*15,00*	*15,00*	*37,00*	*33,00*	
Mean	36,20	25,93	22,65	21,03	105,81
St. Deviation	7,29	3,46	3,53	2,77	
Conf. Interval	2,61	1,24	1,26	0,99	

Table 10: Population size sample averages

The same results, but visually:

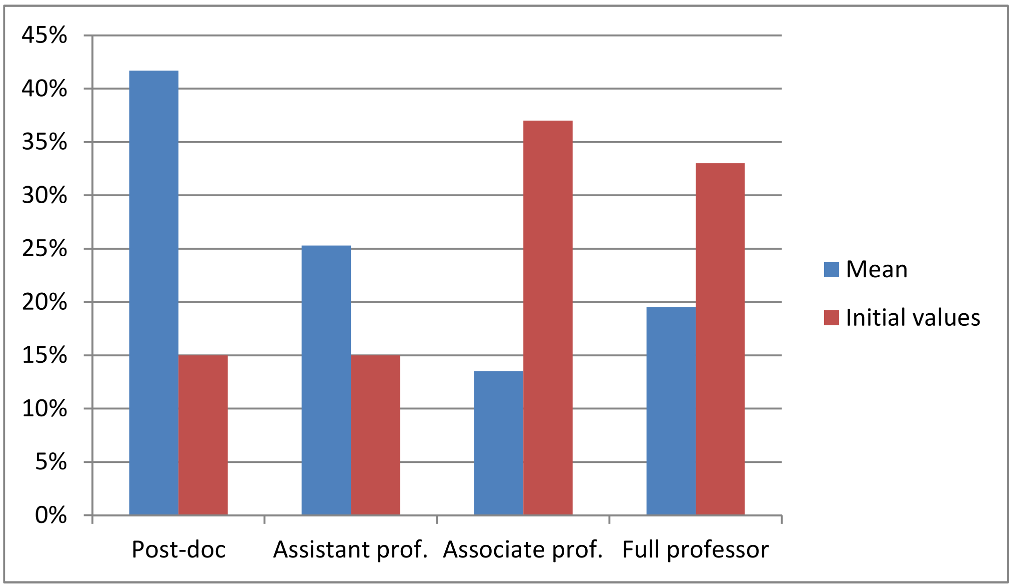

Figure 15: Distribution between positions, confidence interval of 95%

This was anticipated because researchers can only get into organization through bottom, meaning that every researcher goes through post-doc position and are not hired straight to other positions. The logic of the model and reality functions in different ways, so this is examined in conclusions-chapter. The research problem does not extend to population personnel structure, so the conclusion is that the researchers are flowing through organization vaguely as expected, and there are no great bottlenecks on the career path, even though at the last position holds researchers longer than on average.

5.3.5 Amount of resources to be allocated

The amount of resources which is available for researchers each year needs to be determined too. Since the success rates indicate that around a quarter of the researcher's gets the resources they ask for same value will be used (25%). A calculation how many resources 100 researchers will receive in order to fix this value can be shown: Expected resources for each position are set as follows, from post-doc to full professor, respectively: {1.3, 1.2, 1.1, 1.0}. The assumption behind this is that researchers at the beginning of their career (especially outside the tenure) are more dependent on the grants. Population distribution is more or less known (Figure 15), the population is expecting to receive around 115 units worth of resources each year to stay saturated and so a quarter would be 29 resources for each year.

5.3.6 Researchers and citations

First task is to test how well the model (Wang, Song and Barabási 2013) can imitate exceptional real-world data. Number of scientist in the world is around 5.8 million researchers (AAAS 2006). This many researchers are made and the highest research skill is picked. The highest value is ~5.4286.

The highest cited paper in the history of science is Protein Determination by Oliver H. Lowry with 275,669 citations so far. The paper was published in 1951 and count was in 2006. The paper still gets cited roughly around 5,000 times a year only with a slight decay rate (Kresge, Simoni and Hill 2005). This researcher, H. Lowry can represent the scientist who has the highest research skill (5.4286). Samples of citations by this paper are taken from Google Scholar by eye, (1974=5500, 1977=7500, 1983=8000, 1987=7800, and 2003=4500). By these samples a model is generated (figure 16) for the publication using the equation (20).When defining c_{max}^{t}, t=55, the time has set to t, and with cumulative citation count it should approach to that value. Fitness η_i is 5.4286, and c_{max}^{2006} = 275,669.

The model (Wang, Song and Barabási 2013) is fitted to the samples and the cumulative citation count should be close to 275 669 by the year 2006. The variable β is set to 1, as well as A. By varying m it eventually settles to 1055. The shape of the curve is dependent on μ_i and σ_i. Hence these are set to 3.5 and 0.622, respectively. The fit can be seen from the model figure 16 below.

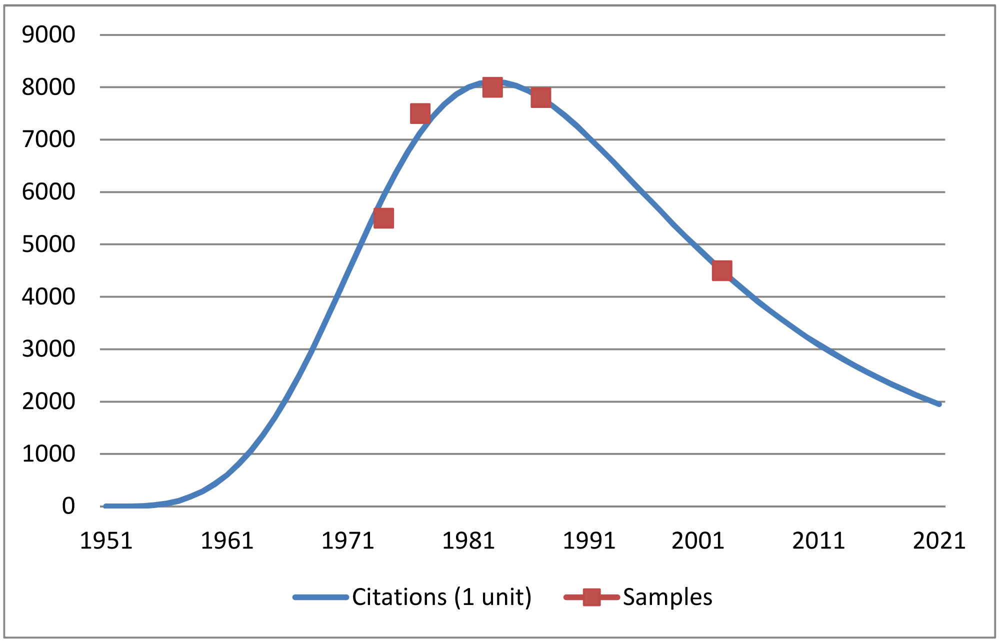

Figure 16: Citations of most cited paper. Note that cumulative is scaled by dividing 20

