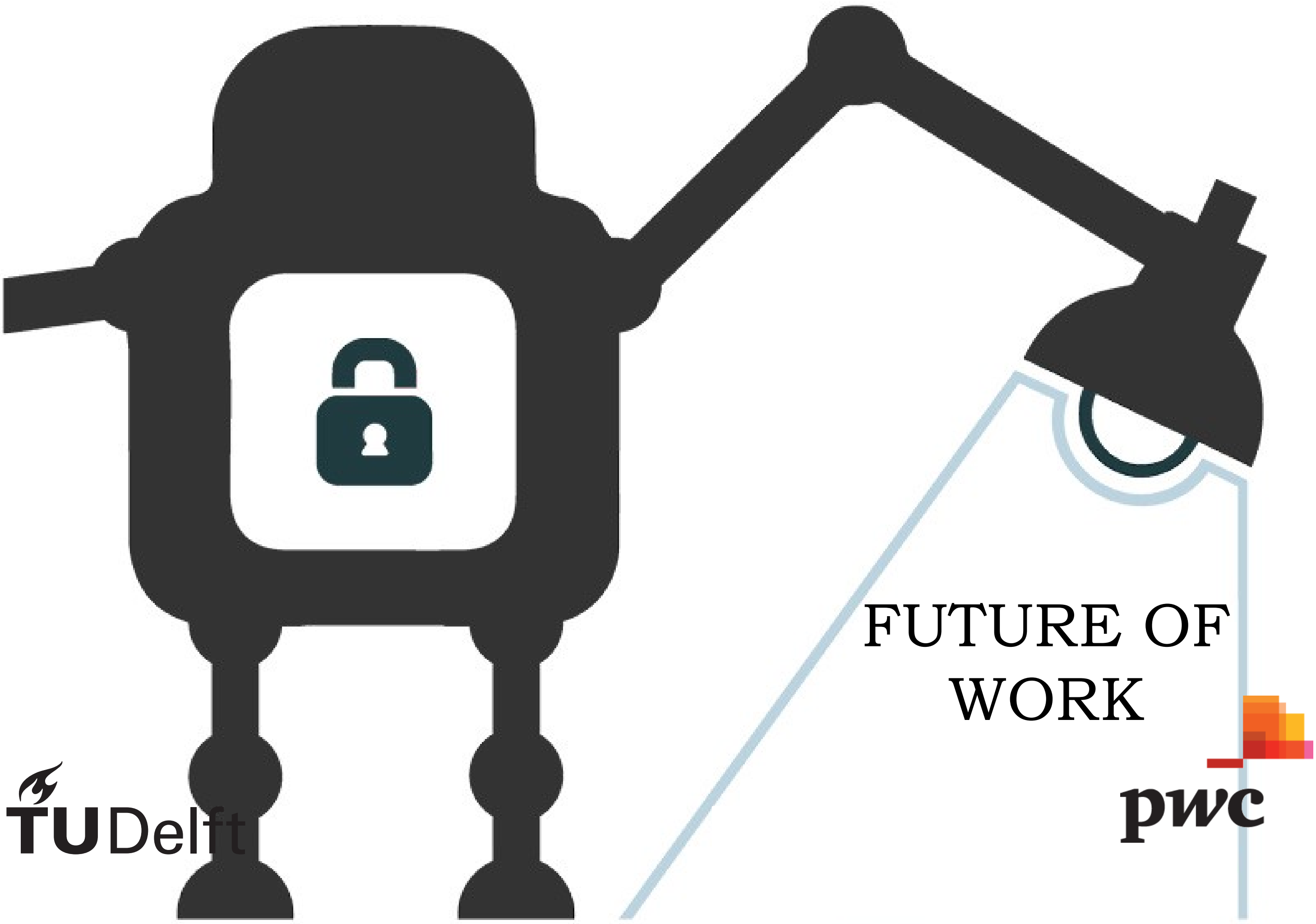
Estimating the effects of technological progress:

A data-analytic and agent-based approach

M.P. Ingwersen

Master Thesis

Delft University of Technology
Faculty Technology, Policy and Management



Estimating the effects of technological progress:

A data-analytic and agent-based approach

by

M.P. Ingwersen

to obtain the degree of Master of Science
at the Delft University of Technology,
To be defended publicly on Friday August 30, 2019 at 15:00 AM.

Student number:	4149459	
Project duration:	February 1, 2019 – August 30, 2019	
Thesis committee:	Prof. dr. C P. van Beers,	TU Delft, Head of committee
	Dr. S. T. H. Storm,	TU Delft, first supervisor
	Dr. Ir. I. Nicolic,	TU Delft, second supervisor
Internship supervisors:	Drs. L. Akkermans,	PwC
	Prof. Dr. R. Wetzels,	PwC

TU Delft

Preface

Dear reader,

This thesis marks the end of my journey at the Delft University of Technology. While the end of a thesis is often experienced as a moment of joy and relief, for me, ending this research leaves me with mixed feelings. Looking back to the months writing this thesis, nothing but warm memories and gratefulness to the many learning experiences arise. Yet, it is this feeling of fortune that makes this end somehow bittersweet.

Like all good stories, the beginning of my thesis started in a bar. Here, my brother told me about his doubts about becoming a surgeon. Knowing the many hours of study and hard work he had put into his career, his hesitation surprised me. However, never could I have known that his motivation for doubting and the underlying reasons would continue to intrigue me for many months. 'Seeing the increasing use of robots on the operating table, how would the content of my job develop in the future' my brother asked. 'Am I putting all this effort into a job that eventually will be replaced by robots?'.

For my thesis, I have analyzed how technological progress can influence the Dutch labour market. Motivated by how the doubts of my brother could eventually influence his decisions, this research distinguishes itself from other related research by including the notion of uncertainty. Building on strong research of other scholars, I was able to perform a very interesting study about how the consequences of advances in technology can affect Dutch jobs in the future.

While writing a thesis is often seen as an individual process, for me, it would be ignorant to ignore all the insights and support that others gave me. For starters, I would like to thank my supervisors at the TU Delft. With his comments and his provision of new angles of approach, Mr Storm has supported me and, above all, inspired me during the whole process. Besides the provision of new insights, with his support, Mr Nikolic has helped me with the many technical aspects of my thesis and made sure that this thesis really is the result of the combination of an economic and technical perspective.

Furthermore, I would like to thank the supervisors of my graduate internship at PwC. Next to the comfortable and inspiring environment for writing my thesis, they gave me valuable support that encouraged me to always take that extra step and to deliver the quality I pursued. Lau Akkermans has been an inspiring and enthusiastic supervisor who, despite his full schedule, always made time to challenge and to stimulate me on a both scientific and personal level. I would also like to thank Ruud Wetzels for his very knowledgeable comments on the data-analysis and his help with my doubts about my future.

Dear reader, I hope that you will enjoy reading this thesis as much as I enjoyed making this. With this thesis, I truly hope that I can contribute to the understanding of the consequences of technological progress and the challenges we face, including that of my brother's.

M.P. Ingwersen
Amsterdam, July 2019

Executive Summary

While much has been written about the possible consequences of technological progress on the labour market, there is still no consensus among scholars about how technological progress and its consequences will manifest themselves on the future of work. This disagreement among scholars reveals itself mostly concerning the question 'how many workers will be replaced due to technological progress?'. While famous research of Frey and Osborne (2013) predicts that 47% of all US workers will be replaced, other research of Arntz, Gregory, and Zierahn (2017) predicted that only 9 % of the US workers would be replaced due to technological progress.

Despite whether it was caused by technological progress, the Dutch labour market already suffers from a gap between labour demand and supply, and when no appropriate action is undertaken, the chances are that this gap will increase in the future. However, when considering the current rapid technological developments, one should incorporate these developments when assessing the right future policies.

This thesis aims at providing more insights into the effects of technological progress on the Dutch labour market. However, as mentioned, these effects –and technological progress itself– are by no means certain. This uncertainty, in turn, can causes firms' and workers to behave differently and thereby influencing the 'real' consequences of technological progress. This research, therefore, tries to make this uncertainty and its effects explicit by answering the following research question:

> *How will uncertainty about technological progress eventually influence the effects of technological progress on the Dutch labour market?*

Estimating the true effects of technological progress: a multi-method approach

To answer the research question, this study applied a so-called multi-method research approach. That is, it first conducted a data-analysis of the competencies of Dutch workers and their possibilities of being automated. This data-analysis was followed by the development and analysis of an agent-based model that was, in contrast to the data-analysis, able to take heterogeneous low-level factors into account –such as the uncertainty and behaviour of firms and workers.

For the data analysis, the model of Frey and Osborne (2013) (FO) was applied to Dutch data. To apply the model of Frey and Osborne (2013), US data about jobs and competencies had to be mapped to Dutch data. Consequently, a data-set was constructed that incorporated data about the competencies of Dutch jobs and workers. Besides the application of the FO model, we used this newly created data-set to analyze the current competencies on the Dutch labour market. Nevertheless, since this data-set was not yet constructed, the creation of this data about the competencies of Dutch workers has opened up great opportunities for further research.

Findings of the FO application turn out not to be that catastrophic as it may have seemed

The application of the FO model has shown that approximately 12% of the Dutch workers are at high risk of being replaced due to technological developments. Moreover, when simulating this risk, this study found that on average, 36% of the Dutch workers can lose their jobs within 20 years due to technological progress.

However, while these numbers are shocking, they do not reveal the whole story. In the FO application, the study did not account for the possibility of retraining of workers and the strategic behaviour of workers and firms stemming from their uncertainty due to technological progress. To take these factors into account, an agent-based model was developed. By experimenting with this model, the average unemployment rate after 21 years dropped to 3%, a difference of 33% with the results of the FO model. However, due to policy and scenario testing, the underlying unemployment rates of this average unemployment rate of 3% varies greatly.

This variance is mostly due to different levels of labour market flexibility. That is, our model found that the higher the proportion of flexible contracts, the higher the final unemployment rate will be. Additionally, this study found that this effect of flexibility is much stronger in determining unemployment due to technological progress than other factors – such as the possibility of retraining. Besides, this study found that the higher the number of years firms and workers consider in their planning, the better they anticipate on future changes due to technological progress and hence; the lower unemployment will be. Finally, this study has shown that uncertainty does indeed matter since it almost determines 90% of the outcomes of the model and thereby shapes the 'true' consequences of technological progress.

Recommendations for the formulation of labour market policies and further research

From a policy perspective, these findings suggest some valuable lessons. At first, this study found that firms are more prone to financial incentives than workers. Next, since workers do consider other non-financial factors into their choices –such as age and working experiences– policies concerning workers –for example stimulate retraining– should take these other factors into account. At last, for all policies to be effective, it should aim at stimulating decision making for the long-term, instead of encouraging firms and workers to formulate their strategy on an ad-hoc basis.

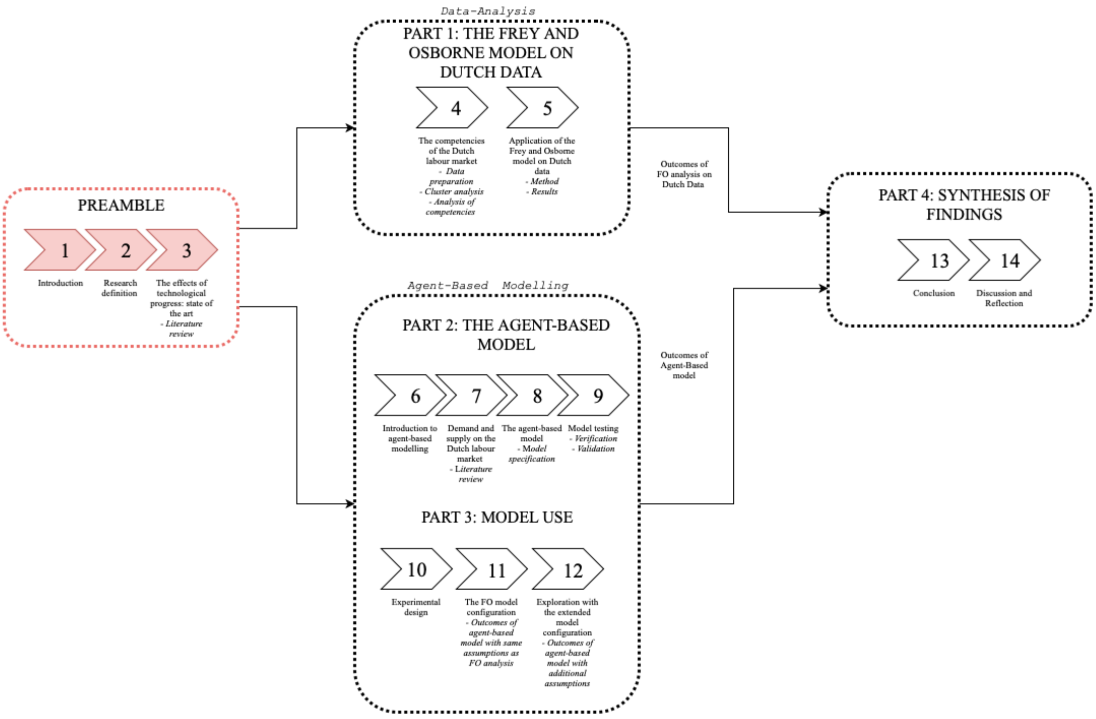
By developing, analyzing and combining both models –the FO model and the agent-based model– this study was able to investigate the real consequences of technological progress on the Dutch labour market. While both models open up many possibilities for further research –such as the analysis of specific labour market activating policies– their limitations should be taken into account. For both models, the most significant limitation is the issue of validity. This study, therefore, recommends further validation of both model–for example through expert opinion– when they are applied for new research.

List of Figures

List of Tables

Contents

Preamble



PREAMBLE
1
Introduction
2
Research definition
3
The effects of technological progress: state of the art
- Literature review
Data-Analysis
PART 1: THE FREY AND OSBORNE MODEL ON DUTCH DATA
4
The competencies of the Dutch labour market
- Data preparation
- Cluster analysis
- Analysis of competencies
5
Application of the Frey and Osborne model on Dutch data
- Method
- Results
Outcomes of FO analysis on Dutch Data
Agent-Based Modelling
PART 2: THE AGENT-BASED MODEL
6
Introduction to agent-based modelling
7
Demand and supply on the Dutch labour market
- Literature review
8
The agent-based model
- Model specification
9
Model testing
- Verification
- Validation
PART 3: MODEL USE
10
Experimental design
11
The FO model configuration
- Outcomes of agent-based model with same assumptions as FO analysis
12
Exploration with the extended model configuration
- Outcomes of agent-based model with additional assumptions
Outcomes of Agent-Based model
PART 4: SYNTHESIS OF FINDINGS
13
Conclusion
14
Discussion and Reflection

Chapter 1

Introduction

Chapter Abstract

This chapter serves as the introduction to this thesis. This chapter will explain briefly how the growing gap between labour demand and supply on the Dutch labour market can have severe societal and economic consequences. Moreover, due to technological development itself and the uncertainty among firms and workers that these developments create, fears are that this gap will continue to increase when no appropriate action is undertaken. For these reasons, the aim of this thesis is to analyze the effects of technological progress by using a multi-method research approach. The outline of this thesis and a reading guide are presented at the end of this chapter.

Will robots take my job? What may seem like a good start for a science fiction movie can suddenly become a reality. In their article, Frey and Osborne (2013) famously stated that almost 47 per cent of the US jobs are at risk due to computerization. Even more surprising, the authors state that not only jobs in logistics and manufacturing together with the bulk of office and administrative support workers, and labour in production occupations, are at risk –an outcome that is consistent with recent technological developments documented in the literature–, but also jobs in the service sector are at high risk.

Fortunately, since the publication of their famous research, scholars –Frey and Osborne included– have tried to put more nuance into these findings or oppose these findings altogether. For example, in their article Arntz et al. (2017) claim that the findings of Frey and Osborne (2013) "are overestimating the share of automatable jobs by neglecting the substantial heterogeneity of tasks within occupations as well as the adaptability of jobs in the digital transformation"(p. 160). By taking these elements into account, Arntz et al. (2017) demonstrated that the automation risk of US jobs drops from 47% to 9%.

Without going into a discussion about who is right or what to believe, the previous example proves at least one thing: we don't know. This uncertainty about the effects of technological progress on economic performance has existed throughout history (Mokyr, Vickers, & Ziebarth, 2015). However, this 'fear' or feeling of uncertainty is not only perceived by workers in the past or by some academic scholars. Today, 76 per cent of the CEO's of the leading multinationals in the world are highly uncertain about the future and fear the consequences of the rapid technological progress (PWC, 2018). These fear for the effects of technological progress mainly involves changes in the labour-side of the economy (Frey & Osborne, 2017; Goos, Manning, & Salomons, 2009; Makridakis, 2017).

As a consequence, the complexity of the problem does not manifest in the *actual* effects of technological progress on the labour market itself. Instead, it is the uncertainty about the effects and *percieved* effects of technological development that makes this problem so interesting.

For example, suppose firm A believes that Artificial Intelligence will have a significant impact on productivity and economic growth in his industry. This firm can decide to alter their strategy –e.g. hiring fewer workers, investing in new machines– to make sure that it will not move behind. However, as game-theory might suggest, other firms, who are also uncertain about the future of technological progress, fear that they will lose the competition with Firm A and will thereby also alter their strategy.

Consequently, the anxiety of firm A towards technological progress has turned into a so-called self-fulfilling prophecy (Reinganum, 1982).Another implication of technological progress uncertainty among firms is that it may slow down aggregate demand. According to Bloom, Bond, and Van Reenen (2007), uncertain firms are more cautious with their employment responses. This can lead to a slowdown of aggregate demand, which, in turn, can make firms more cautious in their investment decisions.

However, these effects of technological progress uncertainty do not only manifest at the demand side of the labour market –e.g. the firms– but also at the supply side of the labour market –e.g. the workers. Technological progress, for example, may reduce the supply of older workers since their career horizon is much shorter. "Hence it is less beneficial for them, or for their employers, to invest in learning new technologies"[p. 107](Bloom et al., 2007).

To sum up, the effects of technological progress are unclear and which therefore creates uncertainty on both the demand and the supply side of the labour market. This uncertainty, in turn, can have an impact on the overall performance of the labour market.

When looking at the Netherlands, currently, a major bottleneck of the Dutch labour market is the mismatch between supply and demand (PWC, 2018). Accordingly, "employers have a difficult time filling their vacancies in these sectors, and the shortage of qualified staff is expected to increase in the upcoming five years" (p.12). This is despite the fact that about 3.6% of the Dutch working population is unemployed (CBS, 2018). Given this already unresolved problem of the gap between labour demand and supply on the Dutch labour market and the effects of uncertainty about technological progress, the chances are that the gap will be wider.

1.1 Societal relevance

As this introduction has made clear, there is a gap between demand and supply on the Dutch labour market and due to current technological developments, this gap can become wider when no appropriate action is taken. Unfortunately, formulating the right policies to close this gap will be difficult since predicting the future capabilities and development pace of technology is difficult and unreliable (Armstrong, Sotala, & hÉigeartaigh, 2014; Frey & Osborne, 2013). This, however, catalyzes the concerns about the future of the Dutch labour market. These concerns manifest themselves in both the demand as the supply side of the labour market.

At the demand side of the labour market, these concerns may express as a decline in investments due to the uncertainty of firms, which in turn can lead to an increase in unemployment (Coccia, 2017; Ravn & Sterk, 2017; Zahra, 1996). Concerning the supply-side of the labour market, the concerns focus mainly about the effects of capital-labour substitution and its impact on society. Accordingly, capital-labour substitution may cause declining wages, higher unemployment, job polarization and higher economic inequality (Autor, Dorn, & Hanson, 2015).

Indeed, these fears for the effects of technological progress may prove to be wrong since technological progress has never resulted in mass unemployment due to the reallocation of labour supply in new tasks and occupations (Mokyr et al., 2015). Nevertheless, the question remains if this would also be true for the upcoming future. Current technological developments -such as Artificial Intelligence (AI) and within robotics– has led scholars –such as Frey and Osborne (2013)– to argue that robots will replace a considerable proportion of the workers. Whether true or not, these predictions have led to even more uncertainty on the demand and supply side of the labour market which, consequently, can turn these feared effects of capital-labour substitution into a self-filling prophecy.

1.2 Scientific relevance

As mentioned, there is a lot of uncertainty about the consequences of technological progress. Moreover, this primary uncertainty is complemented by secondary uncertainty –uncertainty about the actions of others in response to technological progress. In order to gain insight into how technological progress can influence the consequences of technological progress in the Dutch labour market, we have to incorporate these uncertainties. Furthermore, given the complex system of the labour market –characterized by the existence of primary uncertainty, feedback mechanism and a range of possible interaction between agents–, insight in these complex mechanism at work is needed.

An application of the famous Frey and Osborne (2013) study on data about the Dutch labour market is a good starting point for developing these insights. By applying this study on Dutch data, we can develop a baseline future for which we can test the effects of uncertainty about technological progress. As far as we know, this study has never been applied with respect to the Netherlands. Besides from the outcomes of this study, the creation of the data-set that incorporates the competencies of Dutch workers –which is needed for this application of the Frey and Osborne (2013) model– may be precious for further research since it can contribute to new insights about the supply of labour in the Netherlands.

Nevertheless, the execution of a data-analysis is not sufficient for gaining valuable insights about the effects of technological progress because it does not incorporate the forces that are caused by the uncertainty about technological progress or the possibility of retraining. Therefore, the research is extended by the development and analysis of an agent-based model. By applying this so-called multi-method approach, this study moves beyond the traditional econometric data analytic approach and is able to open the 'black box' of how the dynamics on the labour market are created.

1.3 Outline

With respect to the multi-method approach, this research is structured into different parts. This study starts with a theoretical framework of the state of the art literature about the effects of technological progress and uncertainty on the labour market. After identifying the research-gap, the research question and methodology of this study is presented. Hereafter, the document will be divided into four parts:

1. **Application of the Frey and Osborne model on Dutch Data**: In the first part of this study, we perform a data analysis of the competencies of Dutch workers to replicate the FO model on Dutch data. Chapter 4 discusses how data about Dutch jobs is linked to US data about competencies and presents the results of the descriptive analysis to the competencies of Dutch workers. This is followed by chapter 5 that uses these results to replicate the FO model on Dutch data.

2. **The future of work: an agent-based approach**: The second part of the study discusses the development of the agent-based model. This part starts with a general introduction about the agent-based model in chapter 6, followed by a discussion about the theoretical background of the model in chapter 7. The actual development of the model is discussed in chapter 8. Finally, model testing is addressed in chapter 9.

3. **Model use**: The third part will discuss the findings derived from the agent-based model. This part starts with a discussion of the experimental choice in chapter 10. Chapter 11 will present the results of these experiments.

4. **Synthesis and Findings**: This last part will reflect and discuss the results of the research. The conclusion is presented in chapter 10, followed by the discussion in chapter 11.

Given the length of this thesis, we are aware of the fact that this document can be seen as somewhat overwhelming. To avoid confusion and to guide the reader through this document, a so-called 'roadmap' of this thesis is developed (see Figure 1.1).

Before diving into the analyses and conclusions, the reader is recommended to start with the chapters of the *Preamble*. These chapters will provide the reader with the basic information and motivations for this thesis. Furthermore, *Part 1* and *Part 2* can be read independently of each other. That is, to understand the information provided in Part 2, the reader does not have to read Part 1 –and vice versa. Finally, the outcomes of the data-analysis (part 1) and the agent-based model (part 2) are discussed and reflect upon in *Part 4*

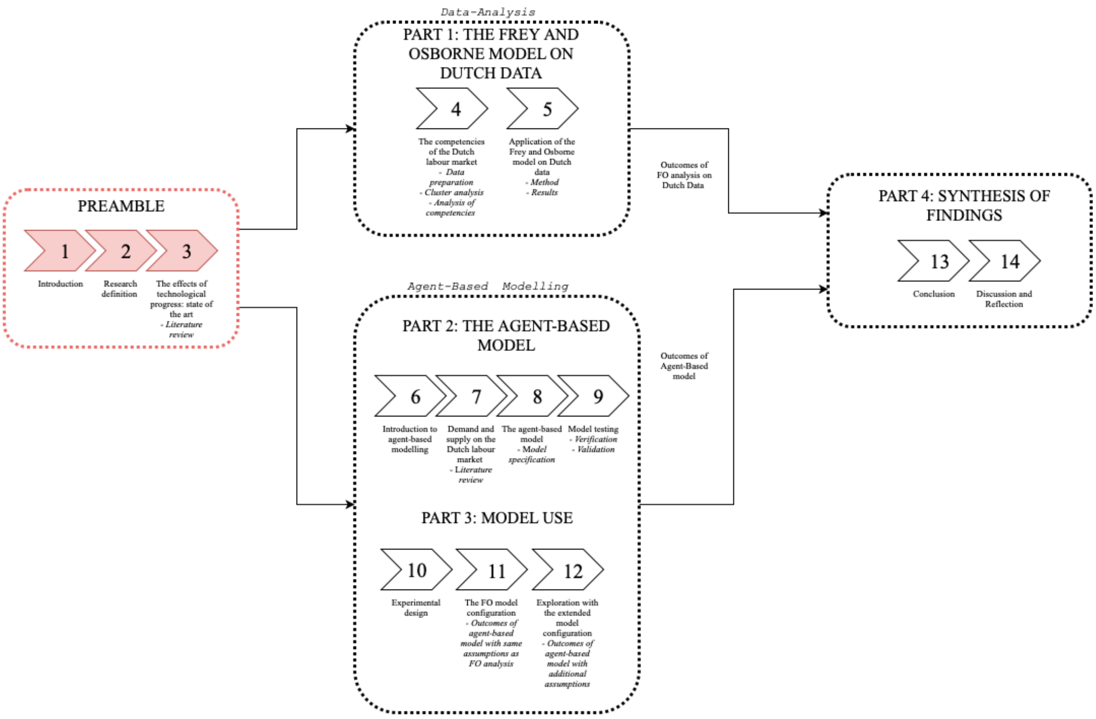


Figure 1.1: Road-map of thesis: The current part is indicated with the colour red (Preamble)

Chapter 2

Research definition

Chapter Abstract

After identifying the research gap, this chapter introduces the following research question: 'How will uncertainty about technological progress eventually influence the effects of technological progress on the Dutch labour market?' To answer this question, a multi-method research approach is proposed. By combining the outcomes of the data analysis of the Dutch labour market with the insights provided by the agent-based model, the aim is to answer the research question as comprehensive as possible.

In this study, we explore the effects of uncertainty about technological progress on the Dutch labour market. More specifically, we explore how outcomes of the Frey and Osborne (2013) application on Dutch data may change due to this uncertainty. In this chapter, the research gap and the research question for this study are defined, and the problem scoped. Furthermore, this chapter discusses the motivation and content of the applied methods.

2.1 Research gap and research question

As mentioned in the introduction, much is written about the effects of technological progress on the labour market, but yet there is no consensus. Moreover, uncertainty about technological progress can shape firms' HR-strategies and therefore, the 'real' effects of technological progress on the labour market. Already suffering from a large gap between labour demand and supply, this uncertainty can further exaggerate the problems on the Dutch labour market.

Consequently, to decrease the gap between labour demand and labour supply, one could start with reducing this uncertainty. Because technological progress itself is difficult to predict, the focus has to be on decreasing the secondary uncertainty about technological progress –e.g. uncertainty about how different agents react in response to technological progress uncertainty. More specifically, we need to explore the uncertain futures of the labour market, considering its complex (feedback) mechanism, heterogeneous agents, their interactions, and the uncertain environment. By doing this, robust policies can be found, which can decrease the effects of technological progress. Note the choice for the word 'robust' and not, for example 'optimal'. According to Van Dam, Nikolic, and Lukszo (2013), "robustness is a measure of how the system performs under stress when confronted by extreme inputs or shocks from the environment, only for particular variables" (p. 48). Applied to policies, here robustness means the set of labour market strategies that under different scenario's and when confronted with extreme shocks from the environment, perform the best.

Besides this robustness check, another advantage of applying this somewhat stringent definition is that it allows for striking off the set of strategies that may have good results for firms and workers individually on the short-term but may harm long-term macro-economic performance. This is necessary since there is often a mismatch between short-term microeconomic interests and long-term macroeconomic performance (Dosi, Napoletano, Roventini, & Treibich, 2016). With a robust policy we, therefore, refer to a labour market strategy that does not only generate desired outcomes on the micro level and short-term under specific circumstances, instead it refers to a strategy that produces the most desired results at the micro- and macro level, on short- and long-term and under all conditions.

To conclude, it cannot be avoided that firms and workers choose their strategy based on their individually optimal outcomes in the short term. To avoid this mismatch between desired results on the micro-level and the macro-level, Dosi et al. (2016) argues that because of the complex feedback between innovation and demand dynamics, there is a strong need for a “coordinated set of policies, going against the peculiar schizophrenia between macro policies, if any, for the short-run and structural policies for the long-run” (p. 21).

2.1.1 Research question

To develop such policies, we need to understand how current technologies will affect the labour market. Since this effect is uncertain, we must incorporate this uncertainty about technological progress. This leads to the following research question:

> *How will uncertainty about technological progress eventually influence the effects of technological progress on the Dutch labour market?*

In the research question, uncertainty about technological progress refers to the primary uncertainty about technological progress, secondary uncertainty and competitive uncertainty. By answering this research question, the aim is to enrich current research about the effects of technological progress on the labour market by incorporating the effects of the behaviour of firms and workers stemming from uncertainty. By identifying these effects, recommendations can be made concerning labour market policy to reduce the gap between labour demand and supply and to reduce the feared consequences of technological developments.

2.1.2 Sub-questions

In order to answer this research questions, the following sub-questions are defined:

1. What are the skills, abilities and knowledge of Dutch workers?
2. How will technological progress affect the Dutch Labour market?
3. How can the behaviour of firms and workers stemming from technological progress uncertainty change these numbers?
4. Which policy measures can be applied to reduce these negative effects of technological progress?

These sub-questions divide the research into four manageable parts and will eventually lead to the answer to the main research question. The first two questions focus on the application of the Frey and Osborne (2013) model on the Dutch labour market. For this, a data set with the competencies of Dutch workers will be constructed for which a descriptive analysis will show how the competencies are distributed among the current Dutch workers (sub-question 1). The second sub-question uses this data set to determine the proportion of jobs being automated in the Netherlands according to the Frey and Osborne (2013) model.

The second and third part of the research will analyze how the uncertainty about technological progress will influence these outcomes of the first part. For this, the third sub-question tries to see how the behaviour of firms and workers alters these number while the fourth sub question takes possible policies of the government into account.

Consequently, a synthesis of these findings will show how technological progress, and its uncertainty about it, influence the consequences of technological progress on the Dutch labour market.

2.2 Research process

This section presents the research approach and the method used for this study. Since this study applies two methods, both methods are discussed separately.

2.2.1 Approach

Currently, there is a widespread disagreement about the effects of technological progress on productivity and macroeconomic growth; the effects of firm behavior on macroeconomic growth; the effect of macroeconomic growth —and thus perhaps of technological progress— on economic policy; and the effect of economic policy uncertainty on firm behavior (see Chapter 3). Moreover, both firms and workers cannot be seen as homogeneous agents in the labour market since they all encompass different features. This *heterogeneity* of firms and workers on the labour market means that different firms and workers can react differently to the technological progress uncertainties. Consequently, we expect that the interaction between these different agents creates different outcomes on the labour market than when these agents are considered homogeneous.

As a result, this study does not only demands a straightforward numerical solution; it also demands a broad understanding of how different elements —e.g. actors, firms, policies— interact with each other. We, therefore, propose a combination of two methods for this research: a data-analysis to explore the current Dutch labour market and an agent-based modelling approach for identifying the emergent patterns that arise as a result of the interaction of heterogeneous firms and workers.

To answer the research question, a mix of two methods is proposed. This combination of two methods within one research is referred to as a multimethod approach:

> The multimethod approach is a strategy for overcoming each method's weaknesses and limitations by deliberately combining different types of methods within the same investigations. (Brewer & Hunter, 1989, p. 11)

For this research, we will use a data-analytical approach to analyze the influence of technological progress on the Dutch labour market. However, given the uncertainty inherent with technological progress, the heterogeneity of agents and thus the complex nature of the labour market, we need to develop an Agent-Based model. By developing this model, we can explore how different behaviours create emergent patterns and how these patterns can influence the labour market. Finally, we can combine both results to induce conclusions of how uncertainty on the labour market may influences the consequences of technological progress on the Dutch labour market.

2.2.2 Method

This section will explain both methods used in this research and how both methods can be combined to develop insights into the effects of technological progress uncertainty. This section, however, only offers a broad overview of the methods used for this research. During this thesis, each method and its sub-methods are explained in more detail.

Data Analysis of the current Dutch labour market

The first step of the research is a data-analysis of the Dutch labour market. Here, we do not incorporate for the effects of uncertainty or other strategic behaviours of firms and workers. In this data analysis, we apply the same methodology as provided by Frey and Osborne (2013). This means that we have to perform an analysis of how skills, abilities and knowledge are distributed within the Dutch labour market. However, research about the distribution of competencies in the labour market is far from extensive and almost non-existent when applied to the Dutch labour market. This means that for this research, we have to start from scratch when we want to analyze the distribution of skills, abilities and knowledge in the Dutch labour market. That is, we first need to create a data set that includes not only the number of people per jobs in the Netherlands but also the skills, knowledge and abilities that are associated with these jobs. Using this data set, we can see how the current Dutch labour force is composed of jobs and their competencies and how technological progress will influence this shape.

Statistics Netherlands (CBS) provides us with data about the number of workers per job. However, their data does not show which skills, abilities and knowledge are associated with these jobs. To extend the CBS data set with this information, we use data from the National Center for O*NET Development (US) that provides the level of each ability, skill and knowledge per job.

When the mapping between the ONet data-base and CBS data is done, we will analyze how technological progress will influence the probability of becoming automated for each job and eventually the consequences of technological progress on the labour market. This is done by using the assumptions of Frey and Osborne (2013) about how the competencies associated with jobs are susceptible to automation.

Exploratory research with Agent-Based Modelling

To analyze how different behaviour on the labour market creates different patterns and eventually affects the consequences of technological progress on the Dutch labour market, we propose to perform an exploratory analysis with the use of agent-based modelling. By developing an agent-based model, we can incorporate all the different dynamics and interactions between agents and their environment.

Nevertheless, the following question may arise: ‘why is this necessary and do we need to know the behaviour on the lower level since we have already valid outcomes on a higher level?’. To answer this question, we refer to the different feedback mechanism on the labour market identified in chapter 3. For example, when there is uncertainty –both primary and secondary– about technological progress, some firms may decide to not invest in retraining or hiring new workers. This can lead to a slow down of aggregate demand, which in turn can lead to more austerity among firms. However, according to Koopmans and Beckmann (1957), some of the uncertainty among firms can be reduced by collaboration. The collaboration will reduce technological progress uncertainty since it reduces the competitive uncertainty. As one can imagine, this feedback mechanism can create very different patterns –ranging from a stable system that has enough corporation to overcome uncertainties to a chaotic system that is locked in a reinforcing pattern of creating more and more employment.

However, data analysis of the effects of technological progress is unable to incorporate these mechanisms. Even under different scenario’s –e.g. different parameters—, the application of the Frey and Osborne (2013) cannot account for such feedback mechanisms, stemming from the interactions between heterogeneous agents and their environment (Caballero, 2010; **?**). However, to answer our research question and to make strong policy recommendations, it is of great importance to gain insights into these mechanisms.

This can be further understood referring to the distinction between a core and periphery in the macroeconomics made by Caballero (2010). Accordingly, the core is the dynamic stochastic equilibrium approach, as it taught in most graduate programs, whereas the periphery approach is based on the combination with macroeconomics and other strands of literature –such as social science– and provide us with frameworks to understand the causes and mechanisms at work of key macro-economic events. While Caballero (2010) argues that the insight-building mode (both past and present) of the periphery of macroeconomics has proven to be more useful than the general equilibrium-mode of to core to help us better understanding macroeconomic patterns and event, he states that both approaches suffer from limitations:

> The ultimate goal of macroeconomics is to explain and model the (simultaneous) aggregate outcomes that arise from the decisions made by multiple and heterogeneous economic

> agents interacting through complex relationships and markets. Neither the core nor the periphery is able to address this incredibly ambitious goal very satisfactorily. The periphery has focused on the details of the subproblems and mechanisms but has downplayed distant and complex general equilibrium interactions. The core has focused on (extremely stylized) versions of the general equilibrium interactions and has downplayed the subproblems. (p.4)

According to S.-H. Chen, Kaboudan, Du, Neugart, and Richiardi (2012), agent-based modelling combines the advantages of both the core and periphery approach. While the outcomes of agent-based modelling can be analyzed as general macro-economic 'aggregate' patterns, it also gives valuable insights into the mechanism at work without suffering from aggregation effects, such as the simple aggregation of individual changes in behaviour. Agent-based modelling offers us an artificial lab for 'what-if' studies on "distant and complex general disequilibrium interactions", rather than forecasting tools as in the traditional econometric tradition (S.-H. Chen et al., 2012). Indeed, aggregate models –such as the model of Frey and Osborne (2013)– that reproduce realistic outcomes of the labour market is a priori a good candidate to investigate the effects of a given policy.

However, the understanding of the causal mechanisms triggered by a policy can often benefit from the insights generated by the agent-based model. To really gain insight in the effects of technological process uncertainty and to search for robust policies, we, therefore, need to perform a 'what-if' study on the complex and heterogeneous interactions of the labour market. For this, agent-based modelling is the most suitable tool for (S.-H. Chen et al., 2012). To summarize, by applying exploratory modelling with agent-based modelling, we can open the black box' and see what *can* happen on the Dutch labour market.

In order to explore the model and to perform the 'what-if' analysis, experiments with the agent-based model need to be conducted. By performing these experiments, an analysis can be done to indicate how different scenario's and policies can affect the outcomes of the model and hence, the consequences of technological progress. Moreover, to compare the outcomes of this agent-based model with the findings of the FO application on Dutch data, we develop an agent-based model that simulates the effects of technological progress by using the automation probabilities as found by the FO application under almost the same assumptions as in the original Frey and Osborne (2013) study. This so-called FO Model Configuration will serve as the benchmark for the agent-based model that incorporates the other mechanism, such as retraining and uncertainty –e.g. *'the Extended Model Configuration'*.

This Extended Model Configuration will thus help us to gain insight into the dynamics of the Dutch labour market. This does not mean, however, that this agent-based model is a full representation of the Dutch labour market. Since we already have a data-analysis of the current Dutch labour market, this is not necessary. The agent-based model instead is an abstract representation of the Dutch labour market. By making it an abstract model, we can deduct the behaviour and patterns on the labour market without the risk of an over-parameterization of the model.

Synthesis of findings: beyond the traditional econometric approach

The combination of the data analysis and the results of the exploratory ABM analysis will eventually lead to answering the research question. By applying this multimethod approach, we combine the best of two worlds. The data-analysis will provide us with a starting point and explains what the effects of technological progress are ceteris paribus. The use of agent-based modelling enables us to "explore the problem, look critically at what might be in that knowledge gap, and provide an incomplete range of possible system futures" (Van Dam et al., 2013, p. 130).

Chapter 3

The effects of technological progress: state of the art

Chapter Abstract

This chapter discusses the theoretical background of the concepts that are used for this thesis. It can, therefore, be seen as a literature review about the state of the art of technological progress, uncertainty about technological progress and possible feedback mechanisms between technological progress uncertainty and economic performance.

Before the application of the Frey and Osborne (2013) model and the development of the agent-based model, this chapter provides a discussion of past research regarding the effects of technological process on the labour market. It also dives deeper into the concept of uncertainty and aims to define the possible effects of (uncertainty about) technological progress on the labour market.

Since this chapter provides the reader with an overview of the relevant literature associated with the research question, this literature review can be seen as a basis for this study. In the following chapters, however, these concepts and others will be explained in more detail.

3.1 The effects of technological progress on the labour market

As the problem statement has made clear, there is much uncertainty about the impact of technological development. At first sight, this seems wrong since it is well known that technological change is seen as one of the most significant driving forces of economic growth (Camagni, 1991; C. Freeman, 1994). However, anxiety about the effects of the technical process on economic growth is generated throughout history. In a comprehensive article about the history of this technological anxiety, Mokyr et al. (2015) discuss the most prominent concerns that are associated with this technological anxiety. Accordingly, one concern of the technical process is that it will cause a widespread substitution of machines for labour and eventually unemployment.

3.1.1 The fear (or no-fear) for capital-labour substitution

This fear for the substitution of labour for capital is not unjustified. In their analysis, Frey and Osborne (2013) suggested that 47% of the jobs in the United States are at high risk of being automated. This claim of capital-labour substitution [1] follows the same line of reasoning as many other economic scholars before. In their historical review and especially the discussion about the industrial revolution, Mokyr et al. (2015) argue indeed that technological progress

[1] Replacing human occupations or tasks by capital —e.g. machines, robots, or computers (Spaanderman, 2018, p. 1).

has caused capital-labour substitution in the past. Along with these findings, many scholars agree on the claim that these changes on the labour market brought by technological progress can bring dramatic consequences for society —such as more unemployment, changing skill demands, growing inequality, and rising polarization of labour market outcomes (Autor et al., 2015).

However, Mokyr et al. (2015) oppose the argument that capital-labour substitution leads to more unemployment since the economic principles will always continue to operate:

> Scarcities will be with us, most notably of time itself. The law of comparative advantage strongly suggests that most workers still have useful tasks to perform even in an economy where the capacities of robots and automation have increased considerably.
> (Mokyr et al., 2015, p. 47).

Autor (2015) further complements this argument by stating that that the economy always will need 'middle-skill jobs' and thus technological progress will not lead to more unemployment. Moreover, in their book, Storm and Naastepad (2012) agree to some extent with this proposition that capital-labour substitution does not automatically generate unemployment. The authors argue that higher productivity growth will lead to higher investments in new machines. Since these new machines embody new technological progress, these investments will lead to capital-labour substitution. Accordingly, the capital-labour substitution will lead to higher productivity growth, which *can* reduce the equilibrium unemployment rate. Nevertheless, since higher productivity growth —and thus lower unemployment— does not only depend on technological progress, the effects of productivity growth on unemployment are ambiguous (Storm & Naastepad, 2012).

In a comparative study of 39 countries and 34 industries, Fagerberg (2000) showed that 'the countries with the technological most growing industries have experienced higher productivity growth than other countries' (p. 393). However, the author opposes the argument that higher productivity growth stemming from the latest technological progress will lead to lower nor higher unemployment. ' New technology, in this case, the technological revolution, has not been linked with structural changes in demand, output and unemployment in the same way as before' (p. 409).

This argument closely relates to the famous growth theory of Joseph Schumpeter, in which he argues that technological innovation is the ultimate source of economic growth. In his work he refers to the effect of technological progress as "Creative Destruction", 'a process that is incessantly at work, as new technologies replace old ones, provoking the growth and entry of some firms (and sectors) in the economy and the fall and exit of others' (Dosi et al., 2016, p. 2). In this sense, the effect of technological progress may lead to more unemployment on the short-term —which, however, can be mitigated by the entry of new job opportunities—, it will lead to more economic growth on the long-term.

Nevertheless, the question remains if this time is different. Indeed, since the Second-World-War, periods of technological progress are often characterized with increases in productivity growth. However, these increases were mainly due to the substitution of unproductive, heavy and dangerous tasks by 'better' technological solutions (Acemoglu & Autor, 2010; Autor et al., 2015; Mit & Restrepo, 2018; Mokyr et al., 2015). Automation mainly occurred at the lower-skilled jobs while the demand for middle- and high-skilled jobs remained stable.

However, with the development of current technological innovations —advanced robotics and Artificial Intelligence (AI)— the situation is different. Current developments in robotics and Information Technology (IT) can replace the routine [2] tasks of many middle-skilled cognitive and manual jobs, such as bookkeeping, clerical work, repetitive production, and monitoring jobs. By way of contrast, jobs that are intensive in either abstract or non-routine manual tasks, are much less susceptible to this process due to the demand for problem-solving, judg-

[2] In here, routine does not mean mundane —e.g., washing dishes— but rather sufficiently well understood that the task can be fully specified as a series of instructions to be executed by a machine —e.g., adding a column of numbers (Acemoglu & Autor, 2010)

ment and creativity in the former case, and flexibility and environmental adaptability in the latter.

Even though their discussion about the replacement of middle-skilled routines tasks by computerization is valid for previous periods, Acemoglu and Autor (2010) tend to underestimate the effects of the rapid development in the field of AI. Recent developments in this domain —such as machine learning— are close to becoming able to replace 'human' task such as rational decision-making, analysis and interaction (Daron Acemoglu & Restrepo, 2018). Consequently, In the future, even higher-skilled work may be vulnerable to technological displacement (*OECD Skills Strategy Diagnostic Report: The Netherlands 2017*, 2017).

As the previous discussion showed, the substitution of high-skilled tasks by new technology may thus increase productivity. Nevertheless, its cumulative effects on the labour market are ambiguous. Moreover, uncertainty about the impact of the technological innovations on the labour market is further complemented by the fact that the technological progress itself is difficult to predict and inherently unreliable (Armstrong et al., 2014; Frey & Osborne, 2017).

While a full literature review about the link between technological progress and unemployment goes beyond of the scope of this chapter, this brief discussion shows that the effects of technological progress on unemployment are still ambiguous and unclear.

3.2 Osborne and Frey: an apocalyptic future of work

In previous sections, we already discussed the influential work of Frey and Osborne (2013) briefly. We also highlighted the fact that in the years following their publications, their results became controversial. Looking at the critique posed by different scholars, we can identify two strands of criticism: criticism on their assumptions regarding technological progress and criticism about their neglecting of specific important mechanism on the labour market.

In their research, conducted in 2013, Frey and Osborne (2013), did not incorporate the recent developments in Artificial Intelligence. Indeed, recent studies incorporating these developments show different results, but the question is whether, at that time, Frey and Osborne (2013) could foresee these developments. In subsequent research, the authors incorporated these developments (Frey & Osborne, 2017).If anything, this strand of criticism only that the effects of technological progress are uncertain since the development of technical progress itself is not predictable.

The second strand of criticism proves to be more relevant. This criticism on the work of Frey and Osborne (2013) focuses more on the assumptions made by the authors regarding the agents within the labour market. One of the most significant points of criticism here is that the authors did not incorporate the possibility of retraining—e.g. learning new skills. Concerning this limitation, Frey and Osborne (2018) themselves acknowledge this as a limitation of their research.

Nevertheless, the overall method applied in their study—looking at skills instead of jobs— has proven to be immune for the bulk of criticisms and became a starting point for many scholars when investigating the effects of technological progress on labour demand and supply.