To perform the second check, if the model correlates with the cumulative citations by 2006 (275 669), the cumulative amount of the model produced in 2006 is 275 747, there is a 0,017% difference between the actual data and predicted data. Conclusion is that the model (Wang, Song and Barabási 2013) can imitate also not typical behavior in data accurately.

The ways researchers work gets cited needs also to be examined. Study (Kelchtermans and Veugelers 2005) states that:

“About 25% achieves top performance at least once, while 5% is persistently top. Analyzing the hazard to first and subsequent top performance shows strong support for an accumulative process. Rank, gender, hierarchical position and past performance are highly significant explanatory factors.”

The quality of researchers work differs greatly in real-life. The explanatory factors in current model are the variables of researchers. Average citations of one paper are 7.9 with standard deviation of 23.4, which is really dependent on the field. For example, average citations in mathematics per paper is 3.32, and in clinical medicine, which is the largest area where papers are published (around 2 million papers), the same value is 12.46 (Reuters 2010). In physics 783.339 papers were cited 6 716 198 times between years 1981-1997. From these 47% were not cited. From the cited papers, 415 229 were cited at least once, 265 281 at least 10 times, over 200 papers were cited 2103 times and 282 over 500 times (ISI 1997), (Redner 1998).

Similar behavior should be found in the simulated (Wang, Song and Barabási 2013) model implementation. From randomly generated sample of 25 936 papers, 32% were not cited whereas from cited papers 17 608 were cited at least once, 8 334 at least 10 times, over 200 papers were cited 26 times and one over 500 times. These values transformed to percentages, table 11 can be presented.

Cited at least	once	10 times	200 times	500 times
Literature	60,804 %	38,847 %	0,308 %	0,041 %
Simulated	67,804 %	32,092 %	0,100 %	0,004 %

Table 11: Effects of factors to system

The created model is more likely to produce papers with at least one citation but higher citation counts are not reached as often, meaning that the model simulates the system quite accurately, but highest peaks occur rarely.

Finally, the simulated system needs to function so that there is a difference between a talented and non-talented researcher.

1935 simulated sample researchers are examined, and the expectation is that the talented researchers will provide higher results, where the amount of citations grows faster than the amount of papers. Table 12 presents the produced papers and citations, depending on research skill.

Research skill		**0.0-0.5**	**Research skill**		**0.5-1.0**
Researchers	Papers	Citations	Researchers	Papers	Citations
505	22	0	517	26	17
Research skill		**1.0-1.5**	**Research skill**		**1.5-2.0**
Researchers	Papers	Citations	Researchers	Papers	Citations
444	56	296	326	56	821
Research skill		**2.0-2.5**	**Research skill**		**2.5-3.3**
Researchers	Papers	Citations	Researchers	Papers	Citations
113	61	2165	24	54	6025

Table 12: Comparison of output averages depending on research skill

The non-talented researchers do not publish almost anything. Partially, the higher outputs on the higher end of research skill are explained though the longer publication time. Talented researcher is more likely to get more recognition through the work, hence staying more satisfied whereas talentless researcher will leave the organization in short period of time. Still, the major difference is clearly depending on research skill. Also note that the last sub-table reaches research skill of 3.3.

6 Experiments

In this chapter the actual simulation experiments and the results are presented. This means that concrete functions, constant, scenarios, lengths, number of runs are also presented.

6.1 Background for experimental design

Reason for experiment is collection of data by the means of process or study, where the results are unknown in advance. Experimental design is a statistical technique for improving the efficiency and effectiveness of experiments with systems. Statistical experiments are conducted in situations where the observer can manipulate the factors and conditions of the experiment. In experimental design we purposefully make changes in the input in order to learn about the output (Kleijnen 1986) (SAS Institute 2005).

Experimental design is the process of planning a study to meet specified objectives. Planning an experiment properly is very important in order to ensure that the right type of data and a sufficient sample size and power are available to answer the research questions of interest as clearly and efficiently as possible (SAS Institute 2005).

The goal of a proper experimental design is to obtain the maximum information with the minimum number of experiments. This saves considerable labor that would have been spent gathering data. A proper analysis of experiments also helps in separating out the effects of various factors that might affect the performance. Also, it allows determining if a factor has a significant effect or if the observed difference is simply due to random variations caused by measurement errors and parameters that were not controlled. Typical approach is to divide the experiment in two phases. First phase has a large number of factors but small levels. This way the significant factors can be recognized. For second phase the factors with higher impact are examined with higher levels (Jain 1991).

The basic concepts are factor and level. the input of the simulation program corresponds to the factors of the experiment; that is, the factors are parameters, variables and behavioral relationships of the simulation model that are changed during the experiment with the simulation model. The factors in experimental design can be either quantitative or qualitative. For example quantitative can be how much resources are available. Qualitative, on the other hand, can be how the researchers are sorted for the resource sharing (Kleijnen 1986) Before doing the actual simulation runs, "2^k-design" evaluates the factors to which the system demonstrates sensitivity. The design also shows the covariance's for the factors that are chosen to observe. Because the research questions are about the resource allocation, amount of overhead and base funding needs to be examined, as well as the effectiveness of the proposal process. The frustration factors affect to the population size, which affects to amount of papers, and to citations. This way the frustration also is wanted to be examined. Similar factor, motivation could be also taken under examination, but the population size is more important factor since only limited amount can be chosen. Also because of the limited amount of factors, the joining to organization does not need to be examined since the overall effect of changes in organization can be seen through frustration which affects directly to population size.

To measure and find correlation between explainable factor *y* and chosen primary factors $x_1, \ldots, x_k$, a *regression model* needs to be done for the experiment results. By starting with a few very high and low levels of primary factors, a linear regression model can be found. By doing more experiments from the middle, it is yet to be seen whether the linear model explains the behavior of the system. If not, a non-linear fit for results needs to be found.

The problem in experimental design is the exponential growth in the number of combinations, explained by this simple example: Let's say that we have two (k=2) factors. Other one is "resources available" which has 8 levels (1...8) and other one is noise level which has 10 levels (10%...100%). Hence these two factors provide 80 different kinds of combinations for measurement. In current experiment the amount of factors for just researcher entity is over 10. Even with their minimum number of levels, 2, this leads to 2^{10} = 1 024 combinations to measure.

The design of experiment is distinguished to three approaches:

- One factor at a time approach
- All factor-level combinations: full factorial design
- Specially selected combinations, 2^{k-p} factorial designs
- When using 2^{k-p} design,

In this experimental design attributes have two levels, so for example for factor "resources available" the levels are $\{1,8\}$. Having two levels is enough to examine how much factor affects to whole system if the levels are set right, and what factors affect the most to the system together. The covariances (or interactions) are exposed when two (or more) factors are examined.

When using 2^{k-p} design, the problem is to choose the factors for experimental design (and which factors can be turned into constants). The factors that have most effect to the system are potential candidates as independent factors for the actual experiment, or at least the effect of these factors needs to be explained. Even if a factor is known to have a great effect for the output, it can be left out from the experimental design. For example, "resources available" factor is known to have a large impact to whole system (shown later on in the 2^k experiment) but if it is exactly wanted to measure how the different resource allocation affects to the output, it is justified to set a fixed value for that factor.

6.2 Regression analysis

When modeling, dependent factor y depends on independent factors $x_1, \ldots, x_k$ through known or unknown function φ:

$$y = \varphi(x_1, \ldots, x_k) \tag{31}$$

Where y is called as response, $(x_1, \ldots, x_k)$ are called response factors. Values of response factors are called levels. Function φ is usually so complex (or unknown) so that it cannot be used as it is. Hence we use empirical model

$$y = f(x1, \ldots, xk) + \epsilon \tag{32}$$

Where function f is an approximation of function φ and ϵ is a random variable.

Now all that is sorted, we can start doing experiments and create data, where each experiment is denoted by its number. The amount of experiments needed is explained more accurately in experimental design section.

Experiment	Levels of factors	response
1	$x_{11}, \ldots, x_{1k}$	y_1
2	$x_{21}, \ldots, x_{2k}$	y_2
…	…	…
N	$x_{N1}, \ldots, x_{Nk}$	y_N

The real examined system can be defined formally as a function

$$y = f_0(V_1, V_2, \ldots, V_n) \tag{33}$$

Where f_0 is unknown and and parameters V_i are only known partially. Simulated model can be defined as follows:

$$y = f_1(Z_1, Z_2, \ldots, Z_n, R_0) \tag{34}$$

Where f_1 is known implicitically and R_0 is the seed value of the randomness. The meta-model can formally be described as follows:

$$y = \sum_{i=0}^{q-1} \beta_i x_i + \epsilon \tag{35}$$

Where x_i are known functions from Z_j (such as powers, logarithms). These kinds of models are generally called as regression models.

6.3 Least-Square fitting methods

The purpose of fitting methods is to find a mathematical equation do describe the collected data. The most commonly used measure of how well an equation describes a set of data is the weighted sum-of-squared-residuals (WSSR, which is proportional to the sample variance-of-fit,s^2)

$$WSSR = \sum_i \frac{(Y_i - Y(X_i))^2}{\sigma_i{}^2} = \sum_i R_i^2 \tag{36}$$

$$s^2 = \frac{WSSR}{NDF} = \frac{WSSR}{n - n_f} \tag{37}$$

Where NDF is the number of degrees of freedom which is normally evaluated as the number of data points, n, minus the number of parameters being estimated, n_f (Johnson 2008). The least square method is used, to find one fit, or equation for the measured results.

6.4 Factorial design

This chapter presents which factors are varied for the experiment. Large amount of factors is caused by different positions in organization, for example, frustration is merged with different factors for each position. To reduce combinations even more, some averages between positions or alternative methods can be used, san average for all positions is used instead different requirements for each position.

Table 13 below demonstrates the factors for the 2^{k-p} design. The environmental factors are examined in the actual experiment, hence these two can be left out since after all, 2^{k-p}-analysis is a stripped version of the actual experiment.

X	**Factors that can be modified**
A	*Overhead*
B	*% of base funding*
C	*Promotional frustration (average)*
D	*Monetary frustration (average)*
E	*Effectiveness of proposal process*
F	*Researchers joining to organization*

Table 13: Factors for simulation

With six factors, full 2^6 (k = 6) factorial design would need $2^6 = 64$ experiments. This being too heavy, selected fractional 2^{k-p} design is used to get a big picture of the system functionality. To be more accurate, 2^{7-4} design is used, meaning only $2^3 = 8$ experiments. However, design has covariances in it, so to examine that there are no joint effects the same design needs to be done twice, but with inversed levels of factors (table 14). Since in the experiment set there is space for one more factor, the covariance of A & B factors is added.

Two levels are set for each factor as follows (All the effects are set to +1 or -1 so, that the higher result expectation is +1):

$$\mathrm{A} = \begin{cases} 1 \text{ if overhead is } 25\% \\ -1 \text{ if overhead is } 75\% \end{cases}$$

$$B = \begin{cases} 1 \text{ if base funding (equally distributed)} is \ 25\% \\ -1 \text{ if base funding (equally distributed)} is \ 75\% \end{cases}$$

$$C = \begin{cases} 1 \text{ if promotional frustration grows 0.08 on average} \\ -1 \text{ if promotional frustration grows 0.3 on average} \end{cases}$$

$$D = \begin{cases} 1 \text{ if monetary frustration grows 0.1 on average} \\ -1 \text{ if monetary frustration grows 0.3 on average} \end{cases}$$

$$E = \begin{cases} 1 \text{ efectiveness of proposal process is 0.3} \\ -1 \text{ if efectiveness of proposal process is 0.9} \end{cases}$$

$$F = \begin{cases} 1 \text{ if researchers joining is 10\% of the population} \\ -1 \text{ if researchers joining is 7.6\% of the population} \end{cases}$$

The simulations are made, and the following table 14 can be presented. The most important information can be found from the last row, effect of factor to Sum of Squares Total.