For their research, Frey and Osborne (2013) started by assigning specific skills to jobs. For example, the authors stated that the most essential skills of lawyers are to be creative and social intelligence. Secondly, the authors identified which skills are at risk to be automated. By looking at the skills and their risks for each job, the authors then predicted which jobs are at high risk to be automated. Staying with the previous example, the authors concluded that the risk of automation for the position of a lawyer is relatively low since, for the work of lawyers to be fully automated, engineering bottlenecks of creativity and social intelligence will need to be overcome [3].

[3] This example shows how the authors did not incorporate the development of Artificial Intelligence. Combining these effects will produce different results —the work of lawyers are at high risk of automation (Daron Acemoglu & Restrepo, 2018).

To conclude, while the results of Frey and Osborne (2013) may not be correct due to recent developments and the neglecting of retraining, its method and line of reasoning prove to be a valid starting point for further research.

3.3 The labour market: mismatch between demand and supply

As mentioned earlier, one of the biggest problems of the Dutch labour market is its mismatch between labour demand and labour supply. Furthermore, expectations are that when no action is taken, this gap will grow (PWC, 2018). However, uncertainty about the effects of technological progress may lead to ineffective strategy's which in turn can lead to cause a mismatch between supply and demand on the labour market.

For example, in their article, Ravn and Sterk (2017) establish an interesting link between employment uncertainty and firms investments. When people are more insecure about their job, consumer demand will drop. As a consequence, firms' profits decreases and these firms will eventually decrease investments. This, in turn, can lead to more unemployment. According to this theory, uncertainty about the development of the labour market may cause a feedback loop in which initial unemployment leads to more unemployment.

Another interesting mechanism that is related to the mismatch between demand and supply in the labour market is the link between over-education and the pace in which technological progress is adopted. According to Croce and Ghignoni (2012), the risk of over-education —e.g. the gap between labour demand and supply— increases when firms do not respond to developments in technology with the same pace as the expansion of skilled labour supply does. To put in other words, when firms are more cautious in hiring people with 'new' technological skills while more and more people are getting educated in this 'new' technical skills, over-education appears. This over-education will widen the gap between labour demand and supply.

As a consequence, this excessive supply of technological skill-biased labour can lead to a decrease in relative wages and, as a result, could force these technical biased skilled workers to prefer or accept a low-skilled job. However, this only can happen as long as we admit heterogeneity in workers' preferences and job characteristics Croce and Ghignoni (2012).

3.3.1 Heterogeneity in the labour market

As discussed earlier, one of the limitations of the work of Frey and Osborne (2013) is that it did not account for the retraining of skills. Nevertheless, the fact that a worker can alter its skills proves to be of great importance to analyze the influence of technological progress on the Dutch labour market (Nedelkoska & Quintini, 2018). However, the question if a worker can be retrained does not only depends on quantified variables such as its current skills set, knowledge or previous education but also on more qualitative assumptions about the worker's behaviour. Bloom et al. (2007), for example, argues that older workers react differently to the uncertainty about the effects of technological progress on their job than younger workers do. Accordingly:

> Since older workers have shorter career horizons, there is a smaller incentive for them or for their employers to invest in learning how to use the new technology. Consequently, they are more likely to stop working. (Ahituv & Zeira, 2010, p. 171)

Moreover, other research of Nedelkoska and Quintini (2018) shows that regarding the likelihood of retraining, the financial position of a worker matters. This finding also highlights an important feedback mechanism: workers with a low economic position are less likely to retrain themselves. However, when they are confronted with the risk of automation and did not gather other skills, they are at high risk of being unemployed. This unemployment will further decrease the economic position of the worker and therefore lowers the chance that the worker will retrain himself. This, in turn, lowers his possibility for a new job.

Acknowledgement of the fact that workers –even with the same observable skills and jobs– can employ different behaviour causes us to acknowledge heterogeneity in the labour market.

Taking this heterogeneity into account will not only produce different results at the individual worker level but will also influence our findings at a higher aggregate level. For example, when predicting the gap between labour demand and supply, most scholars assume that workers are equally good at finding a new job, e.g. search efficiency. However, research that allows for the possibility that some individuals have a higher propensity to form a match than others showed that this might affect the effectiveness of the matching process(Barnichon & Figura, 2015).

This notion of heterogeneity is also of importance when looking at the supply side of the labour market —e.g. the firms. The most apparent heterogeneity of firms are the differences between firms of different industries. However, even within the same sectors, we indicate essential differences between firms. One important aspect is how a firms' strategy –in our case, its HR-strategy– is formed. According to Camagni (1991), the way how firms deal with their uncertainty is also dependent on its internal learning processes (learning-by-doing, by using, by searching and, more indirectly, learning-to-Learn). Their strategy, thus cumulatively builds on tacit, firm-specific know-how and/on 'intangible' assets.

3.4 Uncertainty about technological progress

As the reader may have noticed, we often refer to the effects of technological progress as a function of the uncertainty it creates. However, we are aware of the vagueness this notion of uncertainty. To reduce the uncertainty about the notion of uncertainty, we will, therefore, discuss what we mean by uncertainty in the context of our problem statement.

3.4.1 Breaking down uncertainty

For many decades, the concept of uncertainty has long been a central component of several theories of organization and strategy (Sutcliffe & Zaheer, 1998). March and Simon (1958) stated that uncertainty is a key variable in explaining organizational behaviour. Moreover, researchers in strategic management also have considered uncertainty "to be a major factor affecting key strategic decisions" (Sutcliffe & Zaheer, 1998, p. 1). Notwithstanding this interest, in the context of the effects of uncertainty stemming from technological progress on the labour market, we have to be more specific about how this technological progress creates uncertainty. This means that we have to take a closer look at the sources of this uncertainty about technological progress.

For this, we draw on the typology of (Koopmans & Beckmann, 1957) who distinguished between primary and secondary uncertainty. Primary uncertainty reflects the lack of knowledge about states of nature, such as the uncertainty regarding the development of technological progress itself. Secondary uncertainty demonstrates a lack of knowledge about the actions of other economic actors. Primary uncertainty can thus be seen as uncertainty stemming from exogenous, while secondary uncertainty can be a threat as an endogenous variable. Accordingly, both types of uncertainty affect a firm's investment decision (Koopmans & Beckmann, 1957).

Primary uncertainty about technological progress

As mentioned earlier, technological progress itself is difficult to predict and inherently unreliable (Armstrong et al., 2014; Frey & Osborne, 2017). We refer to this as the primary uncertainty of technological progress. This primary uncertainty incorporates thus the uncertainty about the actual development of technological progress but also the uncertainty about policy measures related to technological progress –in here we treat the policy as an exogenous variable. We refer to this uncertainty about policy measures related to technological progress as Economic Policy Uncertainty (EPU).

According to different studies, EPU harms firm investments: how higher the EPU, the more cautious a firm will be in making investment decisions and hiring people.(Gulen & Ion, 2016; Hadani, Bonardi, & Dahan, 2017; Handley & Limão, 2015; Wang, Chen, & Huang, 2014).

However, we also need to be aware of the reinforcing effect of the component EPU on the total primary uncertainty. Coad, Segarra, and Teruel (2016) argue that policymakers, and thus, economic policy, is influenced by policymakers' perception of the effects of technological progress. For example, policymakers may steer toward more innovative firms because it may boost the economy. Furthermore, EPU can also increase the uncertainty about technological progress since it slows down the innovation itself. (Bhattacharya, Hsu, Tian, & Xu, 2017). We can thus identify the following feedback: uncertainty about technological progress creates uncertainty about economic policy, which in turn can create more uncertainty about technological progress.

Secondary uncertainty about technological progress

In the context of the effects of technological progress, we refer to secondary uncertainty as to the uncertainty about the actions of other economic actors in general. This secondary uncertainty incorporates thus the uncertainty of firms towards other firms, workers towards other workers, firms towards workers and vice versa. However, especially the uncertainty about firms towards other firms can create interesting patterns when dealing with the uncertain effects of technological progress. To make this more specific, we draw on the notion of competitive uncertainty introduced by Sutcliffe and Zaheer (1998):

> We define competitive uncertainty as to the uncertainty arising from the actions of potential or actual competitors, which may be either 'innocent' or 'strategic'. Competitive uncertainty derives from moves or signals by economic actors in current or future competition with the focal firm, which may be 'noisy' and difficult to grasp precisely. (p.4)

3.4.2 Feedback mechanisms between technological progress uncertainty and economic performance

At the supply side of the labour market, primary uncertainty about technological progress can create uncertainty about employment ('Will I lose my job?'). As mentioned earlier, according to the theory of Ravn and Sterk (2017), uncertainty about the development of the labour market may cause a so-called feedback loop in which initial unemployment leads to more unemployment.

Looking at the demand side of the labour market –the firms–, the problem of primary uncertainty about technology can be narrowed down to the behaviour of firms —more specifically, to their strategy on their workforce (Schaal, 2017). As mentioned before, uncertain firms will hire fewer people, which in turn can decrease aggregate demand for widening the gap between labour demand and supply. This primary uncertainty ('What will the developments of technological progress be?') is further complemented by the competitive uncertainty ('What will my competitors do?'). This notion of competitive uncertainty is crucial since it —even in the case with no primary uncertainty – lead to ineffective strategy's which in turn can lead to cause a mismatch between supply and demand on the labour market. In a competitive environment, perceived gains in productivity and thereby a reduction of Unit Labour Costs (ULC) by some firm [4] can cause other firms to fear to lose competitive advantage (Coccia, 2017; Zahra, 1996). So whether true or not, in here, the belief in certain effects of technological progress can be enough to alter firms' strategies. This closely relates to the concept of 'animal spirits' — a term used by John Mayard Keynes in his book *The General Theory of Employment, Interest and Money* to describe the confidence and the 'gut instincts' of businessmen on their future business prospects:

> Most, probably, of our decisions to do something positive, the full consequences of which will be drawn out over many days to come, can only be taken as the result of animal spirits –

[4] ULC measure the average cost of labour per unit of output and are calculated as the ratio of total labour costs to real output.

> a spontaneous urge to action rather than inaction, and not as the outcome of a weighted average of quantitative benefits multiplied by quantitative probabilities. (Keynes, 1937, p. 161-162)

To summarize, primary uncertainty manifests itself on both the supply and demand side of the labour market. This primary uncertainty is further exacerbated by external factors such as EPU and the strategic behaviour of other firms —e.g. animal spirits, which in turn creates competitive uncertainty.

3.5 Activating labour market policies

Since the event of the Great Recession, which hit the world economy in 2007-8, and its significant rise in unemployment across many countries, labour market policies were brought back to the centre stage with a particular focus on the so-called 'activating strategies'. According to Martin (2015), the OECD currently defines these activation labour market policies as strategies aiming:

> To bring more people into the effective labour force, to counteract the potentially negative effects of unemployment and related benefits on work incentives by enforcing their conditionality on active job search and participation in measures to improve employability, and to manage employment services and other labour market measures so that they effectively promote and assist the return to work. (Martin, 2015, p. 2)

In other words, these activating labour market policies *actively* stimulate the labour market and its (unemployed) workers to reduce unemployment. In his article, Dinan (2019) has provided a typology of the current activating labour market policies by categorizing them into policy levers that stimulate the demand side of the labour market or the supply side of the labour market (see Table 3.1).

As mentioned, this thesis aims to aid policy-makers by providing them with valuable insights into the effects of technological progress on the Dutch labour market. For this, we have chosen to discuss the outcomes of our study with respect to activating labour market policies. Moreover, since we did not have the resources to do a full analysis of each type of labour market policy, we see interesting opportunities for more detailed research towards the effects of technological progress on the kinds of labour market policies as given by Dinan (2019).

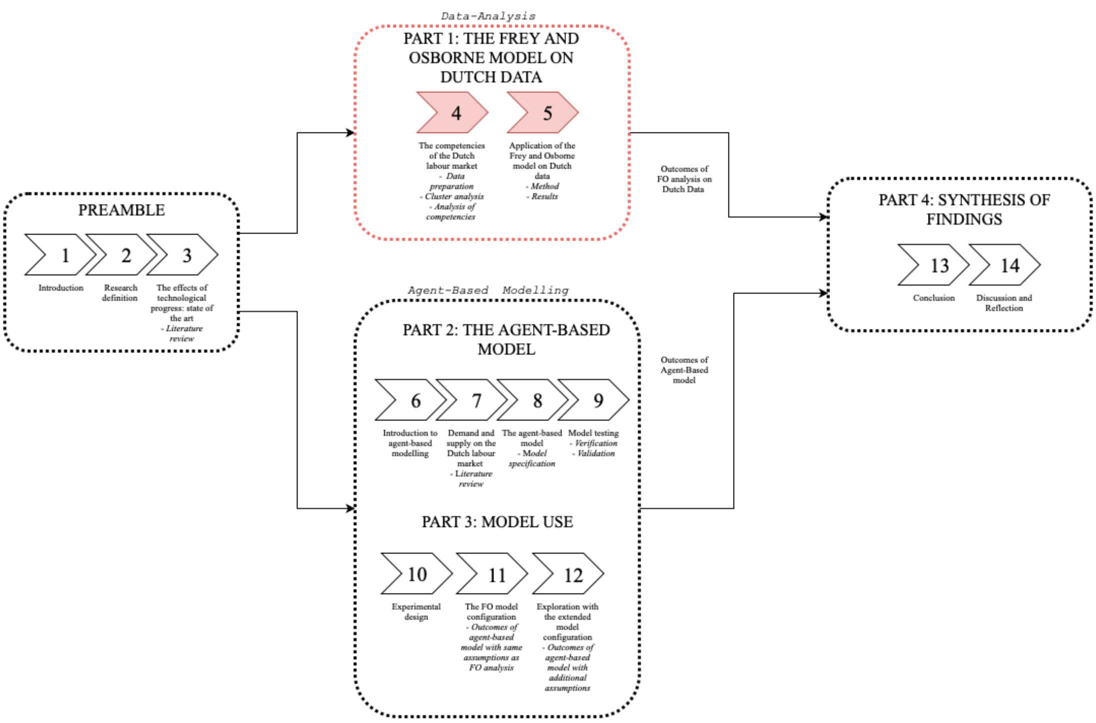
Table 3.1: Typology of activating labour market policies

Labour market lever	Demand-side	Supply-Side
Negative financial incentives	**I. Incentives to encourage employment**: Effective law enforcement for labor market and social protection policies	**II. Increased labor search incentives:** Benefit reduction, Compulsory participation, Increased time limits for recipients
Positive financial incentives	**III. Subsidized employment**: Job subsides below minimum wages; Reduced social contributions; Job creation in public sectors	**IV. Fiscal incentives**: Tax credits; Negative income tax; Financial support
Organizational human capital incentives	**V. Administrative services:** Reinforced cooperation with social partners and local government	**VI. Employment services: Counselling**; Job search programs and placement services
Concrete human capital incentives	**VII. Company training:** State financed company training	**VIII. Upskilling:** Training; Job related VET; Second chance schools

source: Dinan (2019)

I

The future of work in the Netherlands: A data-analytic approach



Chapter 4

The Competencies of the Dutch labour force

Chapter Abstract

In this chapter, we present the method and results of our analysis of the competencies of Dutch workers. This analysis is done as a preparation for the FO application on Dutch data where the automation probabilities of Dutch jobs will be calculated. Descriptive analysis of the competencies of Dutch workers has shown that basic skills such as 'active listening', 'speaking' and 'reading comprehension' are important for almost all Dutch workers. Moreover, we found that abilities that are considered to be important for technical occupations –such as 'explosive strength'– are the least important for most jobs and only workers working in technical occupations possess these abilities. This analysis has also shown that workers working in jobs that are in the category 'authors and artists' have the most difficulty to retrain to other jobs. Furthermore, to use our data for the FO application, we performed a cluster analysis on the competencies of Dutch jobs. Here, we found that we cannot threat the jobs with BRC code 1311 ('occupation group all other'), 0612 ('government officials'), 1214 ('truck drives') and 0221 ('graphic designers and product designers') as one job. Rather we have to split these jobs according to the differences in competencies between their underlying O*NET jobs. Besides the graphical designers, all these jobs could be split into a cluster with managing jobs and a cluster with operational jobs. Consequently, instead of the defined 112 jobs by Statistics Netherlands, we use 116 jobs in our analysis.

To analyze which jobs in the Netherlands may disappear due to technological progress, we first analyzed the distribution of the workforce in the Netherlands. However, the extent to which technological progress influences the rise or fall of a certain job depends on the knowledge, skills and abilities[1] associated with that job. Therefore, only looking at the current amount of workers per job is not enough(Acemoglu & Autor, 2010; Bakhshi, Downing, Osborne, & Schneider, 2017; Frey & Osborne, 2013). In other words, technological progress can raise and change the demand for knowledge, abilities and skills, while investments in knowledge, abilities and skills —e.g. retraining— can satisfy this demand. Applying this 'skill-perspective' on jobs while explaining changes over time in employment across advanced economies, has proven to be successful for many economists (Bakhshi et al., 2017). However as one may argue, the implementation of this perspective "rests on a highly aggregated and conceptually vague measure of skill, typically years of schooling" (Bakhshi et al., 2017, p. 21)[2].

[1] We will refer 'skills, abilities and knowledge' to as 'competencies'

[2] Recent accounts have sought to do a more in-depth analysis of the competencies of workers. Acemoglu and Autor (2010), for example, maps skills to labour to study changes in the returns to skills and the evolution of earnings inequality. In the same line of reasoning, earlier research of Autor, Levy, and Murnane (2003) found that jobs that are associated with non-routine cognitive and manual tasks, grew in importance while the importance of jobs with routine cognitive and manual tasks declined over time. This study was later extended by Levy and Murnane (2004) who "attribute the growth of non-routine cognitive tasks to jobs requiring skills in expert thinking and complex communication" (Bakhshi et al., 2017, p. 21).

Unfortunately, detailed research about the distribution of competencies in the labour market is far from extensive and almost non-existent when applied to the Dutch labour market. With regard to this study, we, therefore, need to construct a new data-set to analyze the distribution of competencies in the Dutch labour market. For this, we need a data-set that includes not only the number of people per jobs in the Netherlands but also the competencies that are associated with these jobs. Using this data-set, we can see how the current Dutch labour force is composed of jobs, skills, knowledge and abilities and how technological progress will influence this shape. In this chapter, we will thus answer the question of how the competencies of Dutch workers are distributed. Besides, since there is almost no information about mapping Dutch jobs with their US counterparts, this chapter aims at giving a clear-cut method in doing so.

4.1 Method

To analyze the competencies of the Dutch labour market, data about the distribution of jobs in the Netherlands need to be mapped with their associated skills, abilities and knowledge. Statistics Netherlands (CBS) provides data about the number of workers per job. However, CBS does not provide us with any information about the competencies that are associated with their jobs.

In order to extend the CBS data-set with this information, we use data from the National Center for O*NET Development (US) that provides the level of each ability, skill and knowledge per job. The reason for using the O*Net database is twofold: first, since this research uses the assumptions of Frey and Osborne (2013) as a starting point, it makes sense to use the same classification of the skills, abilities and knowledge as the authors did. This first reason provides us indirectly with the second and a more sophisticated reason to use the O*net database: the O*Net database is the most complete and useful database about the competencies of workers available. According to Bakhshi et al. (2017), one of the most significant benefits of the O*Net database is that the features are expressed in relatively natural units, which make them easier to understand and address through policy. Nevertheless, the O*Net database has some limitations. More information about the O*Net database, its advantages and its limitations, can be found in Appendix B.1.

Combining the O*Net Database with data about the Dutch labour market from the CBS is, therefore, the first step in this analysis. However, the classification of O*Net occupations is far more detailed than the classification used by the CBS. In other words: each CBS occupation has one or more O*net occupations. Since we only have data about the number of workers in the Netherlands at the aggregation level provided by the CBS, the next step will be the aggregation of O*Net data across the CBS data.

4.1.1 Mapping data

As stated earlier, the first step is to map the CBS data to their corresponding O*Net occupations. Unfortunately, there is no easy way in which the classification code used by the CBS (BRC-code) can be mapped directly to the classification code used by O*Net (O*Net-SOC Code). For this mapping, two intermediate steps are executed: The first step concerns the mapping of BRC-code used by the CBS code with its corresponding ISCO (The International Standard Classification of Occupation) code. After this step, one can link the ISCO code to its corresponding SOC (Standard occupational classification (SOC)) code. For this, the crosswalk table of ISCO with SOC, provided with the National Office of Statistics of the United Kingdom, is used. Note that it is not possible to link the BRC-code *directly* to the SOC code since there is no crosswalk table of BRC to SOC available. This ISCO-SOC classification will then be used to map the SOC-code with the O*Net SOC occupation code. For this, it uses the crosswalk table provided by O*Net. Also, in this case, O*Net only provides the crosswalk between O*Net SOC and SOC.

Table 4.1: Mapping CBS data with O*Net data: Example of BRC code 1022 'psychologist and sociologists'

BRC-code and Name	N	ISCO code	ISCO Name	SOC Code	SOC Name	ONet-SOC code	ONet SOC Name
1022 Psychologen en sociologen	69	2632	Sociologists, anthropologists and related professionals	19-3041	Sociologists	19-3041.00	Sociologists
1022 Psychologen en sociologen	69	2632	Sociologists, anthropologists and related professionals	19-3091	Anthropologists and Archeologists	19-3091.01	Anthropologists
1022 Psychologen en sociologen	69	2632	Sociologists, anthropologists and related professionals	19-3091	Anthropologists and Archeologists	19-3091.02	Archeologists
1022 Psychologen en sociologen	69	2632	Sociologists, anthropologists and related professionals	19-3092	Geographers	19-3092.00	Geographers
1022 Psychologen en sociologen	69	2633	Philosophers, historians and political scientists	19-3093	Historians	19-3093.00	Historians
1022 Psychologen en sociologen	69	2633	Philosophers, historians and political scientists	19-3094	Political Scientists	19-3094.00	Political Scientists
1022 Psychologen en sociologen	69	2633	Philosophers, historians and political scientists	19-3099	Social Scientists and Related Workers, All Other	19-3099.01	Transportation Planners
1022 Psychologen en sociologen	69	2634	Psychologists	19-3031	Clinical, Counseling, and School Psychologists	19-3031.01	School Psychologists
1022 Psychologen en sociologen	69	2634	Psychologists	19-3031	Clinical, Counseling, and School Psychologists	19-3031.02	Clinical Psychologists
1022 Psychologen en sociologen	69	2634	Psychologists	19-3031	Clinical, Counseling, and School Psychologists	19-3031.03	Counseling Psychologists
1022 Psychologen en sociologen	69	2634	Psychologists	19-3032	Industrial-Organizational Psychologists	19-3032.00	Industrial-Organizational Psychologists
1022 Psychologen en sociologen	69	2634	Psychologists	19-3039	Psychologists, All Other	19-3039.01	Neuropsychologists and Clinical Neuropsychologists

An example of this mapping is given in table 4.1. This table shows the name of the job and BRC-code provided by the CBS with the number of workers who exercise this profession —represented by column 'N'. Note that the number of workers is divided by 1000. This table also shows the difference between the aggregation level of the codes. In this example, one BRC-code refers to 12 O*Net-SOC Codes. However, we only have data about the number of Dutch workers at the BRC-code level. This means that we have to aggregate the O*Net data as far as possible. Nevertheless, we have to be aware that no relevant information is being lost.

4.1.2 Aggregation analysis

As shown by the example in table 4.1, the O*Net aggregation level differs from the level used by the CBS. This means that we have to aggregate the O*Net data about skills, abilities and knowledge. The easiest way to do this would be to take the mean of the competencies from the O*Net occupations which have the same BRC-codes. However, this may lead to losing useful information. To stay with the example provided by table 4.1, the importance of the ability 'Inductive Reasoning' could differ greatly between sociologists and clinical psychologists. By averaging this level of 'Inductive Reasoning' across this whole BRC-code, we ignore this difference. Since the levels of this abilities, skills and knowledge are of great importance for this research, we have to be careful when aggregating the data-set. This section will explain briefly the methods used for the aggregation analysis. A more detailed description of these methods and their motivations can be found in appendix B.2

Hierarchical Cluster Analysis

To aggregate the data, we begin with assessing which data can be aggregated and which data points cannot. For this, we use two different methods. The first method involves a cluster analysis to assess if we can identify distinct clusters of skills, knowledge and abilities within a BRC occupation. This is only done for BRC data points that have more than 2 corresponding O*Net occupations.

The cluster method used to aggregate the BRC data points is hierarchical clustering. According to Revelle (1979), hierarchical clustering is proven to be an effective method for forming scales from sets of items. Moreover, "comparisons with other procedures show that hierarchical clustering algorithms can be more useful for scale construction using large item pools than are conventional factor analytic techniques" (p.1).

In short, the hierarchical cluster method builds a cluster hierarchy that is commonly displayed as a tree diagram called a dendrogram. The clustering algorithm starts with each object — here an O*Net occupation — as a separate cluster. At each step, the two most similar clusters are formed into a single new cluster. When the objects are fused, they will never be separated again. To determine the similarity between two clusters, we use five distinct agglomarative methods: Maximum or complete linkage clustering[3]; Minimum or single linkage clustering; Mean or average linkage clustering[4]; Centroid linkage clustering[5]; and Ward's minimum variance method[6]. As a consequence, these different methods can result in different outcomes and thus different clusters. To determine the optimal clustering, the agglomerative coefficient is used. This coefficient "measures the dissimilarity of an object to the first cluster it joins, divided by the dissimilarity of the final merger in the cluster analysis, averaged across all samples. Low values reflect tight clustering of objects; larger values indicate less well-formed clusters". (Holland, 2017).

This cluster analysis is performed on the competencies that are considered to be important to execute the O*net occupations. According to O*Net, skills, abilities or knowledge are only considered to be important for the exercise of an occupation if it scores higher than two on a 5-point scale(Reeder & Tsacoumis, 2017). By only clustering on the important factors, we avoid false clustering due to the overestimation of factors. For example, suppose that the cluster analysis of the skills of a BRC job indicates that three clusters can be found when considering all the skills. However, these clusters are also based on the perhaps large difference between unimportant skills (ranging from 0 to 2). When these skills are eliminated, the cluster analysis may results in different clusters based on the smaller –but more important– differences between the important skills.

When the dendograms of the clusters are made, the next step is to define the optimal number of clusters. For determining the optimal numbers of clusters, several methods can be used. The elbow, silhouette and gap statistic methods are the most commonly used. The elbow and silhouette method can be seen as a direct method that consists of optimizing certain criteria. However, since we have created the data-set from scratch and we have no benchmark information about how to assess the optimal number of clusters, we use a statistical testing method that compares evidence against null hypothesis: the gap statistic method. This method is developed by Tibshirani, Walther, and Hastie (2001) and can be applied to the hierarchical cluster method. The algorithm works as follows:

1. Cluster the observed data, varying the number of clusters from $k = 1, .., k_{max},$, and compute the corresponding total within intra-cluster variation W_k.
2. Generate B reference data sets with a random uniform distribution. Cluster each of these reference data sets with varying number of clusters $k = 1, .., k_{max},$, and compute the corresponding total within intra-cluster variation W_{kb}.
3. Compute the estimated gap statistic as the deviation of the observed Wk value from its expected value W_{kb} under the null hypothesis: $Gap(k) = \frac{1}{b}\sum_{b}^{b=1} \log(W_{kb}) - log(W_k)$. Compute also the standard deviation of the statistics.
4. Choose the number of clusters as the smallest value of k such that the gap statistic is within one standard deviation of the gap at $k + 1 : Gap(k) \geq Gap(k + 1) - s_k + 1$.

(Kassambara, 2018)

[3] It computes all pairwise dissimilarities between the elements in cluster 1 and the elements in cluster 2, and considers the smallest of these dissimilarities as a linkage criterion. It tends to produce long, "loose" clusters

[4] It computes all pairwise dissimilarities between the elements in cluster 1 and the elements in cluster 2, and considers the smallest of these dissimilarities as a linkage criterion. It tends to produce long, "loose" clusters

[5] It computes the dissimilarity between the centroid for cluster 1 (a mean vector of length p variables) and the centroid for cluster 2

[6] It minimizes the total within-cluster variance. At each step, the pair of clusters with minimum between-cluster distance are merged

In this analysis, we first look for BRC-codes, where the optimal number of clusters is larger than 1. When this is the case, it means that more than one distinct cluster of O*net occupations within a BRC-code can be identified and that averaging over these occupations based on their competencies within the BRC-code, ignores the different characteristics of these clusters. On the contrary, when the optimal number of clusters of O*Net occupations within a BRC code is equal to one, it means that we can assume that there are no significant differences between the competencies of the underlying O*Net clusters. That being the case, we can aggregate all the underlying O*Net occupations to its corresponding BRC-code.

When the optimal number of clusters is determined, the actual aggregation of the competencies of O*Net occupations can thus begin. As mentioned, when the optimal number of clusters is one, the competencies of the O*Net occupations within a BRC-code are averaged. When a BRC-code has more than one optimal clusters, the competencies are averaged according to these clusters.

Aggregation of two O*Net occupations per one BRC

When the cluster analysis of the BRC codes with more than two underlying O*Net occupations is finished, we will focus on the BRC-codes that have precisely two O*Net professions. Since there are only 2 O*Net occupations, cluster analysis cannot be performed. This means that we have to find another way for which we can identify if there are relevant differences within a BRC-code. To do this, no standardized method is available. Statistical methods that check for equality of variances between groups across multiple dependent variables –such as Manova– are not sufficient for this since the characteristics of the O*Net data set leads to the violation of several assumptions such as the absence of multivariate outliers and lack of multicollinearity. This means that we have to define our own measures to assess the differences between O*Net occupations. Fortunately, we can use data from the cluster-analysis to make the assessment. The decision to split a BRC code with two underlying O*Net occupations will be based on the outcomes of the previous cluster. For each BRC-code that is split into clusters in the previous cluster analysis, the difference between the clusters will be calculated and will be used as a benchmark:

$$Dif_{brc} = \sum^{Ability} (ClusterA_{ability} - ClusterB_{ability})^2 \tag{4.1}$$

After this, the benchmark value b can be identified for which $b = min(A)$ and where $A = \{Diff_{brc}\}_{brc=1}^{N}$.

This benchmark will be used to determine if one BRC-code with two underlying O*Net occupations has to be split up in two different groups. This is done by calculating the difference between the two O*net occupations within a BRC-code using equation B.1. This difference will then be compared with the benchmark value. When the difference is greater than the benchmark value, we will divide the BRC-code into its two O*Net occupations. As can be seen, a great advantage of this method is that it relies on earlier defined criteria and thereby guarantees consistency in the aggregation progress.

When the aggregation analysis is completed, we combine the aggregated data-set with the level for each competence per BRC-code with actual data from the CBS[7][8] .

4.1.3 Descriptive analysis

To see how the competencies of Dutch workers are distributed, we perform a descriptive analysis of the competence data-set. To gain insights into the distributions, density plots of the distribution of competencies will be made. These density plots are the visual representation of the probability density function (PDF) of a certain competence. Without going into

[7] The data-set with the level of importance for each competence per BRC-code will thus be extended with the actual number of workers in the Netherlands who have that occupation

[8] This combined data-set will be further referred to as ‘the competence data-set’

too much detail about the PDF, the basic idea of the PDF is that it shows the likelihood of a variable falling within a particular range of values[9]. For example, a PDF of the skill 'reading' derived from the competence data-set will show us the likelihood for a worker that its job requires a reading level higher than two. In other words, the PDF will show us how the skill 'reading' is distributed in the Netherlands.

Besides looking at the distribution of the competencies in the Netherlands, we look at the median for each competency. The reason for drawing our analysis on the median and not the mean is that the median is less susceptible for outliers.

The last part of this analysis is the hierarchical cluster analysis of the Dutch occupations based on their competencies. By identifying different clusters, we can identify the possibilities of retraining. For this, we use the hierarchical cluster method. The number of clusters is determined arbitrarily and is based on the job aggregation provided by the CBS. As stated before, we have defined the competencies and number of workers per BRC-code. This BRC-code consists of 4 digits and represents the lowest possible aggregation level of jobs. A BRC-code can thus be aggregated by looking at its 3-digit and a 2-digit code. For example, the BRC-code 0711[10] ('biologen' or 'biologists') belongs to the 3-digit level 071 ('ingenieurs en onderzoekers wis-, natuur- en technische wetenschappen' or 'engineers and scholars mathematics and natural sciences'). This can be further aggegrated to 07 ('technische beroepen' or 'technical occupations'). The 112 unique CBS occupations (4-digit BRC-codes) are aggregated to 41 3-digit occupations and 13 2-digit occupations (CBS Statline, 2019). To summarize, we perform two cluster analyses of the competencies of Dutch workers: one with 13 clusters; and one with 41 clusters.

4.2 Results

In this section, the results of the analysis of the competencies of the Dutch worker are discussed. This section starts with presenting the results of the aggregation/clustering analysis. This is followed by discussing the results of the descriptive analysis.

4.2.1 Results Aggregation Analysis

The cluster analysis of the competencies of the O*Net occupations within one BRC-code group showed that almost all the O*Net Occupations could be aggregated to their corresponding BRC-code groups (see appendix B, table B.4 B.7, B.9). Table 4.2 shows the BRC-codes where the O*Net occupations could not be aggregated. As can be seen in the table, the cluster analysis indicated that each of these BRC groups could be split into two clusters. The column 'Clustered on' shows on which competence the BRC is clustered. For example, the cluster analysis of the skills, abilities and knowledge of the BRC-code 1331 ('beroepsgroup overig' or 'occupation group: all others') showed that both at the analysis of the skills and the analysis of the abilities, two clusters of O*Net occupations could be found. Moreover, the composition of the clusters are the same for skills and abilities –e.g. in both analyses, the same clusters were found. Staying with the same example, we see a clear distinction between the two clusters. Within the first cluster ('occupation group other -managing') the O*Net occupations that are considered to be more 'managing' occupations can be found whereas in the second cluster ('occupation class other -operational') exists of more operational type of O*Net occupations. These two clusters thus differ truly on their important skills and abilities. For more information on which skills, abilities or type of knowledge the clusters of these BRC-codes differ, see appendix B

[9]For more information about the probability density function and its formal definition see: Parzen (1962)
[10]For a description and translation of all BRC codes see Appendix A

Table 4.2: The split of the BRC codes with a large difference between the competencies of their underlying O*Net variables: 1311, 0612, 1214, and 0221

Name	BRC	Clustered on	Underlying O*Net occupations
Occupation group all others - managing	1311	Abilities and Skills	Chief Executives;Chief Sustainability Officers
Occupation group all others - operational	1311	Abilities and Skills	Nursery Workers;Farmworkers and Laborers, Crop;Farmworkers, Farm, Ranch, and Aquacultural Animals;Fishers and Related Fishing Workers; Hunters and Trappers;Laborers and Freight, Stock, and Material Movers, Hand
Government officials- operational	0612	Abilities	Environmental Compliance Inspectors;Licensing Examiners and Inspectors; Government Property Inspectors and Investigators; Coroners;First-Line Supervisors of Police and Detectives;Police Patrol Officers; Sheriffs and Deputy Sheriffs
Government officials- office	0612	Abilities	Equal Opportunity Representatives and Officers;Regulatory Affairs Specialists; Tax Examiners and Collectors, and Revenue Agents;Eligibility Interviewers, Government Programs;Court Clerks;Municipal Clerks;License Clerks
Truck drivers -First-Line Supervisors of Transportation and Material-Moving Machine and Vehicle Operators	1214	Abilities	First-Line Supervisors of Transportation and Material-Moving Machine and Vehicle Operators
Truck drivers -Heavy and Tractor-Trailer Truck Drivers	1214	Abilities	Heavy and Tractor-Trailer Truck Drivers
Graphic designers and product designers -Commercial and Industrial Designers	0221	Knowledge	Commercial and Industrial Designers
Graphic designers and product designers -Graphic Designers	0221	Knowledge	Graphic Designers

4.2.2 The skills, abilities and knowledge of Dutch workers: a descriptive analysis

Looking at the distribution of skills in the Netherlands, we see that basic skills as ‘Active Listening’, ‘Speaking’ , ‘Reading Comprehension’, ‘Critical Thinking’ and ‘Motoring’ are important[11] for almost all the Dutch workers. On the other hand, skills of the category ‘Technical skills’ are the least important for most people. Especially, the skill ‘Installation’ is considered to be important for only a small number of workers with a technical profession (see Appendix B.7.1, Table B.14). If we assume that the Dutch workers only possess the skills that are considered to be important for their job, we can conclude that technical skills are the least common skills across Dutch workers. On the other hand’, skills of the category ‘Process’ are considered to be important for almost all Dutch workers. This result seems plausible since technical skills –such as ‘installation’– are more job-specific than the more general content skills –such as ‘active listening’. In other words, it is not surprising to find that almost all the workers in the Netherlands need skills in the category ‘process’ or ‘content’ and only a few workers need technical skills.

Analysis of the abilities of Dutch workers shows that the abilities within the header ‘cognitive abilities’ are important for most Dutch workers. If we look at the distribution of the level of importance for this cognitive abilities, we found that most Dutch workers need a high level in ‘idea generation’ abilities for their profession (see Appendix B.7.2, Figure B.9

[11] According to O*Net, a competence is important for a job when the competence scores higher than 2 (on a 0 to 5 scale)

c). However, all Dutch workers need a score higher than 2 for the cognitive abilities of categories ‘attractiveness’ and ‘verbal abilities’ (see Appendix B.7.2, Figure B.8 a and Figure B.11 a). This means that verbal cognitive abilities are important for all Dutch workers to exercise their profession. Moreover, we found that the least needed abilities for the Dutch working population are: sensory abilities ‘Night Vision’, ‘Peripheral Vision’, ‘Sound Localization’ and ‘Glare Sensitivity’; and the physical ability ‘Explosive Strength’. Looking at the jobs that need at least one of these abilities to be higher than 2, we found that these are mostly technical occupations (see Appendix B.7.2, Table B.16). The total number of workers exercising professions that needed technical competencies is 17,38800 workers, which are 19% of the total working population. From this number, 18% are women, which lead us to conclude that the workers who need these abilities are mostly men.

Looking at the distribution of knowledge in the Netherlands, we see that a high value for types of knowledge falling in the category ‘Business and Management’ is important for most workers (see Appendix B.7.3, Figure B.13 b). Moreover, knowledge about ‘Education and Training’ and the national language is important for almost all Dutch workers (see Appendix B.7.3, Figure B.13 a and c). If we assume that the workers have the type of knowledge that is required for their job, we can thus assume that almost all Dutch workers have some knowledge about education and training and the national language. Knowledge types that fall into the category ‘Arts and Humanities’ and knowledge about food production are needed for the least number of Dutch workers.

Differences between the competencies of Dutch professions

In order to define the possibilities for retraining, we performed a hierarchical cluster analysis on the competencies of Dutch workers (see Appendix B.7.4). For this, we have defined the same number of clusters as the number of aggregated groups defined by the CBS. The CBS has aggregated the dutch occupations in 41 groups (3-digit BRC) or 13 groups (2-digit BRC). Cluster analysis with 41 groups found that the following occupations are clustered alone:

- 0211 Librarians and curators
- 0214 Visual artists
- 0215 Performing artist
- 0632 Police and firemen
- 0821 User support IT
- 1311 Occupational group other - managing
- 0221a Graphic designers and product designers - Commercial and Industrial Designers

Based on this result, we can say that workers in these professions have the most difficulty in retraining for another job since the competencies of these jobs are the most distinct from other professions. Nevertheless, Cluster analysis with 13 clusters resulted in no clusters with one occupation. Since the differences between these 13 clusters are bigger than the distances between these 41 clusters, we will assume that these clusters represent 13 fields of specializations. For further analysis, we, therefore, assume that workers can retrain themselves to jobs within the same field of specialization. The difficulty to retrain will be defined as the results of the 41-cluster analysis. This means that workers with job 0214, 0215 and 0221 (falling in the 021 category of ‘authors and artists’) do not only have the most difficulty for retraining, they also have the least possibilities to retrain since their cluster has the lowest number of jobs.

Chapter 5

The future of work in the Netherlands: application of the Frey & Osborne model

Chapter Abstract

This chapter presents the results of the FO application on Dutch data. In this analysis, we found that 12 % of the Dutch workers are at high risk of losing their jobs due to automation. The three jobs that have the most risk of becoming replaced by robots are '0434 accounting staff', '0431 administrative staff' and '0333 cashiers'. Moreover, we see that jobs within BRC category 12 (transport and logistics) and 04 (office and administrative occupations) are often highly susceptible to computerization. This is in contrast with jobs that fall in the BRC 2-digits category 05 (management occupations) that often have a low risk of becoming replaced by robots.

This chapter provide an answer to the question of which Dutch jobs are at risk of being automated according to the Frey and Osborne (2013) model. The first section will describe the method and explains how the Frey and Osborne (2013) model is applied to the Dutch competence database. Thereafter, the results of the analysis will be presented. The chapter ends with a discussion of the limitations of this analysis.

5.1 Method

To assess the degree to which jobs are at risk of being automated (automaticity), the FO model uses a machine-learning statistical approach. At the core of the FO's approach is the identification of the current bottlenecks engineering bottlenecks that machine-learning (Ml) and mobile-robotics (MR) developers are facing (Nedelkoska & Quintini, 2018). To identify these bottlenecks, in Frey and Osborne (2013) held a workshop for these experts. They showed these scientist/experts a list with 70 occupations and asked: "Can the task of this job be sufficiently specified, conditional on the availability of big data, to be performed by state of the art computer-controlled equipment". The occupations for which all experts agreed that all tasks could be automated were labelled 1, and the jobs were no tasks could be automated were labelled 0. FO used this data as their training set for their machine learning algorithm.

The second source of data in the FO approach are the competencies for each job given by O*Net. In their research, they focused on nine such O*Net variables that corresponded to the bottlenecks identified by interviews with ML experts. Table 5.1 shows these bottlenecks identified by Frey and Osborne (2013) and their 9 corresponding O*Net variables.

Table 5.1: the O*Net variables and their description corresponding to the identified engineering bottlenecks of Frey & Osborne (2013)

Variable	Definition
Finger Dexterity	The ability to make precisely coordinated movements of the fingers of one or both hands to grasp, manipulate, or assemble very small objects.
Manual Dexterity	The ability to quickly move your hand, your hand together with your arm, or your two hands to grasp, manipulate, or assemble objects.
Cramped workspace, awkward positions	How often does this job require working in cramped work spaces that requires getting into awkward positions?
Originality	The ability to come up with unusual or clever ideas about a given topic or situation, or to develop creative ways to solve a problem.
Fine arts	Knowledge of the theory and techniques required to compose, produce, and perform works of music, dance, visual arts, drama, and sculpture.
Social Perceptiveness	Being aware of others' reactions and understanding why they react as they do.
Negotiation	Bringing others together and trying to reconcile differences.
Persuasion	Persuading others to change their minds or behaviour.
Assisting and caring for others	Providing personal assistance, medical attention, emotional support, or other personal care to others such as coworkers, customers, or patients.

Source: Frey and Osborne (2013), Table 1.

For their analysis, they used these variables as a feature vector denoted by $x \in \mathbb{R}^9$. This feature vector was available for all the 702 O*Net occupations. In order to asses the likelihood of automation, they used a Gaussian process (GP) classifiers, a machine learning technique. The data-set with the 70 occupations and their automation labels are used as the so-called 'training data-set' for the GP classification. In their analysis, FO uses this training data to estimate the underlying latent probability of automation for a job ($P(z_*|f_*)$), assuming that the latent probability distribution is logistic:

$$P(z_*|f_*) = \frac{1}{1 + exp(-f_*)} \tag{5.1}$$

To model this probability for all the 702 occupations, they explored two broad statistical approaches: logistic regression and modelling it as a machine-learning GP classifier. After judging the standard measures for model fit –particularly the area under the curve (AUC) measure–, they concluded that the exponential quadratic (squared exponential) GP model fitted the data the best. They thus used a GP approach with an exponential quadratic covariance function to estimate the probability of automation. A more detailed description of this method and its validation can be found in Appendix C.1.

5.1.1 Application of the FO Model

For our analysis, we used the same research approach as FO. Since no expert information about Dutch jobs was available, we used the expert labels for automation as provided by FO as our training set. However, FO used the O*Net-SOC code in their classification for the occupations. This classification has two implications for our research: (a) the mapping of these O*Net-SOC codes to their corresponding BRC code; and (b) the aggregation of the FO training data to the BRC level since the O*Net-SOC occupation codes are far more detailed –and thus on a lower aggregation level– than the BRC occupation codes used in this analysis. This means that we have to use the results of the previous cluster analysis when labelling our training data according to the FO study.

Since there are more O*NET-SOC-codes for 1 BRC code, we defined for each labelled O*Net occupation the ratio of labelled O*Net occupations per BRC-code. For example, if there are 6 O*Net occupations within 1 BRC occupation and the training set of FO states that 2 of these occupations are labelled, the labelling ratio is $\frac{1}{3}$. In our analysis, we define the training set as BRC-codes with a ratio equal to or higher than 0.5. In other words, if within a BRC-code more than half of the underlying O*Net occupations are labelled by FO, we use this label for the BRC code.

Table 5.2 shows the jobs and labels used in our training-set, D. Our training-set is defined as $D = (X, y)$ where $X \in \mathbb{R}^{19\times 9}$ is the matrix of O*Net variables for the subset of the 19 training/labeled occupations and $y \in \{0, 1\}^{19}$ has the occupational labels of automation. This matrix represents our *training features*.

Table 5.2: Training data for the automation of Dutch occupations based on the training data of Frey and Osborne (2013): 1 represents automated, 0 means not automated

Jobs	Label
0333 Cashiers	1
0411 Accountants	1
0421 Bookkeepers	1
0511 Executive directors	0
0532 Logistics managers	0
0535 Education managers	0
0621 Legal experts	0
0712 Engineers (no electrical engineering)	0
0734 Plumbers and pipe fitters	0
0742 Welders and sheet metal workers	1
1013 Physiotherapists	0
1021 Social workers	0
1022 Psychologists and sociologists	0
1112 Cooks	0
1114 Hairdressers and beauticians	0
1122 Kitchen aids	1
1213 Bus drivers and tram drivers	1
0221 Graphic designers and product designers -Commercial and Industrial Designers	0
0221 Graphic designers and product designers -Graphic Designers	0

We then use classification to predict the probability of automation for all the 116 occupations. We define $X_* \in \mathbb{R}^{116\times 9}$ as the matrix of O*Net variables for all the 116 occupations, this matrix represents our *test features*. Given training-set D, which consists of the aggregated FO labelled occupations, we aim to predict z_* for our test features X_*. Just as FO, we achieve this probabilistic classification by introducing a latent function $f : \underline{x} \to \mathbb{R}$. For this analysis, we also assume that the probability is given by the formula 5.1. According to Frey and Osborne (2013), this approach allows us to use the features of the 19 occupations about which we are the most certain to predict for the remaining 97 occupations. Moreover, the algorithm provides a "smoothly varying probabilistic assessment of automatability as a function of the variables" (p.36). Using a GP classifier, this function is non-linear, "meaning that it flexibly adapts to the patterns inherent in the training data" (p.36). In other words, their approach allows for the incorporation of more complex, non-linear interactions between variables. Just as FO, we report $P(z_*|X_*, D)$ as the probability of computerization henceforth. Figure 5.1 shows indeed that the probability is non-linearly related to the nine O*Net variables selected by FO.