Treatment	I	A	B	C	D	E	F	AB	Y
	1	-1	-1	-1	1	1	1	1	44020,12
	1	1	-1	-1	-1	-1	1	-1	41269,06
	1	-1	1	-1	-1	1	-1	-1	39309,61
	1	1	1	-1	1	-1	-1	1	40928,24
	1	-1	-1	1	1	-1	-1	1	44397,94
	1	1	-1	1	-1	1	-1	-1	48824,68
	1	-1	1	1	-1	-1	1	-1	39946,32
	1	1	1	1	1	1	1	1	72298,94
	1	1	1	1	-1	-1	-1	1	45648,76
	1	-1	1	1	1	1	-1	-1	48167,63
	1	1	-1	1	1	-1	1	-1	60671,85
	1	-1	-1	1	-1	1	1	1	38586,14
	1	1	1	-1	-1	1	1	1	53916,85
	1	-1	1	-1	1	-1	1	-1	45625,57
	1	1	-1	-1	1	1	-1	-1	46592,71
	1	-1	-1	-1	-1	-1	-1	1	35376,96
Sum of Y	745581,4	74720,798	26102,455	51503,141	59824,616	37851,97	47088,333	4766,5354	Total
Sum of Y/16	46598,84	4670,0499	1631,4034	3218,9463	3739,0385	2365,7481	2943,0208	297,90846	Total/16
Sum of Squares Total		1,011E+09							
factor		A	B	C	D	E	F	G	
Portions of SST		348949853	42583633	165785848	223686544	89548226	138581946	1419991	
Effect of factor to SST		34,5%	4,2%	16,4%	22,1%	8,9%	13,7%	0,1%	100,0%
		35 %	4 %	16 %	22 %	9 %	14 %	0 %	

Table 14: Effects of factors to system

Conclusion:

The standard deviations for each factor vary from 2 865 to 5 657, so besides the last two factors the difference can be considered significant. It is obvious that A needs to be examined. However factor B, how equally resources are distributed between researchers, has very low significance, but according to the original research question this factor should also be taken into account. This is problematical since factor B cannot really be analyzed before the rest of the system is more accurate. The more profound examination of factors, especially B is presented in discussions-chapter. The assumption also is that the frustrations (C and D) correlates with the amount researchers are joining to organization (F). Both of the frustration factors are quite significant (explaining 38.5% of the result) so further research about reasons for leaving would be more than welcome. These three values, C, D and F are fixed depending on the behavior of the coming and going researchers (Figure 15) so that the population size varies in a plausible manner. How effective grant applying process (E) is does not affect to results significantly, but it has covariance with B because if base funding is set to 1.0, proposal process becomes obsolete.

The following table can now be presented:

A	B	C	D	E	F	G
Vary	Vary	0.15	0.2	0.6	8.8%	35

Table 15: Levels of factors for experiment

7 Results

This chapter presents background for statistical analysis and provides results as a regression surface model.

7.1.1 Response surface model

Response surface model explains relationships between explanatory variables against one response variable. Response surface models are normal linear regression models regarding to its parameters, whose parameters can be estimated using least-square fitting methods.

7.1.2 Linear and non-linear regression model

In statistics, linear regression is modeling the relationship between dependent variable y and one or more independent variables x_n Suppose that some variable of interest, y, is 'driven by' some other variable x. Then call y the dependent variable and x the independent variable. In addition, suppose that the relationship between y and x is basically linear, but is inexact: besides its determination by *x, y* has a random component, *e*, which can be called the 'disturbance' or 'error' (Cottrell 2011). Now the linear regression model for one independent variable can be written as follows:

$$y = \beta_0 + \beta_1 x_1 + e \tag{38}$$

Where the parameters β_0 and β_1 represent the y intercept and the slope of the relationship, respectively (Cottrell 2011).

The regression model for two independent variables can be expressed as:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \mu_i \tag{39}$$

And so forth, for the higher level variables regression models can be written using the same pattern.

7.2 Response surface model

In this subchapter, two types of results are presented. A big picture is first presented about the behavior of the system, and a smaller window is taken into further examination. 121 samples are taken with 100 runs for each sample point. The result is figure 17 kind of surface.