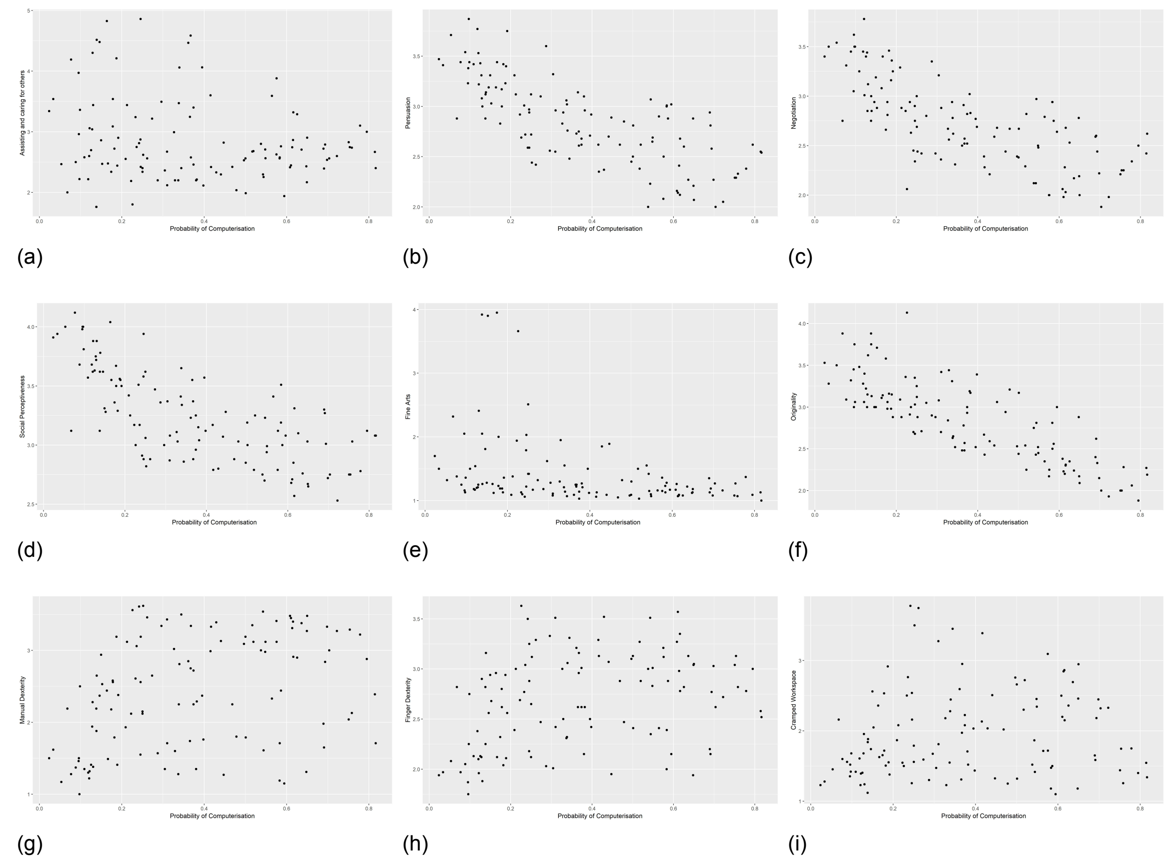


Figure 5.1: The distribution of the occupational variables as a function of the probability of computerization: each Dutch occupation is a unique point.

5.2 Employment in the twenty-first century: Results of the application of the FO model on Dutch data

In this section, we present the results of the FO model applied to occupational data of the Netherlands. Note that just as in the FO study, these results are not an attempt to forecast future changes in the composition of the labour market. These results focus on the impact of computerization on the Dutch jobs that existed in 2018. Moreover, one has to keep in mind that the results of this analysis are based on the assumptions about the automation probability made by Frey and Osborne in 2013. Recent developments –from 2013 onward– in ML and AI are not incorporated in this analysis.

The table with results of the automation probabilities of the Dutch occupations can be found in Appendix C.2.2. As can be seen from this table, 3 of the top-5 jobs that are the most susceptible to automation fall within BRC 2-digits category[1] 04: 'Bedrijfseconomische en administrative beroepen' (Office and administrative occupations). Besides the jobs that fall within BRC-code 04 and the jobs that fall in BRC-code 12 (transport and logistics) are also often highly susceptible to computerization (see Appendix C.2.2, Table C.4). This is in line with the findings of the Frey and Osborne (2013) study, who also founded that most occupations within transportation and logistics, together with the jobs in office and administrative support are highly susceptible to automation.

Looking at the Dutch occupations that are the least susceptible to automation, we see that most jobs are falling in BRC-code 05, the so-called management occupations, are the

[1] For the translation of these 2-digits categories see appendix A, table A.2

least susceptible to automation (see Appendix C.2.2, Table C.4). This is also in line with the Frey and Osborne (2013) study that stated that “most management, business, and finance occupations, which are intensive in generalist tasks requiring social intelligence, are largely confined to the low-risk category” (p.40).

Finally, we these findings for the whole Dutch labour market. Figure 5.2 shows this expected employment impact on the Dutch labour market in 2018. Just as FO, a distinction is made between low, medium and high-risk[2] occupations, depending on their probability of computerization (thresholding at probabilities of 0.7 and 0.3). As can be seen from the figure, 12 % of the Dutch workers are at risk. Most of the workers (49 per cent of total Dutch employment) are at medium risk to their jobs of being automated, and 39 % are at low risk.

Our result is a big difference with the results of the Frey and Osborne (2013) study, which stated that about 47 per cent of total US employment is at risk. A possible explanation for this difference is the fact that the Netherlands has a larger share of business and management related occupations than the United state. Another possible explanation can stem from the differences between how the labour process is organized. For example, the USA has a higher incidence of low-wage work than in the Netherlands (Appelbaum & Schmitt, 2009). With regard to our findings, this high incidence of low-wage work may indicate that more workers perform work for which only a few skills are needed and which therefore is easier to automatize than the work of Dutch workers[3]. This is also reflected in our findings that many Dutch jobs need a high level of the so-called basic-skilled, skills that are often hard to automatize.

However, since our analysis showed no linear behaviour between the variables and the probability of automation, we cannot verify possible explanations for the difference between the US and Dutch results (see Figure 5.1). For this, a comprehensive regression analysis is suggested that will reveal the real relations between the O*Net variables and outcomes while controlling for cross-correlation effects between the variables.

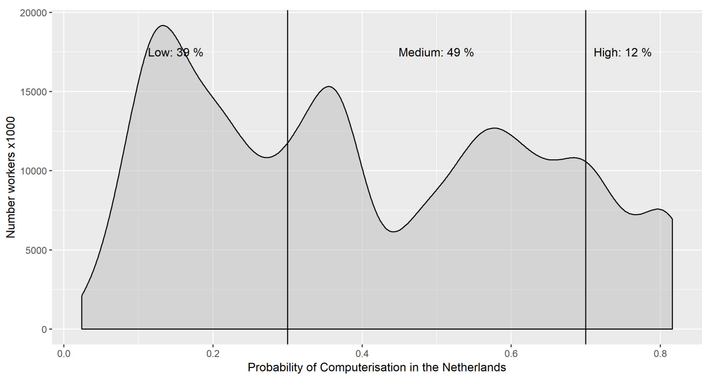


Figure 5.2: The distribution of the Dutch 2018 occupational employment over the probability of computerization, along with the share in low, medium and high probability categories. Note that the total area under the curve is equal to total Dutch employment

5.3 Limitations of the Frey and Osborne model

As mentioned earlier, the application of the FO model to the Dutch model is far from perfect and has many limitations. The first limitations are the limitations inherent to the FO model. For starters, in their model, FO did not compensate for the possibilities of retraining. The model also did not account for new, not existing jobs in the future. Moreover, their analysis is conducted in 2013. This means that recent development in robotics and AI are not incorporated. Moreover, as stated earlier, the FO model does not account for the complex

[2]In later research, we will show how this distinction made by Frey and Osborne (2013) influences their and our conclusions

[3]This discussion closely relates to the notion of the ‘Fordism production progress’ that aims at higher productivity by standardizing and breaking work into small few-skilled task.

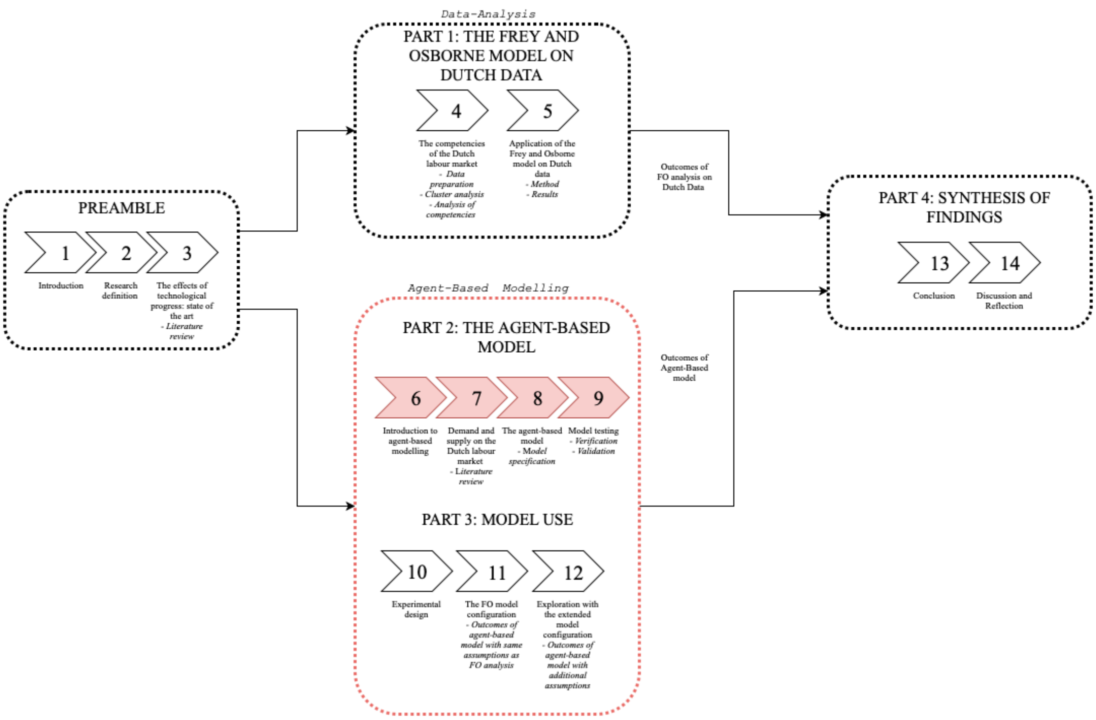
macroeconomic (feedback) mechanism and/or behaviour that may arise when confronted with (the uncertainty about) the changing labour market.

The second strand of limitations of this analysis is due to the application of the model to Dutch data. For starters, we used the same (re-coded) training data-set of FO. However, this training set is defined by US experts for US occupations. Here, we also labelled a BRC code if FO labels more than half of the underlying O*Net occupations. This threshold is set arbitrary, and sensitivity analysis showed that the number of labelled occupations is highly sensitive to this ratio (see Appendix C.2.1). Besides these issues that may arise due to the aggregation of the training data, it could be the case that Dutch experts would label the automation probabilities for Dutch jobs different than US experts would do. This may be due to differences in task description between the US O*Net jobs and Dutch BRC jobs. To overcome these limitations due to the mapping of Dutch data to US data, we, therefore, suggest conducting a new workshop with dutch AI and MR experts to formulate a training-set that could be applied to Dutch data.

To conclude, while the FO model provides us with some insights about the future of employment in the Netherlands, these insights still suffer from too many limitations. To derive more valid conclusions about the consequences of technological progress on the labour market and to account for the complex macroeconomic mechanism, we, therefore, apply a different approach: agent-based modelling. This approach will help us to formulate the specific mechanisms that may occur on the labour market for which we can adjust the outcomes of the FO application.

II

Managing the future of work: An Agent-Based Approach



Chapter 6

Agent-based modelling introduction

Chapter Abstract

This chapter provides an introduction to the development of the agent-based model. Here, the motivation for this method and some general concepts are explained. Moreover, this chapter also presents the conceptual model of the agent-based model.

This chapter will provide an introduction for the agent-based model (ABM) that is developed to gain insights in how uncertainty about technological progress influences behaviour on the labour market which, in turn, affects the real consequences of technological progress[1].

Before turning to the discussion about the development of the agent-based model, one final note has to be made. Since the goal of this agent-based model is to identify how uncertainty may affect the strategy's of firms and workers and thereby influences macro-economic performance on the labour market, a bottom-up approach is applied. By this, we mean that we will analyze how behaviour *on the lower level* generates patterns and outcomes *on the higher level*. By applying this bottom-up perspective, we distance ourselves more or less from the general macroeconomic (equilibrium) approach. This means that to see what really happens at the low level, we must be careful when applying general macro-economic assumptions. Indeed, due to the limited scope of this research, some generalizations are unavoidable, but we still try to apply the bottom-up perspective as much as possible. Consequently, this results in a different and more abstract model than one might is used to.

6.1 Modelling Question

The first step in our ABM process is to formulate the modelling question. This question differs from the research question in the sense that it is often more narrow and specific towards the development of the model. For this research, we formulated the research question 'What is the effect of uncertainty about technological progress on the Dutch labour market?'.

By imitating the study of Frey and Osborne (2013) on Dutch data, we have tried to grasps the macroeconomic effects of technological progress on the Dutch labour market. However, our repetition of the Frey and Osborne (2013) is not without limitations. Besides methodological issues –such as the re-use of their training data (see Chapter 5.3), this analysis has some more fundamental conceptually limitations which are inherent to the Frey and Osborne (2013) analysis. For starters, MacCrory, Westerman, Alhammadi, and Brynjolfsson (2014) argues that the nine variables used by Frey and Osborne (2013) are not enough to capture the diverse economic impact of technological change on skills, especially not when considering the whole magnitude of different occupations within the labour market. In this

[1]From this chapter onward, we may touch upon some concepts or definitions which need some more explanation when applied in the context of agent-based modelling. While we aim to explain these concepts as much as possible, a far more detailed description (and more) about the ABM approach, can be found in Van Dam et al. (2013).

same line of this reasoning, Arntz et al. (2017) applied a so-called 'task-based' approach and observed that within a profession, there are many tasks that cannot be automated. When this task-variation is taken into account, the authors argue that only 9 % of jobs are at risk of being wholly displaced – which is a much lower proportion than found in the original Frey and Osborne (2013). However, this task level approach still raises its own challenges. Indeed, if tasks exist in isolation, one might reasonably infer that similar tasks have similar automation probabilities. However, tasks do not exist in isolation (Bakhshi et al., 2017). By inferring the effects of automation by only looking at the characteristics of tasks within an occupation, scholars of the task-based approach take on a *reductionist approach*:

> Reductionism is the idea that system behaviour is determined by the behaviour of the system components, and is best exemplified by the idea that a thing is nothing more than the sum of its parts. [..] Reductionism is the idea that system behaviour is determined by the behaviour of the system components, and is best exemplified by the idea that a thing is nothing more than the sum of its parts. (Van Dam et al., 2013, p.23)

While explanations of the smaller components are not necessarily wrong; "they are remarkably unhelpful for explaining on the higher level system" (p. 23)[2]. In our case, looking at jobs as a sum of discrete tasks may come at the expense of quality since there is a risk of losing important coordination information which gives occupations their coherence (Bakhshi et al., 2017).

Taking this opposition to reductionism a step further, we argue that while studies such as Frey and Osborne (2013) may give an estimate which occupations are potentially automatable; they do not tell us *how* jobs actually will be automated. Instead of looking at jobs as discrete elements, we have to understand what is really happening with these jobs. As stated earlier, other factors –such as uncertainty about technological process– may influence the choices and behaviour of firms and workers and thereby the actual changes in the workforce. For example, we found that about 12 % of the Dutch workers are at risk of being replaced due to technological progress. However, if all these workers decide to retrain themselves for jobs that have no risks of being replaced by robots, this proportion would be different (in this case, zero).

This need to recognize the interactions embedded in the effect of technological progress on the labour market carries into other spheres. According to Bakhshi et al. (2017):

> Parallel to automation is a set of broader demographic, economic and geopolitical trends, which not only have profound implications for labour markets but are raising challenges for policy in their own right. In some cases, trends are reinforcing one another; in others, they are producing second-order effects which may be missed when viewed in isolation (p. 18).

Consequently, to develop a comprehensive analysis of the real effects of technological progress on the Dutch labour market, one needs to understand the underlying mechanisms that determine the actual consequences of technological progress. For this, we take our FO application on Dutch data as a starting point. While it suffers from limitations, the study of Frey and Osborne (2013) provides us with a baseline to assess the influences of technological process on Dutch jobs *ceteris paribus*. Our agent-based model will complement this analysis by identifying the underlying behaviour that eventually may (or possibly may not) generates the same outcome as in our Frey and Osborne (2013) application. For this, we formulated the following modelling question:

> *How does the reaction of agents in response to technological progress influences the proportion of jobs being automated?*

[2] The opposite, *a holistic approach*, is also not helpful in understanding the system when it is taken too far. Applying an extremely holistic approach –e.g. the whole is more than the sum of its part–, is especially unuseful in addressing current or urgent issues since "an exhaustive study of whole systems is time-consuming and difficult, if not impossible, and extremely inaccessible to anyone who has not studied the same system" (Van Dam et al., 2013, p. 24)

By substituting 'uncertainty' in the original research questions by 'the reaction of agents', we refer to the effects of uncertainty as to the effects of actions from agents. Thereby, this modelling question makes explicit what the goal of our agent-based model is: assess the impact of uncertainty about technological process and its associated actions while applying a bottom-up perspective.

6.2 Conceptual Model

A conceptual model is constructed based on the chosen system, its agents, their behaviour and interactions with each other, themselves, and the environment (see Figure 6.1). This conceptual model depicts the interrelations between these agents and their environment and the connection between labour demand and supply. In the subsequent chapters, these elements are defined and formalized into an agent-based model using the software Netlogo[3]. As can be seen in the conceptual model, our model consists of two kinds of agents, representing the supply and demand side of the labour market. Both types of agents –firms and workers– have states and rules. While the conceptual model shows that workers and companies have different type of states, the figure may presume that the internal rules of both type of agents are the same. However, as can be seen in the subsequent chapters, while the rules of workers and companies follow indeed the same pattern –expect, plan, act and evaluate– they do differ in their formalization.

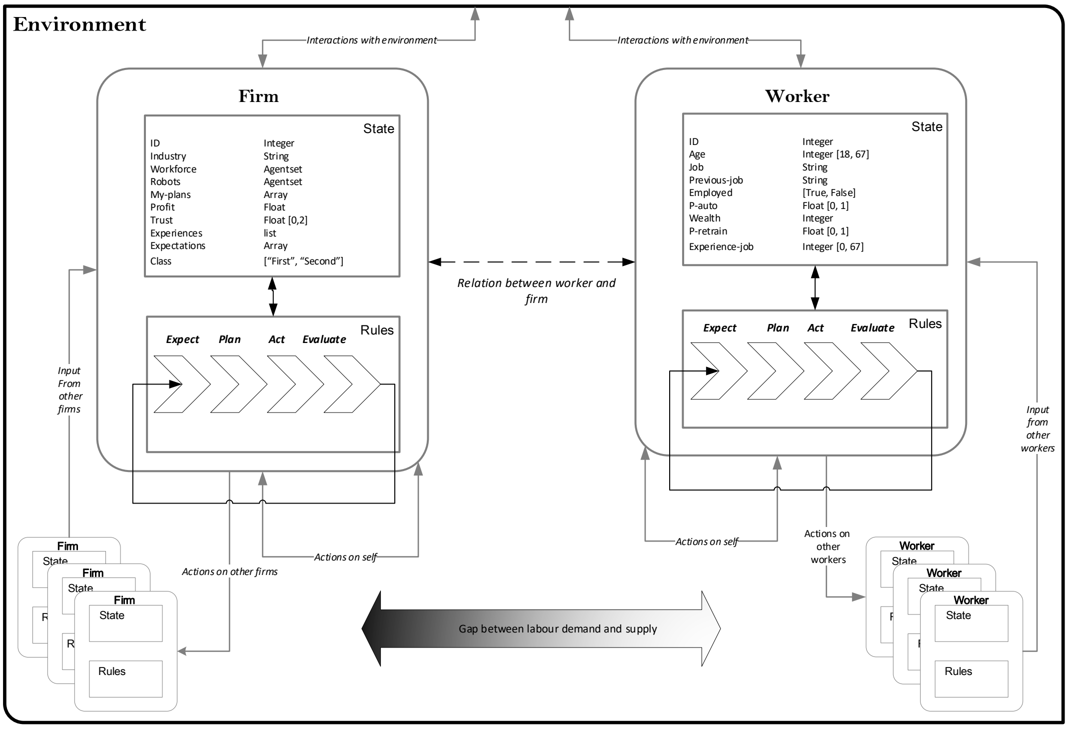


Figure 6.1: Conceptual model of the agent-based model

[3] *NetLogo* is a multi-agent programming language and modelling environment for simulating natural and social phenomena. It is particularly well suited for modelling complex systems evolving over time. Modelers can give instructions to hundreds or thousands of independent "agents" all operating concurrently. This makes it possible to explore connections between micro-level behaviours of individuals and macro-level patterns that emerge from their interactions. (Tisue & Wilensky, 2004, p.16).

Chapter 7

Modelling labour demand and supply

Chapter Abstract

This chapter discusses the theoretical background of the agent-based model. Here, our motivation to model the labour demand and supply as respectively firms and workers will be explained. Furthermore, the possibilities and factors that determine the retraining of workers will be discussed. This possibility of retraining, along with other decisions made by agents in the model, is not the result of perfect foresight and rational expectations. Hence, the notion of 'rational expectations' –or better: giving up rational expectations– will be introduced. This chapter also addresses the motivation for the economic environment and policy levers.

This chapter discusses how we have defined labour demand and supply for our model. Furthermore, it also explains how we have chosen to model the interaction between demand and supply. This chapter, thus, set the stage for the specification of the model, as done in chapter 8.

To analyze the effects of uncertainty about technological progress, we make a clear distinction in our model between labour demand –e.g. the firms– and labour supply –e.g. the workers. Modelling labour supply and demand as firms and workers can be traced back to the roots of agent-based models of the labour market (S.-H. Chen et al., 2012). To understand this choice, we must go back to two early studies that are generally not even recognized as belonging to the agent-based modelling tradition: the micro-simulation of the US economy by Bergmann (1974) and the micro-simulation of the Swedish economy by Eliasson, Olavi, and Heiman (1976). In their studies, they introduced a basic innovation with respect to the standard approach: they explicitly considered the interaction between supply and demand for labour and modelled this as the interactions between firms and workers.

This approach, however, has passed relatively unnoticed in the dynamic micro-simulation literature, which evolved along the lines as probabilistic partial equilibrium models, with limited interaction between the micro units of analysis (S.-H. Chen et al., 2012). On the contrary, the agent-based modelling approach emerged with a focus on general equilibrium feedback and interaction. However, according to S.-H. Chen et al. (2012), this was at the expense of the empirical grounding of AB models, which were developed mainly as analytical tools used to identify and study specific mechanisms that are supposed to work in real systems. "Hence, the work of Bergmann and Eliasson could be interpreted as a bridge between the (older) dynamic micro-simulation literature and the (newer) AB modelling literature, a bridge that however has so far remained unnoticed" (p.4).

For modelling labour demand and supply, we, thus, only model the interactions of two types of agents: firms and workers. This means that the effects of other factors/agents –such as policy– that can influence the dynamics on the labour market are considered to be exogenous[1]. By applying this choice, we, therefore, demarcate the labour market in two basic

[1] However, this does not mean that we consider these factors as not relevant. We only assume that other factors do not influence

blocks of workers and firms and place ourselves along with many other agent-based modellers of the labour market who have built upon the work of Bergmann (1974) and Eliasson et al. (1976) (S.-H. Chen et al., 2012).

7.1 Labour supply

In our model, we refer to the labour supply as the supply of workers. Just as done in the study of Frey and Osborne (2013), we employ a skilled-based approach to assess the effects of automation. We therefore assume that each workers i has a competence-set $x_i \in \mathbb{R}^{127}$. Instead of only the nine skills as indicated by Frey and Osborne (2013), in our model this feature vector includes all the 127 skills, abilities and knowledge as defined by O*Net for each BRC occupation (see Chapter 4).

The supply of labour is thus defined by the number of workers and the competencies of these workers. Note that we define the supply of labour as the supply of *all* workers –thus regardless of the fact if they are employed or unemployed. Chapter 4 gives an overview of the current number of workers (both male and female) and their competencies per occupation. However, since we defined the labour supply as the sum of employed and unemployed workers, we also have to determine the supply of unemployed workers. To determine this supply, we use the same distinction between voluntarily unemployed and involuntarily unemployed as the CBS (CBS, 2019). Just as the CBS (2019), we define the current working population by the sum of employed workers and unwittingly employed workers (3.7 % of the total working population).

In their analysis, Frey and Osborne (2013) estimated the effects of automation by assuming a static supply of labour –e.g the estimated the effects of automation on the labour force by assuming this labour force does not change. This assumption of a never changing labour force is a significant limitation of the Frey and Osborne (2013) study. On the contrary, we argue that workers can improve their skills and can leave or enter the labour market. The agent-based modelling method allows for such dynamic behaviour of heterogeneous workers.

7.1.1 retraining workers

As mentioned, a great limitation of the Frey and Osborne (2013) analysis is that it did not account for the possibility of retraining. To overcome this limitation, we, therefore, add the possibility of retraining as a factor determining labour supply. However, we *do* constraint ourselves by the assumption that workers only retrain themselves out of necessity. This means that workers do not retrain themselves 'just for fun'. Only confronted with an (expected) decrease in demand for their job, workers decide to retrain themselves into jobs with a higher (expected) demand.

However, this is not the whole story. When standing for the choice of retraining, other factors may influence if the worker actually would retrain itself. So what does determine this latent possibility of retraining? For starters, it depends on the extent to which retraining is possible. In order to determine this possibility, we use our cluster analysis of the competencies of Dutch jobs. We formulate the difficulty of retraining a worker of a job into another job as a function of the distance between the competencies of those jobs.

Additional to looking at the current competencies of a worker, we also take other features into account when determining if a worker would retrain himself. In other words, what are the incentives for a worker to invest in retraining? In their article, Bloom et al. (2007) argues that older workers react differently to the uncertainty about the effects of technological progress on their job than younger workers do. Accordingly:

> Since older workers have shorter career horizons, there is a smaller incentive for them or

these factors in our model

> for their employers to invest in learning how to use the new technology. Consequently, they are more likely to stop working. (Ahituv & Zeira, 2010, p. 171)

This relation between the age of the worker and the incentives to invest in retraining is also discussed by Greenhalgh and Mavrotas (1996), who states that very few workers undertake personally funded investments in skills after the age of 25.

Furthermore, other research of Nedelkoska and Quintini (2018) shows that regarding the likelihood of retraining, the financial position of a worker matters – e.g. its income. This finding also highlights an important feedback mechanism: workers with a low economic position are less likely to retrain themselves. However, when they are confronted with the risk of automation and did not gather other skills, they are at high risk of being unemployed. This unemployment will further decrease the economic position of the worker and therefore lowers the change that the worker will retrain himself. This, in turn, lowers his possibility for a new job.

Moreover, Fitzenberger, Osikominu, and Volter (2008) argues that subsidies on training for workers reduce unemployment in the long term since it increases (financial) incentives for a worker to retrain. In our model, we define this subsidy of retraining costs as a policy measure. This policy measure reflects the proportion of training costs that are subsidized by the government. However, Since this notion of 'retrain costs' is a broad measure, we were not able to find any data on the average current proportion of government subsidy in retraining costs. In our initial model settings, we will thus assume that this proportion is 0.

Another essential factor that determines the incentive for a worker for retraining is the number of experience in its currents job (Sanders, 2012). Accordingly, the longer the worker has worked in a particular job, the less willing he is to retrain himself. In the same manner, workers who are entering the labour market are more likely to retrain than workers who are almost leaving the labour market (Sanders, 2012). Note that this assumption should not be confused with the effect of age on workers' preference to retrain. These assumptions are not mutually exclusive but rather complementary.

Giving up rational expectations

Our analysis of the labour market thus extend the research of Frey and Osborne (2013) by adding the possibilities of retraining to it. As mentioned, the question if a worker even considers to retrain himself is based upon his expectations of future demand. For example, if a doctor thinks that his job will be automated and consequently, he will become unemployed, he can consider to retrain himself into a job for which he expects an increase in future demand. If he would actually retrain himself depends on his own characteristics such as his wage, experience or age. However, since there is uncertainty about the effects of technological process and the extent to with jobs are being automated, we cannot model these expectations as 'rational expectations' based on perfect foresight. According to Tesfatsion (2006):

> Rational expectations in its weakest form assume that agents on average make optimal use of their information, in the sense that their subjective expectations coincide on average with objectively true expectations conditional on this information.(p.23)

When assuming that our workers have rational expectations, we implicitly assume that their uncertainty is measurable for all workers in terms of 'objectively true' conditional probability distributions as an anchor for the commonality of beliefs (Tesfatsion, 2006). However, as we have seen from literature about the willingness to retrain of workers, these expectations can differ across workers conditioning on the same information. Rational expectations, thus, rules out the strategic multi-agent situations in which a major source of uncertainty is secondary or competitive uncertainty, –e.g. what will the other agents do?'. Since the goal of our model is to model the effects of uncertainty –both primary and competitive–, we have to give up the rational expectation assumption. The fact that our agent-based model allows for this 'irrational' heterogeneous behaviour strengthened us in the choice for the agent-based paradigm.

Concerning the retraining of workers, an essential element is the question of what kind of new job a worker should retrain. When assuming perfect information and rational decision-making, this would be an easy question: workers retrain themselves into the best possible job with the highest demand. Consequently, this implicitly suggests that workers always know the jobs with the highest labour demand. However, in reality, there is no full information (Ahituv & Zeira, 2010). This means that workers are limited in the ability to make an informed decision to which job it should retrain to. With respect to the effect of retraining on unemployment, this so-called information gap can hinder the positive effects of retraining (Hillage & Pollard, 1998).

Fortunately, by executing active policies that focus on the provision of valid information, the government can remove this barrier. According to Hillage and Pollard (1998), the government can support workers in making informed decisions by providing information since "individuals need relevant and usable labour market information to help them make informed decisions about the labour market options available to them" (p.4). In our model, we, therefore, define the provision of job information as a policy measure. Here, the government can choose to invest in actively steer the decisions of workers by providing them with useful labour market information.

7.2 Labour demand

In our model, we refer to labour demand as to firms, more specifically: to the interactions and behaviour of firms. To avoid confusion, we will use the term labour demand strictly for the demand for labour of *all* firms at a given time. While we refer to labour demand as to the demand of firms at the aggregate level, our bottom-up approach implies that we do not model this labour demand as an aggregate function over firms.

To analyze how technological progress influences the labour demand, several studies have shown that investments in new forms of technology can be seen as substitutes for hiring workers to perform routine jobs or are complements to hiring workers for non-routine jobs (Acemoglu & Autor, 2010; Autor, 2015; Autor et al., 2003; Goos et al., 2009). With respect to the effects of technological progress on labour demand, these studies all refer to the same topic, namely job polarization[2]: The simultaneous growth of jobs in a skill category (such as high-skills jobs) at the expense of jobs with a different skill category (such as middle-skilled jobs). However, by explaining this job polarization, recent literature has not explicitly taken into account the role of firms in this process (Heyman, 2016).

To bridge this gap, Heyman (2016) performed an analysis of how firms influence the current process of job polarization and thus, labour demand. Using Swedish labour market data, his study built on the work of Goos, Manning, and Salomons (2014) that showed that that overall job polarization is driven by the polarization of jobs within and between industries. To look at the role of firms specifically, he extends this analysis by addressing firms and the corresponding within-firm and between-firm components of job polarization (Heyman, 2016). Using the probability of automation and provided by Frey and Osborne (2013) low-level firm data, this analysis found that indeed, overall job polarization stems from both within- and between-firm reallocation. It argues that the proportion of jobs with routine tasks within a firm's workforce is most important in explaining within-firms job polarization.

While this study tried to make the role of firms more explicit in explaining labour demand, it does not necessarily explain how. Indeed, by performing a regression analysis, the author explained the change in labour demand by the initial workforce of the firm. However, the deeper reason behind these changes in labour demand remains unclear[3].

[2] Note that according to Autor et al. (2015) job polarization does not necessarily bring about a wage polarization.

[3] In his article, the author did give us some explanation. He states that "firms with a high initial share of routine workers have more opportunities over time to reallocate their workforce in favour of more non-routine jobs, compared to firms that initially have low shares of routine jobs (i.e. an initial high share of non-routine workers)".(Heyman, 2016, p. 6)

The article of Graetz and Michaels (2018) gives a more intuitively explanation of the effects of automation on labour demand, that is: "Firms' technology choice is simple: adopt robots when profits from doing so exceed profits from using the labour only technology by at least the fixed cost of using robots" (p .757). While more suitable for our model, this assumption implicitly assumes that all firms have rational expectations. However, the goal of our model is to assess whether uncertainty about technological process influences the effects of automation. Additionally, this assumption also assumes that the firms always know the *true* costs of production. On the contrary, this notion of 'perfect information' does not hold in the real world (Stiglitz & Rotschild, 1978). Consequently, since the aim of our model is to include uncertainty, we cannot assume perfect information.

In his discussion on firms decisions under on technological uncertainty, Hoppe (2000) argues that firms can either adopt new technology quickly (first wave) and thereby expose itself to the high cost of uncertainty, or later after the first adopters (second wave). The advantage of being a second-wave adopter translates itself into the fact that firms can gather information about the effects of this new technology and thereby reduce the costs of uncertainty. However, a disadvantage of being a late-adapter is that when the adoption of a certain technology has proven to be a success, the profitability of it decreases with the number of adopters increasing. This decision strategy about technological innovations used by firms under uncertainty can be used to formulate the capital-labour substitution decision rule of individual firms. Suppose m firms that are divided into K types of firms where $k \in \{Certain, Uncertain\}$. If a firm is certain about the effect of automation and its competitive advantage, it will use its estimations about these effects and act accordingly.

On the other hand, firms that are uncertain about the effects will wait and only react to their estimation if the results of the first wave have proven to be successful. However, if these second-wave firms wait too long, they will experience a competitive disadvantage. This more-or-less irrational behaviour of firms about their investments is also argued by Dixit (1992). In his article, he argues that firms behave contrary to the standard theory of investments. That is, instead of investing as soon a price rises above long-term average costs; Dixit (1992) observed that firms wait until prices rise substantially above long term average costs and reducing uncertainty. "On the downside, firms stay in business for lengthy periods while absorbing operating losses, and price can fall substantially below average variable cost without inducing disinvestment or exit"(Dixit, 1992, p.).

This decision rule for investing in new technologies under uncertainty has also proven to be true for other kinds of investments, such as investments in labour (Coccia, 2017). When firms are confronted with uncertainty about technological process, they will first wait to assess the outcomes from other, less uncertain, firms. The degree of which the market is competitive will then determine this reaction. That is, in a highly competitive environment – characterized by high competitive uncertainty–, firms will quicker fear to lose competitive advantage and will more quickly decide on than firms in a less competitive environment (Coccia, 2017; Zahra, 1996).

In our model, we will, therefore, apply this decision rule for firms for both investments in technology as investments in labour decisions. Here, we define investments in labour as to investments in new technologies, hiring new workers or retrain workers.

By using this decision rule, we incorporate both the competitive as the primary uncertainty that firms may have about technological progress. Additionally, this decision rule allows us to observe relevant (feedback) patterns that may follow from these uncertainties. For example, according to Bloom et al. (2007), firms that are uncertain about technological progress may slow down aggregate demand because they are too cautious –e.g. they wait too long– in making investments. This cautious behaviour of firms leads to more cautious behaviour of other firms, which in turn leads to more uncertainty. This mechanism on the higher level is elegantly captured by the lower-level decision rule of Hoppe (2000). Here, when firms are uncertain they wait, do not make investments, which in turn leads other firms to wait. This waiting of firms may create a decrease in the wages for workers and austere hiring policies.

Workers, confronted with lower income or unemployment, may not have enough money to react to technological progress and retrain themselves (Nedelkoska & Quintini, 2018). When technological progress has happened, firms find themselves unable to find sufficient workers. To conclude, this uncertainty *can* lead to a so-called 'lock-down' of the labour market in which cautious behaviour of firms has led to too little qualified workers.

Our model, thus, defines labour demand as the observed pattern created by certain or uncertain firms. The initial distribution of firms being certain or uncertain is determined exogenously. However, as firms can experience bad or good results, its uncertainty about its actions can change. If for example, a firm has experienced bad results, it will be less sure to take immediate action and prefers more to wait than before. Moreover, since the degree of cooperation determines the pace in which the real value of the response from another firm is revealed, how firms interact with each other also determines the uncertainty of firms.

7.2.1 New jobs

Another important factor related to labour demand is the creation of new jobs by technological progress. However, the exact characteristics of new jobs that will emerge are off-course unknown. Indeed, one can use the estimations of the most important skills for future job demand of Nedelkoska and Quintini (2018) to assess how that new job should look like. Nevertheless, the proportion of these new jobs for all jobs is unknown. This means that even if we can estimate the characteristics of new jobs, we cannot estimate how this will affect future demand.

7.2.2 Retraining by firms

According to Frey and Osborne (2013), 47% of the workers in the US that may be replaced by robots in the future. While this number is shocking, their article does not cover what can happen to these workers. Will these workers be fired or be retrained to a new job? As mentioned, the possibility for workers to retrain may give less chocking outcomes than the original Frey and Osborne (2013) study (Nedelkoska & Quintini, 2018). However, according to Bassi, Ludwi, and Mcmurrer (2003), when firms are uncertain about the (economic) effects of retraining, they are less willing to take these investments in retraining. This argument is in line with the previous discussion about the relation of uncertainty and investments. In our model, we, therefore, incorporate this possibility of retraining by firms. Instead of firing the worker when he's replaced due to technological progress, the firm has the possibility to retrain the worker.

7.3 Modelling the interaction between labour demand and supply

To see how uncertainty about technological process influences the Dutch labour market, we must model the interaction between labour demand and supply. For this, we formulate a so-called 'matching-function'. In general, the fundamental idea behind a matching function is "that it summarizes in a neat way the behaviour of firms trying to fill vacancies and workers looking for jobs, and relates it to job creation"(Neugart, 2004, p. 2). This matching function is used to match the firms with the workers. However, one should keep in mind that this matching function is formulated by using a bottom-up approach. This means that we do not implement macro-economic assumptions directly.

In our model, we have to make sure that our matching function includes heterogeneity on both the labour demand as the labour supply side. Extending the model of Burdett, Shi, and Wright (2001) –which defined a matching function for an agent-based model allowing heterogeneity on the demand side– Smith and Zenou (2003) defined a matching function that allowed for this heterogeneity on both sides. In their model, jobs having different skill requirements and workers have different skills. By doing this, the authors explicitly allow for skill mismatch. This matching function, however, matches workers directly on their skills: workers have no choice between firms that offer the same jobs which match their requirements. Since a study towards the preferences of workers is beyond the scope of this analysis, we have nevertheless chosen to use this matching function. However, we suggest that the expansion of the study with the preferences of workers would be an exciting topic for further research.

While this matching function is straightforward in its core, we believe that this simplicity allows us to model the effect of uncertainty about technological progress as direct as possible. By keeping the aggregated model as simple as possible and using a bottom-up approach, we can see how technological progress influence the Dutch labour market[4].

7.4 The economic environment

As explained in the previous sections, the labour market is modelled as the interaction between firms and workers through a matching process. Firms post vacancies, including the type of jobs and associated skills. Workers with the lowest investment costs for firms qualify themselves for the job and will be matched to the firms. Workers who do not succeed in finding a job will be unemployed.

The labour market is thus modelled as a decentralized market. In contrast to a centralized market, there is no central location through which these transactions take place. This does not mean, however, that there is no central mechanism that influences the behaviour on the market. According to Autor et al. (2015), the degree of trade exposure within sectors reduces overall employment (resulting from technological progress) and shift the distribution of employment between sectors:

> Local labour markets with greater exposure to trade competition experience differentially large declines in manufacturing employment, with corresponding growth in unemployment and non-employment. The employment decline is not limited to production jobs but instead affects all major occupation groups, including a notable decline in managerial, professional and technical jobs. (p.644).

According to Gerritsen and Høj (2013), this is also true for the Netherlands. Here, the labour market has benefited from trade and skilled-biased technical change over recent decades, as reflected in higher economic growth, low and declining unemployment and modest structural mismatches. Accordingly, one of the biggest indicators of this notion of trade on the labour market is labour market flexibility. Gerritsen and Høj (2013) argue the high rigidity of the Dutch labour market hinders the labour market to adapt well to changes in the labour market due to technological development and globalization. This argument is shared among many mainstream economists who believe that the high social protection on the labour market, along with low labour market mobility, prevents a smooth adaption of the labour market to external shocks and can lead to more unemployment. However, these effects are not empirically confirmed, and other scholars argue that the effects of labour market flexibility on unemployment are the opposite or at least ambiguous (R. B. Freeman, 2005; Storm & Naastepad, 2012).

In order to develop meaningful insights into the effects of technological progress on the Dutch labour market, a measure of labour market flexibility is incorporated. Here, labour market flexibility reflects the ease of firing for firms. By including this policy measure, we can see how a labour market with flexible contracts and low employment protections benefits the adaptability of the Dutch labour market to technological progress.

To summarize, our agent-based model will model the labour market as a decentralized economy in which firms and workers interact. In this market, firms have the ability to retrain

[4] A comprehensive overview about the macro-economics impacts of technological progress is provided by the thesis of (Spaanderman, 2018). Besides the results of his own model, he presented a full review of the state of the art literature about the effects of skill-biased technological change.

workers, hire workers or replace workers with robots. This will be influenced by the incentives of workers for retraining, the level of uncertainty –of both workers and firms– and policy measures as the provision of information by the government, labour market flexibility and retrain subsidy.

Chapter 8

The Agent-based model

Chapter Abstract

This chapter presents the specification of the agent-based model. This chapter starts with the discussion of the model narrative, which explains the base narrative of 'expect, plan, act and evaluate' of all agents in the model. Moreover, this chapter also discusses the motivation and specification for modelling technological progress as a logistic growth curve. Furthermore, we discuss how we have specified other global variables, such as aggregate demand and policy levers. This chapter also addresses the implications of some important assumptions that are made in the model.

After identified the possible interactions and behaviour on the labour market, an agent-based model is developed. In this chapter, the specification of the model will be discussed. Moreover, since we have chosen to apply the bottom-up perspective as much as possible, this chapter starts to explain the model from the perspective of the agent itself in a so-called model-narrative.

8.1 Model narrative

With respect to the agent-based modelling paradigm, the model narrative can be understood as "an informal description of the generative theory of the system under study, leading to emergent patterns we are interested in exploring. For an agent-based model, the behaviour of each of the agents can be captured in a story which explains which agent does what with whom and when" (Van Dam et al., 2013, p. 88). As explained in chapter 7, we define the labour demand and supply side by modelling two types of agents: firms and workers. While both types of agents are highly dependent and connected to each other, their model narrative is developed more or less separately from each other. Nevertheless, as already introduced in the conceptual model in chapter 6, their rules are developed according to more or less the same pattern: expect, plan, act and evaluate. For both firms and workers, this pattern is repeated every tick.

Figure 8.1 shows the model narrative for both workers and firms. The subsequent sections will explain these phases for both types of agents –workers and firms– in more detail.

8.2 Specification of the model

This section will give the specification of the model. It will start with our demographic and time specification, followed by the specification of agents. Furthermore, the specification of the global variables and most important made assumptions will be discussed.

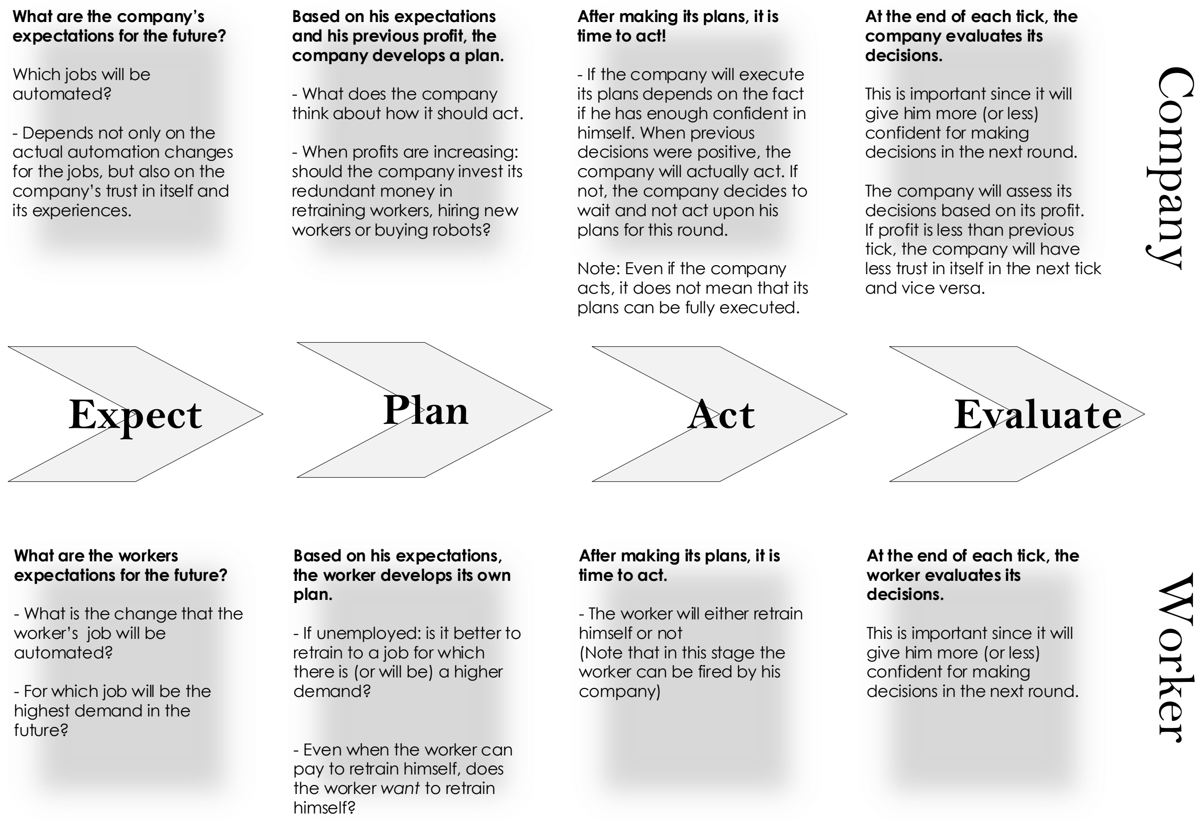


Figure 8.1: Model narrative for firms and workers

8.2.1 Time and space

As mentioned earlier, the model goal of this research is to gain insights into the final effects of the behaviour of firms and workers in the Netherlands due to technological progress uncertainty. Our application of the FO model on Dutch data provided us with a good starting point. However, in their analysis, Frey and Osborne (2013) did not explicitly specify a time interval for these probabilities. In other words, our FO probabilities of Dutch jobs being automated, are the probabilities of automation for some undefined point in the future. Nevertheless, according to Arntz et al. (2017), the automation probabilities of the FO study, are specified for the next 10 to 20 years. For our base model, we, therefore, have chosen to model the effects within 20 years with a temporal resolution of the model of one year. In other words, our model will run 20 ticks, where each tick represents one year.

As explained in the subsequent section, the conditional probability of a job being automated within a tick depends on the time interval (or end-time) of its cumulative probability. Consequently, we expect the conditional probability of a job being automated at some point in time, depends on the choice of the length of the time interval. For this reason, additional analyses with the length of the time interval varying from 10 till 20 are performed (see Chapter 9).

The model aims at representing, albeit in a very simplified and stylized way, the labour market of the Netherlands. To reflect this, we use labour market statistics as provided by the CBS for our initial set-up. However, since we aim to make an abstract model, we have chosen to only implement the real distributions of workers (and their jobs) within industries. This is done by linking UWV[1] data – the number of jobs within a sector– with CBS data about the

[1]UWV (Employee Insurance Agency) is an autonomous administrative authority (ZBO) and is commissioned by the Ministry of

number of workers per each job in the Netherlands. By linking this data, we were able to get a distribution of employment in each Dutch sector.

Since we model the labour market for 20 years, we have to incorporate a demographic structure. For this, we use data of the CBS about the working population that divide workers within each job into the following age groups:18-25; 25-35; 35-45; 45-55; 55-65; and 65-67. Note that in our model, we only model the working population from age 18 to 67 (the Dutch retirement age). This data about the generation of workers for each occupation can be found in appendix D.1.1.

In our model, we assume that when a worker turns 67, he will retire and will leave the model. For each retired worker, one[2] a new worker with the age of 18 will be created. However, the job of this new worker is determined randomly and does not depend upon the job of the replaced retiree.

In addition to this demographic structure, the agents are linked by a socio-economic network reflecting firms and workers. Within each industry, five firms are created that have a workforce according to the distribution of their sector. Every firm, thus, has the same distribution of jobs within his workforce. This assumption enables us to avoid strange outcomes, –where, for example, a firm within the financial sector only has 20 cleaners within its workforce.

Furthermore, in our model, workers depend on their retraining choice for a job on where they believe the highest demand is. We, therefore, create one working network for each worker. By using his network, the worker can determine which jobs are the most successful in the future. This working network consists of a worker's (former) colleagues. Moreover, in choosing a new career, the worker will also depend on his choice upon general information about job demand. This 'general information' will be modelled in the form of a global policy variable that calculates for each tick the number of new vacancies per job. For our initial set-up, we assume that the worker will assess both forms of information –network and general– the same.

8.2.2 Agents

Figure 8.1 shows the model narrative of our agent-based model. In this section, we will specify this narrative in more detail. This section discusses the most important rules for both firms and workers. The complete specification of the agents can be found in Appendix D.

Firms

As can be seen in figure 8.1, at the beginning of each tick, the firm starts to make its expectations about how technological progress will develop. These expectations are not only based on the actual automation probabilities (as provided by our application of the FO analysis) but also depend on the firm's trust in himself. For specifying this 'self-confident' of firms, we use a variance of the so-called 'dynamic trust formula' as developed by Jaffry and Treur (2013):

$$T_c(t + \Delta t) = Tc(t) + (X_c(t) - (T_c(t)) * \Delta t \tag{8.1}$$

Here, $T_c(t)$ and $X(t)$ are trust level of a firm c in itself and the experience level given by the firm at time in point t. Both values $T_c(t)$ and $X_c(t)$ are in the interval $[0, 1]$. "Where 0 value for $X_c(t)$ and $T_c(t)$ means absolute bad experience and no trust at all while 1 for $X_c(t)$ and $T_c(t)$ means absolute positive experience and absolute trust respectively" (p.4). The experience X of a firm is defined by his proportion of good decisions over all his decisions made.

Social Affairs and Employment (SZW) to implement employee insurances and provide labour market and data services.(UWV, 2018)

[2]Since we do not parameterize the whole labour market, we have assumed the replacement ratio retired worker/new worker to be 1. It would, however, be an interesting topic for further research to take the effects of other replacement ratio's into account.

However, to use this trust value in our model, we scale this trust value to the interval $[0, 2]$. By using this scale, a firm can modify control expectations about the automation of jobs for its own trust-level:

$$Exp_c^{job}(t + \Delta t) = P_{t+\Delta t}^{job}(Y = 1) * T_c(t + \Delta t) \tag{8.2}$$

Here, $P_{t+\Delta t}$ depicts the chance that a job will be automated in the time $t + \Delta t$. This chance per job is determined globally and is the same for each worker and firm. However, by multiplying it by its own trust-level, each firm develops its own different expectation about this change.

When each firm has made its expectations about the automation change of job, the firms will develop a plan. These plans are based on the money each firm has to spend and its expectations about the automation of jobs. The plan of a firm consists of the following elements:

1. To-hire: *The workers that the firm wants to hire*
2. To-replace: *The workers that the firm wants to replace for robots (and thus fire)*
3. Robots-to-buy: *The robots that the firm wants to buy*
4. To-retrain: *The workers that the firm wants to retrain*

Algorithm 1 shows the pseudo-code of the algorithm that is followed by the firms to make their plans.

```
while money-to-invest < 0 do
    check replace;
    for each job → job do
        if expected-unit-costs-robot < unit-costs-worker + fire costs then
            set worker to my-plans:replace;
            set money-to-invest money-to-invest - robot-costs;
            set robot to my-plans:robots-to-buy;
            if unit-costs + retrain costs < unit-costs-robot then
                set worker to my-plans;retrain;
                set robot to my-plans: robots-to-buy;
            else
                do nothing
            end
        else
            check hire or buy robot;
            if unit-costs-worker < expected-unit-costs-robot then
                if any unemployed workers with [job = job] then
                    hire worker;
                    set money-to-invest money-to-invest - wage-worker ;
                    set worker to my-plans:to-hire
                else
                    do nothing;
                end
            else
                set robot to my-plans:robots-to-buy;
            end
        end
    end
end
```

Algorithm 1: Algorithm for making investments plans by firms

When each firm has developed its plan, it will act according to his plans. This means that a firm can replace, hire and fire workers and/or can buy robots. However, even when each firm has made plans, this does not mean that a firm will act upon his plans. As revealed in the literature about behaviour and investment decisions of firms under uncertainty, an uncertain firm will prefer to wait than act. In our model, we consider firms that have a trust-level lower than one as uncertain and therefore classify them as 'waiters'. On the contrary, firms that have a trust level higher than 1 are considered as certain and are classified as 'first movers'.

However, it may happen that a firm cannot act upon his plans even if he is a first-mover.

if *trust-level* >= *1* **then**
  | set class "first-movers"
else
  | set class "waiters"
end
if *class = "first-mover"* **then**
  | execute-plan;
else
  | dismiss plan;
end

Algorithm 2: Algorithm of the acting process of firms

This is the case when a firm wants to hire certain workers, and there are no more workers left to hire since other firms have hired these workers. In that case, the firm will not hire these workers and will have a bad experience.

At the end of each tick, the firms will evaluate their decisions. This evaluation results in bad or good experiences from the firms. As stated earlier, a firm can have a bad experience when it wanted to hire certain people but was unable to do this. Other bad experiences for a firm are:

- When current profit is less than the profit of the previous tick (decrease in profit)
- when the firm anticipated on an automation event to happen, but this event didn't happen
- when the firm did not expect an automation event to happen, but this event did happen

On the other hand, a firm has a good experience when:

- profit had increased relative to the firm's profit last year
- The firm anticipated an automation event to happen, and this event indeed did happen.
- the firm did not expect on an automation event to happen, and this event indeed didn't happen

For each of these evaluation points, the firm will add his experience to a list. Here, bad experiences are represented by zero's and good experiences are represented by a one. Using this updated experience list, the firm will calculate its new trust for the next tick. The complete specification of firms can be found in appendix D.1.2.

Workers

Having explained the narrative of the firms, we will now discuss the narrative of the workers. In our model, we only make expectations for unemployed workers. This is because we assume that workers who are employed are happy with their work and do not consider retraining. Moreover, as explained earlier, we do not consider workers who want to be unemployed out of the free will and therefore take time off to re-educate themselves.

At the beginning of each tick, an unemployed worker will thus make expectations about future job demand. He will make expectations about his own job (will I be automated -if not already?) and future job demand (where is the highest demand?). This future job demand will be determined by what the worker expectations are about which jobs in the future will have the highest demand. This job demand expectation of a worker, in turn, depends on the extent to which he uses government information or information derived from his own social network.

When the worker has made its expectations, it is time to make a plan. This plan depicts if the worker wants to retrain and if so, to what job. If a worker plans to retrain depends on multiple elements. For starters, this depends on the training costs. These training costs, $C_{training}$, are determined by multiplying the absolute distance between the skills of a worker S_i and the skills required $S_i^{required}$ for the new job. This 'skill bridge' is then multiplied with the unit-training costs C_{unit}. Finally, these costs are multiplied with the factor $1 - subsidy$, the factor that represents the proportion of training costs that are not paid by the government and has to be paid by the worker.

$$C_{training} = \sum_{i=1}(|S_i^{required} - S_i|) * C_{unit} * (1 - Subsidy) \tag{8.3}$$

The worker will then decide if these retrain costs will be worth it by determining the net costs and benefits (NCB). This value will thus depends on the actual training costs and his expectation about the automation probability of its job:

$$NCB = (p_{auto} * profit_{retrain}) - ((1 - p_{auto}) * train - costs) \tag{8.4}$$

However, even if these costs are less than the wealth of a worker, the worker still has to decide if he wants to retrain at all. As discussed in chapter 7, this probability of retraining depends on the age and work-experience of the workers. In our model, we assume that both of these relations are linear. Moreover, we consider these changes to be independent from each other so that $p(age|jobexperience) = p(age)$ and $p(jobexperience|age) = p(experience)$ and thus $p(age \cap experience) = p_{age} * p_{jobexperience}$ This probability of retraining, $p(retrain = yes)$, is calculated by:

$$p_{retrain=Yes} = p_{jobexperience} * p_{age} \tag{8.5}$$

where:

$$p_{jobexperience} = 1 - \frac{1}{18 - 67} Y_{jobexperience} \tag{8.6}$$

$$p_{age} = 1 - \frac{1}{18 - 67}(Y_{age} - 18) \tag{8.7}$$

At the end of the planning period, the unemployed worker will act and either retrain himself or not. During this phase, firms also act. This means that the worker can be hired or not. This result (did my decision lead to hiring?) will be evaluated in the evaluation phase at the end of the tick. The complete specification of workers can be found in appendix D.1.1.

8.2.3 Global variables

In this section, the specification of our global variables will be discussed. This specification can also be found in appendix D.2.

Aggregate demand growth

Since we aim to apply a bottom-up perspective as much as possible, we refrain from implementing macro-behaviour as much as possible. Nevertheless, we have specified a global variable that computes the aggregate-demand. This aggregated-demand is then used to control for the firm's profit. The somehow practical reason for using this 'macro-economic variable' to specify the 'micro-economic' profit of firms, stems from the limited time and scope available for this thesis. Indeed, a sound study towards demand, output and profit of firms would be better than using one predefined function. However, this would imply to develop a micro-economic agent-based model which incorporate these mechanisms of supply and demand.

Given the time available, this is not possible and reaches far beyond the scope of this study.

To calculate the firm's profit, we thus use aggregate demand. More precisely, we use the increase in aggregate demand, r_{ad}, to determine a firm's revenues.

$$profit = (output * r_{ad}) - costs \tag{8.8}$$

Here, the aggregate demand growth is determined by the ratio of unemployed and employed workers. In our model, we assume that workers who are employed have a full demand, which is equal to 1. This means that when all workers are employed, the aggregate demand will be the highest. Unemployed workers, however, do not have full demand. Their level of demand is thus lower than 1. Moreover, the level of demand for unemployed workers depends on the number of years that they are unemployed. In the Netherlands, when a worker is less than two years unemployed, it qualifies for an unemployment benefit that constitutes 70% of its last earned wage (Juridisch Loket, 2019). After two years, the employed worker receives a general unemployment benefit that is independent of its last earned wage. In our model, we, therefore, specify the demand of workers who are less than two years unemployed as 0.7 and 0.5 for workers are longer unemployed.

The aggregate demand ad_t is calculated by summing the individual demands of the workers. The aggregate demand growth r_{ad} is calculated as followed:

$$r_{ad} = 1 + \frac{ad_t - ad_{t-1}}{ad_{t-1}} \tag{8.9}$$

Automation chance

In our model, jobs are being automated independent of the expectations of workers and firms. Whenever a job is automated is considered as a one-time event. This means that we do not consider gradual automation of jobs. A job is either fully automated or not. This assumption is consistent with the assumptions of Frey and Osborne (2013) study, that also does not consider jobs being partly automated.

In our model, we use the automation probabilities for each job as derived from our application of the Frey and Osborne (2013) study. However, these probabilities only indicate the probability of automation after a certain period. We thus consider these probabilities as *cumulative probabilities*. Or in other words, the FO probability of a job being automated after an interval of T is the sum of all his *conditional* probabilities during this interval. To determine if a job is being automated at a specific time, we, thus, use the conditional automation probability of this job. This conditional probability can be seen as the probability that a job is being automated on time t *given* the fact that this job has not been automated before.

In order to determine these conditional probabilities, we use the model of technological change as developed by Fisher and Pry (1971). This model does not assume that technological progress evolves linearly but instead has the shape of an s-curve. Accordingly, "experience shows that substitutions tend to proceed exponentially (i.e.., with a constant percentage annual growth increment) in the early years, and to follow an S-shaped curve" (p.76). By defining our conditional probabilities about the automation of jobs, we will use this same assumption.

We, therefore, define the cumulative change of a job being automated, p_j^{cum} as the logistic function:

$$f(x) = \frac{L}{1 + e^{-\kappa(x-x_0)}} \tag{8.10}$$

Here, L can be seen as the curve's maximum value – e.g. the cumulative automation change of a job. The value of κ determines the logistic growth of the curve. Experiments with different values for κ showed that the model is the most stable at a κ of 2 (see Appendix F.4)[3]. We,

[3]Nevertheless, sensitivity analysis for the value of κ showed that the model is not sensitive for changes in κ (see Chapter 9.2.4)

therefore, use the value 2 as our logistic growth rate. The value x_0 is the Sigmoid midpoint of the curve. Since our model will run for 20 years, this value will be 10 Note, however, that when running the model for a longer period, this value has to be modified. Figure 8.2 shows an example of the cumulative and conditional change of being automated for the job ”0434 Accounting staff”.

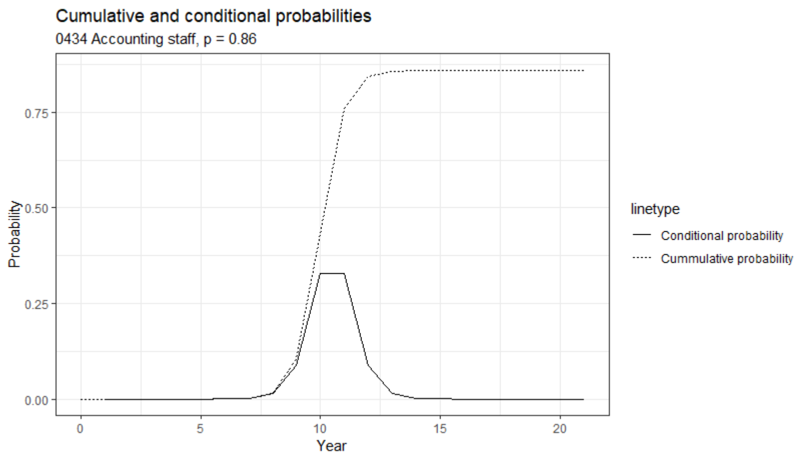


Figure 8.2: Example cumulative and conditional automation probability

During each tick, for each job will be determined if this job is being automated. This will be done by using conditional probabilities. To summarize, the cumulative automation probability can be interpreted as the probability of a job being automated after a certain time interval. The conditional probability, $p_{j,t}^{con}$ can be interpreted as the probability of a job being automated at a certain time. In our model, we assume that agents make their expectations for a specified time interval. The ‘foresight parameter’ represents this time interval. Both firms and workers have this foresight-variable. When taken this foresight variable into account, the conditional probability of a job being automated *within this foresight time-interval* is specified as:

$$p_{j,t}^{con} = \frac{p_j^{cum}}{1 + e^{(-t-foresight)-p_j^{cum}}} - \frac{p_j^{cum}}{1 + e^{-t-p_j^{cum}}} \tag{8.11}$$

8.2.4 Policy measures

In chapter 7, we have defined three important policy measures that may influence the outcomes of the model. The first measure is the extent to which the government actively provides information about labour demand. As explained in section 8.2.2, workers depend their choice of retraining to which they believe is the best option for retraining. This belief depends on their expectation about technological progress and the proportion of information gathered through their network and from the government. Initially, we set the weight of both information sources on 0.5, meaning that both information sources contribute the same in the decision making progress of the worker. However, if the government decides to intensify its information policy, the relative contribution of government information will be higher. On the contrary, when the government decides to relax its policy, the proportion of government information will be lower. Table 8.1 (a) shows the different information levels used in the model.

The second policy measure is labour market flexibility. Here, we define labour market flexibility as the extent of how easy it is to fire employees. In a total flexible market, there will be no barrier for firms to fire employees. On the contrary, in a full rigid market, the barrier for firms will be the highest. In our agent-based model, we have translated this barrier to costs, more specifically: the costs to fire a worker. These costs depend on how long the firm is obliged to pay the worker after his resignation. In the Netherlands, when a worker is fired without legitimate reasons[4], the firms are obliged to pay the worker the amount left on his contract. Firing costs, thus, depends on the length left on the contract of the worker.

In our agent-based model, we, therefore, define labour market flexibility as the proportion of flexible contracts to fixed-term contracts. Here, it is cost-less for firms to fire workers with a flexible contract, whereas the firm is obliged to pay the rest of the wage left when firing a worker with a fixed-term contract (of one year).

According to CBS, at the end of 2018, approximately 27 % of the workers had a flexible contract compared to 73 % of the workers who had a fixed contract. Table 8.1 (b) shows the defined flexibility measures for our agent-based model. The percentage of flexible contracts determines the proportion of flexible contracts among newly hired people. When a worker has a flexible contract, it is cost-less for the firm to fire him after 1 tick (year). When, however, the contract is fixed, the firm has to pay the worker the wage of 1 year[5].

The third policy measure is the extent to which the government subsidized the training costs of the worker. We define this measure as a proportion of training costs paid by the government. To this end, a value of 0.25 means that the government will pay 25% of the retraining costs of the worker. As already mentioned, we do not have any data on the current proportion of retraining costs paid by the government. We, therefore, assume that initially, the government does not subsidize any training costs. Table 8.1 (c) shows the specification of the retrain subsidy policy measure.

Table 8.1: Polices measures

(a) Information policies

Information	Weight government information
None	0
Low	0.25
Medium	0.50
High	0.75
All	1

(b) Labour-market flexibility policies

Flexibility	% Flexible contracts
None	0
Low	25
Medium	50
High	75
All	100

(c) Retrain Subsidy

Subsidy	% paid by the government
None	0
Low	25
Medium	50
High	75
All	100

8.2.5 Assumptions

As with almost every model, we cannot avoid making assumptions about some elements in the model. As explained earlier, the aim of this agent-based model is to *gain insights* in how the behaviour of workers and firms under uncertainty alters the outcomes of the FO automation probabilities. Given this modelling goal and the time available for this study, we, therefore, have chosen to keep other economic factors as abstract as possible. By modelling this abstract version of the economy, we can see how the uncertainty about technological process influences the behaviour of firms and workers and thereby the 'true' automation probability. Note, however, the quotation marks around the word true, indicating our aware-

[4] In our model, we assume that all workers are fired without legitimate reasons. According to article 6:681 DCC (Dutch Civil Code or Bugerlijk Wetboek), legitimate reasons for firing is must be urgent and highly convincing of nature, such as fraud and theft. We consider resignation due to replacement by robots not as a highly convincing or urgent reason.

[5] In our model, we assume that firms make expectations for the next year. Workers who thus are hired in tick t are thus hired for tick $t + 1$.

ness that our model is off-course also limited in its ability to reveal the true probability of automation. To cite a famous statistician:

> Now it would be very remarkable if any system existing in the real world could be exactly represented by any simple model. However, cunningly chosen parsimonious models often do provide remarkably useful approximations. For example, the law PV = RT relating pressure P, volume V and temperature T of an "ideal" gas via a constant R is not exactly true for any real gas, but it frequently provides a useful approximation and furthermore its structure is informative since it springs from a physical view of the behaviour of gas molecules.
>
> for such a model, there is no need to ask the question "Is the model true?". If "truth" is to be the "whole truth" the answer must be "No". The only question of interest is "Is the model illuminating and useful?". (Box, 1979, p. 74)

To conclude, due to the scope of our study and our modelling goal, we have made some strong assumptions in our model. While our model will thus never be 'true', we believe that even with the incorporation of these abstract assumptions, our modelling goal can be achieved: gain insights in the effects of uncertainty about the technological process.

Price, wage, productivity and unit-costs

To calculate the costs, revenues and profit of firms and the wealth of workers, we have made assumptions about the variables that determine these factors; Price, wage, productivity – both workers' and robots'– and unit-costs. To keep our model as abstract as possible, we have normalized these factors.

In our model, each worker has productivity, λ, of 1 unit of goods per tick. The wage of each worker, w, is also 1 per tick. Following equation 8.12, each worker has a unit costs (costs per unit), UC, of 1. In order to formulate the unit price of goods, UP, we specify a markup[6], for firms. This markup is mapped to the aggregate demand growth, r_{ad}. In other words, when aggregate demand on time t has decreased compared to total demand at time $t-1$, $r_a d$ will be less than zero; hence the unit-price will be less than unity.

$$UC = \frac{w}{\lambda} \tag{8.12}$$

Robots

In our model, firms can to produce goods by either robots or workers. Initially, every robot has the same productivity, λ, as workers: one unit per tick. The costs, however, for a robot are higher: 3 per tick. This means that initially, robots have a UC of 3. In our model, the supply of robots is endless, and so the price of robots is not a function of the supply of robots.

According to Acemoglu and Autor (2010), once a job us being automated, a robot can perform the work of 6.5 workers. This means that when an automation event happens for a certain job, robots specialized in that job will have a λ of 6.5. Following equation 8.12, the unit-costs of the robot declines from 3 to 0.46. By applying this assumption made by Acemoglu and Autor (2010), we, therefore, assume that when a job is being automated, a robot can replace the worker for approximately half of its costs. Note that we thus only look at the effect of replacement of workers by robots, the so-called *replacement effect*: "holding prices and output constant, robots displace workers and reduce the demand for labour, because with robots it takes fewer workers to produce a given amount of output" (p.9). Another effect of robots on productivity is the so-called *price productivity effect*: as automation (the further deployment of robots) lowers the cost of production in an industry, that industry expands and thus increases its demand for labour. For our model, however, we have chosen to only model the replacement effect. The reason for this choice is that we use our application of the Frey and

[6] the amount added to the cost price of goods to cover overheads and profit.

Osborne (2013) study as a starting point. For this, we only look at the effects of replacement by robots. Nevertheless, we consider an expansion of our model with the mechanisms of the price-productivity effects to be a good start for further research.

In our model, we have specified robots as agents. This is done out of a pragmatic reason to capture the heterogeneity of specific jobs, productivity and prices between robots. Robots, however, differ from the 'real' agents in our model since they do not have rules or behaviour.

Unemployment

For our model, we use the unemployment percentage as provided by CBS at the end of 2018 (see Table 8.2). Since we use different age-groups, we differentiate in an age when specifying unemployment. However, we were unable to find any useful data about unemployment per job type. Consequently, initial unemployment is randomly distributed among the jobs.

Table 8.2: Unemployment in the Netherlands 2018 per age cohort

Age	Unemployment percentage %
15 - 75 year	3,8
15 - 25 year	7,2
25 - 45 year	2,8
45 tot 75 year	3,6

Source: Statistics Netherlands, 2018

Chapter 9

Model testing

Chapter Abstract

This chapter discusses the results of the verification of the model and the motivation for our validation of results. Verification of the model has shown that all important mechanisms and assumptions are implemented in the model. Furthermore, sensitivity analysis of the model has found that the model is highly sensitive for the value of the foresight parameter. For this reason, we have chosen to include different values of the foresight parameter in our model. This chapter also addresses the choice of the validation of our results. Because of limited time and resources, we have chosen to validate our results through a literature review.

This chapter discusses the verification and validation of the model. The chapter starts with a discussion of the verification of the model, followed by the sensitivity analysis. Furthermore, it presents the choice and design for the validation of the model.

9.1 Verification

According to Van Dam et al. (2013), once we have a working model in computer code, “we must ask ourselves an important question: did we correctly translate the conceptual model into the model code?” (p.98). To answer this question, we perform two types of tests: single-agent testing and multi-agent testing (test of the whole model). Both tests will verify if our model implementation corresponds with the model design.

9.1.1 Single-agent testing

To check for the correct implementation of agent-rules, we perform a single-agent test. This single-agent test consists of following one agent through a model run. This single-agent testing helps us to test whether our model narrative and general predictions about agent behaviour –as defined in chapter 8–, are rightfully implemented (Van Dam et al., 2013). We perform a single-agent test by making the predictions explicit of what we theoretically expect the agent to do. We then test if the agents in our model act according to these predictions. Note, however, that by predictions, we mean behaviour as specified in the specification. During the verification, we do test any predictions about the outcomes of the model (hypotheses). In other words, we only check whether the model specification is rightfully implemented.

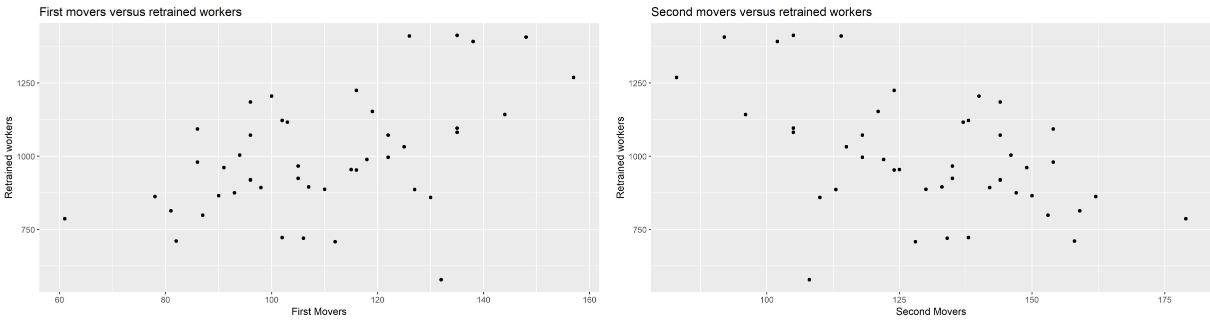
Within a run of 20 ticks, every decision phase within a tick –expect, plan, act, evaluate– is checked. Single-Agent testing is done for both workers and firms. For each type of agent, the single-agent testing has confirmed that all rules of the agent are implemented in the model and correspond to the specification of the model discussed in chapter 8. The full single-agent testing verification can be found in Appendix E.1.

9.1.2 Multi-agent testing

Once we have tested the agents in the model individually, we should test if all the implemented rules corresponded to our theoretical prediction of the entire model. For this, we vary some parameters for which we can make certain predictions about their effects. When varying the parameters result in the expected outcomes, we can confirm that we have rightfully implemented the model.

First and second movers

In our model specification, we discussed that when firms are uncertain, they wait with making the investment decision. In the model, we refer to this waiters as to 'second movers', firms that are certain and do invest are called 'first movers'. All things being equal, when there are more second-movers, fewer investments are made. To test this behaviour, we have looked at how the number of first or second mover influences the number of workers retrained by firms. Figure 9.1 shows indeed that the model produces outcomes that correspond with our theoretical specification which stated that the more uncertain firms are, the more they wait and hence the fewer actions are undertaken by firms (here: retraining workers).



(a) First movers (b) Second movers

Figure 9.1: Verification First and second movers

9.2 Sensitivity Analysis

In our specification, we have made some important assumptions (see Chapter 8). This section will discuss how these assumptions influence the outcomes of our model. For this, we have chosen the unemployment rate as our test variable. We have chosen this variable because the unemployment rate represents the whole behaviour of the model. In other words, every element, agent and rule affects the unemployment rate. By choosing the unemployment rate as the test measure, we can measure the sensitivity of the whole model to certain parameters.

9.2.1 Tick end

As already mentioned in the model specification, we defined the automating probabilities of the Frey and Osborne (2013) model for a time interval of 20 years. However, this time interval is not confirmed by Frey and Osborne (2013). We, therefore, have tested how the outcomes of our model vary when a larger or smaller time frame is chosen. The outcomes of the sensitivity analysis showed that the unemployment rate is not sensitive to small changes in the parameter 'tick-end' (see Appendix F.1).

9.2.2 Foresight

In our agent-based model, we assume that agents make decisions for the next coming three years. This 'foresight' time-interval is specified in the model through the foresight parameter.

Sensitivity analysis of this parameter shows that the model is (slightly) sensitive for small changes in this parameter (see Appendix F.3). The longer the time interval for which firms and workers make a decision, the lower the unemployment rate.

Since our model is sensitive to changes in this foresight parameter and this cannot be explained through the model mechanisms, we include several foresight values in our experimental setup (see Chapter 10.4).

9.2.3 Number of firms

For our agent-based model, we model five firms per sector. This number is determined arbitrary and is therefore not consistent with the actual number of firms per sector. The choice for modelling only five firms per sector is a pragmatic one. Modelling more firms per sector increases the memory usage of the model significantly. In order to conduct experiments with different runs and repetitions, the number of firms should not be too large. Given the technical specifications of our resources, we found that five firms per sector are the limit to have a manageable run-time.

To see how this assumption influences the outcomes of our model, we performed a sensitivity analysis on the number of firms. For this, we have looked at the outcomes of the model when chosen four, five, six and ten number of firms per sector. This sensitivity analysis showed that the unemployment rate is sensitive to our choice of the number of firms (see Appendix F.2). That is, the higher the number of firms, the lower the unemployment rate. For example, the average unemployment rate of 4 firms is 0.04 % whereas the average unemployment rate of 6 firms is 0.02%.

This sensitivity of the unemployment rate to the chosen number of firms can be explained. As mentioned earlier, when firms are uncertain, they will wait with making an investment decision. This means that these firms will not hire any new employees. However, every tick, new workers enter the model. Moreover, the higher the unemployment rate, the lower the aggregate demand will be. This can make firms more insecure and thereby causes them to wait even more. This effect is the most visible when there are fewer firms in the model. Since every firm has a share of workers in the sector, the number of workers per firms is higher when there are fewer firms. This means that with the same ratio certain and uncertain firms, a model with more firms will have a lower employment rate than a model with fewer firms. In other words, the lower the number of firms, the fewer places a fired worker can go to and hence, the higher the unemployment. Additionally, this result of the sensitivity of unemployment to the number of firms also indicates that the model is consistent with our theoretical expectations and specification.

However, this effect of the number of firms in our model cannot be seen as an 'outcome' of our model (analysis). That is, we cannot draw conclusions on these findings. Instead, we state that the effect of the number of firms on the unemployment rate can be seen as an artefact of the model[1].

9.2.4 Logistic Growth rate

In our model, we have defined the probability of a job being automated at each step as the conditional probability of a logistic growth function:

$$f(x) = \frac{L}{1 + e^{-\kappa(x - x_0)}} \tag{9.1}$$

Here, κ depicts the logistic growth rate. This logistic growth rate can be seen as the parameter that determines the 'steepness' of the S-curve. Figure 9.2 shows the difference curves by different values of κ. To test if the model is sensitive for different growth rates, we have performed a sensitivity analysis with different values of κ: 1, 2, 3, 4, and 5. The result of this sensitivity analysis showed that the model is not sensitive to different growth rates (see

[1] *Model artefact*: significant phenomena caused by accessory assumptions in the model that are often irrelevant to the significant results of the model (Gajan et al., 2009)

Appendix F.4). However, since the variation of the unemployment rate with a growth rate, κ, of 2 is the smallest, our model uses the value 2 for κ

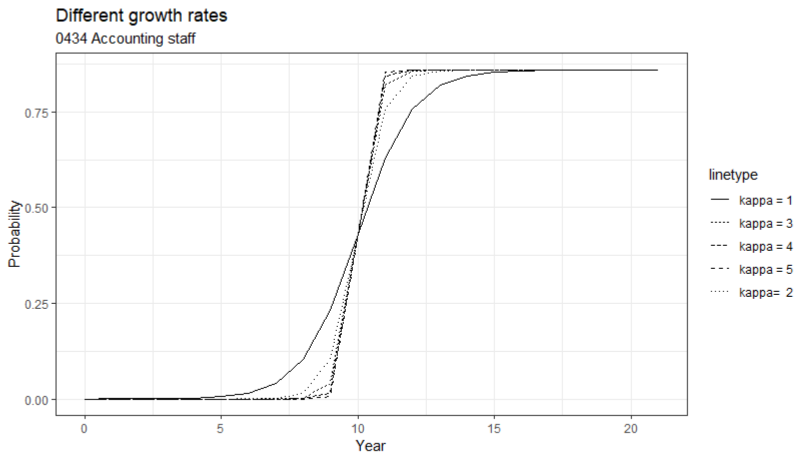


Figure 9.2: Different curves of the Logistic growth function for different values of κ

9.3 Validation design

Whereas verification deals with the question "did we built *the thing right*", validation concerns the question "did we built the *right thing*" (Van Dam et al., 2013). However, a traditional validation of an agent-based model is difficult and often not applicable to these types of models. As explained by Louie and Carley (2008):

> From a history of science perspective, it is important to note that the most advanced methods of validation were developed in engineering fields for assessing models of technical systems that followed fundamental physical laws. In contrast, these large-scale multi-agent systems are used for examining socio-political systems where the fundamental underlying laws are not known. Multi-agent models of social systems are difficult to validate because these models represent a new approach to simulation for which traditional validation methods are not always applicable (p. 1-2).

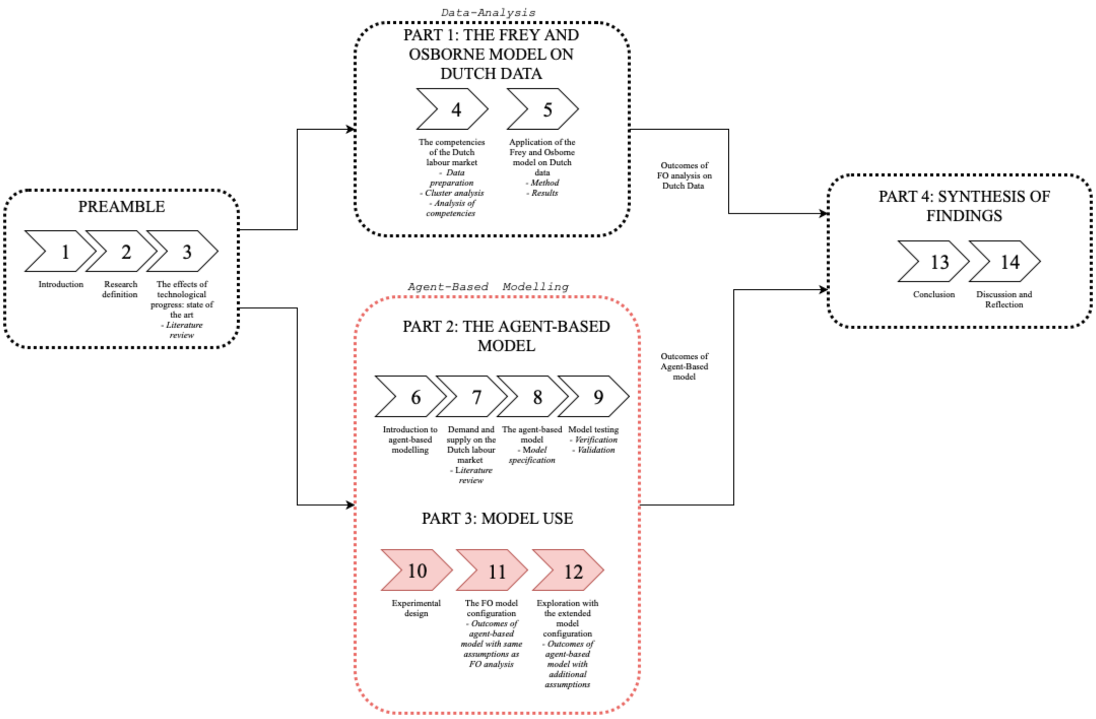
Taking these validation challenges into account, Van Dam et al. (2013) propose several methods that are suitable for the validation of agent-based models. Accordingly, validation cannot "simply compare computed behaviour to "real" system behaviour if there is no "real" system available for comparison or if the model is exploring possible future states" (p. 127). Rather, validation of agent-based models has to focus on whether the model is useful and convincing in gaining valuable insights. Consequently, appropriate agent-based validation methods as proposed by Van Dam et al. (2013, p. 127) include:

- Historic replay;
- Face validation through expert consultation;
- Literature validation; and
- Model replication

Due to the limited time, resources available and scope of this study, model validation is done through literature validation. For this, we will check whether the generated behaviour can be explained by literature. By doing this, we test how the higher level model outcomes generated by model behaviour on the individual lower level is consistent with the existing literature. Nevertheless, according to Van Dam et al. (2013, p. 129), "when performing such comparison experiments, we should not focus on replicating the exact outputs, but instead on the general outcomes and recommendations". This means that we will not check if the model reproduces the exact outputs as given by the literature. Rather we check *how* and *if* our model produces outcomes that can be *explained* through the literature.

To summarize, the validation of the model will thus be done through assessing how the gained insights by the model can be explained by literature. This validation will be done after the experimentation and synthesis of outcomes from the model.

III
Model use



Chapter 10

Experimental design

Chapter Abstract

This chapter discusses the methods and motivations for the experimentation with the agent-based model. Since the aim of the agent-based model is to analyze how additions to the original FO model can give us more insights in the effects of technological progress, an FO configuration of the agent-based model will be developed and be used as a benchmark. This FO model configuration only encompasses the assumptions made by Frey and Osborne (2013) and can, therefore, be compared with the extended model configuration that incorporates the possibility of retraining, strategic behaviour and uncertainty. Furthermore, due to policy analysis, scenario analysis and the number of repetitions per run, the total number of runs will be 982.

This chapter will provide an outline and explanation for the design of the experiments with the ABM model. This chapter will start with a discussion of the different model configurations and will introduce the Extended Model Configuration. Hereafter, the (choice for) scenario selection, policy levers and outcomes of interest will be discussed. Finally, this chapter will give the complete experimental design and method.

10.1 Model configuration

As mentioned, this study aims to investigate the effects of uncertainty about technological progress influences on the Dutch labour market. For this, we use the Frey and Osborne (2013) study as our starting point. The application of this study on Dutch data has shown that about 12% of the Dutch workers are at risk of being automated (see Chapter 5). This analysis, however, does not take the uncertainty about technological progress and retraining possibilities into account. To include these elements, an agent-based model is developed that is considered to be an 'extension' of the original FO model. By analyzing the possible outcomes of this model, we hope to gain more insight into the *true* effects of technological progress. To compare the original FO model with the Extended model[1], we conduct experiments with two distinct configurations: one *with* and one *without* these additional features. Table 10.1 gives a high-level overview of the differences between the discussed configurations.

[1] This extended model is the model as developed in 8. Since we now also have an FO Model Configuration, we will refer to this model as the Extended Model Configuration

Table 10.1: Differences between features of the two model configurations: FO Model Configuration and Extended Model Configuration

Mechanism	FO Model Configuration	Extended Model Configuration
Time	Yes	Yes
Technological progress as function over time	Yes	Yes
Feedback mechanism between Aggegrate Demand Growth and investments	No	Yes
Relation between uncertainty and investments	No	Yes
Replacing workers by robots	Yes	Yes
Retraining workers by firms	No	Yes
Retraining workers by themselves	No	Yes

10.1.1 The Frey & Osborne model configuration

As mentioned, the Frey and Osborne (2013) study towards the effects of technological progress on the labour market did not account for any feedback mechanism, uncertainty, or retraining possibilities. In order to see how the addition of these features affects our outcomes, we experiment with the agent-based model that does not incorporate these features. Nevertheless, to compare both models, this model also is an agent-based model with a time-limit (20 years).

The FO-MC configuration will thus be used as a benchmark model for which we test the effects of the additions of uncertainty and the possibility for retraining. This means that this configuration will not be used for scenario or policy testing. The exact specification of the Frey & Osborne model configuration can be found in appendix G.2.

10.1.2 Extended Model Configuration

As mentioned, the purpose of the agent-based model is to see how uncertainty about technological progress influences the effects of technological progress on the labour market. Here, the possibility of retraining is also be taken into account. To gain insights into these effects, an agent-based model is developed. The specification of this model is explained in chapter 8 and appendix D. This model will be referred to as the Extended Model Configuration or E-MC. The E-MC will be compared with the FO-MC in order to see how the addition of retraining and uncertainty affects the outcomes of the Frey and Osborne (2013) application on Dutch data.

Since we believe that the addition of these features will give more relevant insights, this E-MC will also be used for scenario and policy testing. The exact specification of the E-MC can be found in appendix G.3.

10.2 Scenario selection

In the previous section, we have discussed the E-MC that we use for our experiments. This model configuration can be seen as a certain set of model parameters and is often referred to as a scenario (Van Dam et al., 2013). To gain insights into the effects of technological progress uncertainty on the Dutch labour market, additional scenario's are created. These scenarios are constructed along the axes of two elements: economic growth and the level of automation.

10.2.1 Economic growth

To determine how our model behaves under different economic growth scenario's, we defined two scenario's: a scenario with high economic growth (scenario High) and a scenario with low economic growth (scenario Low).

These scenarios are based on the two scenario's used for the economic forecast of the Netherlands by the PBL[2]. In their forecasting of the Dutch economy, Manders and Kool (2015) analyzed the future trends and uncertainties that are important for the assessments of (future) policies. As a result, Manders and Kool (2015) defined these two possible scenario's that may prove to be important:

1. *High*: High Economic growth (2 %)

2. *Low*: Low economic growth (1 %)

To take these two economic growth scenario's into account, the aggregate demand rate (increase or decrease in aggregate demand), r_{AD}, will be multiplied with one of these growth rates:

if *Scenario* = *Low* **then**
  $r_{AD}^{final} = r_{AD} \times 1$
end
if *Scenario* = *High* **then**
  $r_{AD}^{final} = r_{AD} \times 2$
end

Algorithm 3: Determining of the aggregate demand rate of both scenario's: High and Low

10.2.2 Endogenous or exogenous technological growth

So far in our model, we have assumed that technological growth is exogenous: e.g. it is not influenced by any factor in our model. This, however, is not necessarily true. For instance, in his theory about economic growth, Romer (1990) argues that technological change arises from intentional investment decisions made by profit-maximizing agents. Here, economic growth is driven by technological change. This notion of endogenous technological change is also found in Schumpeter's theory of economic growth which states that "innovations result from entrepreneurial investments that are themselves motivated by the prospect of monopoly rents" (Akcigit & Howitt, 2014, p. 3). By developing an evolutionary agent-based model that nested macroeconomic dynamics into heterogeneous bounded rational firms, Dosi, Fagiolo, and Roventini (2006) were able to observe this pattern of endogenous technical change in combination with Keynesian features:

> Technical progress is machine-specific and diffuses in the economy via the time-consuming investment by users. In turn, investment and production decisions induce demand propagation effects much alike Keynesian 'multiplier' effects. Conversely, adaptive expectations on demand-driven investments in manners closely resembling the Keynesian 'accelerator'. (p. 22)

To see how the uncertainty about technological progress affects the labour market, we consider two different scenario's. The first scenario is the scenario where technological change is exogenous, and in the second scenario, technological process is considered to be endogenous.

The endogenous scenario is based on the work of Dosi et al. (2006), where firms' investments in R&D are considered to be drivers for technological change. Here, firms investments in robots will increase technological change and thereby the initial automation probability as given by Frey and Osborne (2013). In contrast to the exogenous technological change scenario, firms can influence the pace of technological development. Here, the more firms invest in robots, the higher the possibility that these robots will replace workers. This mechanism is specified by controlling the cumulative automation probabilities for firms investments. At

[2] PBL: Netherlands Environmental Assessment Agency is the national institute that executes strategic policy analysis in the fields of the environment, nature and spatial planning (PBL, 2019)

the beginning of each tick, the investments of firms will be evaluated, and for each job, a so-called technological change factor, $TCF_{j,r}$, will be calculated:

$$TCF_{j,r} = \frac{I_{j,r}}{I_{all}} \tag{10.1}$$

Here, I_{all} represents the total investments made by all firms in both new workers and robots of all jobs during the last ticks. The variable $I_{j,r}$, represents the investments made in robots of a certain job. The factor, $TCF_{j,r}$, is thus the proportion of the investments in a certain type of technology overall investments made by firms. The final automation probability of a job will be controlled by this proportion as follows:

$$Pauto^{j}_{t+1} = (1 + TCP_{j,r}) \times Pauto^{j}_{t} \tag{10.2}$$

In this endogenous-change scenario, the automation probability of a job increases when more investments in these robots are made. Here, the initial values of these automation probabilities are the automation probabilities as calculated by the FO application. This adjustment of the FO application probabilities differs with the exogenous technological change scenario that uses these FO application probabilities the entire model run.