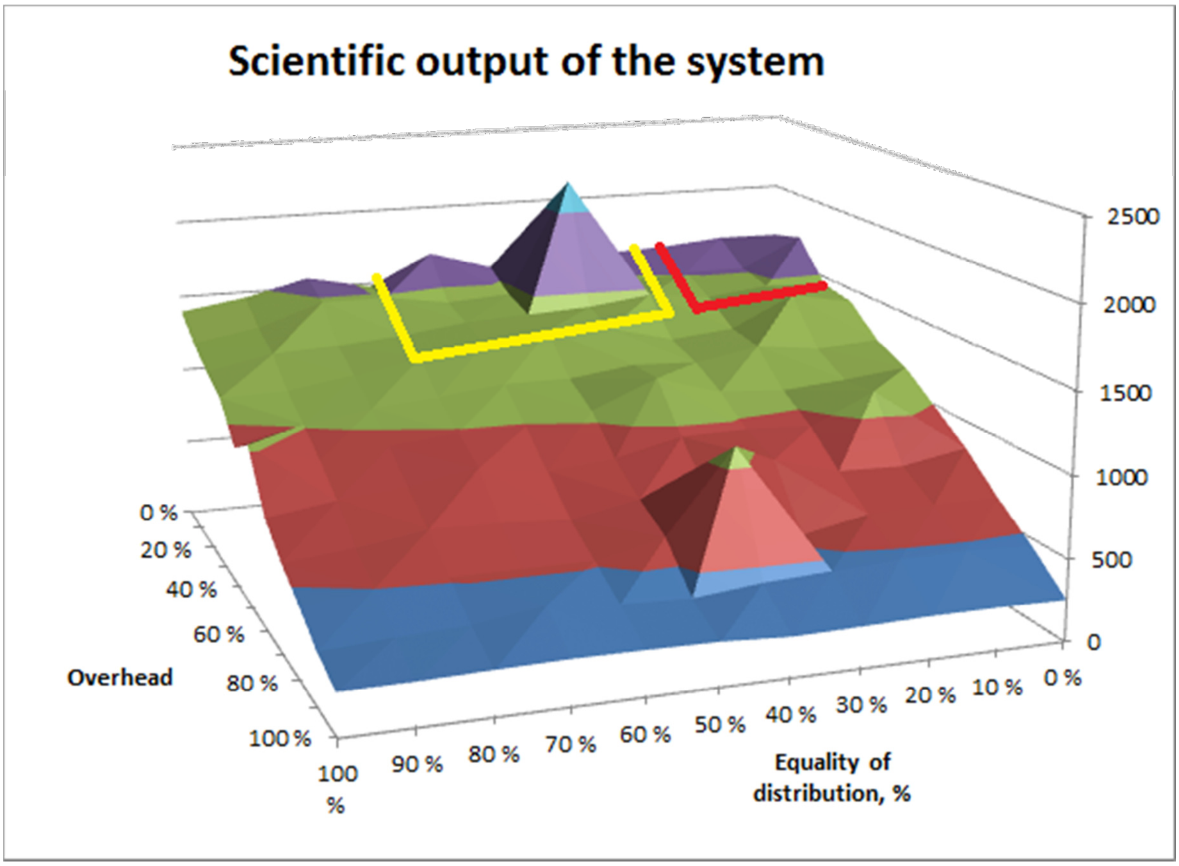

Figure 17: Measured output of the system of 11x11 data points

As expected, overhead affects to the results the most. Now, the area that we are particularly interested in is marked with red (Figure 17). This area needs to be examined, with higher amount of sample rates since clearly the system has a tendency to produce outliers (two “pyramids” at 40% of equality in distribution, as seen in figure 17), meaning peaks whose would even out when sample rate would be set to infinity. The equality of distribution does seem to have a little effect, but according to these results it seems that when resources are concentrated for the few, the results are slightly better.

The area marked with yellow in figure 17 also needs to be examined more carefully in case of outliers. Further examination reveals that indeed lower results are produced within the yellow area, in comparison to red area.

More samples are taken from 20% x 20% area, with 5000 years and 32 runs per sample point, following table 16 can be presented.

Ovehead	**0.0**	**0.05**	**0.10**	**0.15**	**0.2**
0.0	1642,61	1603,79	1567,73	1508,68	1451,74
0.05	1622,79	1597,81	1553,24	1531,70	1460,06
0.10	1612,80	1580,57	1556,72	1509,18	1453,26
0.15	1622,41	1586,98	1528,62	1484,27	1449,33
0.20	1608,22	1582,35	1548,21	1468,64	1433,82
Distribution equality					

Table 16: Effects of factors to system, citations per year

In order to know which measurements are plausible, the standard deviations need also be present (table 17). Generally, the darker red the color, the more accurate the results are (less outliers). The following table reveals that the most weighted references for response surface model are in area overhead 0.2, as well as 0.0 area gives quite good reference.

Standard deviations for measurements				
91,05	56,51	53,32	85,47	50,07
66,50	40,98	45,90	116,84	50,30
82,14	37,68	92,05	72,77	74,95
151,80	70,97	45,26	55,73	45,69
83,07	59,09	70,79	35,63	49,71

Table 17: Standard deviations for table 16 values

From these samples following regression model can be presented:

$$y = 1\,643 - x_1 172 - x_2 954$$

where x_1 is the equality in distribution, $\{0..20\%\}$ and x_2 is the overhead $\{0..20\%\}$. This equation is examined within observed area 0-20% for both factors. The fit of the model is not tested outside the observed area. In comparison to research questions, it seems that it is better to withdraw the resources from administration and use them for research instead. Figure 18 presents the surface visually.

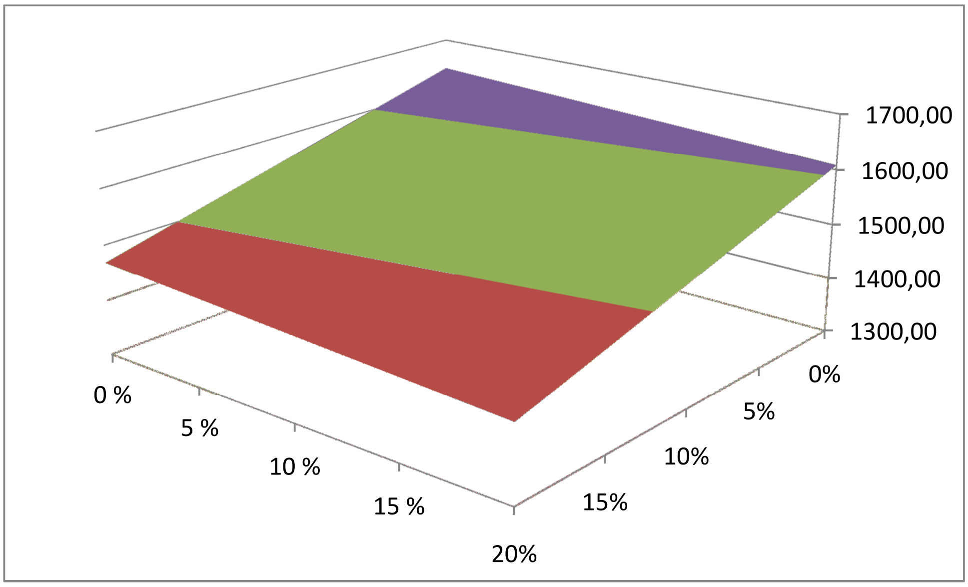

Figure 18: Response surface model 20% x 20% area

7.2.1 Model validation

Confidence intervals of 95% are calculated using the standard deviations on table 17 and equation 29.

Overhead	0 %	5 %	10 %	15 %	20 %
0 %	28,22	17,51	16,52	32,85	15,52
5 %	20,61	12,70	14,23	44,91	15,59
10 %	25,45	11,68	28,52	27,97	23,23
15 %	47,04	21,99	14,03	21,42	14,16
20 %	25,74	18,31	20,41	13,70	15,40
Eq. Distribution					

Table 18: Confidence intervals for measured samples

Almost all (92%) of the measured samples are within the confidence interval. In other words, there is a 95% probability that the model predicts the measured value. For example, when x_1 and x_2 are set to 10% in regression model, result is 1537.84 which is within confidence interval 1556.72 ± 28.52.