10.3 Policy levers

In chapter 7 and chapter 8 we have identified three policy parameters that may affect the behaviour on the labour market with respect to the automation of jobs of workers. Table 10.2 shows the explanation and specification of these parameters. To see how these policy variables or so-called ‘policy-levers’ can influence the performance of our model, we vary these parameters according to their specification. This means that we experiment with the possible combinations of these policy levers (3^3 = 27 combinations).

Table 10.2: Policy levers

Policy Lever	Explanation	Specification
Information Policy	The extent to which workers take government information about job demand into account when determining retrain possibilities. 0.25 means that workers have 25% chance to retrain to the job where the highest demand is (information that is provided by the government) and 75% chance that the worker makes his decision on experiences within his social circle (which does not have to reflect the actual demand)	Factor [0.25 0.5 0.75]
Retrain subsidy	The proportion of retraining costs paid by the government. These retrain costs are initially derived from the skill bridge between the current job of the workers and the workers preferred job	Factor [0.25 0.5 0.75]
Labour market flexibility	Proportion of flexible contracts. 0.75 means that 70% of all the employed workers have a flexible contract	Factor [0.25 0.5 0.75]

10.4 Foresight variable

The sensitivity analysis of the foresight variable showed that the model is sensitive to the foresight variable (see Chapter 9.2.2). We, therefore, have chosen to vary the foresight variable during our experiment. For this, perform the experiments with three different values for the foresight variable: 1, 3, and 6.

10.5 Outcomes of interest

Table 10.3 shows the outcomes of interest that are specified for this experiment. These outcomes will be collected for each tick and will be analyzed afterwards. Note, however, that the goal of these experiments –and the development of the agent-based model– is to *gain insights* in how uncertainty may influence the outcomes of the FO application of the Netherlands. This means that we do not only look for quantitative measures such as the unemployment rate. Rather, we look for the overall explanation of how uncertainty influences the true automation probabilities. We, therefore, look for patterns in our results. These patterns can tell us how the low-level individual behaviour of workers and firms may affect the automation probability as given by the FO application, under a variety of circumstances.

Table 10.3: Outcomes of interest: explanation and specification

Outcome	Explanation	Specification
Unemployment rate	Percentage of unemployed workers	$U_r = \frac{n_{unemployedworkers}}{n_{workers}}$
Unfilled demand	For each job: the number of vacancies unfilled	$D_U^j, j \in J$
Total demand	For each job: the total demand for that job from firms	$D_T^j = D_U^j + D_F^j, \quad j \in J$
Endogenous TC probability	The adjusted automation probability	if scenario = endogenous-TC: $\rightarrow$ $p_{auto,t} = f(p_{auto,t-1})$

10.6 Final experiment set-up

Figure 10.1 shows the final experiment setup. Here, we see that the FO configuration model will be replicated 10 times with no variation in the parameters. Since we have defined 324 experiments[3] and replicate each run 3 times, the total number of runs for the experiments will be 972. This means that the total number of runs for the whole experiment will be 982 runs.

10.7 Method

After running all the experiments, the results of these experiments will be analyzed. In this section, we will explain the used methods for the analysis of both the FO-MC and E-MC. Note, however, that since the FO Model Configuration has no variations or changes in parameters, the experiments of the FO configuration will only be used as a benchmark for the E-MC.

10.7.1 FO Model Configuration analysis

To analysis what will happen when we run the model with the same assumptions about the effects of technological progress of Frey and Osborne (2013), we analyze the unemployment rate at the end of each model run. Results of the FO application on Dutch data revealed that about 12 % of the Dutch workers are at high risk to lose their job due to automation (see Chapter 5). However, if we look closely to the definition of 'High risk' –the same definition as used by Frey and Osborne (2013)– this number could be much higher.

10.7.2 Extended model analysis

In all these analyses, we use the unemployment rate as the final output measure. Moreover, since we also want to know the percentage of retrained workers, replaced workers and retrained workers by firms, we include these variables in our analysis. While these three

[3] 4 scenario's × 27 parameter combinations × 3 possible values of the foresight variable = 324 experiments

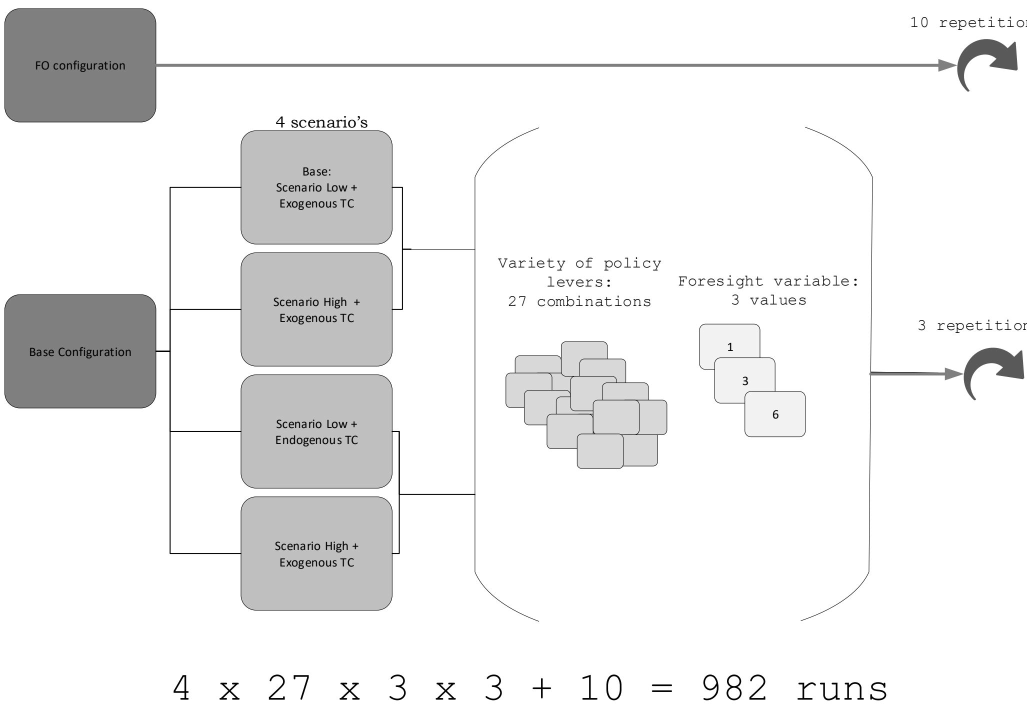


Figure 10.1: Conceptual model of the the experiment set-up

variables can be treated as independent outcome variables, we know that these variables influence the final unemployment rate. Consequently, we will treat the parameters percentage 'replaced workers', 'percentage retrained workers' and 'percentage retrained workers by firms' as partly endogenous variables. That is, when these correlations matters, such as in the regression analysis, we will treat them as endogenous variables.

Behaviour of the model

Before analyzing the effect of different policy levers and scenario's, we first start with a general model exploration that reveals the behaviour of the model. This model exploration starts with an analysis of the different states of the model. For this, we conduct a t-Distributed Stochastic Network Embedding (t-SNE) analysis on the outcomes of the model. By using the t-SNE analysis technique, we are able to do a multidimensional reduction[4] and visualize the different 'states' of the model (van der Maaten & Hinton, 2008). A detailed explanation of the t-SNE technique can be found in appendix I.1.

Effect of uncertainty

To test how uncertainty affects the outcomes of our model, we analyze to what extent the percentage of uncertain firms act as 'intermediary' variable. In short, we test how the addition of the variable 'percentage of uncertain firms' can explain the effect of the policy levers on the unemployment rate. For this, we have divided the data set in two sets: one where the percentage of uncertain firms is equal to zero (control set) and one where the percentage of uncertain firms is higher than zero (test set). We then perform a mediation analysis to test whether there is a difference between the relationships in the data sets and to what extent

[4]Dimensionality reduction methods are able to convert a high dimensional data set, $X = \{x_1, x_2 ... x_n\}$, into two-dimensional data set. $Y = \{y_1, y_2 ... y_n\}$, that can be easily visualized –for example in a scatterplot (van der Maaten & Hinton, 2008)

the percentage of uncertain firms can explain this (see Appendix J for more details about this method).

Scenario and policy analysis

To see how the different policy levers, the foresight variable and scenario's influences our results, we will compare the medians and perform a regression analysis. Since we do not vary these parameters over time –only at the beginning of each run– we only compare the influence of these parameters after 20 years. This means that for each variation of a parameter, we calculate the median of an output variable (for example the unemployment rate) and compare this mean with the means of the other variations. For example, in our experiments, we have used three different values for the foresight variable: 1, 3 and 6. To see how the foresight variable affects our model, we calculate the median for the unemployment rate for the runs that have a foresight variable of 1, 3 and 6 after 20 years. Next, construct a box plot to see how these different values for the foresight variable influence our model.

Furthermore, we perform a multivariate analysis to see how these variations in different affect different parts in our model and hence, the final unemployment rate. Accordingly, multivariate analysis is a method that can be used when several measurements are made on each individual or object in one or more samples (Rencher & Christensen, 2012). For this, we perform a General Linear Regression Analysis (GML). Equation 10.3 shows the formula for this regression analysis. Here, $\beta_1 PL$ depicts the policy lever (or scenario parameter) and $\beta_2 p_{replaced} \times \beta_3 p_{retrained} \times \beta_4 p_{retrainedfirm}$ depicts respectively: the percentage of workers replaced, the percentage of workers retrained and the percentage of workers retrained by firms. Moreover, we have set the unemployment rate, U, as the final depended variable.

$$g(u) = \beta_0 + \beta_1 PL \times \beta_2 p_{replaced} \times \beta_3 p_{retrained} \times \beta_4 p_{retrainedfirm} + e \tag{10.3}$$

In our regression analysis, we, therefore, analyze how a policy lever or scenario *in combination* with other model variables influence the final unemployment rate. Furthermore, we use 'Gaussian' as the regression model family. In other words: we assume that our depended variable is normally distributed *and* we assume a 'identity link' function such that $g(U) = U$ and the default error distribution is a normal distribution.

Note, however, that the goal for this regression analysis is not to make model predictions. Instead, it only is used as a tool to gain some more insights into the effects of policy levers and scenarios. These regression models will thus not be used to make predictions or to forecast specific relations. Consequently, we will not comment on the properties of these models and whether or not they violate the linear regression assumptions [5]

Unfilled labour demand

Finally, we will analyze how labour demand for each job will change over time and to what extent this labour demand is unfilled. While this analysis is more or less straightforward, some definitions need to be explained before presenting the results.

In this analysis, we refer to 'total (labour) demand' as to the sum of filled and unfilled vacancies for all jobs. Since this analysis works with proportions (of total labour demand), we assume that this aggregate demand is equal to 1. The total demand for a specific job is thus the proportion of demand for that job to total demand. For example, if aggregate demand for all jobs is 500 workers, and Job A needs 10 workers, the total demand for job A will be 1%. Furthermore, we refer to 'unfilled (labour) demand' as to the proportion of unfilled vacancies to total labour demand. To stay with the previous example, if the total labour demand is 500, total demand for job A is 1% and Job A has 5 unfilled vacancies, the unfilled demand for job A will be 0.5%.

[5] The following assumptions must hold for linear regression models: (1) the residuals are independent; (2) the residuals are normally distributed; (3) the residuals have a mean of 0 at all values of x; and (4) the residuals have constant variance

Chapter 11

The FO model configuration

Chapter Abstract

This chapter presents the results of the analysis of the FO model configuration. Here, we see that simulation has shown that the percentage of workers who will lose their jobs due to technological progress is higher than initially predicted by the FO application on Dutch data. This can be explained by the way how Frey and Osborne (2013) have defined a 'high risk'. In their analysis, only jobs that have an automation probability higher than 0.7 are considered as a high risk. However, when simulation these automation probabilities, we see that at the end of 20 years, the unemployment rate increases to 36%. For this reason, we, therefore, modify our conclusion by stating that 36% of the workers are at risk to become unemployed due to technological progress whereas 12% of the workers are considered as a high risk.

This chapter explores the behaviour of the agent-based model under the assumptions of Frey and Osborne (2013). For this, both the retraining probability for firms as for workers are removed from the model which resulted in a 'naked' model were firms immediately replace all workers whose jobs are automated through technological progress (for more information see chapter 10.1.1).

In the application of the FO model on Dutch data, we found that 12% of the Dutch workers are at *high* risk of becoming automated. Note, however, the emphasis on the word 'high'. Just as in the original Frey and Osborne (2018) study, we have considered a 'high' risk as a probability above 0.7. Table 11.1 shows the other risk categories as defined by Frey and Osborne (2013) and used in our study.

By stating that 12% of the Dutch workers are at high risk of being automated, we, therefore, only point out that 12% of the workers have a change of more than 70% of being automated. However, this does not mean that eventually (only) 12% of the workers are at risk. Since most Dutch jobs have a 'medium' risk of being automated –which ranges from 0.3 to 0.7–, the actual number of workers that that see their jobs being replaced by robots can be higher (or lower). Hence, we expect the final unemployment rate of the FO model configuration to be higher than 12 %.

Table 11.1: Defined risks of automation by Frey and Osborne (2013)

Risk Category	Probability
High	$p_{auto} > 0.7$
Medium	$0.3 < p_{auto} < 0.7$
Low	$p_{auto} < 0.3$

source: Frey and Osborne (2013)

11.1 Automated jobs in the model

Before discussing the results of the FO model configuration, we compare the percentage jobs that are automated in the model with the actual FO probabilities, p_j^{cum}. Just as in the extended model version, the probability of a job being automated within each step, t, is calculated by taking the difference of the S-shaped cumulative probability function between t and $t-1$:

$$p_{j,t}^{con} = \frac{p_j^{cum}}{1+e^{-2((t-1)-10)}} - \frac{p_j^{cum}}{1+e^{2(t-10)}} \tag{11.1}$$

This probability for each tick, $p_{j,t}^{con}$, is then used to determine if a job will be automated. This determination is implemented in the model by the following algorithm:

if *random-float* < $p_{auto,t}$ **then**
| set automated $TRUE$
else
| set automated $FALSE$
end

Algorithm 4: Automation algorithm

By using algorithm 4, we aim to approach the automation probability as given by the FO application. To check for this, we calculate for each job the percentage of runs in which the job is being automated. For example, if a job has been automated in 10 runs out of 100 runs, this percentage will be 10%. Next, this percentage will be compared with the FO probabilities by calculating their difference:

$$Diff_j = P_j^{model} - p_j^{cum} \tag{11.2}$$

Table 11.2 gives the summary of the difference between the model percentage of runs automated and the FO probabilities. Moreover, figure 11.1 shows the five jobs with the most positive differences and the five jobs with the most negative differences. The differences between the FO probabilities and the model outcomes of all jobs can be found in Appendix H.1.

Looking closely at these differences, we see that the highest differences are found for the jobs in the high or medium risk category. A possible explanation for this can be derived from basic probability theory. As mentioned, at each step, the fact whether a job will be automated is determined by algorithm 4. Suppose, for example, that we have two jobs A and B. Each job has a $p_{auto,t}$ of respectively 0.1 and 0.6. According to algorithm 4, this means that job A has 10% probability of being automated within a step, whereas this probability for job B is 60%. To determine if that job is being automated, we ask the computer to draw a random number between 0 and 1. When that number is less than $p_{auto,t}$, the job is automated. Conversely, the probability of the job not being automated is 90% for job A and 40% for job B.

One could probably expect that if we ask the computer to draw a random number for a number of times (in our experiment 100 times), the frequency of automation will be 10% for job A and 60% for job B. This assumption, however, will not hold and is also known as the *Gambler's fallacy*[1]. Since the computer draws the random floats independent from its previous draws, there is no certainty that eventually the frequencies of jobs being automated will balance towards its probability.

This fact can be considered 'truer' for job B than for job A. Since job a has a probability of 10 %, each time the computer draws a random number; there is 90% chance that the job is not automated. For job B, there is a 60% chance that the computer draws a number of less than 0.6. Because the numbers are drawn independently, there is thus more chance that the differences between probability and actual frequencies are higher at job B than by job A.

[1]Gambler's fallacy or Monte Carlo fallacy: "The belief that, for random events, runs of a particular outcome (e.g., heads on the toss of a coin) will be balanced by a tendency for the opposite outcome (e.g., tails)." (Ayton & Fischer, 2004, p. 1396).

Table 11.2: Summary: Difference model outcomes with FO probabilities

	Min	1st Qu.	Median	Mean	3rd Qu.	Max
Difference	-0.08	0.02	0.06	0.07	0.13	0.24

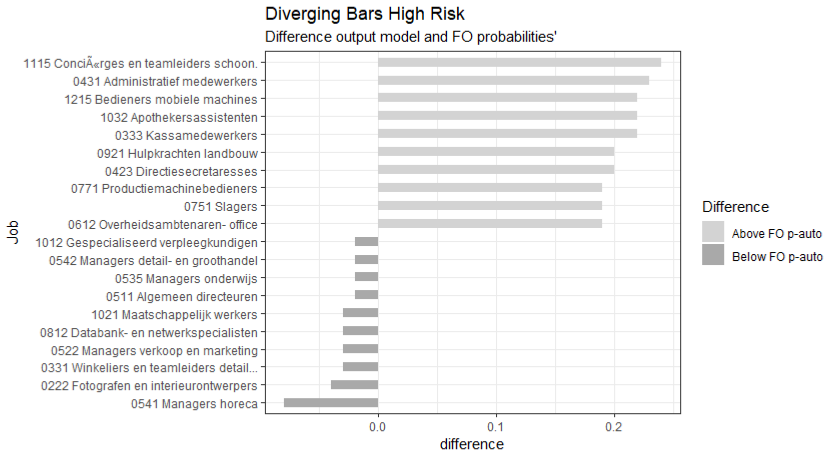


Figure 11.1: Top 5 most positive and negative differences between the FO-MC simulation and the actual automation probabilities

11.2 Unemployment in the Frey & Osborn model configuration

To analyze the unemployment rate due to technological progress under the assumptions of the model of Frey and Osborne (2013)[2], we conducted experiments with the FO configuration of the model.

Table 11.3 shows the statistics of the unemployment rate at the end of every run. As one can see –and as expected– the average unemployment rate includes much more than only the 12% of workers who are at high risk, namely 0.36%. Moreover, figure 11.2 shows a pattern of the unemployment rate across all runs. Here, we see that this pattern resembles an S-shape. Because an S-curve probability function determines the probability of automation for each job (see Formula 11.1), this result makes sense.

Above all, the results of this simulation show that the effects of technological progress can be more significant than our conclusion of the FO model application. This is explained by the definition of 'high risk' as used by Frey and Osborne (2013). Nevertheless, we do not change our conclusion derived from the FO model application. Instead, we add these insights to these predictions. This leaves us with the conclusion that when the FO model assumptions hold, after 20 years, on average, 36% of the workers will lose their jobs due to technological progress and become unemployed. Considering the uncertainty about which jobs are going to be automated, we can only say that 12% of all Dutch workers are at high risk of being automated -e.g. there is a high probability that these workers become part of the 36% total unemployment.

Table 11.3: Summary: Unemployment rate under the FO model configuration

	Min	1st Qu.	Median	Mean	3rd Qu.	Max
Unemployment rate	-0.23	0.32	0.36	0.36	0.39	0.50

[2]Where there is no uncertainty, no strategic behaviour of firms, no retraining possibilities of workers and where technological change is determined endogenously

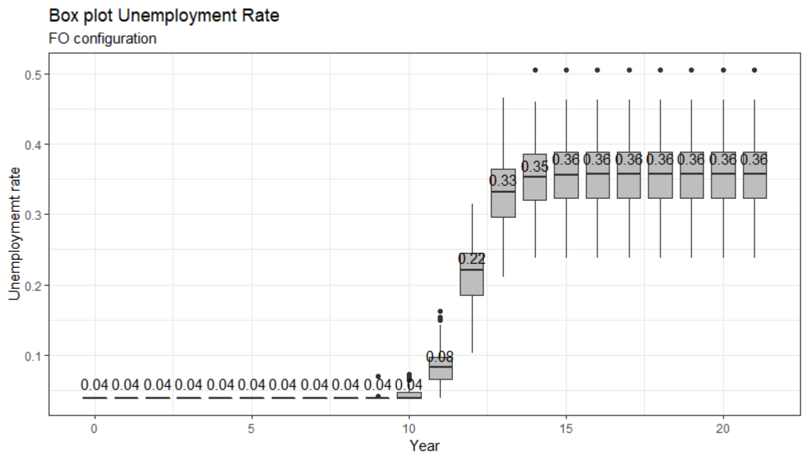


Figure 11.2: Unemployment rate for each step in the FO-MC: unemployment follows and S-curve and stabilizes after 16 years at 0.36%

Chapter 12

Exploration and policy identification with the Extended Model Configuration

Chapter Abstract

This chapter presents the results of the analysis of the extended model. The t-SNE analysis of the states of the model has found that the level of flexibility is the most powerful in explaining the different states of the model. Here, the higher the proportion of flexible contracts, the higher the percentage of replaced workers and the higher the unemployment rate will be. This result is also confirmed by the policy identification where a regression analysis towards the effects of the three policy parameters has found that the level of flexibility can explain the difference in unemployment rates the best. Moreover, we have found that policies regarding the retraining of workers –information policy and retrain subsidy– do not have a significant effect on the unemployment rate or even the number of retrained workers. Instead, we found that, next to flexibility, the foresight parameter can affect the unemployment rate: the larger the number of years firms plan ahead, the lower the unemployment rate. While not a policy parameter in the first place, this result gives some valuable insights into the effect of 'planning' on the consequences of technological progress. Another important finding is the extent to which uncertainty influences our model outcomes. Mediation analysis has shown that the model variable 'percentage of uncertain firms' act as an intermediary variable in the model and influences more than 85% of the model outcomes. Furthermore, no significant differences between the scenario's are found. In this chapter, we also present the results of the analysis of job demand in the Netherlands. Here, we state that even though there is technological progress, the total labour demand increases. This is especially true for technical professions were not only the total demand for these jobs increased but also the proportion of unfilled vacancies. Finally, we validated the model results through literature. Nevertheless, for more complete and exhaustive validation of the model, we recommend further validation through expert opinions.

This chapter discusses the outcomes of the experiments with the Extended Model Configuration[1]. At first, this chapter shows the results of the model exploration. Here, no distinction between different scenarios or policy levers is made. Here, we also analyze to what extent the percentage of uncertain firms influence the outcomes of our model. Next, this chapter discusses how the policy levers influence the model outcomes. At last, we evaluate how the results differ between the four scenarios.

[1]The E-MC is the model configuration that incorporates uncertainty about technological progress, strategic behaviour of firms, the possibility for workers to retrain, both endogenous and exogenous technological change and two scenario's of low and high economic growth (see Chapter 10).

12.1 Model exploration

This section presents the results of the model exploration. For this, we performed a t-SNE analysis that exposed the general states and patterns of the model. A brief explanation of this method can be found in chapter 10. Next, we compare the unemployment rate(s) of the model with the results of the FO-MC. This result will show how the 'doom scenario' of Frey and Osborne (2013) can be mitigated.

12.1.1 States of the model

Figure 12.1 shows the result of the t-SNE multidimensional reduction output. This figure shows that the model does not generate distinctive states[2] throughout the experiments. The line plot represents the 'route' of the model and indicates that when the model runs, the model changes it states very often. Nevertheless, the scatter plot of the two embedded dimensions gives us the most information. Here, every dot represents a different state of the model. In this scatter plot, one can see clearly that there are many different clusters of states –or so-called tropic levels. This means that next to the fact that the model changes quickly to another state, there is also a lot of difference between the states. Given the fact that we

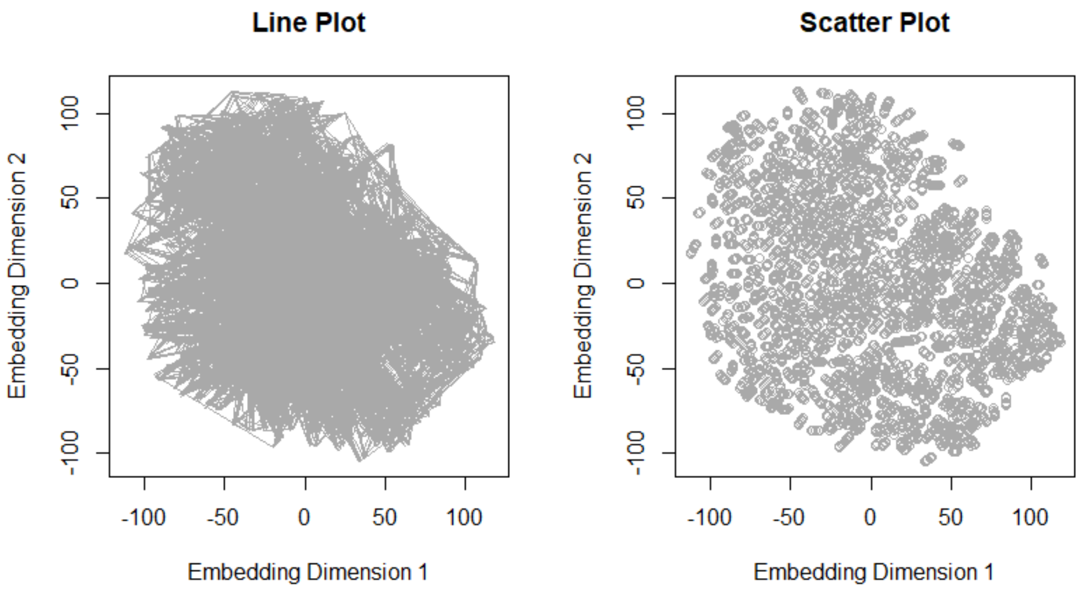


Figure 12.1: t-SNE States of the model. Line plot represents the path between the different states whereas the scatterplot shows the clustering of these states.

formulated multiple model outcomes and that the experiments are done in such a manner that every combination of policy levers is used, this results makes sense. However, more insight is gained when we look at the model outcomes[3] separately. Here, we found that the policy lever 'flexibility' has the most influences on the clustering of the different model outcomes. This is shown in figure 12.1. This figure shows how different values influence the clustering of different states of the model for flexibility. Here, the larger the size of the circle, the higher the value for a particular outcome.

For example, in the figure, we can see that high flexibility (0.75) correlates to a high unemployment rate. This is also true for the number of replaced people: the higher the

[2]Here, we refer to the state of the model as to the combinations of different parameter and outcome values

[3]Regardless the fact that 'the percentage replaced workers', 'the percentage retrained workers' and 'the percentage retrained workers by firms' can be seen as endogenous variables, in this analysis, we treat them as model outcomes. Nevertheless, by performing this t-SNE analysis, we can look for the patterns between these model outcomes (and thereby implicitly consider them as dependent on each other)

flexibility, the higher the number of replaced people. Conversely, the higher the flexibility, the lower the percentage retrained workers and retrained workers by firms compared to lower flexibility.

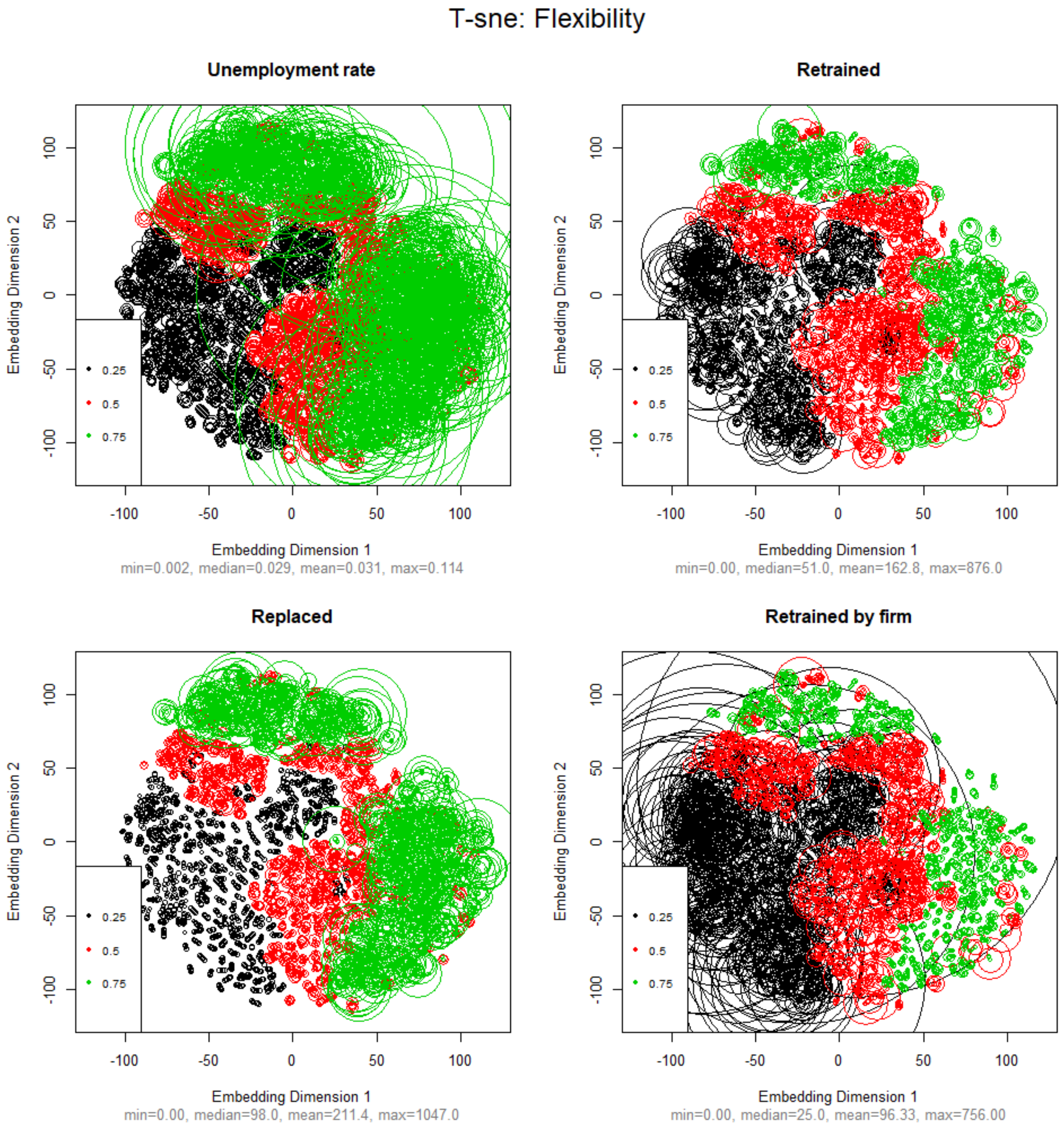


Figure 12.2: t-SNE States of the model coloured by different levels of flexibility

Furthermore, our t-SNE analysis revealed that this correlation between flexibility and different model outcomes is the most important for the clustering of different states. In other words: the t-SNE study of flexibility clearly reveals a pattern between flexibility and the combination of outcomes variables. As the t-SNE analysis of other policies levers has revealed, such a strong pattern cannot be found by the t-SNE analysis of other policy levers[4](See Appendix I).

Finally, the t-SNE results have shown us a clear and important pattern between the different outcome variables: in states where the unemployment rate is high, the number of replaced people is also high. Moreover, the states that have a high unemployment rate also have often a high number of retrained people (which makes sense given the fact that besides

[4]This, however, does not mean that there is no pattern or correlation between the different states and policy levers. The effect of the policy lever flexibility can be so strong that it may 'oppress' other correlations.

firms only unemployed workers have the decision to retrain themselves). Conversely, we found that a high unemployment rate often corresponds to a low number of workers that are retrained by firms. This pattern or correlation between the different states of the model (and thus, outcomes variables), is also confirmed by simple correlation analysis (see Appendix I.5 and Figure 12.3).

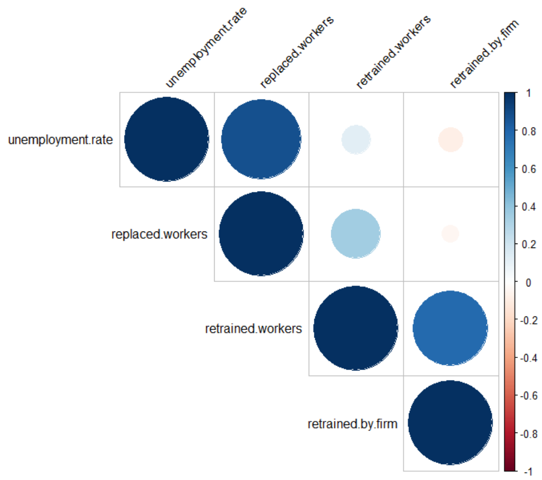


Figure 12.3: Correlation between different outcomes of the model: strongest and most relevant correlation between unemployment rate and the number of replaced workers.

12.1.2 Unemployment compared to the FO Model Configuration

Table 12.1 shows the statistics of the unemployment rate of the E-MC. Here, we see a big difference with the unemployment rate with the FO-MC. For example, in the FO-MC, the average unemployment rate is 36%. This is a 33% difference with the average unemployment rate of the E-MC, which is 3%. This means that considering the initial unemployment rate of 3.6%, unemployment could be reduced with 0.06%.

The box plot of the unemployment rate per step of the E-MC shows a similar unemployment rate pattern as was found in the FO-MC(see Chapter 11, Figure 11.2). For this, we can conclude that the pattern of the unemployment rate for both model configurations is strongly influenced by the formulation of the probability of technological progress and its distribution. In our model, we have chosen to model this as a so-called S-curve. Motivated by this result, we, therefore, recommend experimentation with different distributions when using this model for further research.

Nevertheless, 12.4 shows an interesting pattern. Before approximately nine years, unemployment decreases with a small variance between the different runs. After nine years, we see that the unemployment rate has a much larger variance –there is a large difference in the values of the unemployment rate between different runs. This finding, therefore, suggests that technological progress (or the anticipating on it by firms and workers) occurs after nine years. The spread in the values of unemployment can be explained by the different policies that start to generate effect when technological progress occurs.

Table 12.1: Statistics unemployment: E-MC

Min.	1st Qu.	Median	Mean	3rd Qu.	Max
0.001	0.012	0.028	0.030	0.039	0.126

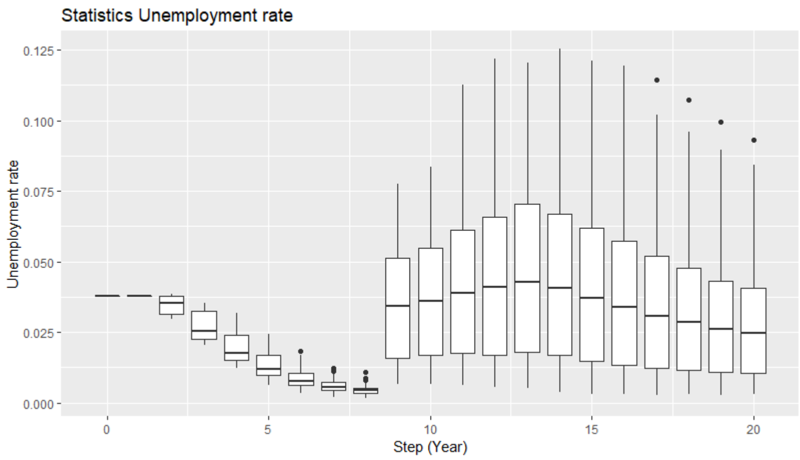


Figure 12.4: Boxplot of unemployment rates per step E-MC: unemployment follows an s-curve over time

12.1.3 The effect of uncertainty

As explained earlier, one of the reasons to develop the E-MC was to account for uncertainty about technological progress among firms and workers. In this section, we will analyze how this uncertainty can have an impact on the model outcomes. Since the t-SNE analysis revealed the pattern between the different model outcomes, we only explain how uncertainty affects the unemployment rate. For this, we first performed a regression analysis between the percentage of uncertain firms at $t = 1$ and the difference in the unemployment rate between time $t = 1$ and $t = 2$. Figure 12.5 presents the results of this analysis. Here, we see that there is a small positive correlation between the percentage of uncertain firms and the unemployment rate (R = 0.1).

Note, however, that this result only tells us the correlation between the percentage of uncertain firms and the unemployment rate when no other effects are taken into account. These 'other effects' are illustrated by the help of figure 12.6. This figure depicts the relationship between the model parameters and the model outcomes (here: unemployment rate). As one can see, there is a *direct* effect and *indirect* effect through endogenous variables –such as the percentage of uncertain firms. This means that to have 'complete' insight in how uncertainty affects model behaviour; we need –besides only looking at the effect of uncertainty on the unemployment rate directly (in the figure depicted by x)– also look at the effect of these intermediaries.

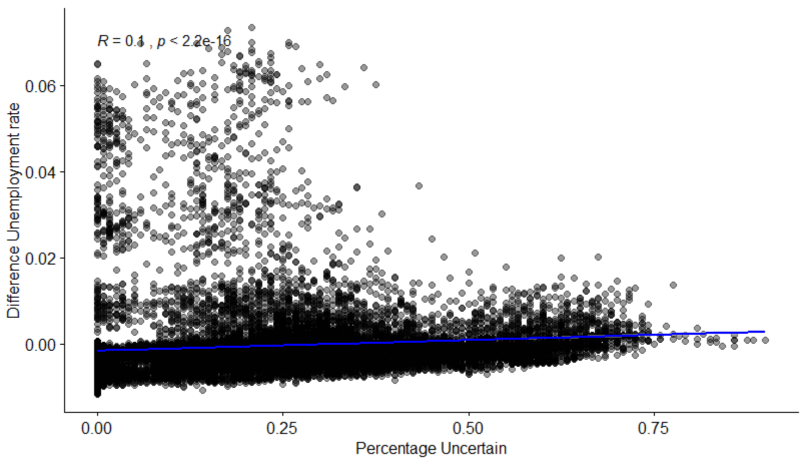


Figure 12.5: Regression between the percentage of uncertain firms and the difference in unemployment rate: no strong direct correlation between number of uncertain firms and the unemployment rate, R = 0.1

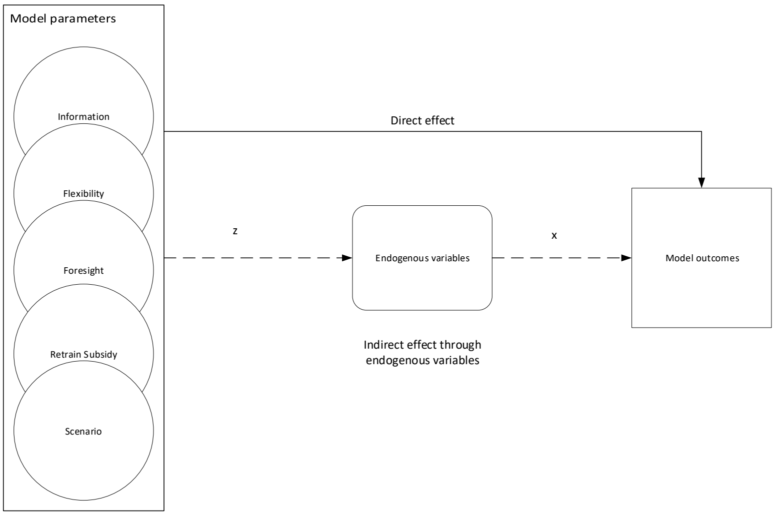


Figure 12.6: Indirect effect on model outcomes through endogenous model variables

Mediation Analysis

In order to see how the percentage of uncertain firms truly influence the unemployment rate, we conducted a causal mediation analysis. Accordingly, "causal mediation analysis is widely used across many disciplines to investigate possible causal mechanisms. Such an analysis allows researchers to explore various causal pathways, going beyond the estimation of simple causal effects" (Imai, Keele, Tingley, & Yamamoto, 2009, p. 1). A brief explanation about Causal Mediation Analysis can be found in Appendix J.1.

Causal mediation analysis towards the indirect effect of the percentage of uncertain firms on the unemployment rate found that 87.8 % per cent of the impact of the policy levers on the unemployment rate is controlled by the variable 'percentage of uncertain firms'. Given this result, we can say that the addition of uncertainty has a significant impact on the behaviour of the model.

12.2 Policy levers

In this section, we present the results of the policy identification. For this, we look at how the different policies –retrain subsidy, flexibility, information policy– influences the effect of technological progress.

12.2.1 Flexibility of the labour market

The t-SNE analysis of policy parameter 'flexibility' showed that this parameter has a strong influence on the model outcomes (see Figure 12.2). Figure 12.7 reveals that there is a positive relationship between the number of flexible contracts and the unemployment rate. This box plot shows that even when there is technological progress, the unemployment rate reduces. Conversely, when the labour market flexibility is high, unemployment will increase due to the effects of technological development.

Moreover, the histogram of the model runs, and unemployment rate grouped by the level of flexibility shows that the model runs are normally distributed around the average unemployment rate for each level of flexibility (see Figure K.2, Appendix K.1).

Furthermore, regression analysis to the effects of the 'percentage replaced workers'[5], 'percentage of retrained workers'[6], the 'percentage of retrained workers by firms'[7] *and* the level of labour market flexibility on the unemployment rate, showed us that there is strong positive correlation between the replaced workers and the unemployment rate (see Figure 12.8 and Appendix K). Here, we also see that the higher the level of flexibility, the higher the replaced workers and the higher the unemployment rate. In combination with the results of the T-SNE analysis of the effect of Flexibility on the model states and outcomes (see Figure 12.2), this results proves the strong positives correlations –or pattern– between the level of labour market flexibility, percentage replaced workers and the final unemployment rate.

A somewhat different pattern is found when we look at the outcomes of our regression model that depicts the relationship between the unemployment rate, retrained workers by firms and the level of flexibility (see Figure 12.9). Here, we see that when flexibility is low, there is a positive correlation between the retrained workers by firms and the unemployment rate. As can be seen from the figure, this correlation decreases when the level of flexibility increases. At a labour market flexibility of 0.75, this correlation is almost nonexistent. Nevertheless, there is a small positive correlation between the unemployment rate and the percentage of workers who are retrained by firms at low market flexibility (0.25). However, the regression analysis of the effects on the total percentage retrained workers (workers retrained by themselves and by firms) shows that there is a negative correlation between the percentage of the total retrained workers and the unemployment rate (see Appendix K.1, Figure K.4). In other words, while the retrained workers by firms can have a positive effect on the unemployment rate –e.g. the higher this percentage, the higher the unemployment rate–, the cumulative effect of retraining is still negative. This negative correlation indicates that the negative correlation between workers retrained by themselves and the unemployment rate is stronger than the positive correlation between the percentage of workers retrained by firms and the unemployment rate.

Furthermore, the generalized linear models of both the relationship between unemployment and the retrained workersand unemployment and the percentage workers retrained by firms showed that the higher the level of flexibility, the lower the unemployment rate. This is in contrast with the relationship found in the GLM of the effect of the replaced workers on the unemployment rate (see Figure 12.8) and the cumulative effect of the level of flexibility on the unemployment rate (see Figure K.1). Given these results, we can, therefore, conclude that the level of flexibility has the most effect on the unemployment rate through the replaced workers. Here, the higher the level of flexibility, the higher the number of replaced workers and hence, the higher the unemployment rate. The retraining of workers by themselves can strengthen this effect. In other words, a low unemployment rate because of a small replaced workers can become lower when also the retrained workersis high. However, this is especially true for the percentage of workers who are retrained by themselves since the percentage of workers retrained by firms can increase unemployment when labour market flexibility is low.

To conclude, experiments with different values for the proportion of flexible contracts has shown that when flexibility is low (0.25), after 21 years, technological progress has caused the unemployment rate to fall under its current level.

12.2.2 Information Policy

Analysis towards the effect of information policy –the proportion of true government information that workers use to decide to which job they should retrain– on the unemployment rate, confirmed our first conclusions derived from the t-SNE analysis: the information policy has no effect on the model outcomes and unemployment rate (see Figure 12.10 and Appendix K.2). This, however, does not mean that these results are useless. The GLM models of the relations between unemployment and the percentage replaced workers, the percentage retrained workers and the percentage retrained workers by firms, showed similar patterns as

[5] Referred to as ' retrained workers'.
[6] Referred to as 'retrained workers'.
[7] Referred to as 'retrained workers by firms'

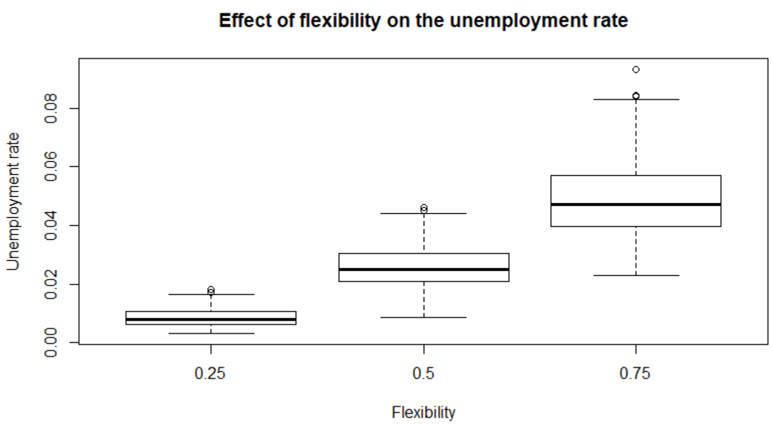


Figure 12.7: Box plot labour market flexibility and unemployment rate: strong effect of flexibility on unemployment rate

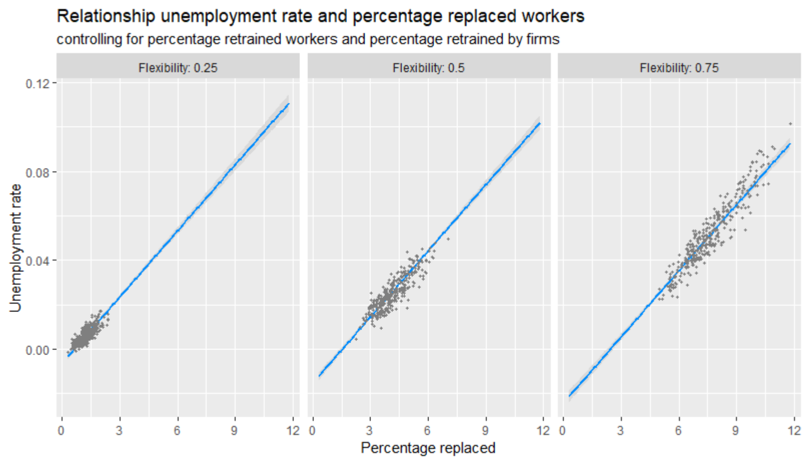


Figure 12.8: GML of the relation between percentage replaced workers and unemployment rate categorized by level of labour market flexibility: strong correlation between unemployment and flexibility for all levels of labour market flexibility

found in the analysis of the level of flexibility: the strongest correlation can be found in the relationship between unemployment and the replaced workers –the higher the replaced workers the higher the unemployment rate–; and there is a weak negative and positive correlation between the unemployment rate and respectively the percentage of total retrained workers and the retrained workers by firms (see Appendix K.2, Figure K.10, K.8 and K.9).

12.2.3 Retrain subsidy

Analysis towards the effects of the retrain subsidy – the percentage of retraining costs paid by the government– showed that the retrain subsidy does not affect the unemployment rate (see Appendix K.4). Indeed, one could argue that the subsidy of retraining does not have to influence the unemployment rate since many other (stronger) factors influence this. But even only looking at the direct effect of retraining subsidy to the total percentage of retrained workers, showed that there is no relationship (see Figure 12.11). This result, therefore, indicates that instead of retraining costs, other factors –such as age and previous working experience– determine the fact whether a worker would retrain or not. Since the choice of retraining is the same for all workers –thus also for the workers who are planned to be retrained by firms–, we can state that these findings are true for all workers in the model.

Our findings are in line with the findings of Jacobson, LaLonde, and Sullivan (2005), who stated that the individual characteristics of a person are far more important in determining

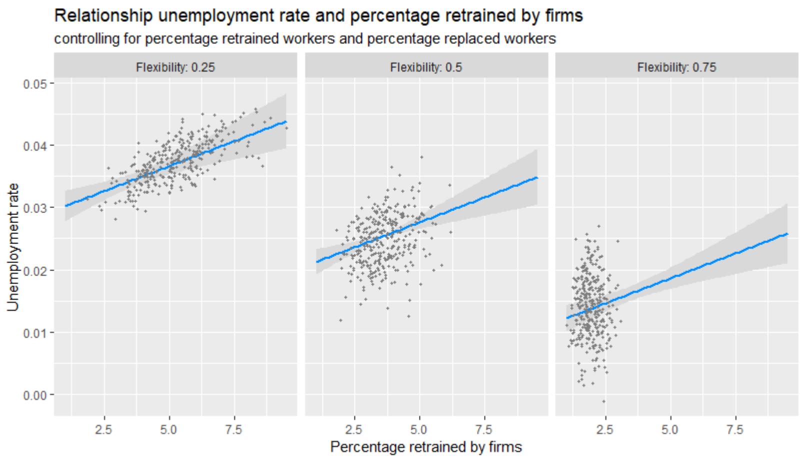


Figure 12.9: GML of the relation between percentage retrained workers by firms and unemployment rate categorized by level of labour market flexibility: strongest correlation at a flexibility of 0.25

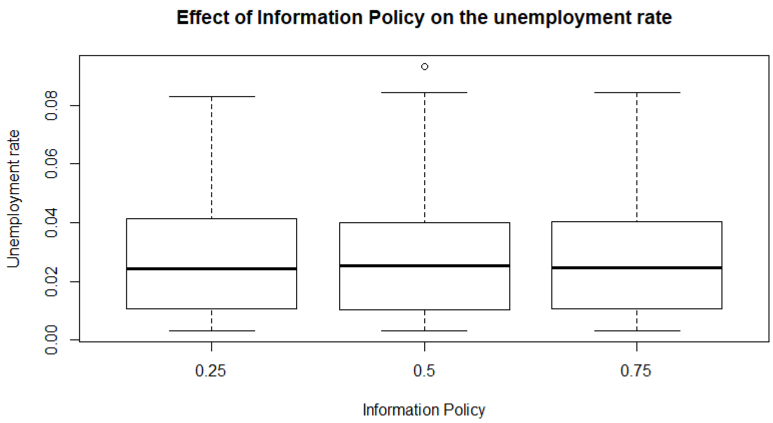


Figure 12.10: Box plot: Information Policy and unemployment rate. No differences in the unemployment rate between different values for information policy

whether workers would retrain than the retraining costs alone.

> Even if it were the case that individuals induced to participate in retraining by increased public subsidies experienced the same gains as those who already participate, it is not clear that increasing subsidies is the best policy. If an investment in retraining is optimal from the point of view of society, it is very likely also optimal from the point of view of the individual considering retraining. (Jacobson et al., 2005, p. 61).

Given all these results, we conclude that the policies with respect to retraining (information policy and retrain subsidy) do not affect the effects of technological progress. Moreover, even when these policies would work, the unemployment rate will still be mostly affected by the level of labour market flexibility. This is because the labour market flexibility has a strong influence on the replaced workers, which, in turn, has a strong influence on the unemployment rate. In our analysis, we only tested the labour market flexibility as a policy that has a strong influence on the number of replaced workers but similar policies that affect the 'easiness' of the replacement of workers can have the same effect.

Nevertheless, the retrained workersdoes have a small influence on the unemployment rate –and thus the effects of technological progress. Note, however, that only the workers who are retrained by themselves have a negative effect on the unemployment rate while firm retraining can have a positive influence on the unemployment rate –although this effect is very weak or even nonexistent. However, our model founds that to stimulate the retraining of workers; subsidy policies will not work since other factors are more critical in determining the choice of retraining of workers.

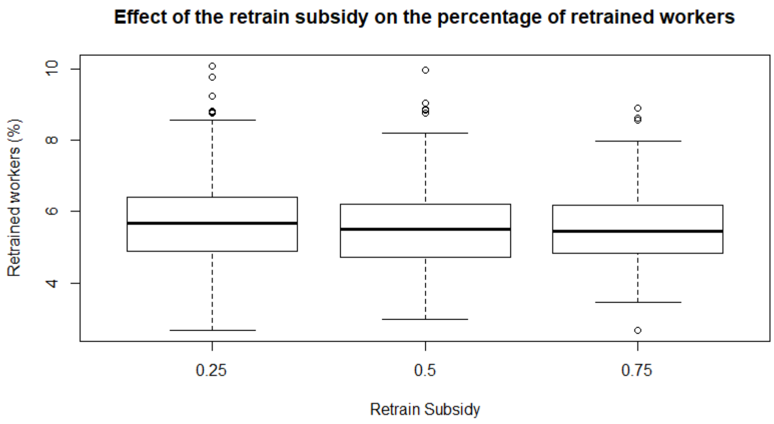


Figure 12.11: Box plot: No difference in the retrained workersfor different values of retrain subsidy

12.2.4 Foresight parameter

While not a policy, we analyzed the effect of the foresight parameter –the number of years of which firms and workers based their strategy– on the model outcomes. This is done because the sensitivity analysis of the foresight parameter showed that the model is sensitive for this parameter (see Chapter 9). Further analysis found that the higher the number of years firm encounter in their planning, the lower the unemployment rate (see Figure 12.12 and Appendix K.6). Moreover, the general linear regression model also found that the higher the foresight parameter, the higher the number of retrained workers and finally, the unemployment rate (see Figure 12.13). This means that for policies concerning retraining to have a negative effect on the unemployment rate, the outlook of firms and workers should also be high. For this reason, we argue that even when it was not considered as a policy parameter initially, this finding can be valuable for policymakers. Whereas policies that are oriented to stimulate retraining on the short-term –such as retraining subsidy– are not effective, other policies that stimulate workers and firms to formulate their planning on the long-term may be proven to be more successful when dealing with the consequences of technological progress.

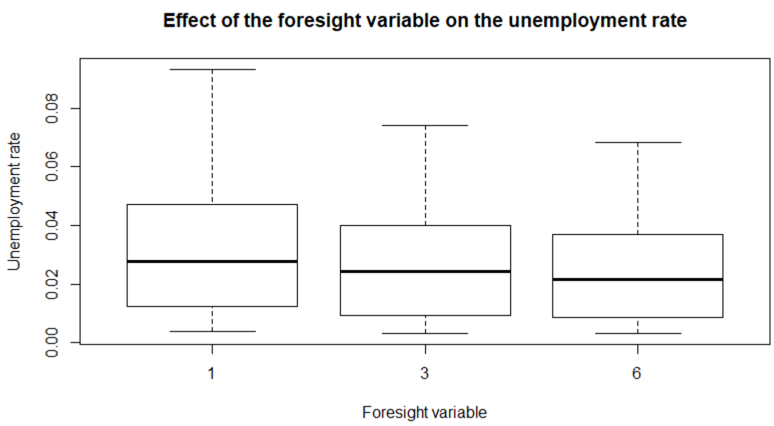


Figure 12.12: Box plot: The higher the foresight parameter, the lower the average unemployment rate

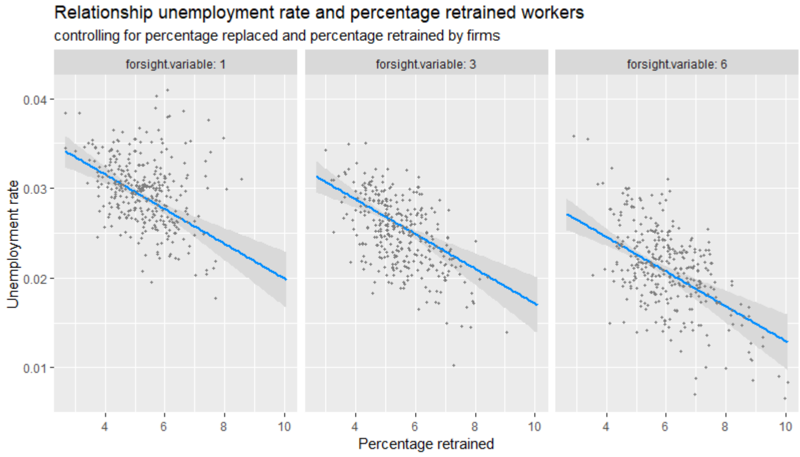


Figure 12.13: GML of the relation between percentage retrained workers and unemployment rate categorized by different values of the foresight parameter: weak correlation between the percentage retrained and the unemployment rate for all levels of the foresight parameters and a negative correlation between the foresight parameter and the unemployment rate

12.3 Scenario discovery

This section presents the results of the scenario analysis. In our scenario analysis, we considered four scenarios along two axes:

1. Economic growth scenario: High (2%) and Low (1%) economic growth
2. Technological change (TC): Endogenous technological change and Exogenous technological change

In this section, we present the results of the scenario analysis for the two axes separately. This means that we first discuss the results of both scenarios of economic growth, followed by the discussion of the scenario analysis of the two mechanisms of technological growth.

12.3.1 Economic growth

Analysis of the effect of economic growth showed that economic growth does not have an effect on the unemployment rate with respect to the consequences of technological progress (see Appendix L.1). Note the addition of 'with respect to the consequences of technological change' to illustrate that we are aware that economic growth does have an impact on the unemployment rate in the real world. However, our model is built to analyze only the effects of technological progress, and here, the effects of economic growth on the unemployment rate are minimal. This result is, therefore, likely due to the way of how we have defined economic growth in our model (see chapter 8.2.3 for the complete specification).

In the E-MC, economic growth is an endogenous variable that is not only determined by the scenario. Instead, it is determined by the unemployment rate. Since mediation analysis has shown us that the number of uncertain firms is an important intermediary variable, we have looked at how economic growth affects the percentage of uncertain firms (see Figure 12.14). Here, we see that, in general, when economic growth is high, the percentage of uncertain firms is higher. This can be explained by the fact that more economic growth also means more possibilities. When economic growth is high, firms have more resources to act upon the effects of technological change. This also means that there is more chance for firms to make the wrong decisions –which makes them more uncertain. Put in other words, when there is low economic growth, firms cannot invest in the first place, so there is simply no opportunity to become uncertain. Nevertheless, even when there are more uncertain firms when there is high economic growth, the unemployment rate is not affected by the value of economic growth (see Appendix L.1, Figure L.1). Considering the fact that uncertain firms don't act upon their expectations, we can conclude that the effects of economic growth on the unemployment rate –whether this will be an increase or decrease– are crowded out by the effect of economic growth on the number of uncertain firms.

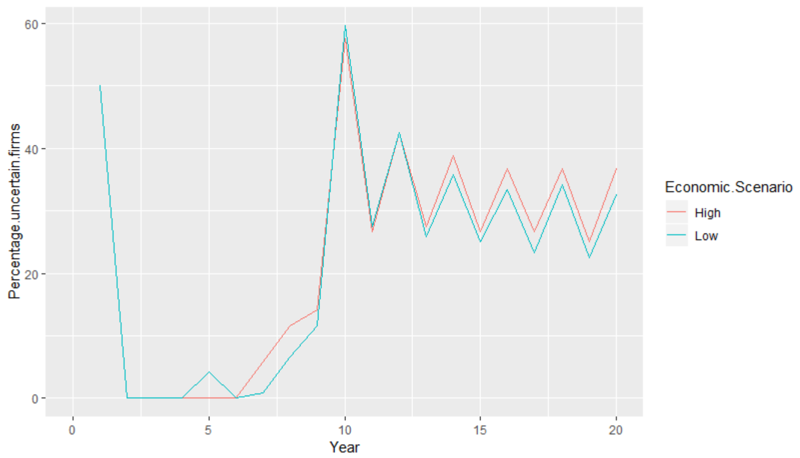


Figure 12.14: Plot median of the percentage uncertain firms per scenario over time: scenario 'High' corresponds to a higher percentage of uncertain firms

12.3.2 Endogenous or exogenous technological change

The scenario analysis of the difference between endogenous technological change and exogenous technological change found that these different mechanisms do not lead to different model outcomes (see Appendix L.2). Here, we found that the percentage of jobs that were automated after 20 years differs (very slightly) between the two scenarios. When technological change is endogenous, on average more jobs are automated (14.54%) than when technological change is exogenous (14.37%). The difference between these percentage is, however, so small that for our results, we do not consider it relevant.

Overall, the scenario analysis did not provide us with significant results. This, however, does not have to mean that these results are not meaningful. For starters, we can state that our model is robust for the different specified mechanisms of technical change in literature. Moreover, this robustness is also true for different scenarios of economic growth. Since the aim of this model is to provide insights in the possible effects of technological change, no significant differences between scenarios –or robustness of the model for these scenarios– can be considered as a solid ground to build on conclusions.

12.4 Unfilled demand: the development of Dutch jobs in the future

This last section discusses the results of the analysis of the effects of technological progress on the jobs individually. For this, we analyzed the demand for jobs[8] and the unfilled demand[9] for these jobs.

Analysis of the unfilled demand in total showed that the proportion of unfilled vacancies to the total labour demand for jobs increases over time (see Figure 12.15). In other words:

[8]The demand for each job is presented as a percentage of the total demand for jobs. So when job A has a demand of 1%, this means that 1% of the total job demand is for job A

[9]Unfilled demand for each job is defined as the percentage vacancies for that job of total demand. So if job A has an unfilled demand percentage of 0.5%, this means that of all total job demand, 2% is unfilled for job A. Consequently, a high demand percentage for job A does not have to mean that a high number of workers have job A. It can also mean that there is a high demand, but there are no workers to fill these vacancies (unfilled demand)

the gap between labour demand and supply widens over time. Figure 12.15 presents this increase of the proportion unfilled labour demand to total labour demand. As one can see, the slope of this curve is the steepest between approximately four and eight years. This means that in these years, the fraction of unfilled labour demand to the total labour demand for jobs increases the most. An explanation of this can be the fact that in these years, firms and workers do not yet suffer from the consequences of technological progress. Since earlier analysis showed that the 'start' of technological progress –e.g. the automation of jobs– happens after approximately nine years, real changes in job demand are expected to be found only a couple years before this 'start'. Note, however, that this also depends on the foresight parameter. If, for example, the foresight parameter is set on ten years, changes in job demand can be seen earlier. As mentioned, in this analysis, the foresight parameter varies from 1 to 6.

Nevertheless, after eight years, we still see a (large) increase in the proportion of unfilled labour demand to total demand. This means that even when due to technological progress robots replace workers, firms increasingly suffer to fill their vacancies. To gain some more insights into this pattern, we looked at the jobs in more detail. For this, we calculated for each job the difference in unfilled demand (open vacancies) between four and twenty years (see Table M.2, Appendix M). Figure 12.16 shows the increase in the proportion of unfilled demand to total labour for the top 15 jobs with the largest difference. The change in unfilled demand over time for all jobs can be found in appendix M, table M.2.

As can be seen from the figure, the job 'pr and marketing advisors' with BRC code 0311 has the largest percentage of unfilled demand. Here, after 20 years, the unfilled vacancies for this job will make up approximately 2.3 % of the total labour demand –e.g. all the filled *and* unfilled vacancies together. While job 0311 has the largest percentage of unfilled labour demand, job 'executive directors' with BRC code 0511 has the most substantial increase in unfilled demand after 12 years: within eight years, the proportion of unfilled demand from job 0511 to total labour demand increases with almost 1 %. Moreover, analysis of the total demand for these jobs showed that both job 0311 and job 0511 experience a significant increase in their total demand (see Appendix M, Table M.1). This means that, within 16 years, demand for these jobs increases substantially. However, this demand cannot be filled, and the number of unfilled vacancies for these jobs increases with almost the same pace as its total demand.

On the contrary, the jobs with BRC code 0811, 0332, 1113 and 0435 have a decrease in their proportion unfilled demand to total demand after 12 years. In other words, during or after the start of technological progress, these jobs will have less trouble to fill their vacancies. Moreover, analysis of the total demand for these jobs shows that the need for these jobs (filled and unfilled) declines over time (see Appendix M, Table M.1).

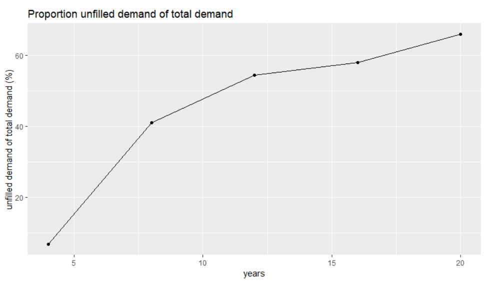


Figure 12.15: The increase of the proportion of unfilled labour demand to total labour demand over time

12.4.1 Unfilled demand per job group

Finally, we analyzed the labour demand and unfilled labour demand per 2-digits BRC codes[10]. Figure 12.17 shows the total demand for each 2-digits BRC code and Figure M.3 shows the

[10] The translations of all the 2-digits BRC codes can be found in appendix A, table A.2

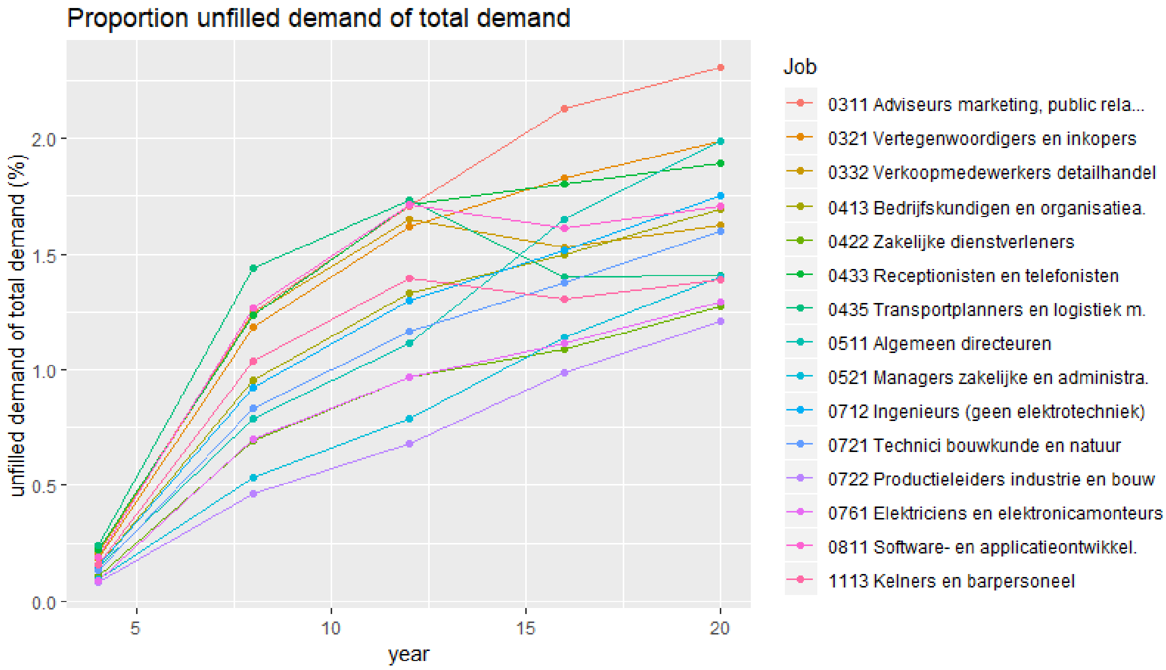


Figure 12.16: Plot: The proportion of unfilled labour demand to the total labour demand for the 15 jobs with the largest differences after 16 years

proportion of unfilled demand over time for each 2-digit BRC codes.

As can be seen from both figures, the total labour demand for job group 04 remains the largest. Furthermore, as time increases, the number of unfilled vacancies for this job group also increases. This job group 04 are the so-called business and administrative professions. Earlier descriptive analysis towards the Dutch workforce has shown that most workers in the Netherlands work in professions that fall within this category (see Appendix B.7, Table **??**). Nevertheless, our analysis states that after 20 years, firms still have trouble to fill their vacancies for these professions. Moreover, both figures show that this increase in unfilled and total labour demand peaks at 12 years. Here, technological progress is happening, and firms will anticipate on this. We, therefore, also state that even when the total demand for these administrative or business professions is the largest, it decreases due to technological progress.

The fact that this demand did not increase substantially –and even largely increased before 12 years– can be explained by the concept of uncertainty. Whereas in the FO model, this decrease would be much larger[11], these results indicate that because of the uncertainty of firms, demand for these jobs will not fall substantially.

If we look at the proportion of unfilled demand to total demand for each 2-digit BRC code, we see that there is a large increase in unfilled vacancies of the 2-digits BRC codes 07, the technical occupations. Here we see that regardless of technological progress, the total labour demand for these jobs increases. So whereas the labour demand for job-group 04 declines due to technological progress, the demand for job-group 07 increases. This is also true for 2-digits BRC job groups 03 and 05. Nevertheless, over time, the proportion of unfilled demand for technical occupations (07) increases the most. The reasons for this decline is two-fold: First, as descriptive analysis about the Dutch workforce has indicated, competencies associated for these professions are very specific and have little overlap with other professions (see Appendix B). If demand for these technical professions increases, it will be hard for firms to

[11] If we look at the FO automation probabilities of jobs within category 04 we see that many 04 jobs are classified as 'high' or 'medium' risk (see Appendix C.2.2, Table C.4)

find workers with technical or similar competencies. The second reason is the fact that these jobs are hard to automate (see Appendix C.2.2, Table C.4). This means that when 'robots' are taking over other jobs, demand for labour will shift to the jobs that are not being automated. In combination with the fact that the Dutch workforce has a limited amount of people with (specific) technical skills, firms will have more and more trouble in filling their vacancies for technical professions.

But which jobs will be taken over by robots? According to these results, for most professions, there will be no significant decline in job demand. Nevertheless, jobs within the 2-digits BRC code 10 ('care and welfare professions') will experience a large decline in their total demand after already four years. This decline, however, stabilizes after 12 years and thus we cannot conclude that this decline is due to technological progress. Instead, we may say that after and *because* technological progress, the decline for jobs in these professions stagnates. Jobs that do suffer from technological progress in their demand are jobs within the BRC 2-digits codes 11 (service professions) and 12 (Transport and Logistics professions). By these jobs, the proportion of unfilled demand also declines.

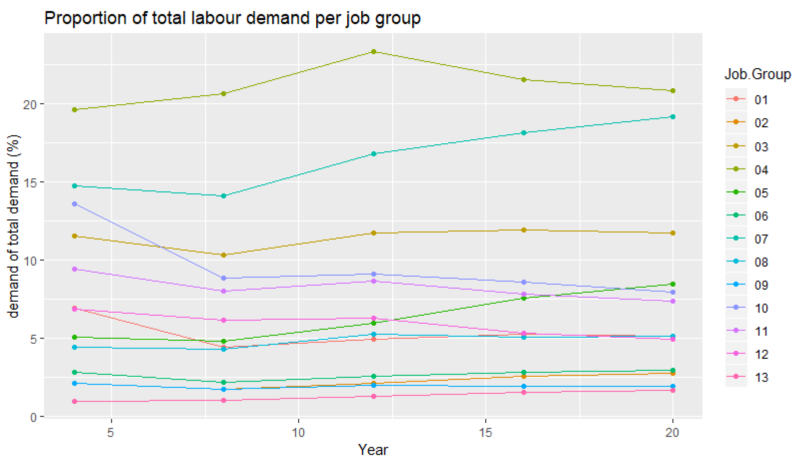


Figure 12.17: Proportion of labour demand for each 2-digit BRC code to total labour demand

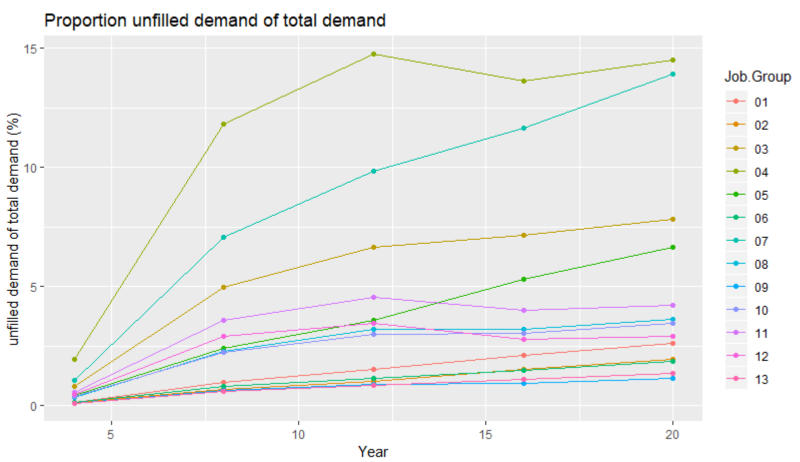


Figure 12.18: Proportion of unfilled demand for each 2-digit BRC code to total labour demand

12.5 Validation of results

As mentioned in chapter 9, for the model validation, we compare the model results with established literature. This section discusses our findings with respect to the validation of the model.

12.5.1 Frey and Osborne assumptions

Since we have used the Frey and Osborne (2013) model as a starting point for the development of the agent-based model, we use their insights as our validation. For starters, we have calculated the probability of automating for the Dutch BRC-jobs according to the method as used in Frey and Osborne (2013). While the name and specifics of jobs differ between the US and the Netherlands, we used the same training data in our machine learning algorithm as in the research of Frey and Osborne (2013). With respect to the validation of this data, we, therefore, rely on their validation of this training data. Nevertheless, as mentioned earlier, to validate the mapping process (the mapping of Dutch and US jobs) and to make the automation probabilities more specific for Dutch occupations, an expert validation or re-constructing of the training data is recommended.

Furthermore, as our analysis of the E-MC has shown, technological progress will mostly affect the unemployment rate through the number of replaced workers. While the goal of the E-MC was to show that the addition of retraining and uncertainty could mitigate the FO results, this found pattern –automation leads to the replacement of workers which in turn leads to an increase in unemployment – was predicted by Frey and Osborne (2013). This, therefore, contributes to the validation of the model[12].

12.5.2 Discovered patterns

Besides the fact that our model can be validated by relying on its background of the Frey and Osborne (2013) study, we also validate the model by observing patterns that were not implemented explicitly but could, either way, be validated through the established literature. Nevertheless, one should note that this method of validation –to our perspective– has a perilous pitfall: the risk of confirmation bias. Accordingly, "confirmation bias, as the term is typically used in the psychological literature, connotes the seeking or interpreting of evidence in ways that are partial to existing beliefs, expectations, or a hypothesis in hand" (Nickerson, 1998, p. 176). For this reason, we only discuss the validation of some 'obvious' found patterns in the model and will compare them with only well-established literature or well-known facts.

To validate the demand for jobs, we use the labour market prognoses as developed by de Grip (2017) in collaboration with The Research Centre for Education and the Labour Market (ROA). Comparing their prognoses and our findings regarding labour demand, we find that our found labour demand follow more or less the same pattern as stated by de Grip (2017). For starters, in both analyses –our model and the prognoses of de Grip (2017)– the total labour demand and unfilled vacancies will increase over time. Furthermore, in the report of de Grip (2017) research, technical occupations (BRC code: 07) and IT occupations (BRC code: 08) will have the highest increase in demand whereas occupations in agriculture will decrease in demand (BRC code 09) (de Grip, 2017). While our model has indeed the highest increase in demand for technical occupations and also an increase in demand for IT occupations, their argument that agricultural jobs will decline the most is not found by our model. In our model, jobs of BRC code 12 (transports and logistics) and BRC code 10 (care and welfare professions) will have the largest decrease in demand. This difference can be explained by how 'total demand' is defined and whether the effects of technological progress are included. For starters, de Grip (2017) define the labour demand for a job as the total of persons working in that job. Here, a decrease in labour demand is a decrease in people working in that profession. As in our model, we define the total labour demand as

[12] Likewise, one could also argue that our model contributes to the validation of the Frey and Osborne (2013) model

the number of new vacancies (filled and unfilled) each year. The fact that in our model the demand for job group 10 declines, can thus also be explained by the fact that the number of vacancies for vacancies (filled and unfilled declines). The number of people working in that profession can remain the same. Moreover, in their prognoses, it is not clear whether the authors included the effects of technological progress.

Another striking finding of our model is the fact that instead of the actual training costs, other factors such as age and work experiences matter for workers when they consider retraining. This fact is also proven and discussed in detail by the research of Jacobson et al. (2005) and may, therefore, serve as validation of the model.

Furthermore, if we look at the percentage of uncertain firms, we see that this variable follows a somewhat cyclic pattern (see Appendix L, figure 12.14). Before the start of technological progress (at year 10), the percentage of uncertain firms drops to almost 0. This means that firms are optimistic and will invest more. The fact that this percentage drops at a high pace before the consequences of technological progress occurs and after that continues as a sinus pattern can be explained by the concept of 'Animal spirits' as defined by Keynes (1937). In his work, Keynes stated that in times of economic upheaval, irrational thoughts might influence peoples' decision making. In other words, even when the effects of technological progress can be estimated, and firms may act upon them, bad decisions in the past can make firms uncertain and thereby influences the decisions of firms. This can lead to the scenario where uncertain firms wait and make decisions, which in turn increases their uncertainty, even more, leading to a downward spiral of bad choices. As our analysis has found, in our model, this uncertainty indirectly influences the unemployment rate by almost 89% and is, therefore, an important determining factor when assessing the consequences of technological progress. Since this uncertainty is only based on the firm's experiences in the past, we cannot say that it is a product of rationality. Instead, this uncertainty –and therefore the behaviour of our model– is determined by 'animal spirits'.

Finally, we look at the found results of the level of 'flexibility'. While many scholars argue that an increase in the number of flexible contracts will lead to less unemployment, we found that the higher the proportion of flexible contracts, the higher the unemployment rate. This can be understood by the positive correlation between the number of replaced workers and the unemployment rate. Here, we have applied a variance of the same assumption as Frey and Osborne (2013): if a robot replaces your job, you will lose your job and become unemployed. The only difference between the FO assumption and the assumption of our model is the possibility of retraining: if a robot replaces your job *and* the firm does not want to retrain you, you will lose your job and become unemployed.

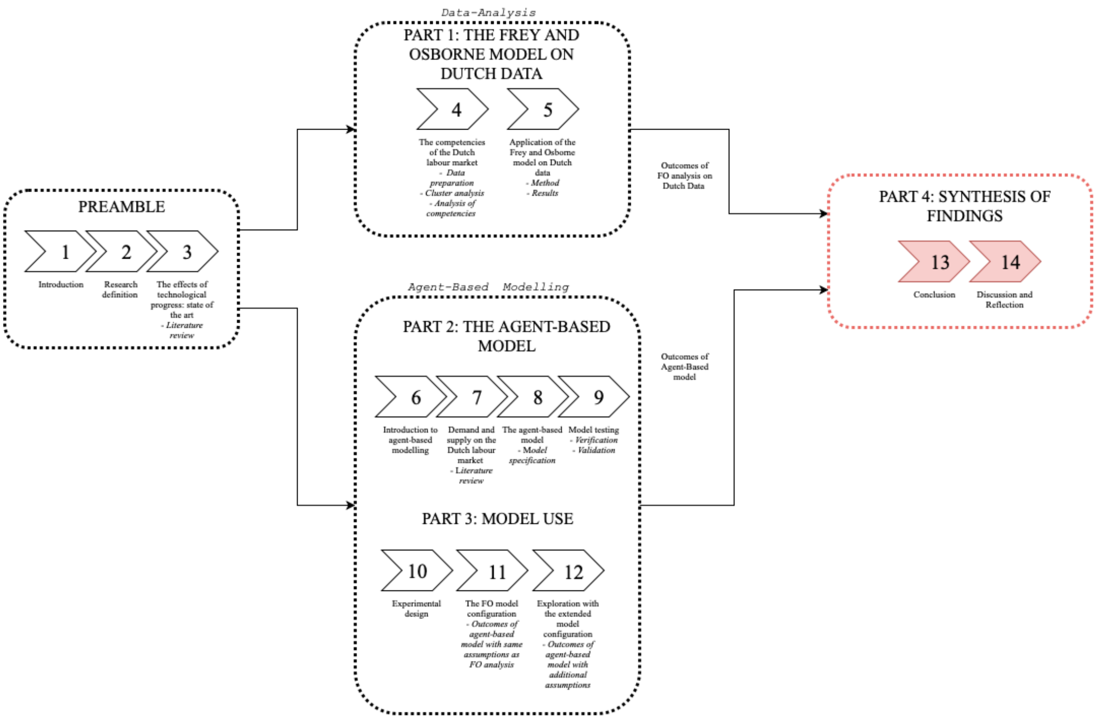
In our model, when workers have a flexible contract, it is cheaper for a firm to fire those workers. Although these workers can be retrained by the firm or by himself, we found that this 'easiness' of firing workers leads to more unemployment. While an analysis about the effects of the flexibility of the labour market on unemployment asks for more exhaustive analysis and even another method, our model fosters more doubts to the effects of labour market flexibility by stating that these effects are the opposite or at least ambiguous.

In this section, we have presented our validation of the model. Admittedly, we are aware of the fact that this validation is far from exhaustive and by no means complete. While the validation of agent-based models is inherently difficult, we argue that two steps can substantially increase the quality of this validation:

- Expert validation of the agent-based model
- Expert validation of the FO model application on Dutch data (including the revisiting of the training data)

These steps can easily be performed and will increase the quality of the validation process. Moreover, by performing these additional validation steps, some valuable insights about the consequences of technological progress can be gained. Unfortunately, due to limited time, these steps could not be performed for this research.

IV
Synthesis of findings



Chapter 13

Conclusion

As the literature about the effects of technological progress revealed, the effects of technological progress are unclear and therefore creates uncertainty on both the demand and the supply side of the labour market. This uncertainty, in turn, can have an impact on the overall performance of the labour market and may affect the behaviour on the labour market even more than the actual technological progress itself.

If we consider the Dutch labour market, we see that one of the major bottlenecks is the mismatch between labour demand and supply. This current state of the Dutch labour market in combination with the uncertainty about the future consequences of technological progress creates a complex problem whereby possible feedback mechanisms and a range of possible interactions between agents adds up to this complexity. In order to gain insights into this problem and to create a framework for policy-testing, a study is performed that has tried to expose the 'true' effects of technological progress as much as possible. By performing both a data-analysis where the Frey and Osborne (2013) study is replicated on Dutch data and by developing an agent-based model that was able to include the effects of uncertainty, retraining and strategic behaviour of firms and workers, this study was able to answer the following research question:

> *How will uncertainty about technological progress eventually influence the effects of technological progress on the Dutch labour market?*

To answer the research question, we first analyzed the effects of technological progress without uncertainty and other complex mechanisms. By replicating the research of Frey and Osborne (2013) on Dutch data, we found that 12 % of the Dutch workers are at a so-called 'high risk' of losing their job due to technological progress. However, according to Frey and Osborne (2013), a job is considered a high risk when the probability of becoming automated is higher than 0.7. Given the fact that our analysis has classified many Dutch jobs as a 'medium risk', where the probability ranges from 0.3 to 0.7, simulation of this FO application has shown that, on average, almost 36% of the Dutch workers will lose their jobs due to technological progress.

While shocking, the same question arise as by the original Frey and Osborne (2013) study of how these results should be understood. Will firms really replace all workers immediately when their jobs are being automated, and to what extent will uncertainty and the possibility of retraining affects these findings? Since data analysis is insufficient to include these 'low-level' interactions and feedback mechanisms, an agent-based model is developed. Experiments with this agent-based model have shown that uncertainty, indeed, matters. Whereas the average unemployment rate after 20 years is 36% in the FO model, the average unemployment rate in the 'extended' model has dropped to 3 %. These findings suggest that, under the right circumstances and policies, unemployment could reduce to under its initial rate of 3.6%.

With respect to other factors that may affect the value of this final unemployment rate –such as policies–, our findings suggest that the level of uncertainty can be seen as a gate of last resort. That is, whereas policies, economic growth and other contextual factors shape the *content* of firms' and workers' expectations, uncertainty eventually determines if these expectations will lead to real behaviour. Analysis of this 'indirect' effect of uncertainty has found that 89% of the effects of technological progress on the unemployment rate are mediated through the percentage of uncertain firms.

Moreover, this uncertainty is slightly influenced by economic growth. In times of high economic growth –in our case 2 %– the level of uncertainty among firms and workers will also be higher. This can be explained by the fact that in times of economic prosperity, there are more opportunities for firms and workers to become uncertain. To put it in other words, when there is low economic growth, firms and workers have less money to decide upon and hence have fewer 'opportunities' to make wrong decisions and to eventually become uncertain. Since the increase in economic growth slightly increases the level of uncertainty, the expected associated increase in the unemployment rate is crowded out by this indirect effect on the level of uncertainty.

To overcome the limitation of the Frey and Osborne (2013) study which did not account for the possibility of retraining, the opportunity for workers to retrain themselves or for firms to retrain their workers was included in the model. However, we found that even though both firms and workers can retrain themselves, this does not affect the final unemployment rate very much. An explanation for this can be found in the results of our analysis of the competencies of Dutch workers. For example, we discovered that competencies associated with technical occupations –such as explosive strength– are only found among workers who work in these professions. Furthermore, due to technological progress, our study showed that the labour demand for these professions would increase. This increase in demand for technical jobs is accompanied with even stronger growth in the number of unfilled vacancies for these jobs, indicating that this demand does not match its supply. These results show that although there is high unemployment, workers do not retrain themselves to the jobs where there is a shortage of demand. This can be explained by the large number of specific competencies needed for these technical jobs, which makes it more difficult for workers to retrain to these jobs.

To stimulate the retraining of workers, we tested a policy that reduces the retraining costs for workers and firms. This policy, however, does have any significant effect on the number of retrained workers and eventually the unemployment rate. Instead, we argue that other factors, such as age, working experiences and the differences in needed competencies are more critical in determining the choice for workers to retrain.

That being the case, our analysis has shown that the replacement of workers due to technological progress can be considered as the most significant consequence of technological progress, which in turn affects unemployment the most. Indeed, this effect can be slightly mitigated through the retraining of workers, but policies that engage in the replacement of workers will be the most effective in reducing the unemployment rate. For example, we have tested how the flexibility of the labour market –here the number of flexible contracts– influences the effects of technological progress. We found that the higher the proportion of flexible contracts, the fewer workers will be replaced, the more workers will be retrained and hence, the lower the final unemployment rate will be. Moreover, we found that by a weak labour market flexibility, the unemployment rate can drop under its current level of 3.6% after 20 years of technological developments. Only in this context of low labour market flexibility, we found that the retraining of workers can affect reducing unemployment.

Furthermore, this study also showed that the number of years that firms and workers consider in their planning is essential for the consequences of technological progress. The larger the number of years that are considered in the planning, the lower the final unemployment rate, and the weaker the effects of technological development will be.

From a policy perspective, these outcomes highlight that even though one may consider retraining as a policy option to mitigate the effects of technological progress, the most success can be gained by policies that recognize the replacement of workers. While many scholars and policymakers argue for more flexibility in the labour market, one should first consider how this relaxation of the labour market interacts with the current technological development. Admittedly, technological progress is happening, and its effects cannot be stopped altogether. However, one could mitigate these effects by imposing policies that stimulate firms and workers to base their decisions on long-term expectations –such as reducing the number of flexible contracts and thereby making it less easy for firms to make ad hoc decisions. Additionally, by imposing these long-term policies, uncertainty about the effects of technological progress can also be reduced.

In our study, we found that uncertainty is a crucial factor in determining the behaviour of firms and workers. While low uncertainty and a short-term orientation can lead to high unemployment, we found that low uncertainty in combination with a long-term orientation –here reflected in a small number of flexible contract–, stimulates the retraining of workers and reduces the ad-hoc replacement of workers. This mechanism, thereby, eventually reduces unemployment below its current level. This means that with respect to our findings of FO application on Dutch data, the combination of the right policies and long-term outlook can turn this doom scenario into a real gloom scenario.

Chapter 14

Discussion

In this chapter, we discuss the findings of this study with regard to already existing or proposed labour market policies. For this, we will take a close look at the so-called 'activating labour market policies' as defined by the OECD. Next, this chapter address the limitations of this study, discuss possibilities for the re-use of the developed models and provide recommendations for further research.

Finally, this chapter also reflect on the societal and social context of the findings of this study.

14.1 Policy implications: activating labour market policies

This section will discuss the outcomes of our study with respect to activating labour market policies. Chapter 3.5 provides a typology of activating labour market policies created by Dinan (2019). Using this typology, we discuss the impact of the outcomes of our model on the activating labour market policies that are based on financial incentives; organizational human capital incentives; and concrete human capital incentives.

14.1.1 Financial incentives

One category of labour market activation strategies are the policies that aim at stimulating the labour market through financial incentives. Here, these financial incentives can be seen as negative financial incentives –e.g. the discouragement of 'bad behaviour' with regard to reducing unemployment– or positive financial incentives –by rewarding behaviour that reduces unemployment. This perspective of rewarding and punishment trough financial means can be seen as a result of how economic theory predicts how a *rational* actor should react to a change in cost-benefits of joining the labour market or increasing demand for labour (Dinan, 2019).

In an analysis of the effectiveness of activating labour market policies in OECD countries, Martin (2015) shows the vital role that cost-benefit policies can play in reducing unemployment. However, the author also states that while these financial incentives can be successful, we should not overestimate them in their effectiveness. That is, the effectiveness of these fiscal policies depends on other, non-financial, factors too.

With regard to the findings of our study, these arguments partly hold. In our research, we have modelled the labour market as a pool of *irrational* firms and workers. Here, we have shown that positive financial policies for unemployed workers do not have a significant effect on the unemployment rate. Rather, non-financial factors such as age, working experience and uncertainty play a more critical role in the choice of workers. Nevertheless, at the demand-side of the labour market, we have found that negative financial incentives –such as increasing the costs for firing workers– are effective in reducing unemployment.

Given these findings, in combination with the results of Martin (2015), two valuable lessons with respect to financial incentives as activating labour market policies can be drawn.

First, since we found that financial incentives do work for the demand side of the labour market, we argue that firms –or better: their strategies– are more rational in their decision making than individual workers are. This, in turn, explains the findings of Martin (2015) that financial incentives are in general very effective but that other factors are important too. In his argument, the author referred to these other factors as to policies that involve workers on the individual level, such as monitoring and controlling the job-search process.

The second lesson closely relates to the first lesson and states that, when considering financial incentives as activating labour market policies, one should consider the different forms of rationality between firms or organizations and workers individually. For example, according to Martin (2015), labour market activating policies often include negative financial incentives for (unemployed) workers such as cutting of unemployment benefits. These policies are often motivated by stating that negative financial incentives avoid the poverty trap and policy dependence. Without going into a discussion about the causes and mechanisms of a poverty trap, our findings suggest that more success in reducing unemployment can be gained by posing financial incentives on firms rather than on workers.

14.1.2 Organizational human capital incentives

A different type of activating labour market policies are the policies that involve organizational human capital incentives. These policies are aimed at creating and reinforcing the links between business and local regional governments (Dinan, 2019). In other words, these policies simplify the matching process between firms and workers. With respect to the findings of our study, these policies could be successful when considering the consequences of technological progress. As our study has shown, the gap between labour demand and supply on the Dutch labour market will grow. This is true for both low and high economic growth and when policies are included that are aimed at the upskilling of workers. For this reason, we, therefore, argue that matching the right workers to firms can be effective in reducing the gap between labour demand and supply.

However, one final note has to be made, and that is that we did not explicitly test for these (combination of) policies. Moreover, since these types of incentives are not commonly included in public employment services, no empirical conclusions can be drawn on the effectiveness of these types of policies (Dinan, 2019).

14.1.3 Concrete human capital incentives

The final type of activating labour market policies involves policies that include concrete human capital incentives. These policies mostly involve investment in human capital through training. In our study, we have analyzed the effect of retraining. Here, we found that only retraining does not have a large effect on the unemployment rate. Besides, we found that while the total number of retrained workers could reduce unemployment, retraining workers by firms can –in a market with low labour market flexibility– increase unemployment. These somewhat contradicting results can be explained by our findings concerning the flexibility of the labour market –e.g. the percentage of flexible contracts– and brings us back to the discussion about financial incentives and the difference between the rationality of firms and workers. When considering a market that consists of a high percentage of workers who have a flexible contract, it is relatively cheap for firms to fire these workers and to hire new workers. Besides, by including the notion of 'opportunity costs', we also argue that in a flexible market, it is relatively cheaper for firms to hire workers. This means that when considering a market with relatively many fixed contracts, firms will sooner consider the possibility of retraining instead of replacing when workers' jobs are being automated. That being the case, firms will hire less new workers since they can fill their vacancies with their workers. This, in turn, explains the slight positive relation between firm retraining and unemployment.

Note, however, that this relation is only found when labour market flexibility is low and that the effect of retraining in total still is effective in reducing unemployment. More importantly, our study has shown that, in either case, when labour market flexibility is low –e.g. it is cheaper for firms to fire workers–, unemployment is lower than in a market with more flexible contracts. This, in turn, strengthens our argument that financial activating labour market policies are the most effective for firms.

Since we have argued that workers are less rational than firms and that their choice of retraining is determined by more factors other than only a cost-benefit analysis, activating labour market policies concerning human capital incentives should include these other contextual factors.

To conclude, our study has shown the importance of considering the true rationality of firms and workers. Whereas firms are often more rational in their decision-making and more prone to financial incentives, workers are often less rational and include more often non-quantitative factors into their decision-making process. With respect to policies that are aimed at reducing unemployment, we, therefore, recommend incorporating these findings when combining a set of labour market activating policies. However, since our study did not explicitly analyze specific activating labour market policies, no one-to-one mapping can be made with our findings and other conclusions about particular labour market policies. Nevertheless, when considering activating labour market policies in a broader perspective and with the help of the framework of Dinan (2019), our findings were able to give some valuable insights.