8 Discussion

The overhead is a radical parameter, since all the factor levels at least over 50% are highly theoretical since in the current system resources used for grant applications and their administration is around 20% (even though studies indicate a rise on this factor). Hence this parameter should not be examined on the very high end (chapter 4.6). It is enough to state that the simulation functions as it should, causing lowest results when no resources are allocated for research at all (100% to overhead) and the resource distribution does not have effect since there are no resources to distribute.

The overall trend seems to be that the less surveillance and more concentrated the resources are, as seen on figure 18, the better is the citation output. The equality in distribution does not have a lot of weight on results, but the amount of resources is the dominating factor. The results indicate that they they're not very dependent on which researcher is receiving the resources, hence the conclusion can be drawn that the simulated system does not make a broad difference between skilled and unskilled researcher, as long as the researcher has time for the actual research. Even when some tests were made where the effectiveness of the funding process (*E*) was varied radically, equality in distribution was zero and the resources remained constant, only slight differences (~23%) were observed. A raw interpretation of the system (table 14) variables is that not much conclusion can be drawn from factor funding distribution (B) before the rest of the observed variables are examined more carefully since they have significantly larger weight in the system.

It seems to be that no matter how much resources are there, it is not effective to divide the resources equally, but when concentrated even for researchers of lower quality they produce better results with satisfaction. This could be a consequence of the system creating a tail for citations, since they occur over longer period of time. This means that even if the paper would be good quality, if the researcher frustrates too soon the effect of the good paper cannot be seen but it is taken away with the researcher. This hypothesis can be tested with setting the parameters immediacy and longevity (equation 26) so that citations are formed on one instance. If the regression model is interpreted excessively, 20% change in overhead changes the results as much as 100% change in equality of distribution.

The current results are showing evidence of too heavy resource usage for grant-review process (less administration provides better results). The administrative cost plays small role (as seen on table 4) but the time used for applying being around 20% indicates that indeed too much time is used for writing and reviewing the applications. There is no distinguish between these two, so further question could be which one is more time consuming: is it the reviewing or too many and too complete applications researchers need to write.

The other question is whether the results are credible. Clearly the debate around grant providing did not just came out of thin air, but the system is indeed failing on some level. Any kind of skewing and leading to conclusions has been tried to avoid during this thesis, and the results could be condensed into sentence “capitalistic system with no surveillance who actually deserves the resources provides highest output”. The plausibility can be argued, but I didn’t expect this result. Of course, it is common in this kind of experiments that phenomena’s that would lead the results into different direction are not examined so further research or experiments are indeed necessary.

9 Conclusions

So is simulation generally a good method for this type of problem-solving? It has seen that system's behavior with individual agents can be predicted. Good examples being validated man-made models about animal population (rabbit population growth, predator-prey models) or spreading of a virus. This kind of experiments would take decades (or too expensive) to see the results in real-life so clearly social simulations can be used as a tool into which direction decisions should be made. The problem of social simulations is that they try to quantize human behavior with mathematical means, often in too simple way since humans and our actions are difficult to predict, at least on individual scale. However, when enough samples are taken the majority's behavior can be predicted.

The validity of the created model is debatable. The model's purpose is to capture research organizations capability to produce scientific output. Beginning from the scientific output, by having already citation production model (Wang, Song and Barabási 2013) the system can imitate produced citations quite well. Also, the amount of produced papers is rather easy to capture, even though this is highly dependent on the examined field of research (chapters 5.3.6 and 5.3.7). Chapter 5.3.6 also presents the correlation between scientific output and research skill.

The real challenge is to interpret the researcher's citation and paper producing features. Even though the model captures the citation and paper occurrence, which personal characteristics cause researcher to be successful is yet to be examined properly. The characteristics are recognized, but their levels as attributes yet remain unknown (chapter 3.3.3). There might be also non-personal attributes which would be dependent on the environment, not the researcher.

The motivational attributes, (frustration and motivation) and how they extend to researchers capability to produce research are rather hypothetical. Starting from frustration, the consequences of not received grants need to be examined more implicitly, to be more accurate, how this causes. As seen in table 14, the frustration factors together causes almost 40% effect to results. This kind of effect clearly needs to be examined more carefully since this can change the results completely. Good approach would be to examine the researchers who do not get the funding subsequently, and have quantified information about leaving the research organization.

The literature (Link, Swann and Bozeman 2008) presents quite well the time usage for research, reviewing and evaluating papers, and how much there are resources spent in administration. The challenge is to define how all these factors affect to the quality of research. Now the assumption is that the time just affects to the amount of done research, which is rather hypothetical. It would also worth to do further research about how the actual grant application writing (amount of time) and motivation helps with the quality of the application. Also now it is just assumed that the current grant-system fails 30% of the time, meaning gives resources to wrong researchers. Similar assumptions were made about how the current funds are actually distributed. The assumption was that around 40% gets the resources and the system is calibrated by using this value.

The modeling of peer-review could be done for getting a more broad view (Paolucci and Grimaldo 2014), after examining more vital mechanisms, such as how the received grant affects to scientific productivity (Jacob and Lefgren 2011), this unimplemented model, when functional, could increase the trustworthiness of the results drastically.

Moving to research organization, the measurement how long the researchers stayed in organization, on which position and what was the reason for leaving could be examined. This should be examined together with frustration to get the idea which is a consequence from frustration and what is affected by other reasons (other better positions etc.). This improvement, with simulation modified so that the researchers could join also to all positions would improve the predictability power of the model. Information about researcher population size should be routine-kind of information for universities HR departments which they follow as a routine, so the data should exist.

In reality the research is often not done by one person, but a group. Usually one researcher (professor) leads a handful of PhD-students and post-docs and this professor is the person who applies for the resources, at the end of the day. The dynamic organization does not present functionally real-life model, but this model is better than just static amount of seats available. The amount of seats changes, but really talented researchers can work their way up in the organization, meaning that positions are made for them if necessary. A small change in researchers joining (table 14) can make a significant change in results, since the assumption is that researchers struggle in organization a few years even without external funding and causing some results during this short existence.