14.2 Limitations and recommendations

While the combination of a data-analysis and an ABM model did provide us with interesting insights, the study has several limitations. This section will address these limitations of the two models separately. By taking these limitations into account, we discuss the re-usability of the models and the possibilities for further research.

14.2.1 The FO model application

The limitations of the FO model application can be divided into two categories: (1) limitations about the mapping process of Dutch data; and (2) the limitations that are inherent to the original Frey and Osborne (2013) study.

The first strand of limitations involves the mapping process of Dutch datasues regarding the validation of this mapping. For starters, to perform the machine learning analysis, we used the same training data as Frey and Osborne (2013). For this, we had to re-code their US jobs into Dutch jobs with mapping tables provided by O*Net (2019) and Statistics Netherlands (2018). To make sure that the jobs were mapped according to their needed competencies, a hierarchical cluster analysis was performed. However, a well-known limitation of cluster analysis is that it is inherently subjective. That is, the results of any cluster analysis depend heavily on its defined measures ex-ante. During a hierarchical cluster analysis, clusters are based on an arbitrarily defined measure of interconnections. Changes in the value of this measure change the size and shape of clusters (T.-S. Chen et al., 2005).

Besides, by re-coding the training data-set with US jobs to a training data-set with Dutch jobs, we assumed that a Dutch job is considered to be automated when Frey and Osborne (2013) labelled more than half of the underlying O*Net occupations. However, this threshold of 0.5 is set arbitrary, and sensitivity analysis towards this threshold has shown that the labelling process is sensitive to small changes in this threshold. Moreover, it could also be the case that Dutch professions are labelled differently than there US counterparts due to differences in content and tasks between US and Dutch occupations.

Consequently, when using this data and model for further research, we strongly recommend more validation of the mapping process and the training data set. For this, experts can be used in order to assess both the results of the cluster analysis ('are the underlying US jobs clustered to the right aggregated Dutch jobs?') and the labelling of the training data ('are the labels of the training data-set consistent with the content of Dutch jobs?').

Moreover, there are three limitations that are inherent to the limitations of the Frey and

Osborne (2013) model. First, the model does not account for the creation of new jobs in the future or the possibility of retraining. Moreover, their analysis is conducted in 2013, which means that the used training data set can be outdated. Finally, as the simulation with the found FO probabilities of Dutch jobs has shown, their main conclusion is very depended on their definition of 'high risk'. Consequently, these limitations served as the motivation to develop an agent-based model as a mean to overcome these limitations and gain more insights into the real consequences of technological progress.

14.2.2 The agent-based model

While the agent-based model is developed to control for some of the limitations of the FO model, this model also suffers from its imperfections. For starters –and as always the case by agent-based models–, the developed model relies heavily on assumptions. That is, almost every aspect of the agent-based model is based on some assumption. While most assumptions are grounded in literature, no promises can be made on the rightness of our model. For example, even though we modelled specific mechanisms based on literature and empirical evidence, we do not know how these mechanisms should interact with each other. Admittedly, an agent-based model is, in general, developed to assess these interactions, but this also adds up to the difficulty of validating such a model. With regard to our findings, our unemployment rate over time has the shape of an S-curve since we modelled the technological growth curve as a logistic growth function. While modelling the technological change as a logistic growth curve is based on existing well-known research, its direct connection with the shape of the unemployment rate over time may not be realistic. When this model is used in further research, we, therefore, suggest varying with the curve of technological change.

Furthermore, by using data about the number of workers and their jobs, our model may suggest that it resembles more or less the Dutch economy. This, by all means, is not true. When interpreting these findings, one should always keep the goal of this agent-based model in mind: To gain more insights into the consequences of technological progress on Dutch job, not to make precise and quantitative predictions about labour demand in the Netherlands. For this, an exhaustive data-analysis is a more suitable method. Nevertheless, our findings could be used well as a starting point for further quantitative analysis.

At last, we would like to address the limitations of the policy and scenario analysis. For this analysis, we have chosen three policies that are based on literature. While this analysis provided us with some deep insights about the effect of policies and scenario's, the limited number of policies and scenarios also limit us in drawing conclusions about the effect of policies on the Dutch labour market. As our discussion about activating labour market policies revealed, it would be very interesting to include more specific policies in the model to assess their combinations and their impact on society. Moreover, our scenario analysis has shown that the effects of technological progress are robust for economic growth. Getting back to the point of the predictive value of our model, this results further confirms the fact that the agent-based model is not a replication of the Dutch economy. We, therefore, should be cautious by interpreting these findings of the interaction between economic growth, technological progress and unemployment since many other parts of the Dutch economy are not included in the model.

14.2.3 Model re-usability

As been made clear by the previous discussion, both models suffer from some limitations. This, however, does not mean that these models cannot be used for further research. Instead, we believe that by developing both models, we have opened up a lot of opportunities for further research.

For starters, by replicating the FO study, a data-set was constructed that contains all the competencies of Dutch workers. In this study, we performed a small descriptive analysis of the distribution of competencies in the Netherlands, but much more insights can be gained by a more in-depth analysis of this data-set. Since our data-set contains the competencies of Dutch workers on the same level used by Statistics Netherlands, existing information about Dutch workers can be combined. For example, one could study how competencies of Dutch workers have evolved over time and if there are differences in competencies between men and women, workers with different ethical background or variations on the educational level.

Next, to the fact that this data-set can be very interesting on a scientific level, this data can also be used by the formulation of policies or workers individually. For instance, the US government provides the original US data-set to their citizens to help people in their job or education choice. To summarize, by developing this data-set, we opened up plenty of possibilities on both the scientific and policy level.

With respect to the consequences of technological progress, we have seen that the combination of the data analysis and the development of the agent-based model has proven to be useful in gaining insights. Moreover, our model can be used to analyze the consequences of other labour market policies in combination with the effects of technological progress. However, as our discussion about the limitations of the models has revealed, one of the most significant limitations is the issue of validation. To overcome these limitations, we, therefore recommend the validation through expert opinion.

At last, besides of the preference of retraining, the agent-based model did not include the preferences of workers. That is, the model did not account for the personal considerations of workers when they were matched to firms. However, it would be interesting to see how these preferences of workers affect the labour market and hence, the real consequences of technological progress. Furthermore, the agent-based model did not include the possibility of the increase in the productivity of workers because of technological developments. It would, therefore, interesting to see how including these effects make the ‘real’ consequences of technological progress more realistic.

14.3 Reflection: societal and social context

In this study, we have analyzed the consequences of technological progress in the Dutch labour market. For this, we have built upon the work of Frey and Osborne (2013) for estimating the percentage of Dutch workers who are at risk of losing their jobs due to technological progress. By stating that almost 47% of the US citizens are at high risk, Frey and Osborne (2013) ruffled many feathers across the world. Even today and despite some controversy about the content of their study, their findings remain subject of talks. In a recent interview with a Dutch newspaper, Carl Frey again points to the possible catastrophic effects of technological progress. Here, he points explicitly towards the anger that may arise when workers are confronted with the fact that robots are replacing their jobs. As argued in this interview by Frey, “while cashiers are not yet demolishing self-service cash registers, technological progress is not safe for the anger of workers who cannot connect with the modern economy” (Witteman, 2019).

While this may sound a bit radical, this statement of Carl Frey truly exposes a vital point: the issue of the short-term consequences of technological progress. Indeed, by altering the FO model and by including the behaviour of bounded-rational heterogeneous agents, we found that long-term unemployment due to technological progress can be avoided when considering the right policies. Nevertheless, from an ethical perspective, we cannot make predictions of how technological progress may influence the lives and well-being of individual workers. Moreover, our study did not account for differences within the social or economic backgrounds of the workers. This means that our findings do not say anything about which part of society will be affected the most by technological progress. As one may argue, unemployment due to technological progress may be worse for people who have less financial means than for the people who can afford some period with no income. To get back to the point made by Carl Frey in his interview, this can lead to more inequality and anger among the victims of technological progress. This dissatisfaction, in turn, can lead to new political movements and policies that may impact society in a way that could be included in a model.

While our study has shown that long-term unemployment may not reduce through technological progress, this finding should not be used to underestimate the short-term effects of technological development which may impact the economy and society in a way that is not captured by our model. For these reasons, we recommend integrating the societal impacts of technological progress in the model for further research.

In our model, we did not account for the possibility of off-shoring labour. According to Autor et al. (2015), technological progress facilitates globalization by relieving location restraints to move the production process elsewhere. Consequently, work can be off-shored to places with lower production costs –which is often reflected in lower labour costs. As one can imagine, this reallocation of jobs can have a significant impact on the Dutch labour force. According to Statistics Netherlands, in the period between 2014 and 2016, already six per cent of companies with a workforce larger than 50 people moved some of their business operations abroad (Statistics Netherlands, 2018).

Moreover, since several scholars have pointed towards the co-emerge of labour off-shoring and technological progress –mainly developments in IT–, there is a good chance that the off-shoring of jobs in the Netherlands will increase (DeCanio, 2016; Goos et al., 2009). According to Autor et al. (2015) this will mostly affect the jobs of low-wage labour. Getting back to the argument of the short-term consequences of technological progress, this off-shoring of jobs is an important factor to be considered. Therefore, in combination with the societal impacts of technological development, we recommend to include the off-shoring of jobs in the model for further research.

References

Acemoglu, D., & Autor, D. (2010, 2). Skills, Tasks and Technologies: Implications for Employment and Earnings. *National Bureau of Economic Research, Inc, NBER Working Papers*, *4*. doi: 10.1016/S0169-7218(11)02410-5

Ahituv, A., & Zeira, J. (2010, 11). Technical Progress and Early Retirement. *The Economic Journal*, *121*(551), 171–193. Retrieved from `https://dx.doi.org/10.1111/j.1468-0297.2010.02392.x` doi: 10.1111/j.1468-0297.2010.02392.x

Akcigit, U., & Howitt, P. (2014, 1). What Do We Learn From Schumpeterian Growth Theory? *Handbook of Economic Growth*, *2*, 515–563. Retrieved from `https://www.sciencedirect.com/science/article/pii/B978044453540500001X` doi: 10.1016/B978-0-444-53540-5.00001-X

Appelbaum, E., & Schmitt, J. (2009, 10). Review article: Low-wage work in high-income countries: Labor-market institutions and business strategy in the US and Europe. *Human Relations*, *62*(12), 1907–1934. Retrieved from `https://doi.org/10.1177/0018726709349200` doi: 10.1177/0018726709349200

Armstrong, S., Sotala, K., & hÉigeartaigh, S. S. □. (2014). The errors, insights and lessons of famous AI predictions – and what they mean for the future. *Journal of Experimental & Theoretical Artificial Intelligence*, *26*(3), 317–342. Retrieved from `https://doi.org/10.1080/0952813X.2014.895105` doi: 10.1080/0952813X.2014.895105

Arntz, M., Gregory, T., & Zierahn, U. (2017, 10). Revisiting the risk of automation. *Economics Letters*, *159*, 157–160. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0165176517302811` doi: 10.1016/J.ECONLET.2017.07.001

Autor, D. H. (2015, 9). Why Are There Still So Many Jobs? The History and Future of Workplace Automation. *Journal of Economic Perspectives*, *29*(3), 3–30. Retrieved from `http://www.aeaweb.org/articles?id=10.1257/jep.29.3.3` doi: 10.1257/jep.29.3.3

Autor, D. H., Dorn, D., & Hanson, G. H. (2015, 5). Untangling Trade and Technology: Evidence from Local Labour Markets. *The Economic Journal*, *125*(584), 621–646. Retrieved from `https://doi.org/10.1111/ecoj.12245` doi: 10.1111/ecoj.12245

Autor, D. H., Levy, F., & Murnane, R. J. (2003, 11). The Skill Content of Recent Technological Change: An Empirical Exploration*. *The Quarterly Journal of Economics*, *118*(4), 1279–1333. Retrieved from `https://dx.doi.org/10.1162/003355303322552801` doi: 10.1162/003355303322552801

Ayton, P., & Fischer, I. (2004). The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness? *Memory & Cognition*, *32*(8), 1369–1378. Retrieved from `https://doi.org/10.3758/BF03206327` doi: 10.3758/BF03206327

Bakhshi, H., Downing, J. M., Osborne, M. A., & Schneider, P. (2017). *The future of skills : employment in 2030* (Tech. Rep.). London. Retrieved from `https://www.nesta.org.uk/report/the-future-of-skills-employment-in-2030/`

Barnichon, R., & Figura, A. (2015, 10). Labor Market Heterogeneity and the Aggregate Matching Function. *American Economic Journal: Macroeconomics*, *7*(4), 222–249. Retrieved from `http://www.aeaweb.org/articles?id=10.1257/mac.20140116` doi: 10.1257/mac.20140116

Bassi, L. J., Ludwi, J., & Mcmurrer, D. P. (2003). Profiting from Learning: Firm-Level Effects of Training Investments and Market Implications. *Singapore Management Review*, *24*(3), 61–76. Retrieved from `https://search.proquest.com/openview/`

9c9c254b7d5576f3db6f747b190a0db2/1?pq-origsite=gscholar&cbl=30516
Bergmann, B. R. (1974). Occupational Segregation, Wages and Profits When Employers Discriminate by Race or Sex. *Eastern Economic Journal*, *1*(2), 103–110. Retrieved from http://www.jstor.org/stable/40315472
Bhattacharya, U., Hsu, P.-H., Tian, X., & Xu, Y. (2017, 10). What Affects Innovation More: Policy or Policy Uncertainty? *Journal of Financial and Quantitative Analysis*, *52*(05), 1869–1901. Retrieved from https://www.cambridge.org/core/product/identifier/S0022109017000540/type/journal_article doi: 10.1017/S0022109017000540
Bloom, N., Bond, S., & Van Reenen, J. (2007, 4). Uncertainty and Investment Dynamics. *The Review of Economic Studies*, *74*(2), 391–415. Retrieved from https://dx.doi.org/10.1111/j.1467-937X.2007.00426.x doi: 10.1111/j.1467-937X.2007.00426.x
Boehmke, B. (2017). *Hierarchical Cluster Analysis.* Retrieved from https://uc-r.github.io/hc_clustering
Box, G. (1979, 1). Robustness in the Strategy of Scientific Model Building. *Robustness in Statistics*, 201–236. Retrieved from https://www.sciencedirect.com/science/article/pii/B9780124381506500182?via%3Dihub doi: 10.1016/B978-0-12-438150-6.50018-2
Brewer, J., & Hunter, A. (1989). *Multimethod research: A synthesis of styles.* Thousand Oaks, CA, US: Sage Publications, Inc.
Burdett, K., Shi, S., & Wright, R. (2001). Pricing and Matching with Frictions. *Journal of Political Economy*, *109*(5), 1060–1085. Retrieved from http://www.jstor.org/stable/10.1086/322835 doi: 10.1086/322835
Caballero, R. J. (2010). Macroeconomics after the Crisis: Time to Deal with the Pretense-of-Knowledge Syndrome. *National Bureau of Economic Research Working Paper Series, No. 16429.* Retrieved from http://www.nber.org/papers/w16429 doi: 10.3386/w16429
Camagni, R. P. (1991). Technological Change, Uncertainty and Innovation Networks: Towards a Dynamic Theory of Economic Space. In D. E. Boyce, P. Nijkamp, & D. Shefer (Eds.), *Regional science: Retrospect and prospect* (pp. 211–249). Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved from https://doi.org/10.1007/978-3-642-76311-3_10 doi: 10.1007/978-3-642-76311-3{_}10
CBS. (2019). *Werklozen.* Retrieved from https://www.cbs.nl/nl-nl/visualisaties/dashboard-arbeidsmarkt/werklozen
CBS Statline. (2019). *Werkzame beroepsbevolking; beroep.* Retrieved from Werkzameberoepsbevolking;beroep
Chen, S.-H., Kaboudan, M., Du, Y.-R., Neugart, M., & Richiardi, M. (2012). *Agent-Based Models of the Labor Market.* Oxford University Press. Retrieved from http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199844371.001.0001/oxfordhb-9780199844371-e-44 doi: 10.1093/oxfordhb/9780199844371.013.44
Chen, T.-S., Tsai, T.-H., Chen, Y.-T., Lin, C.-C., Chen, R.-C., Li, S.-Y., & Chen, H.-Y. (2005). A combined K-means and hierarchical clustering method for improving the clustering efficiency of microarray. In *2005 international symposium on intelligent signal processing and communication systems* (pp. 405–408). doi: 10.1109/ISPACS.2005.1595432
Coad, A., Segarra, A., & Teruel, M. (2016, 3). Innovation and firm growth: Does firm age play a role? *Research Policy*, *45*(2), 387–400. Retrieved from https://www.sciencedirect.com/science/article/pii/S0048733315001687 doi: 10.1016/J.RESPOL.2015.10.015
Coccia, M. (2017). Sources of technological innovation: Radical and incremental innovation problem-driven to support competitive advantage of firms. *Technology Analysis & Strategic Management*, *29*(9), 1048–1061. Retrieved from https://doi.org/10.1080/09537325.2016.1268682 doi: 10.1080/09537325.2016.1268682
Croce, G., & Ghignoni, E. (2012). Demand and Supply of Skilled Labour and Overeducation in Europe: A Country-level Analysis. *Comparative Economic Studies*, *54*(2), 413–439. Retrieved from https://doi.org/10.1057/ces.2012.12 doi: 10.1057/ces.2012.12
Daron Acemoglu, & Restrepo, P. (2018). *Artificial Intelligence, Automation and Work *.* Re-

trieved from http://www.nber.org/papers/w24196 doi: 10.3386/w24196

DeCanio, S. J. (2016, 9). Robots and humans – complements or substitutes? *Journal of Macroeconomics*, *49*, 280–291. Retrieved from https://www.sciencedirect.com/science/article/pii/S016407041630043X doi: 10.1016/J.JMACRO.2016.08.003

de Grip, A. (2017). *De arbeidsmarkt naar opleiding en beroep tot 2022* (Tech. Rep.). Maastricht: Researchcentrum voor Onderwijs en Arbeidsmarkt (ROA). Retrieved from http://www.hbomonitor.nl/application/files/5914/7549/1224/ROA_R_2015_6.pdf

Dinan, S. (2019, 1). A typology of activation incentives. *Social Policy & Administration*, *53*(1), 1–15. Retrieved from https://doi.org/10.1111/spol.12456 doi: 10.1111/spol.12456

Dixit, A. (1992, 3). Investment and Hysteresis. *Journal of Economic Perspectives*, *6*(1), 107–132. Retrieved from http://www.aeaweb.org/articles?id=10.1257/jep.6.1.107 doi: 10.1257/jep.6.1.107

Dosi, G., Fagiolo, G., & Roventini, A. (2006). An Evolutionary Model of Endogenous Business Cycles. *Computational Economics*, *27*(1), 3–34. Retrieved from https://doi.org/10.1007/s10614-005-9014-2 doi: 10.1007/s10614-005-9014-2

Dosi, G., Napoletano, M., Roventini, A., & Treibich, T. (2016). *Micro and macro policies in the Keynes+Schumpeter evolutionary models.* doi: 10.1007/s00191-016-0466-4

Eliasson, G., Olavi, G., & Heiman, M. (1976). *A Micro-Macro Interactive Simulation Model of the Swedish Economy: Preliminary Model Specification* (Tech. Rep. No. 7). Stockholm. Retrieved from http://hdl.handle.net/10419/94789

Fagerberg, J. (2000). *Technological Progress, Structural Change and Productivity Growth: A Comparative Study* (Tech. Rep. No. 5). Centre for Technology, Innovation and Culture, University of Oslo. Retrieved from https://EconPapers.repec.org/RePEc:tik:wpaper:5

Fisher, J., & Pry, R. (1971, 1). A simple substitution model of technological change. *Technological Forecasting and Social Change*, *3*, 75–88. Retrieved from https://www.sciencedirect.com/science/article/pii/S0040162571800057 doi: 10.1016/S0040-1625(71)80005-7

Fitzenberger, B., Osikominu, A., & Volter, R. (2008). Get Training or Wait? Long-Run Employment Effects of Training Programs for the Unemployed in West Germany. *Annales economie et de Statistique*(91/92), 321–355. Retrieved from http://www.jstor.org/stable/27917250 doi: 10.2307/27917250

Fleisher, M. S., Tsacoumis, S., & Human Resources Research Organization. (2018). *O*NET Analyst Occupational Skills Ratings: Procedures Update* (Tech. Rep.). Raleigh: National Center for O*NET Development. Retrieved from HumanResourcesResearchOrganization(HumRRO)

Freeman, C. (1994). The economics of technical change. *Cambridge Journal of Economics*, *18*(5), 463–514. Retrieved from http://www.jstor.org/stable/24231814

Freeman, R. B. (2005, 6). Labour market institutions without blinders: The debate over flexibility and labour market performance. *International Economic Journal*, *19*(2), 129–145. Retrieved from https://doi.org/10.1080/10168730500080675 doi: 10.1080/10168730500080675

Frey, C. B., & Osborne, M. (2018). *Automation and the future of work: understanding the numbers.* Retrieved from https://www.oxfordmartin.ox.ac.uk/opinion/view/423

Frey, C. B., & Osborne, M. A. (2013). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, *114*(C), 254–280. Retrieved from https://EconPapers.repec.org/RePEc:eee:tefoso:v:114:y:2017:i:c:p:254-280

Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, *114*(C), 254–280. Retrieved from https://EconPapers.repec.org/RePEc:eee:tefoso:v:114:y:2017:i:c:p:254-280

Gajan, J. M., Izquierdo, L., Izquierdo, S., Santos, J., Olmo, R. d., Lopez-Paredes, A., & Edmonds, B. (2009, 1). Errors and Artefacts in Agent-Based Modelling. *Journal of Artifi-*

cial Societies and Social Simulation, *12*(1). Retrieved from https://EconPapers.repec.org/RePEc:jas:jasssj:2008-10-2

Gerritsen, M., & Høj, J. (2013). *The Dutch Labour Market: Preparing for the Future.* Paris. Retrieved from https://doi.org/10.1787/5k4dlff4wcwl-en.

Goos, M., Manning, A., & Salomons, A. (2009, 2). Recent changes in the European employment structure: the roles of technology, globalization and institutions.

Goos, M., Manning, A., & Salomons, A. (2014, 8). Explaining Job Polarization: Routine-Biased Technological Change and Offshoring. *American Economic Review*, *104*(8), 2509–2526. Retrieved from http://www.aeaweb.org/articles?id=10.1257/aer.104.8.2509 doi: 10.1257/aer.104.8.2509

Graetz, G., & Michaels, G. (2018, 7). Robots at Work. *The Review of Economics and Statistics*, *100*(5), 753–768. Retrieved from https://doi.org/10.1162/rest_a_00754 doi: 10.1162/rest{_}a{_}00754

Greenhalgh, C., & Mavrotas, G. (1996, 3). Job Training, New Technology and Labour Turnover. *British Journal of Industrial Relations*, *34*(1), 131–150. Retrieved from https://doi.org/10.1111/j.1467-8543.1996.tb00474.x doi: 10.1111/j.1467-8543.1996.tb00474.x

Gulen, H., & Ion, M. (2016). Policy uncertainty and corporate investment. *Review of Financial Studies*. doi: 10.1093/rfs/hhv050

Hadani, M., Bonardi, J.-P., & Dahan, N. M. (2017). Corporate political activity, public policy uncertainty, and firm outcomes: A meta-analysis. *Strategic Organization*, *15*(3), 338–366. Retrieved from https://doi.org/10.1177/1476127016651001 doi: 10.1177/1476127016651001

Handel, M. J. (2016). The O*NET content model: strengths and limitations. *Journal for Labour Market Research*, *49*(2), 157–176. Retrieved from https://doi.org/10.1007/s12651-016-0199-8 doi: 10.1007/s12651-016-0199-8

Handley, K., & Limão, N. (2015, 11). Trade and Investment under Policy Uncertainty: Theory and Firm Evidence. *American Economic Journal: Economic Policy*, *7*(4), 189–222. Retrieved from http://www.aeaweb.org/articles?id=10.1257/pol.20140068 doi: 10.1257/pol.20140068

Heyman, F. (2016, 8). Job polarization, job tasks and the role of firms. *Economics Letters*, *145*, 246–251. Retrieved from https://www.sciencedirect.com/science/article/pii/S0165176516302361 doi: 10.1016/J.ECONLET.2016.06.032

Hillage, J., & Pollard, E. (1998). *Employability: Developing a framework for policy analysis* (Vol. 107).

Hinton, G., & Roweis, S. (2002). Stochastic Neighbor Embedding. In *Proceedings of the 15th international conference on neural information processing systems* (pp. 857–864). Cambridge, MA, USA: MIT Press. Retrieved from http://dl.acm.org/citation.cfm?id=2968618.2968725

Holland, S. (2017). *Data Analysis in the Geosciences: Cluster Analysis.* Retrieved from http://strata.uga.edu/8370/lecturenotes/clusterAnalysis.html

Hoppe, H. C. (2000, 2). Second-mover advantages in the strategic adoption of new technology under uncertainty. *International Journal of Industrial Organization*, *18*(2), 315–338. Retrieved from https://www.sciencedirect.com/science/article/pii/S0167718798000204 doi: 10.1016/S0167-7187(98)00020-4

Imai, K., Keele, L., Tingley, D., & Yamamoto, T. (2009). *Causal Mediation Analysis Using R.* doi: 10.1007/978-1-4419-1764-5{_}8

Jacobson, L. S., LaLonde, R., & Sullivan, D. (2005). Is retraining displaced workers a good investment? *Economic Perspectives*(Q II), 47–66. Retrieved from https://EconPapers.repec.org/RePEc:fip:fedhep:y:2005:i:qii:p:47-66:n:v.29no.2

Jaffry, S. W., & Treur, J. (2013). *Agent-Based and Population-Based Modeling of Trust Dynamics* (Vol. 7770). doi: 10.1007/978-3-642-36815-8{_}6

Juridisch Loket. (2019). *Hoogte van de WW-uitkering.* Retrieved from https://www.juridischloket.nl/werk-en-inkomen/werkloosheid-en-bijstand/duur-ww-uitkering/

Kassambara, A. (2018). *Determining The Optimal Number Of Clusters: 3 Must Know Meth-*

ods. Retrieved from https://www.datanovia.com/en/lessons/determining-the-optimal-number-of-clusters-3-must-know-methods/

Keynes, J. M. (1937, 2). The General Theory of Employment. *The Quarterly Journal of Economics*, *51*(2), 209–223. Retrieved from https://dx.doi.org/10.2307/1882087 doi: 10.2307/1882087

Koopmans, T. C., & Beckmann, M. (1957). Assignment Problems and the Location of Economic Activities. *Econometrica*, *25*(1), 53–76. Retrieved from http://www.jstor.org/stable/1907742 doi: 10.2307/1907742

Levy, F., & Murnane, R. J. (2004). *How Computers Are Creating the Next Job Market.* Princeton University Press. Retrieved from http://www.jstor.org/stable/j.ctt1r2frw

Louie, M. A., & Carley, K. M. (2008, 2). Balancing the criticisms: Validating multi-agent models of social systems. *Simulation Modelling Practice and Theory*, *16*(2), 242–256. Retrieved from https://www.sciencedirect.com/science/article/pii/S1569190X07001542 doi: 10.1016/J.SIMPAT.2007.11.011

MacCrory, F., Westerman, G., Alhammadi, Y., & Brynjolfsson, E. (2014). Racing With and Against the Machine: Changes in Occupational Skill Composition in an Era of Rapid Technological Advance. In *Icis.*

Makridakis, S. (2017, 6). The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms. *Futures*, *90*, 46–60. Retrieved from https://www.sciencedirect.com/science/article/pii/S0016328717300046 doi: 10.1016/J.FUTURES.2017.03.006

Manders, T., & Kool, C. (2015). *Nederland in 2030 en 2050: twee referentiescenario's* (Tech. Rep.). Den Haag: Planbureau voor de Leefomgeving en Centraal Planbureau. Retrieved from https://www.wlo2015.nl/wp-content/uploads/PBL_2015_WLO_Nederland-in-2030-en-2050_1558.pdf

March, J. G., & Simon, H. A. (1958). *Organizations.* New York: John Wiley and Sons. Retrieved from SSRN:https://ssrn.com/abstract=1496194

Martin, J. (2015). Activation and active labour market policies in OECD countries: stylised facts and evidence on their effectiveness. *IZA Journal of Labor Policy*, *4*(1), 1–29. Retrieved from https://EconPapers.repec.org/RePEc:spr:izalpo:v:4:y:2015:i:1:p:1-29:10.1186/s40173-015-0032-y

Mit, D. A., & Restrepo, P. (2018). *Artificial Intelligence, Automation and Work *.* Retrieved from http://www.nber.org/papers/w24196 doi: 10.3386/w24196

Mokyr, J., Vickers, C., & Ziebarth, N. L. (2015, 9). The History of Technological Anxiety and the Future of Economic Growth: Is This Time Different? *Journal of Economic Perspectives*, *29*(3), 31–50. Retrieved from http://www.aeaweb.org/articles?id=10.1257/jep.29.3.31 doi: 10.1257/jep.29.3.31

Nedelkoska, L., & Quintini, G. (2018). *Automation, skills use and training.* Retrieved from http://dx.doi.org/10.1787/2e2f4eea-en doi: 10.1787/2e2f4eea-en

Neugart, M. (2004). *Endogenous matching functions: An agent-based computational approach* (Vol. 07). doi: 10.1142/S0219525904000147

Nickerson, R. S. (1998, 6). Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Review of General Psychology*, *2*(2), 175–220. Retrieved from https://doi.org/10.1037/1089-2680.2.2.175 doi: 10.1037/1089-2680.2.2.175

OECD Skills Strategy Diagnostic Report: The Netherlands 2017. (2017). OECD. Retrieved from https://www.oecd-ilibrary.org/education/oecd-skills-strategy-diagnostic-report-the-netherlands-2017_9789264287655-en doi: 10.1787/9789264287655-en

O*Net. (2019). *Disclaimer.* Retrieved from https://www.onetonline.org/help/disclaimer/

Parzen, E. (1962). On Estimation of a Probability Density Function and Mode. *The Annals of Mathematical Statistics*, *33*(3), 1065–1076. Retrieved from http://www.jstor.org/stable/2237880

PBL. (2019). *About PBL.* Retrieved from https://www.pbl.nl/en/aboutpbl

PWC. (2018). *Future of work 2030: A wake-up call for organisations, people and government* (Tech. Rep.). Amsterdam: PriceWaterhouseCoopers. Retrieved from https://www.pwc.nl/nl/dienstverlening/people-and-organisation/documents/pwc-future-of-work-2030.pdf

Rasmussen, C. E. (2004). Gaussian Processes in Machine Learning. In O. Bousquet, U. von Luxburg, & G. Rätsch (Eds.), *Advanced lectures on machine learning: Ml summer schools 2003, canberra, australia, february 2 - 14, 2003, tübingen, germany, august 4 - 16, 2003, revised lectures* (pp. 63–71). Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved from https://doi.org/10.1007/978-3-540-28650-9_4 doi: 10.1007/978-3-540-28650-9{_}4

Ravn, M. O., & Sterk, V. (2017, 10). Job uncertainty and deep recessions. *Journal of Monetary Economics*, *90*, 125–141. Retrieved from https://www.sciencedirect.com/science/article/pii/S0304393217300806 doi: 10.1016/J.JMONECO.2017.07.003

Reeder, M., & Tsacoumis, S. (2017). *O*NET Analyst Ratings of Occupational Abilities: Analysis Cycle 17 Results* (Tech. Rep.). National Center for O*NET Development. Retrieved from https://www.onetcenter.org/dl_files/Wave17.pdf

Reinganum, J. F. (1982). A Dynamic Game of R and D: Patent Protection and Competitive Behavior. *Econometrica*, *50*(3), 671–688. Retrieved from http://www.jstor.org/stable/1912607 doi: 10.2307/1912607

Rencher, A., & Christensen, W. (2012). *Methods of Multivariate Analysis* (Vol. 93). John Wiley & Sons. doi: 10.1002/9781118391686

Revelle, W. (1979, 1). Hierarchical Cluster Analysis And The Internal Structure Of Tests. *Multivariate Behavioral Research*, *14*(1), 57–74. Retrieved from https://doi.org/10.1207/s15327906mbr1401_4 doi: 10.1207/s15327906mbr1401{_}4

Romer, P. M. (1990, 10). Endogenous Technological Change. *Journal of Political Economy*, *98*(5, Part 2), S71-S102. Retrieved from https://doi.org/10.1086/261725 doi: 10.1086/261725

Sanders, C. (2012). *Skill Accumulation , Skill Uncertainty , and Occupational Choice.* Retrieved from https://ideas.repec.org/p/red/sed012/633.html

Sargent, R. G. (2013, 2). Verification and validation of simulation models. *Journal of Simulation*, *7*(1), 12–24. Retrieved from https://doi.org/10.1057/jos.2012.20 doi: 10.1057/jos.2012.20

Schaal, E. (2017, 11). Uncertainty and Unemployment. *Econometrica*, *85*(6), 1675–1721. Retrieved from https://doi.org/10.3982/ECTA10557 doi: 10.3982/ECTA10557

Smith, T. E., & Zenou, Y. (2003, 1). *Spatial Mismatch, Search Effort and Urban Spatial Structure* (Tech. Rep. No. 692). Retrieved from https://EconPapers.repec.org/RePEc:iza:izadps:dp692

Spaanderman, k. (2018). *Economies of the future: Robots and Artificial Intelligence, the new economic motor or downfall of the working class?* (Doctoral dissertation, Delft University of Technology). Retrieved from http://resolver.tudelft.nl/uuid:f30b5d2e-d48d-4c90-8780-3d55a2e1a7d3

Statistics Netherlands. (2018). *Internationalisation Monitor 2018, second quarter* (Tech. Rep.). The Hague: Statistics Netherlands. Retrieved from https://www.cbs.nl/en-gb/publication/2018/26/internationalisation-monitor-2018-second-quarter

Stiglitz, J., & Rotschild, M. (1978, 1). Equilibrium in competitive insurance market: An essay on the economics of imperfect information. *Uncertainty in Economics*, 257–280. Retrieved from https://www.sciencedirect.com/science/article/pii/B9780122148507500243 doi: 10.1016/B978-0-12-214850-7.50024-3

Storm, S., & Naastepad, C. W. M. (2012). *Macroeconomics Beyond the NAIRU.* Harvard University Press. Retrieved from http://www.jstor.org/stable/j.ctt24hhs5

Sutcliffe, K. M., & Zaheer, A. (1998, 1). Uncertainty in the transaction environment: an empirical test. *Strategic Management Journal*, *19*(1), 1–23. Retrieved from https://doi.org/10.1002/(SICI)1097-0266(199801)19:1<1::AID-SMJ938>3.0.CO2-5 doi: 10.1002/(SICI)1097-0266(199801)19:1<1::AID-SMJ938>3.0.CO;2-5

Tesfatsion, L. (2006, 1). Chapter 16 Agent-Based Computational Economics: A Constructive Approach to Economic Theory. *Handbook of Computational Economics*, *2*, 831–880. Retrieved from https://www.sciencedirect.com/science/article/pii/S1574002105020162 doi: 10.1016/S1574-0021(05)02016-2

Tibshirani, R., Walther, G., & Hastie, T. (2001, 1). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *63*(2), 411–423. Retrieved from https://doi.org/10.1111/1467-9868.00293 doi: 10.1111/1467-9868.00293

Tisue, S., & Wilensky, U. (2004). NetLogo: A simple environment for modeling complexity. In *in international conference on complex systems* (pp. 16–21). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.117.949

UWV. (2018). *About UWV.* Retrieved from https://www.uwv.nl/overuwv/english/about-us-executive-board-organization/detail/about-us

Van Dam, K., Nikolic, I., & Lukszo, Z. (2013). *Agent-Based Modelling of Socio-Technical Systems* (9th ed.). Springer Netherlands. doi: 10.1007/978-94-007-4933-7

van der Maaten, L., & Hinton, G. (2008). Visualizing Data using t-SNE. *Journal of Machine learning research*, *9*, 25790–2605. Retrieved from http://www.jmlr.org/papers/v9/vandermaaten08a.html

Wang, Y., Chen, C. R., & Huang, Y. S. (2014, 1). Economic policy uncertainty and corporate investment: Evidence from China. *Pacific-Basin Finance Journal*, *26*, 227–243. Retrieved from https://www.sciencedirect.com/science/article/pii/S0927538X13000966 doi: 10.1016/J.PACFIN.2013.12.008

Weel, B., & Kok, S. (2013). *De Nederlandse arbeidsmarkt in taken Eerste bevindingen uit de Nederlandse Skills Survey* (Tech. Rep.). Den Haag: Centraal Planbureau. Retrieved from https://www.cpb.nl/publicatie/de-nederlandse-arbeidsmarkt-in-taken-eerste-bevindingen-uit-de-nederlandse-skills-survey

Witteman, J. (2019, 6). *Interview Carl Frey.* Amsterdam. Retrieved from https://www.volkskrant.nl/economie/economisch-historicus-carl-frey-onvrede-over-technologie-kan-nieuwe-voedingsbodem-voor-populisten-worden~bd046751/?referer=https%3A%2F%2Fwww.google.com%2F

Zahra, S. A. (1996, 5). Technology strategy and financial performance: Examining the moderating role of the firm's competitive environment. *Journal of Business Venturing*, *11*(3), 189–219. Retrieved from https://www.sciencedirect.com/science/article/abs/pii/0883902696000018 doi: 10.1016/0883-9026(96)00001-8

Zhang, B., Langfelder, P., & Horvath, S. (2007, 11). Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics*, *24*(5), 719–720. Retrieved from https://dx.doi.org/10.1093/bioinformatics/btm563 doi: 10.1093/bioinformatics/btm563

Appendices

Appendix A

Occupations translation tables

Table A.1: BRC code. job name Dutch and job name English

Brc	Job	Job english
0111	Docenten hoger onderwijs en hoog onderwijs	Higher education teachers
0112	Docenten beroepsgerichte vakken	Profession-oriented teachers
0113	Docenten algemene vakken secundair onderwijs	Teachers of general subjects in secondary education
0114	Leerkrachten basisonderwijs	Primary education teachers
0115	Onderwijskundigen en overige docenten	Educationalists and other teachers
0121	Beroepsgroep sportinstructeurs	Professional group of sports instructors
0131	Leidsters kinderopvang	Childcare leaders
0211	Bibliothecarissen en conservatoren	Librarians and curators
0212	Auteurs en taalkundigen	Authors and linguists
0213	Journalisten	Journalists
0214	Beeldend kunstenaars	Visual artists
0215	Uitvoerend kunstenaars	Performing artists
0221a	Grafisch vormgevers en productontwerpers -Commercial and Industrial Designers	Graphic designers and product designers - Commercial and Industrial Designers
0221b	Grafisch vormgevers en productontwerpers -Graphic Designers	Graphic designers and product designers - Graphic Designers
0222	Fotografen en interieurontwerpers	Photographers and interior designers
0311	Adviseurs marketing, public relations	Marketing, public relations consultants
0321	Vertegenwoordigers en inkopers	Representatives and buyers
0331	Winkeliers en teamleiders detailhandel	Retailers and team leaders retail
0332	Verkoopmedewerkers detailhandel	Retail sales staff
0333	Kassamedewerkers	Cashiers
0334	Callcentermedewerkers outbound	Call center staff outbound
0411	Accountants	Accountants
0412	Financieel specialisten en economen	Financial specialists and economists
0413	Bedrijfskundigen	Business Administration
0414	Beleidsadviseurs	Policy advisers
0415	Specialisten personeels- en loopbaanbegeleiders	Personnel and career counselors
0421	Boekhouders	Accountants
0422	Zakelijke dienstverleners	Business service providers
0423	Directiesecretaresses	Executive secretaries
0431	Administratief medewerkers	Administrative staff
0432	Secretaresses	Secretaries
0433	Receptionisten en telefonisten	Receptionists and telephone operators
0434	Boekhoudkundig medewerkers	Accounting staff

Table A.1: BRC code. job name Dutch and job name English

Brc	Job	Job english
0435	Transportplanners en logistiek medewerkers	Transport planners and logistics employees
0511	Algemeen directeuren	Managing Directors
0521	Managers zakelijke en administratie	Business and administration managers
0522	Managers verkoop en marketing	Sales and marketing managers
0531	Managers productie	Managers production
0532	Managers logistiek	Logistics managers
0533	Managers ICT	IT managers
0534	Managers zorginstellingen	Healthcare institution managers
0535	Managers onderwijs	Education managers
0536	Managers gespecialiseerde dienstverlening	Managers specialized services
0541	Managers horeca	Managers horeca
0542	Managers detail- en groothandel	Retail and wholesale managers
0543	Managers commercieel	Commercial managers
0611	Overheidsbestuurders	Government administrators
0612a	Overheidsambtenaren- office	Government officials -office
0612b	Overheidsambtenaren- operational	Government officials - operational
0621	Juristen	Lawyers
0631	Politie-inspecteurs	Police inspectors
0632	Politie en brandweer	Police and fire department
0633	Beveiligingspersoneel	Security personnel
0711	Biologen en natuurwetenschappers	Biologists and natural scientists
0712	Ingenieurs (geen elektrotechniek)	Engineers (no electrical engineering)
0713	Elektrotechnisch ingenieurs	Electrical engineers
0714	Architecten	Architects
0721	Technici bouwkunde en natuurkunde	Engineering and physics technicians
0722	Productieleiders industrie en bouw	Production leaders industry and construction
0723	Procesoperators	Process operators
0731	Bouwarbeiders ruwbouw	Construction workers structural work
0732	Timmerlieden	Carpenters
0733	Bouwarbeiders afbouw	Construction workers finishing
0734	Loodgieters en pijpfitters	Plumbers and pipe fitters
0735	Schilders en metaalspuiters	Painters and metal sprayers
0741	Metaalbewerkers en constructiewerkers	Metalworkers and construction workers
0742	Lassers en plaatwerkers	Welders and sheet metal workers
0743	Automonteurs	Car mechanics
0744	Machinemonteurs	Machine mechanics
0751	Slagers	Butchers
0752	Bakkers	Bakers
0753	Productcontroleurs	Product inspectors
0754	Meubelmakers, kleermakers en stomerij	Furniture makers, tailors and dry cleaners
0755	Medewerkers drukkerij en kunstnijverheid	Printing and arts and crafts employees
0761	Elektriciens en elektronicamonteurs	Electricians and electronics engineers
0771	Productiemachinebedieners	Production machine operators
0772	Assemblagemedewerkers	Assembly staff
0781	Hulpkrachten bouw en industrie	Auxiliary construction and industry
0811	Software- en applicatieontwikkeling	Software and application development
0812	Databank- en netwerkspecialisten	Database and network specialists
0821	Gebruikersondersteuning ICT	User support IT
0822	Radio- en televisietechnici	Radio and television technicians
0911	Land- en bosbouwers	Farmers and foresters
0912	Hoveniers, tuinders en kwekers	Gardeners and growers

Table A.1: BRC code. job name Dutch and job name English

Brc	Job	Job english
0913	Veetelers	Livestock farmers
0921	Hulpkrachten landbouw	Agricultural assistants
1011	Artsen	Artsen
1012	Gespecialiseerd verpleegkundigen	Specialist nurses
1013	Fysiotherapeuten	Physiotherapists
1021	Maatschappelijk werkers	Social workers
1022	Psychologen en sociologen	Psychologists and sociologists
1031	Laboranten	Laboratory workers
1032	Apothekersassistenten	Pharmacist's assistants
1033	Verpleegkundigen (mbo)	Nurses (MBO)
1034	Medisch praktijkassistenten	Medical practice assistants
1035	Medisch vakspecialisten	Medical specialists
1041	Sociaal werkers, groeps- en woonbegeleiders	Social workers, group and housing counselors
1051	Verzorgenden	Caretakers
1111	Reisbegeleiders	Tour guides
1112	Koks	Cooks
1113	Kelners en barpersoneel	Servers and bar staff
1114	Kappers en schoonheidsspecialisten	Hairdressers and beauticians
1115	Conciërges en teamleiders schoonmakers	Janitors and team cleaners
1116	Verleners van overige persoonlijke diensten	Providers of other personal services
1121	Schoonmakers	Cleaners
1122	Keukenhulpen	Kitchen aids
1211	Dekofficieren en piloten	Deck officers and pilots
1212	Chauffeurs auto's, taxi's en bestuurders	Drivers cars, taxis and drivers
1213	Buschauffeurs en trambestuurders	Bus drivers and tram drivers
1214a	Vrachtwagenchauffeurs -First-Line Supervisors of Transportation and Material-Moving Machine and Vehicle Operators	Truck Drivers -First-Line Supervisors of Transportation and Material-Moving Machine and Vehicle Operators
1214b	Vrachtwagenchauffeurs -Heavy and Tractor-Trailer Truck Drivers	Truck Drivers -Heavy and Tractor-Trailer Truck Drivers
1215	Bedieners mobiele machines	Mobile machine operators
1221	Laders, lossers en vakkenvullers	Loaders, unloaders and compartment fillers
1222	Vuilnisophalers en dagbladenbezorging	Garbage collectors and newspaper delivery
1311a	Beroepsgroep overig - managing	Occupational group other - managing
1311b	Beroepsgroep overig - operational	Occupational group other - operational

Table A.2: Translation: 2 digit BRC occupations groups

BRC-code	Dutch name	English name
01	Pedagogische beroepen	Pedagogical professions
02	Creatieve en taalkundige beroepen	Creative and linguistic professions
03	Commerciële beroepen	Commercial professions
04	Bedrijfseconomische en administratieve beroepen	Business and administrative professions
05	Managers	Managers
06	Openbaar bestuur, veiligheid en justitie	Public administration, security and justice
07	Technische beroepen	Technical professions
08	ICT beroepen	IT professions
09	Agrarische beroepen	Agricultural professions
10	Zorg en welzijn beroepen	Care and welfare professions
11	Dienstverlenende beroepen	Service professions
12	Transport en Logistiek beroepen	Transport and Logistics professions
13	Beroepklasse overig	Occupational class other

Appendix B

Abilities, skills and knowledge of Dutch workers

This appendix present the results of the aggregation analysis and the analysis of the competencies of Dutch workers. It will also discuss the content of the O*Net database and its limitations.

B.1 The O*Net database

In order to assign the competencies to Dutch workers, we use the O*Net database[1]. This O*Net database provides information about both the level and importance of a given skill, ability or knowledge of an occupation. This 'level' or 'importance' of a certain competence is derived from a survey that O*Net conducted among occupational analysts. Accordingly:

> Sixteen trained occupational analysts were responsible for rating the importance and level of the 35 skills for each of the O*NET occupations. A minimum of eight raters per occupation was required. This number was based on the number of raters needed to ensure the target level of interrater reliability. Fleisher, Tsacoumis, and Human Resources Research Organization (2018, p. 9)

The process is as followed: for each occupation, an occupational analyst is asked to rate the level and importance for a given skill, ability or knowledge in order to exercise that particular occupation[2]. In the end, responses are averaged across the respondents.