Finally, it is needed to state that grant provider and the system around it needs to be evaluated more carefully. The model should be at least extended to direction where different types of grants would be available. The current system just feeds the resources to organization, without any judgments about the actual needs of the researchers.

Before verification of the resource distribution policy it should be rebuilt for accurate results. The more equally the resources are distributed with as little control as possible, the higher are the results. One option for verification would also be do that in smaller pieces. These pieces could be organization behavior, productivity etc. Citation production model for example, has been verified quite well, but the controlling parameters for the model need to be clarified. However, the predictability of the simulation can drastically be improved with further research.

To examine the system more widely, new factors should also be considered. For example, how do we know that when moving to direction of freedom from surveillance, the behavior of agents would not change? When there is no surveillance, resources would be easier to gain no matter what kind of scientific output would be produced, so researchers would produce weaker science. The factor could be laziness, which grows when surveillance lowers. This kind of factor could play role in non-surveying end of the model.

As mentioned before, frustration and researchers joining behavior could be investigated further. Since the current model focuses just on observing one organization, there could be multiple organizations and their researcher-swapping dynamics involved. This would also make it possible to observe individual researcher's career.

Current simulation advances in pre-defined steps of time. Actual event-based simulation might be more functional, where the funders, organization and researchers would run in their own processes and would communicate between each other within intervals. This kind of application is heavy to implement compared to current method, so it might be too much for thesis. Technically, the agent-based simulation is clearly a good approach: Artificial intelligence can be built, and behavior controlled as the researchers now. The downside is that the amount of parameters is drastically larger, and grows even more if, for example, funder would carry some kind of intelligence.

References

AAAS. *American Association for the Advancement of Science.* 2006.

Afonso, Alexandre. *The natural order of things.* November 21, 2013. http://alexandreafonso.wordpress.com/2013/11/21/how-academia-resembles-a-drug-gang/ (accessed July 24, 2014).

Andersson, Kent. *The scholarly kitchen.* April 22, 2013. http://scholarlykitchen.sspnet.org/2013/04/22/more-review-costs-more-the-dynamics-of-a-complex-and-varied-expense-for-journals/ (accessed July 17, 2014).

Arbib, Michael A. *Theories of abstract automata.* New Jersey: Prentice-Hall, Inc., 1969.

Banks, Jerry. "Introduction to simulation." *Winter Simulation Conference* (AutoSimulations, Inc.), 1999: 7-13.

Banks, Jerry, John S. Carson, Barry L. Nelson, and David M. Nicol. *Discrete-Event System Simulation.* Prentice Hall, 2010.

Barnes, Laura L. B., Menna O. Agago, and William T. Coombs. "Effects of job-related stress on faculty intention to leave academia." *Research in Higher Education, Vol 39. No 4*, 1998.

Barnett, Adrian, Danielle Herbert, and Nicholas Graves. *The Conversation.* November 3, 2013. theconversation.com/the-end-of-written-grant-applications-lets-use-a-formula-19747 (accessed July 25, 2014).

Baxt, WG, JF Waeckerle, JA Berlin, and ML Callaham. "Who reviews the reviewers? Feasibility of using a fictitious manuscript to evaluate peer reviewer performance." *US National Library of Medicine National Institutes of Health*, 1998.

Bazeley, Pat. ”Defining ‘early career’ in research.” *Higher Education 45*, 2003.

Becker, K., and J. R. Parker. *The guide to computer simulations and games.* Wiley, 2011.

Bourguignon, François, and Amedeo Spadaro. "Microsimulation as a tool for evaluating redistribution policies." *ECINEQ*, 2006.

Boyack, Kevin W., and Katy Börner. "Indicator-assisted evaluation and funding of research: Visualizing the influence of grants on the number and citation counts of research papers." *Journal of the American Society for Information Science and Technology*, 2003.

Canterbury University HR. *Human resources Department of Canterbury University, New Zealand* (April 15, 2013).

Carayol, Nicolas, and Mireille Matt. "Does research organization influence academic production?: Laboratory level evidence from a large European university." *Elsevier Research Policy*, 2004: 1081-1102.

Carson, Lydia, Christoph Bartneck, and Kevin Voges. "Over-Competitiveness in Academia: A literature review." *Disruptive Science and Technology*, 2013: 183-190.

Chen, Chun-Hung. ”An Effective Approach to Smartly Allocate Computing Budget for Discrete Event Simulation.” *Proceedings of the 34th Conference on Decision & Control.* 1995.

Cottrell, Allin. "Regression Analysis: Basic consepts." 2011.

Curtis, John W., and Saranna Thornton. "The Annual Report on Economic Status of the Profession." *Academe*, April 2013: 4-19.

Day, Nancy Evelyn. ”The silent Majority: Manuscript rejection and its impact on scholar.” *Academy of Management Learning & Education*, 2011: 704.

Economist, The. *How science goes wrong.* October 19, 2013. http://www.economist.com/news/leaders/21588069-scientific-research-has-changed-world-now-it-needs-change-itself-how-science-goes-wrong (accessed July 24, 2014).

EPSRC. *Engineering and Physical Sciences Research Council.* 2013. http://www.epsrc.ac.uk/funding/fundingdecisions/successrates/201213/ (accessed July 4, 2014).

ESRC. *Economic and Social Research Council.* 2010. http://www.esrc.ac.uk/about-esrc/mission-strategy-priorities/demand-management/success-rates.aspx (accessed July 17, 2014).

Foundation Center. *Foundation growth and giving estimates.* 2005. http://www.grantspace.org/Tools/Knowledge-Base/Funding-Research/Statistics/Percentage-of-funded-grant-proposals Foundation cente (accessed July 4, 2014).

Franceschet, Massimo. ”A cluster analysis of scholar and journal bibliometric indicators.” *Journal of the American Society for Information Science and Technology*, 2009.

Garfield, Eugene. ”The history and meaning of the journal impact factor.” *Thomson Scientific*, 2006.

Geard, Nicholas, and Jason Noble. "Modelling academic research funding as a resource allocation problem." *School of Electronics and Computer Science, University of Southampton, UK*, 2010.

Hattie, John, and H. W. Marsh. "The relationship between research and teaching: A meta-analysis." *Education & Educational Research*, 1996.

Herbert, Danielle L, John Coveney, Philip Clarke, Nicholas Graves, and Adrian G Barnett. "On the time spent preparing grant proposals: an observational study of Australian researchers." *BMJ Open*, 2013.

Herbert, Danielle L, John Coveney, Philip Clarke, Nicholas Graves, and Adrian G Barnett. "On the time spent preparing grant proposals: an observational study of Australian researchers." *BMJ Open*, 2013.

Herbert, Danielle L, John Coveney, Philip Clarke, Nicholas Graves, and Adrian G Barnett. "The impact of funding deadlines on personal workloads, stress and family relationships: a qualitative study of Australian researchers." *BMJ Open*, 2014.

Hirsch, J. E. ”An index to quantify an individual’s scientific research output.” 2005.

ISI. *Institute of Scientific Information.* 1997.

Jacob, Brian A., and Lars Lefgren. "The impact of research grant funding on scientific productivity." *Journal of public economics*, 2011: 1168-1177.

Jain, Raj. *Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling.* Wiley Computer Publishing, John Wiley & Sons, Inc., 1991.

Johnson, Michael L. ”Nonlinear Least-Squares Fitting Methods.” *PubMed*, 2008.

Johnsrud, Linda K., ja Vicki J. Rosser. ”Faculty Members' Morale and Their Intention to Leave: A Multilevel Explanation.” *The Journal of Higher Education, Volume 73*, 2002: 518-542.

Jones, Anita. ”The explosive growth of postdocs in computer science.” *Communications of the ACM*, 2013: 37-39.

Kelchtermans, Stijn, and Reinhilde Veugelers. "Top research productivity and its persistence. A survival time analysis for a panel of Belgian scientists." *DTEW Research Report 0576*, 2005: 1-31.

Kemeny, John G., Mirkil J. Hazleton, Snell Laurie, and Gerald L. Thompson. "Finite mathematical structures." *Prentice-Hall, Inc.*, 1959.

Kleijnen, Jack P. C. *Statistical tools for simulation practitioners.* New York: M. Dekker, 1986.

Koelman, Jos, and Richard Venniker. "Public funding of academic research: the Research Assessment Exercise of the UK." 2011.

Kreiman, Gabriel, and John H. R. Maunsell. "Ninecriteriaforameasureofscientificoutput." *Frontiers in Computational Neuroscience*, 2011.

Kresge, Nicole, Robert D. Simoni, and Robert L. Hill. *The Most Highly Cited Paper in Publishing History: Protein Determination by Oliver H. Lowry.* The Journal of Biological Chemistry, 2005.

Kärkkäinen, Olli. *Taloustieteilijän silmin maailma on kannustimia täynnä (In Finnish).* April 2, 2013. Secondbestworld.wordpress.com (accessed July 17, 2014).

Link, Albert N., Christopher A. Swann, and Barry Bozeman. "A time allocation study of university faculty." *Economics of Education Review 27*, 2008: 363–374.

Lyubomirsky, Sonja. *The Myths of Happiness.* Penguin Press, 2014.

Mandavilli, Apoorva. ”Peer review: Trial by Twitter.” *Nature*, 2011.

Matloff, Norm. ”Introduction to Discrete-Event Simulation and the SimPy Language.” *N. S. Matloff*, 2008.

Moses, Ingrid. "Promotion of academic staff: Reward and incentive." *Higher Education 15*, 1986: 135-149.

NHMRC. ”National Health and Medical Research Council of Australia: Annual report 2012-2013.” 2013.

NIH. *National Institute of Healt: Success Rates.* 2013. http://report.nih.gov/success_rates/ (accessed July 4, 2014).

OECD. *Frascati Manual: Proposed Standard Practice for Surveys on Research and Experimental Development, The Measurement of Scientific and Technological Activities.* OECD Publishing, 2002.

Page, Bernd, and Wolfgang Kreutzer. *The Java Simulation Handbook - Simulating Discrete Event Systems with UML and Java.* Shaker Verlag GmbH, 2005.

Paolucci, Mario, and Francisco Grimaldo. "Mechanism change in a simulation of peer review: from junk support to elitism." *Scientometrics*, 2014.

Peters, Douglas P., and Stephen J. Ceci. "Peer-review practices of psychological journals: The fate of published articles, submitted again." *Behavioral and Brain Sciences*, 1982: 187-195.

Redner, S. ”How popular is your paper? An empirical study of the citation distribution.” *The European Physical Journal B*, 1998.

Reuters, Thomson. *Thomson Reuters' Essential Science Indicators database.* 2010.

Ross, Sheldon M. *Simulation.* San Diego: Academic press, 1997.

SAS Institute. *Concepts of experimental design.* Design Institute for Six Sigma, 2005.

Schembri, Patrick J. ”The different types of scientific literature.” 2007.

Shaw, Debora, and Liwen Vaughan. "Publication and citation patterns among LIS faculty: Profiling a “typical professor”." *Library & Information Science Research, Volume 30, Issue 1*, 2008: 47-55.

Smith, Richard. ”Peer Review: a flawed process at the heart of science and journals.” *Journal of the Royal Society of Medicine*, 2006.

Tertiary Education Comission. "Performance-Based Research Fund: Evaluating research excellence." 2012.

Tiihonen, Timo. *Simulation course, University of Jyväskylä.* 2014. http://users.jyu.fi/~tiihonen/simul/ (accessed July 4, 2014).

Troitzsch, Klaus G., U Müller, Gilbert G. N, and Doran. *Social sciences microsimulation.* Springer, 1996.

University Alliance. ”Funding research excellence: research group size, critical mass & performance.” 2011.

Wang, Dashun, Chaoming Song, and Albert-László Barabási. "Quantifying Long-term Scientific Impact." *Science Magazine*, 2013.

Wolfram. *Wolfram Mathworld.* 2014. http://mathworld.wolfram.com/RandomNumber.html (accessed 9 18, 2014).

Yang, Otto O. *Guide to Effective Grant Writing: How to Write a Successful NIH Grant Application.* 2005.

Attachments

A Dependency map of variables

Attachment A