Figure B.1 shows an example of a question about the skill 'negotiation'. As can be seen in the figure, the occupation analysts are asked to scale the importance of a skill/ability/knowledge on a 5-point scale. Moreover, if an analyst states that the level of importance for a given skill to occupation is 1, the skill is considered not important for the performance of that job. This assumption –ratings below 2 are considered not important– is also used in our analysis[3].

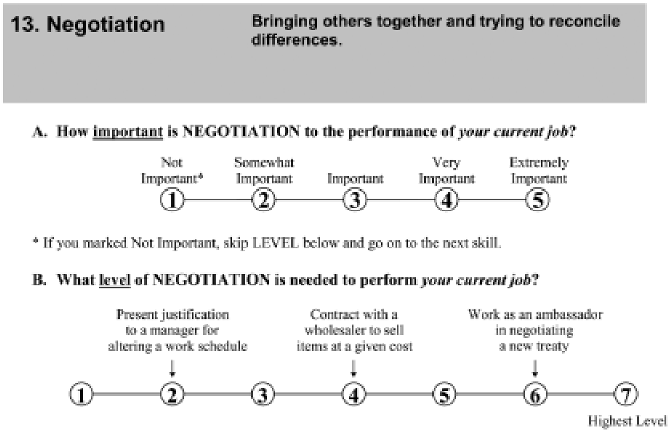
Since the O*Net survey is conducted in the US, and not in the Netherlands, we have decided to only use the scale of importance of competencies and not the level. Besides the fact that the level scale does is too specific for the US labour force, our analysis strategy departs from Frey and Osborne (2013) who also only rely on the O*Net 'importance' rating. Moreover, analyses of O*NET data suggest that there is substantial overlap between the importance

[1] O*NET is an application that was created for the general public to provide broad access to the O*NET database of occupational information. The site is maintained by the National Center for O*NET Development, on behalf of the U.S. Department of Labor, Employment and Training Administration (USDOL/ETA) (O*Net, 2019); see https://www.onetonline.org/

[2] Since there are at least 16 respondents per occupation, Handel (2016) estimates that "currently approximately 125,000 incumbent questionnaires were collected toward the end of the first round of original data collection . This would mean there are roughly 31,000 respondents per item because each of the questionnaires is completed by only one-quarter of the sample, and there is an average of 39 respondents per item within each of the 809 occupations in the first complete database" (p.160)

[3] For example, the cluster analysis of O*Net occupations according to their competencies is only done for the competencies that are considered to be important to the performance of that occupation

and level ratings and so this modelling choice does not lead to vastly different predictions in practice (Bakhshi et al., 2017; Handel, 2016).



13. Negotiation **Bringing others together and trying to reconcile differences.**

A. How <u>important</u> is NEGOTIATION to the performance of *your current job*?

Not Important* ① Somewhat Important ② Important ③ Very Important ④ Extremely Important ⑤

* If you marked Not Important, skip LEVEL below and go on to the next skill.

B. What <u>level</u> of NEGOTIATION is needed to perform *your current job*?

Present justification to a manager for altering a work schedule (2)

Contract with a wholesaler to sell items at a given cost (4)

Work as an ambassador in negotiating a new treaty (6)

① ② ③ ④ ⑤ ⑥ ⑦ Highest Level

Figure B.1: Example: O*Net Importance and Level Scales
source: Handel (2016)

B.1.1 Possible limitations of the O*Net database

While the O*Net database is impressive in its substantive scope, sampling and level of detail, the use of this database will also give us some limitations. These limitations are due to the limitations inherent to the O*Net content model itself or to the mapping process of O*Net data to Dutch labour statistics. The limitations of the O*Net content model are discussed in a comprehensive article about the strength and limitations of the O*Net model by Handel (2016). In his article, the author argues that the O*Net model encompasses some significant gaps and duplication in content. Moreover, he argues that the "underlying constructs, item wording, and response options are often vague or overly complex" (p.157). However, despite these concerns, validity analysis of the O*Net database has proven that it appears that the implications of O*NET's limitations for the validity of its measures are smaller than might be expected (Handel, 2016).

While the first set of limitations concerns the whole O*Net database in general, the second set of limitations concerns our research explicitly; the mapping process of O*Net data to Dutch occupations. In our analysis towards the competencies of the Dutch workforce, we map the O*Net ratings to the occupations defined by the CBS. The first concern that arises when mapping the O*Net data to Dutch data from the CBS, is the question of how well the O*Net occupations (indicated by an O*Net SOC-code) correspondent to the Dutch occupations (indicated by a BRC-code). In order to map CBS data (the BRC-codes) with the O*Net data (O*Net SOC-codes), we have used crosswalk tables provided by the CBS (from CBS to ISCO), the International Labour Office (ISCO to SOC), and O*Net (SOC to O*Net SOC). All these crosswalks are published and tested by these institutions [4]. For the verification of the crosswalks, we will rely on the verification of the crosswalks done by these institutions.

However, The assessment of the *validity*[5] of this mapping poses a more serious concern: do the dutch Occupations and its competencies correspondent to their US counterparts? As one may argue, the O*Net database is developed for and by an institution in the United States and can create validity problems if the competencies, job-related tasks and/or even job descriptions differ between the USA and the Netherlands. However, literature review towards the validity of the mapping of European job data to O*Net data, revealed no such concerns. Moreover, in their analysis towards different databases of skill sets, the CPB (Netherlands Bureau for Economic Policy Analysis) poses no such concerns about the O*Net database (Weel & Kok, 2013)[6].

Nevertheless, we need to be careful in mapping the O*Net data to Dutch professions. This means that we have to perform an analysis before we can simply aggregate the O*Net data to their BRC (Dutch) counterparts. For this, we performed a cluster analysis. This analysis found that the O*Net occupations (and their competencies) of the corresponding BRC-codes 1311, 0612, 1214 and 0221 could not be directly aggregated. According to the cluster-analysis, aggregation of the O*Net occupations of the remaining 108 BRC-codes is possible. While this hierarchical cluster-analysis provides us with a good starting point, we still have to be aware of the validity issues that may raise due to cross-country differences. Since there is no similar data-set about the competencies of Dutch occupations available, no quantitative validation is possible. We, therefore, recommend to assess the validity of our database –how well do the competencies of the dutch occupations (derived from O*Net data) agree with the actual competencies of Dutch occupations– by Dutch occupational experts.

[4]for more information about the crosswalks see: https://www.cbs.nl/nl-nl/onze-diensten/methoden/classificaties/onderwijs-en-beroepen/beroepenclassificatie–isco-en-sbc–, https://www.bls.gov/soc/soccrosswalks.htm and https://www.onetcenter.org/crosswalks.html

[5]Model verification is defined as 'ensuring that the computer program of the computerized model and its implementation are correct'. Model validation is defined as the 'substantiation that a model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model (Sargent, 2013).

[6]This does not mean that they have no concerns about the O*Net database. They point out that the O*Net database is susceptible to measuring errors since it relies on the response of respondents (Weel & Kok, 2013)

B.1.2 The classification of Skills, Abilities and Knowledge by O*Net

As mentioned, O*Net scales the importance of the abilities, skills, and knowledge needed to for a job on a 5-points scale. For this, O*Net has defined 31 different skills, 52 different abilities and 33 different types of knowledge. Furthermore, O*Net has aggregated its skills and abilities within headers (main-level) and categories (sub-levels). The different types of knowledge is only divided into categories. The following section will present the definition and classification of skills, abilities and knowledge as defined by O*Net.

Classification of abilities

According to O*Net, abilities are "enduring attributes of the individual that influence performance" (Fleisher et al., 2018, p. 19). O*Net has divided its 52 abilities under four headers, namely: 'Cognitive Abilities', 'Psychomotor Abilities', 'Physical Abilities', and 'Sensory Perceptual Abilities'. The abilities are further divided into different categories. Table B.1 shows the abilities with their description and classification as defined by O*Net.

Table B.1: Classification of Abilities

Ability	Description	Header	Category
Oral Comprehension	Listening to and understand information and ideas presented through spoken words and sentences	Cognitive Abilities	Verbal Abilities
Written Comprehension	Reading and understand information and ideas presented in writing	Cognitive Abilities	Verbal Abilities
Oral Expression	Communicating information and ideas in speaking so others will understand	Cognitive Abilities	Verbal Abilities
Written Expression	Communicating information and ideas in writing so others will understand	Cognitive Abilities	Verbal Abilities
Fluency of Ideas	Coming up with a number of ideas about a topic (the number of ideas is important, not their quality, correctness, or creativity)	Cognitive Abilities	Idea Generation and Reasoning Abilities
Originality	Coming up with unusual or clever ideas about a given topic or situation, or to develop creative ways to solve a problem	Cognitive Abilities	Idea Generation and Reasoning Abilities
Problem Sensitivity	Telling when something is wrong or is likely to go wrong. It does not involve solving the problem, only recognizing that there is a problem.	Cognitive Abilities	Idea Generation and Reasoning Abilities
Deductive Reasoning	Applying general rules to specific problems to produce answers that make sense	Cognitive Abilities	Idea Generation and Reasoning Abilities
Inductive Reasoning	Combining pieces of information to form general rules or conclusions (includes finding a relationship among seemingly unrelated events)	Cognitive Abilities	Idea Generation and Reasoning Abilities
Information Ordering	Arranging things or actions in a certain order or pattern according to a specific rule or set of rules	Cognitive Abilities	dea Generation and Reasoning Abilities
Category Flexibility	Generating or using different sets of rules for combining or grouping things in different ways	Cognitive Abilities	Idea Generation and Reasoning Abilities
Mathematical Reasoning	Choosing the right mathematical methods or formulas to solve a problem.	Cognitive Abilities	Quantitative Abilities
Number Facility	Adding, subtracting, multiplying, or dividing quickly and correctly.	Cognitive Abilities	Quantitative Abilities
Memorization	Remembering information such as words, numbers, pictures, and procedures	Cognitive Abilities	Perceptual Abilities

Table B.1: Classification of Abilities

Ability	Description	Header	Category
Speed of Closure	Quickly making sense of, combining, and organizing information into meaningful patterns	Cognitive Abilities	Perceptual Abilities
Flexibility of Closure	identifying or detecting a known pattern (a figure, object, word, or sound) that is hidden in other distracting material	Cognitive Abilities	Perceptual Abilities
Perceptual Speed	Quickly and accurately comparing similarities and differences among sets of letters, numbers, objects, pictures, or patterns. The things to be compared may be presented at the same time or one after the other. This ability also includes comparing a presented object with a remembered object	Cognitive Abilities	Perceptual Abilities
Spatial Orientation	Knowing your location in relation to the environment or to know where other objects are in relation to you	Cognitive Abilities	Spatial Abilities
Visualization	Imagining how something will look after it is moved		
around or when its parts are moved or rearranged	Cognitive Abilities	Spatial Abilities	
Selective Attention	Concentrating on a task over a period of time without being distracted	Cognitive Abilities	Attentiveness
Time sharing	Shifting back and forth between two or more activities or sources of information (such as speech, sounds, touch, or other sources).	Cognitive Abilities	Attentiveness
Arm-hand Steadiness	Keeping your hand and arm steady while moving your arm or while holding your arm and hand in one position.	Psychomotor Abilities	Fine Manipulative Abilities
Manual Dexterity	Quickly moving your hand, your hand together with your arm, or your two hands to grasp, manipulate or assemble objects	Psychomotor Abilities	Fine Manipulative Abilities
Finger Dexterity	Making precisely coordinated movements of the fingers of one or both hands to grasp, manipulate or assemble very small objects.	Psychomotor Abilities	Fine Manipulative Abilities
Control Precision	Quickly and repeatedly adjusting the controls of a machine or a vehicle to exact positions	Psychomotor Abilities	Control-Movement Abilities
Multi-limb Coordination	Coordinating two or more limbs (for example two arms, two legs, or one leg and one arm) while sitting, standing, or lying down. It does not involve performing the activities while the whole body is in motion	Psychomotor Abilities	Control-Movement Abilities
Response Orientation	Choosing quickly between two or more movements in response to two or more different signals (lights sounds, pictures). It includes the speed with which the correct response is started with the hand, foot, or other body part	Psychomotor Abilities	Control-Movement Abilities

Table B.1: Classification of Abilities

Ability	Description	Header	Category
Rate Control	Timing your movements or the movement of a piece of equipment in anticipation of changes in the speed and/or direction of a moving object or scene	Psychomotor Abilities	Control-Movement Abilities
Reaction Time	Quickly responding (with the hand, finger, or foot) to a signal (sound, light, picture) when it appears.	Psychomotor Abilities	Reaction Time and Speed Abilities
Wrist-finger Speed	Making fast, simple, repeated movements of the fingers,	Psychomotor Abilities	Reaction Time and Speed Abilities
Speed of Limb Movement	Quickly moving the arms and legs	Psychomotor Abilities	Reaction Time and Speed Abilities
Static Strength	Exerting maximum muscle force to lift,push, pull, or carry objects	Physical Abilities	Physical Strength Abilities
Explosive Strength	Using short bursts of muscle force to propel oneself (as in jumping or sprinting) or to throw an object.	Physical Abilities	Physical Strength Abilities
Dynamic Strength	Exerting muscle force repeatedly or continuously over time. This involves muscular endurance and resistance to muscle fatigue	Physical Abilities	Physical Strength Abilities
Trunk Strength	Using your abdominal and lower back muscles to support part of the body repeatedly or continuously over time without "giving out" or fatiguing	Physical Abilities	Physical Strength Abilities
Stamina	Exerting yourself physically over long periods of time without getting winded or out of breath	Physical Abilities	Endurance
Extent Flexibility	Bending, stretching, twisting, or reaching with your body, arms, and/or legs.	Physical Abilities	Flexibility Balance, and Coordination
Dynamic Flexibility	Quickly and repeatedly bending, stretching, twisting or reaching out with your body, arms, and/or legs.	Physical Abilities	Flexibility Balance, and Coordination
Gross Body Coordination	Coordinating the movement of your arms, legs and torso together when the whole body is in motion.	Physical Abilities	Flexibility Balance, and Coordination
Gross Body Equilibrium	Keeping or regaining your body balance or stay upright when in an unstable position.	Physical Abilities	Flexibility Balance, and Coordination
Near Vision	Seeing details at close range (within a few feet of the observer).	Sensory Abilities	Visual Abilities
Far Vision	Seeing details at a distance	Sensory Abilities	Visual Abilities
Visual Color Discrimination	Matching or detect differences between colors including shades of color and brightness	Sensory Abilities	Visual Abilities
Night Vision	Seeing under low-light conditions.	Sensory Abilities	Visual Abilities
Peripheral Vision	Seeing objects or movement of objects to one's side when the eyes are looking ahead	Sensory Abilities	Visual Abilities
Depth Perception	Judging which of several objects is closer or farther away from you, or to judge the distance between you and an object.	Sensory Abilities	Visual Abilities
Glare Sensitivity	Seeing objects in the presence of a glare or bright lighting.	Sensory Abilities	Visual Abilities

Table B.1: Classification of Abilities

Ability	Description	Header	Category
Hearing Sensitivity	Detecting or telling the differences between sounds that vary in pitch and loudness	Sensory Abilities	Auditory and Speech Abilities
Auditory Attention	Focusing on a single source of sound in the presence of other distracting sounds Sensory Abilities	Auditory and Speech Abilities	
Sound Localization	Telling the direction from which a sound originated.	Sensory Abilities	Auditory and Speech Abilities
Speech Recognition	Identifying and understand the speech of another person.	Sensory Abilities	Auditory and Speech Abilities
Speech Clarity	Speaking clearly so others can understand you.	Sensory Abilities Auditory and Speech Abilities	

Source: Fleisher et al. (2018)

Classification of skills

According to O*Net, skills are "proficiencies that are developed through training or expertise" (Fleisher et al., 2018, p. 18). In the O*Net database, there are 35 skills that are divided into 'Basic Skills' and 'Cross-functional Skills'. Accordingly, basic skills facilitate the acquisition of new knowledge. Examples of basic skills are reading and writing. Cross-functional skills are skills that extend across several domains of activities. Examples of cross-functional skills are Negotiation and Complex-problem solving. Under these two headers, the skills within the O*Net database are further divided into smaller categories (Fleisher et al., 2018). Table B.2 presents the description and classification of the 35 skills as defined by O*Net.

Table B.2: Classification of skills

Skill	Description	Header	Category
Reading Comprehension	Understanding written sentences and paragraphs in work related documents	Basic Skill	Content
Active Listening	Giving full attention to what other people are saying, taking time to understand the points being made, asking questions as appropriate, and not interrupting at inappropriate times	Basic Skill	Content
Writing	Communicating effectively in writing as appropriate for the needs of the audience	Basic Skill	Content
Speaking	Talking to others to convey information effectively	Basic Skill	Content
Mathematics	Using mathematics to solve problems	Basic Skill	Content
Science	Using scientific rules and methods to solve problems	Basic Skill	Content
Critical Thinking	Using logic and reasoning to identify the strengths and weaknesses of lternative solutions, conclusions or approaches to problems	Basic Skill	Process
Active Learning	Understanding the implications of new information for both current and future problem-solving and decision-making.	Basic Skill	Process
Learning Strategies	Selecting and using training/instructional methods and procedures appropriate for the situation when learning or teaching new things	Basic Skill	Process

Table B.2: Classification of skills

Skill	Description	Header	Category
Monitoring	Monitoring/Assessing performance of yourself,other individuals, or organizations to make improvements or take corrective action.	Basic Skill	Process
Social Perceptiveness	Being aware of others' reactions and understanding why they react as they do	Cross-Functional Skills	Social Skills
Coordination	Adjusting actions in relation to others' actions.	Cross-Functional Skills	Social Skills
Persuasion	Persuading others to change their minds or behavior.	Cross-Functional Skills	Social Skills
Negotiation	Bringing others together and trying to reconcile differences	Cross-Functional Skills	Social Skills
Instructing	Teaching others how to do something.	Cross-Functional Skills	Social Skills
Service Orientation	Actively looking for ways to help people	Cross-Functional Skills	Social Skills
Complex Problem-Solving	Identifying complex problems and reviewing related information o develop and evaluate options and implement solutions	Cross-Functional Skills	Technical Skills
Operations Analysis	Analyzing needs and product requirements to create a design.	Cross-Functional Skills	Technical Skills
Technology Design	Generating or adapting equipment and technology to serve user needs	Cross-Functional Skills	Technical Skills
Equipment Selection	Determining the kind of tools and equipment needed to do a job	Cross-Functional Skills	Technical Skills
Installation	Installing equipment, machines, wiring, or programs to meet specifications	Cross-Functional Skills	Technical Skills
Programming	Writing computer programs for various purposes.	Cross-Functional Skills	Technical Skills
Operation Monitoring	Watching gauges, dials, or other indicators to make sure a machine is working properly	Cross-Functional Skills	Technical Skills
Operation and Control	Controlling operations of equipment or systems	Cross-Functional Skills	Technical Skills
Equipment Maintenance	Performing routine maintenance on equipment and determining when and what kind of maintenance is needed	Cross-Functional Skills	Technical Skills
Troubleshooting	Determining causes of operating errors and deciding what to do about it	Cross-Functional Skills	Technical Skills
Repairing	Repairing machines or systems using the needed tools	Cross-Functional Skills	Technical Skills
Quality Control Analysis	Conducting tests and inspections of products, services, or processes to evaluate quality or performance	Cross-Functional Skills	Technical Skills
Judgment and Decision-Making	Considering the relative costs and benefits of potential actions to choose the most appropriate one	Cross-Functional Skills	Systems Skills

Table B.2: Classification of skills

Skill	Description	Header	Category
Systems Analysis	Determining how a system should work and how changes in conditions, operations, and the environment will affect outcomes	Cross-Functional Skills	Systems Skills
Systems Evaluation	Identifying measures or indicators of system performance and the actions needed to improve or correct performance,relative to the goals of the system	Cross-Functional Skills	Systems Skills
Time Management	Managing one's own time and the time of others	Cross-Functional Skills	Resource-Management Skills
Management of Financial Resources	Determining how money will be spent to get the work done, and accounting for these expenditures	Cross-Functional Skills	Resource-Management Skills
Management of Material Resources	Obtaining and seeing to the appropriate use of equipment, facilities, and materials needed to do certain work	Cross-Functional Skills	Resource-Management Skills
Management of Personnel Resources	Motivating, developing, and directing people as they work, identifying the best people for the job	Cross-Functional Skills	Resource-Management Skills

Source: Fleisher et al. (2018)

Classification of knowledge

According to O*Net, knowledge are organized sets of principles and facts applying in general domains (Fleisher et al., 2018). In their database, O*Net has specified 33 types of knowledge. They categorized these types of knowledge into 10 categories: 'Business and Management'; 'Manufacturing and Management'; 'Engineering and production'; 'Engineering and Technology'; 'Mathematics and Science'; 'Health Services'; 'Arts and Humanities'; 'Law and Public Safety'; 'Communications'; 'Transportation'; and 'Education and Training'. Table B.3 shows the types of knowledge, their description and corresponding category as defined by O*Net.

Table B.3: Knowledge

Knowledge	Description	Category
Administration and Management	Knowledge of business and management principles involved in strategic planning, resource allocation, human resources modeling, leadership technique, production methods, and coordination of people and resources	Business & Management
Clerical	Knowledge of administrative and clerical procedures and systems such as word processing, managing files and records, stenography and transcription, designing forms, and other office procedures and terminology	Business & Management
Economics and Accounting	Knowledge of economic and accounting principles and practices, the financial markets, banking the analysis and reporting of financial data	Business &Management
Sales Marketing	Knowledge of principles and methods for showing, promoting, and selling products or services. This includes marketing strategy and tactics, product demonstration, sales techniques, sales control systems	Business & Management

Table B.3: Knowledge

Knowledge	Description	Category
Customer Personal Service	Knowledge of principles and processes for providing customer and personal services. This includes customer needs assessment, meeting quality standards for services, evaluation of customer satisfaction.	Business &Management
Personnel Human Resources	Knowledge of principles and procedures for personnel recruitment, selection, training, compensation and benefits, labor relations negotiation, and personnel information systems	Business &Management
Production Processing	Knowledge of raw materials, production processes, quality control, costs, and other techniques for maximizing the effective manufacture distribution of goods	Manufacturing production
Food Production	Knowledge of techniques and equipment for planting, growing, and harvesting food products (both plant and animal) for consumption, including storage/handling techniques	Manufacturing Production
Computers Electronics	Knowledge of circuit boards, processors, chips electronic equipment, and computer hardware software, including applications and programming.	Engineering Technology
Engineering Technology	Knowledge of the practical application of engineering science and technology. This includes applying principles techniques, procedures, and equipment to the design production of various goods and services.	Engineering Technology
Design	Knowledge of design techniques, tools, and principles involved in production of precision technical plans blueprints, drawings, and models.	Engineering Technology
Building Construction	Knowledge of materials, methods, and the tools involved in the construction or repair of houses buildings, or other structures such as highways and roads	Engineering Technology
Mechanical	Knowledge of machines and tools, including their designs, uses, repair, and maintenance	Engineering Technology
Mathematics	Knowledge of arithmetic, algebra, geometry calculus, statistics, and their applications.	Mathematics Science
Physics	Knowledge and prediction of physical principles laws, their interrelationships, and applications to understanding fluid, material, and atmospheric dynamics, and mechanical, electrical, atomic sub- atomic structures and processes	Mathematics Science
Chemistry	Knowledge of the chemical composition, structure, properties of substances and of the chemical processes transformations that they undergo. This includes uses of chemicals and their interactions, danger signs production techniques, and disposal methods	Mathematics Science
Biology	Knowledge of plant and animal organisms, their tissues cells, functions, interdependencies, and interactions with each other and the environment.	Mathematics Science

Table B.3: Knowledge

Knowledge	Description	Category
Psychology	Knowledge of human behavior and performance; individual differences in ability, personality, and interests ; learning and motivation; psychological research methods ; and the assessment and treatment of behavioral affective disorders.	Mathematics Sience
Sociology and Anthropology	Knowledge of group behavior and dynamics, societal trends influences, human migrations, ethnicity, cultures their history and origins.	Mathematics Science
Geography	Knowledge of principles and methods for describing the features of land, sea, and air masses, including their physical characteristics, locations, interrelationships, distribution of plant, animal, and human life.	Mathematics Science
Medicine Dentistry	Knowledge of the information and techniques needed to diagnose and treat human injuries, diseases, and deformities This includes symptoms, treatment alternatives, drug properties and interactions, and preventive health-care measures.	Health Services
Therapy Counseling	Knowledge of principles, methods, and procedures for diagnosis, treatment, and rehabilitation of physical mental dysfunctions, and for career counseling guidance	Health Services
National Language	Knowledge of the structure and content of the national language including the meaning and spelling of words, rules of composition, grammar.	Arts Humanities
Foreign Language	Knowledge of the structure and content of a foreign language including the meaning and spelling of words, rules of composition and grammar, and pronunciation.	Arts Humanities
Fine Arts	Knowledge of the theory and techniques required to compose produce, and perform works of music, dance, visual arts, drama, sculpture.	Arts Humanities
History Archeology	Knowledge of historical events and their causes, indicators, effects on civilizations and cultures.	Arts Humanities
Philosophy Theology	Knowledge of different philosophical systems religions. This includes their basic principles, values ethics, ways of thinking, customs, practices, and their impact on human culture	Arts Humanities
Public Safety Security	Knowledge of relevant equipment, policies, procedures and strategies to promote effective local, state or national security operations for the protection of people data, property, and institutions	Law Public Safety
Law and Government	Knowledge of laws, legal codes, court procedures precedents, government regulations, executive orders agency rules, and the democratic political process	Law Public Safety
Telecommunications	Knowledge of transmission, broadcasting, switching control, and operation of telecommunications systems	Communications
Communications Media	Knowledge of media production, communication, dissemination techniques and methods. This includes alternative ways to inform and entertain via written, oral, visual media.	Communications

Table B.3: Knowledge

Knowledge	Description	Category
Transportation	Knowledge of principles and methods for moving people or goods by air, rail, sea, or road, including the relativen costs and benefits	Transportation
Education Training	Knowledge of principles and methods for curriculum training design, teaching and instruction for individuals groups, and the measurement of training effects.	Education Training

Source: Fleisher et al. (2018)

B.2 Method aggregation analysis

As shown by the example in chapter 4 table 4.1, the O*Net aggregation level differs from the level used by the CBS. This means that we have to aggregate the O*Net data about skills, abilities and knowledge. The easiest way to do this would be to take the mean of the competencies from the O*Net occupations which have the same BRC-codes. However, this may lead to losing useful information. To stay with the example provided by table 4.1, the importance of the ability 'Inductive Reasoning' could differ greatly between sociologists and clinical psychologists. By averaging this level of 'Inductive Reasoning' across this whole BRC-code, we ignore this difference. Since the levels of this abilities, skills and knowledge are of great importance for this research, we have to be careful when aggregating the data-set.

B.2.1 Hierarchical Cluster Analysis

To aggregate the data, we begin with assessing which data can be aggregated and which data points cannot. For this, we use two different methods. The first method involves a cluster analysis to assess if we can identify distinct clusters of skills, knowledge and abilities within a BRC occupation. This is only done for BRC data points that have more than 2 corresponding O*Net occupations.

The cluster method used to aggregate the BRC data points is hierarchical clustering. According to Revelle (1979), hierarchical clustering is proven to be an effective method for forming scales from sets of items. Moreover, "comparisons with other procedures show that hierarchical clustering algorithms can be more useful for scale construction using large item pools than are conventional factor analytic techniques" (p.1).

In hierarchical clustering, the objects are organized and represented into a dendrogram whose branches are the desired clusters. This hierarchical process of cluster detection is referred to as tree cutting (Zhang, Langfelder, & Horvath, 2007). After the dendogram is made, the optimal number of clusters can be determined.

In this analysis, we first look for BRC-codes, where the optimal number of clusters is larger than 1. When this is the case, it means that more than one distinct cluster of O*net occupations within a BRC-code can be identified and that averaging over these occupations based on their competencies within the BRC-code, ignores the different characteristics of these clusters. On the contrary, when the optimal number of clusters of O*Net occupations within a BRC code is equal to one, it means that we can assume that there are no significant differences between the competencies of the underlying O*Net clusters. That being the case, we can aggregate all the underlying O*Net occupations to its corresponding BRC-code.

Moreover, the cluster analysis is performed on the competencies that are considered to be important to execute the O*net occupations. According to O*Net, skills, abilities or knowledge are only considered to be important for the exercise of an occupation if it scores higher than two on a 5-point scale(Reeder & Tsacoumis, 2017). By only clustering on the important factors, we avoid false clustering due to the overestimation of factors. For example, suppose that the cluster analysis of the skills of a BRC job indicates that three clusters can be found when considering all the skills. However, these clusters are also based on the perhaps large difference between unimportant skills (ranging from 0 to 2). When these skills are eliminated, the cluster analysis may results in different clusters based on the smaller –but more important– differences between the important skills.

Before discussing the method for determining the optimal number of clusters, it is important to choose the basic procedure of the hierarchical cluster method used for this analysis. The method of hierarchical clustering can be divided into two main types: agglomerative [7] and divisive clustering. [8] Since the agglomerative method is good at identifying small clusters, we use the agglomerative method for this analysis. However, more important is the determination of how this analysis will measure the dissimilarity between two clusters of observations. For this, several methods can be used:

- **Maximum or complete linkage clustering**: It computes all pairwise dissimilarities between the elements in cluster 1 and the elements in cluster 2, and considers the largest value (i.e., maximum value) of these dissimilarities as the distance between the two clusters. It tends to produce more compact clusters.

- **Minimum or single linkage clustering**: It computes all pairwise dissimilarities between the elements in cluster 1 and the elements in cluster 2, and considers the smallest of these dissimilarities as a linkage criterion. It tends to produce long, “loose” clusters.

- **Mean or average linkage clustering**: It computes all pairwise dissimilarities between the elements in cluster 1 and the elements in cluster 2, and considers the average of these dissimilarities as the distance between the two clusters.
- **Centroid linkage clustering**: It computes the dissimilarity between the centroid for cluster 1 (a mean vector of length p variables) and the centroid for cluster 2.

- **Ward’s minimum variance method**: It minimizes the total within-cluster variance. At each step, the pair of clusters with minimum between-cluster distance are merged.

(Holland, 2017)

To choose between these agglomeration methods, we use the agglomerative coefficient. This coefficient “measures the dissimilarity of an object to the first cluster it joins, divided by the dissimilarity of the final merger in the cluster analysis, averaged across all samples. Low values reflect tight clustering of objects; larger values indicate less well-formed clusters” (Holland, 2017). In this analysis, for each BRC-code, all agglomeration methods are executed. Nevertheless, only the results of the cluster analysis with the highest agglomeration method will be used. This means that for different BRC-codes, different agglomeration methods are used. In contrast, the method for calculating the distance matrix will be the same across all BRC-codes, namely the Euclidean distance.

When the dendograms of the clusters are made, the next step is to define the optimal number of clusters. For determining the optimal numbers of clusters, several methods can be used. The elbow, silhouette and gap statistic methods are the most commonly used. The elbow and silhouette method can be seen as a direct method that consists of optimizing certain

[7] It’s, also known as AGNES (Agglomerative Nesting). It works in a bottom-up manner. That is, each object is initially considered as a single-element cluster (leaf). At each step of the algorithm, the two clusters that are the most similar are combined into a new bigger cluster (nodes). This procedure is iterated until all points are a member of just one single big cluster (root) (see figure below). The result is a tree which can be plotted as a dendrogram (Boehmke, 2017).

[8] It’s also known as DIANA (Divise Analysis), and it works in a top-down manner. The algorithm is an inverse order of AGNES. It begins with the root, in which all objects are included in a single cluster. At each step of the iteration, the most heterogeneous cluster is divided into two. The process is iterated until all objects are in their own cluster (Boehmke, 2017).

criteria. However, since we have created the data-set from scratch and we have no benchmark information about how to assess the optimal number of clusters, we use a statistical testing method that compares evidence against null hypothesis: the gap statistic method. This method is developed by Tibshirani et al. (2001) and can be applied to the hierarchical cluster method. The algorithm works as follows:

1. Cluster the observed data, varying the number of clusters from $k = 1, .., k_{max}$, , and compute the corresponding total within intra-cluster variation W_k.
2. Generate B reference data sets with a random uniform distribution. Cluster each of these reference data sets with varying number of clusters $k = 1, .., k_{max}$, , and compute the corresponding total within intra-cluster variation W_{kb}.
3. Compute the estimated gap statistic as the deviation of the observed Wk value from its expected value W_{kb} under the null hypothesis: $Gap(k) = \frac{1}{b}\sum_{b}^{b=1} \log(W_{kb}) - log(W_k)$. Compute also the standard deviation of the statistics.
4. Choose the number of clusters as the smallest value of k such that the gap statistic is within one standard deviation of the gap at $k + 1 : Gap(k) \geq Gap(k + 1) - s_k + 1$.

(Kassambara, 2018)

When the optimal number of clusters is determined, the actual aggregation of the competencies of O*Net occupations can begin. When the optimal number of clusters is one, the competencies of the O*Net occupations within a BRC-code are averaged. When a BRC-code has more than one optimal clusters, the competencies are averaged according to these clusters.

B.2.2 Aggregation of two O*Net occupations per one BRC

When the cluster analysis is finished, we will focus on the BRC-codes that have precisely two O*Net professions. Since there are only 2 O*Net occupations, cluster analysis cannot be performed. This means that we have to find another way for which we can identify if there are relevant differences within a BRC-code. To do this, no standardized method is available. Statistical methods that check for equality of variances between groups across multiple dependent variables –such as Manova– are not sufficient for this since the characteristics of the O*Net data set leads to the violation of several assumptions such as the absence of multivariate outliers and lack of multicollinearity. This means that we have to define our own measures to assess the differences between O*Net occupations. Fortunately, we can use data from the cluster-analysis to make the assessment. The decision to split a BRC code with two underlying O*Net occupations will be based on the outcomes of the previous cluster. For each BRC-code that is split into clusters in the previous cluster analysis, the difference between the clusters will be calculated and will be used as a benchmark:

$$Dif_{brc} = \sum^{Ability} (ClusterA_{ability} - ClusterB_{ability})^2 \tag{B.1}$$

After this, the benchmark value b can be identified for which $b = min(A)$ and where $A = \{Diff_{brc}\}_{brc=1}^{N}$.

This benchmark will be used to determine if one BRC-code with two underlying O*Net occupations has to be split up in two different groups. This is done by calculating the difference between the two O*net occupations within a BRC-code using equation B.1. This difference will then be compared with the benchmark value. When the difference is greater than the benchmark value, we will divide the BRC-code into its two O*Net occupations. As can be seen, a great advantage of this method is that it relies on earlier defined criteria and thereby guarantees consistency in the aggregation progress.

When the aggregation analysis is completed, we combine the aggregated data-set with the level for each competence per BRC-code with actual data from the CBS[9][10] .

B.3 Results aggregation analysis

This section will show the results of the aggregation analysis. It will present these results for the analyses of abilities, skills and type of knowledge separately.

B.4 Abilities

This section presents the results of the aggregation analysis[11] of the abilities from the 1012 O*Net occupations to their corresponding 113 BRC occupations.

B.4.1 Results cluster analysis abilities

Table B.4 presents the results of the cluster analysis of the abilities. This table shows that only BRC-codes 0612 ('Overheidsambtenaren' or 'Civil Servants') should be divided into two clusters. Figure B.2 shows the clustering of both 0612 and 1311.

Looking at the type of jobs within the clusters of 0612 (see Figure B.2) and the differences between the abilities (see Table B.5, we see that cluster one is composed of more office related O*Net occupations whereas the second cluster is composed of more 'on-the-street' operational O*Net occupations. Differences between the clusters can mainly be found in the physical and sensory abilities for which the operational cluster group score higher.

Figure B.2 shows which clusters are found within the code 1311. Looking at the O*Net occupations of the clusters, we see that the first cluster correspondent with more managing type of occupations whereas the second cluster considers the more operational type of occupations. This can be further explained by looking at the actual differences in abilities between the two clusters (see table B.6. Here one can see that the managing cluster scores higher at the most cognitive abilities than the cluster with the operational jobs. The cluster with the operational O*Net occupations scores higher at the physical and sensory abilities.

Table B.4: Results cluster analysis abilities

BRC	Ratio	Method	ac	nc
0611 Overheidsbestuurders	14	ward	0.60	1
1311 Beroepsgroep overig	8	ward	0.79	2
0511 Algemeen directeuren	3	ward	0.22	1
0521 Managers zakelijke en administra.	19	ward	0.70	1
0522 Managers verkoop en marketing	9	ward	0.56	1
0531 Managers productie	19	ward	0.70	1
0532 Managers logistiek	3	ward	0.23	1
0533 Managers ICT	1	NAP	NAP	1
0534 Managers zorginstellingen	4	ward	0.071	1
0535 Managers onderwijs	4	ward	0.25	1
0536 Managers gespecialiseerde dienst.	12	ward	0.61	1
0541 Managers horeca	2	n=2	n=2	n=2
0542 Managers detail- en groothandel	1	NAP	NAP	1

[9]The data-set with the level of importance for each competence per BRC-code will thus be extended with the actual number of workers in the Netherlands who have that occupation

[10]This combined data-set will be further referred to as 'the competence data-set'

[11]r-scripts used for this analysis: *abilityclustering.R*, *clusterfunction.R* and *finalframe.R*

Table B.4: Results cluster analysis abilities

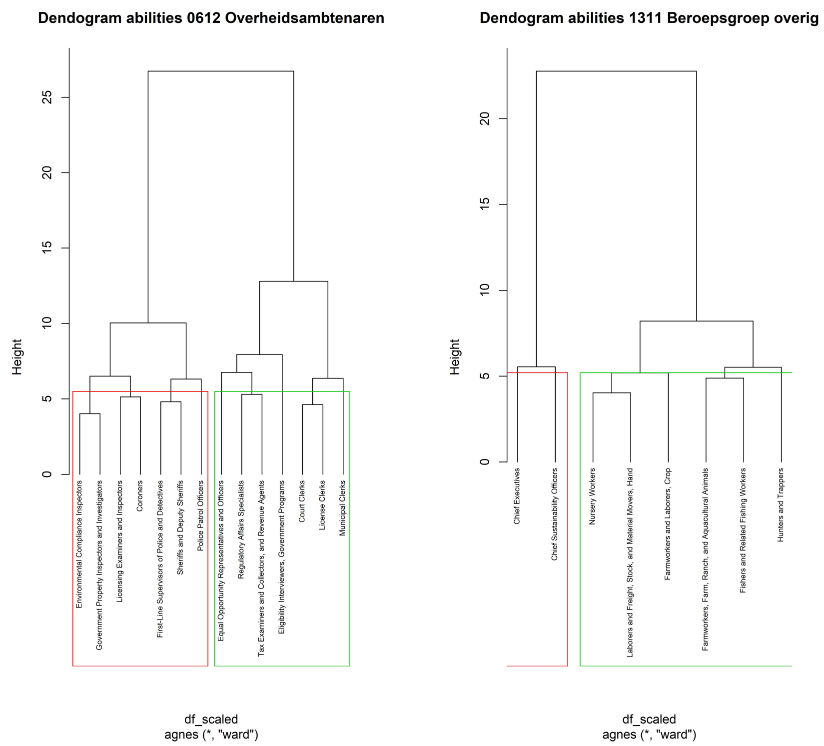
BRC	Ratio	Method	ac	nc
0543 Managers commerciÃ«le en persoonl.	10	ward	0.57	1
0711 Biologen en natuurweten-schappers	27	ward	0.76	1
0712 Ingenieurs (geen elektrotech-niek)	23	ward	0.65	1
0713 Elektrotechnisch ingenieurs	3	complete	0.17	1
0714 Architecten	6	ward	0.52	1
0221 Grafisch vormgevers en pro-ducton...	2	n=2	n=2	n=2
1011 Artsen	21	ward	0.69	1
1012 Gespecialiseerd ver-pleegkundigen	6	ward	0.48	1
1013 Fysiotherapeuten	11	ward	0.49	1
0111 Docenten hoger onderwijs en hoog...	35	ward	0.67	1
0112 Docenten beroepsgerichte vakken ...	2	n=2	n=2	n=2
0113 Docenten algemene vakken secunda...	1	NAP	NAP	1
0114 Leerkrachten basisonderwijs	7	ward	0.44	1
0115 Onderwijskundigen en overige doc...	5	complete	0.38	1
0411 Accountants	4	ward	0.38	1
0412 Financieel specialisten en economen	6	complete	0.28	1
0413 Bedrijfskundigen en organ-isatiea.	4	ward	0.25	1
0414 Beleidsadviseurs	6	ward	0.46	1
0415 Specialisten personeels- en loop.	2	n=2	n=2	n=2
0311 Adviseurs marketing	5	ward	0.32	1
0811 Software- en applicatieon-twikkel.	18	ward	0.54	1
0812 Databank- en netwerkspecial-isten	15	ward	0.54	1
0621 Juristen	4	ward	0.29	1
0211 Bibliothecarissen en conser-vatoren	4	ward	0.26	1
1022 Psychologen en sociologen	13	ward	0.63	1
1021 Maatschappelijk werkers	4	complete	0.27	1
0212 Auteurs en taalkundigen	4	ward	0.21	1
0213 Journalisten	1	NAP	NAP	1
0214 Beeldend kunstenaars	2	n=2	n=2	n=2
0215 Uitvoerend kunstenaars	11	ward	0.55	1
0721 Technici bouwkunde en natuur	45	ward	0.80	1
0722 Productieleiders industrie en bouw	2	n=2	n=2	n=2
0723 Procesoperators	13	ward	0.64	1
1211 Dekofficieren en piloten	10	ward	0.60	1
1031 Laboranten	9	ward	0.46	1
1032 Apothekersassistenten	1	NAP	NAP	1
1035 Medisch vakspecialisten	20	complete	0.60	1
1033 Verpleegkundigen (mbo)	2	n=2	n=2	n=2
1034 Medisch praktijkassistenten	6	complete	0.35	1
0421 Boekhouders	14	ward	0.50	1

Table B.4: Results cluster analysis abilities

BRC	Ratio	Method	ac	nc
0321 Vertegenwoordigers en inkopers	9	ward	0.50	1
0422 Zakelijke dienstverleners	18	ward	0.68	1
0423 Directiesecretaresses	5	complete	0.28	1
0612 Overheidsambtenaren	14	ward	0.80	2
0631 Politie-inspecteurs	6	ward	0.49	1
1041 Sociaal werkers	6	ward	0.48	1
0121 Beroepsgroep sportinstructeurs	5	ward	0.37	1
0222 Fotografen en interieurontwerpers	4	complete	0.35	1
1112 Koks	6	ward	0.41	1
0821 Gebruikersondersteuning ICT	2	n=2	n=2	n=2
0822 Radio- en televisietechnici	4	ward	0.30	1
0431 Administratief medewerkers	17	ward	0.70	1
0432 Secretaresses	3	ward	0.20	1
0433 Receptionisten en telefonisten	8	ward	0.39	1
0434 Boekhoudkundig medewerkers	6	ward	0.44	1
0435 Transportplanners en logistiek m.	2	n=2	n=2	n=2
1111 Reisbegeleiders	5	ward	0.28	1
1113 Kelners en barpersoneel	1	NAP	NAP	1
1114 Kappers en schoonheidsspecialisten	6	ward	0.48	1
1115 ConciÃ«rges en teamleiders schoon.	1	NAP	NAP	1
1116 Verleners van overige persoonlij.	7	ward	0.36	1
0334 Callcentermedewerkers outbound e...	7	ward	0.55	1
0331 Winkeliers en teamleiders detail...	2	n=2	n=2	n=2
0332 Verkoopmedewerkers detailhandel	2	n=2	n=2	n=2
0333 Kassamedewerkers	1	NAP	NAP	1
0131 Leidsters kinderopvang en onderw...	2	n=2	n=2	n=2
1051 Verzorgenden	6	ward	0.46	1
0632 Politie en brandweer	8	ward	0.57	1
0633 Beveiligingspersoneel	10	ward	0.56	1
0911 Land- en bosbouwers	10	ward	0.56	1
0912 Hoveniers	2	n=2	n=2	n=2
0913 Veetelers	7	ward	0.55	1
0731 Bouwarbeiders ruwbouw	13	complete	0.63	1
0732 Timmerlieden	2	n=2	n=2	n=2
0733 Bouwarbeiders afbouw	11	ward	0.55	1
0734 Loodgieters en pijpfitters	3	ward	0.12	1
0735 Schilders en metaalspuiters	3	complete	0.08	1
0741 Metaalbewerkers en constructiewe.	22	ward	0.61	1
0742 Lassers en plaatwerkers	5	ward	0.47	1
0743 Automonteurs	8	complete	0.48	1
0744 Machinemonteurs	7	ward	0.42	1
0755 Medewerkers drukkerij en kunstni.	19	ward	0.72	1
0761 Elektriciens en elektronicamonteurs	23	ward	0.62	1

Table B.4: Results cluster analysis abilities

BRC	Ratio	Method	ac	nc
0751 Slagers	3	complete	0.19	1
0752 Bakkers	1	NAP	NAP	1
0753 Productcontroleurs	11	complete	0.58	1
0754 Meubelmakers, kleermakers en sto.	14	ward	0.71	1
0771 Productiemachinebedieners	53	ward	0.79	1
0772 Assemblagemedewerkers	7	ward	0.29	1
1215 Bedieners mobiele machines	20	ward	0.67	1
1212 Chauffeurs auto's	5	ward	0.34	1
1213 Buschauffeurs en trambestu-urders	4	ward	0.30	1
1214 Vrachtwagenchauffeurs	2	n=2	n=2	n=2
1121 Schoonmakers	5	ward	0.42	1
0921 Hulpkrachten landbouw	5	ward	0.32	1
0781 Hulpkrachten bouw en indus-trie	15	ward	0.59	1
1221 Laders, lossers en vakken-vullers	6	ward	0.53	1
1122 Keukenhulpen	4	complete	0.18	1
1222 Vuilnisophalers en dag-bladenbezo.	11	ward	0.55	1



(a) 0612 (b) 1311

Figure B.2: Dendograms of 0612 and 1311

For both BRC-code 0612 and BRC-code 1311, the sum-squared differences between the clusters are calculated according to formula B.1. The sum-squared difference between the clusters in 0612 is 0.75 and for 1311 it is 1.60. This means that we use 0.75 as the benchmark for the aggregation analysis of the BRC-codes with ratio 2. The results of these analyses give that only BRC-code 1214 ('Vrachtwagenchauffeurs' or 'Truck drivers') should be split according to its two O*Net occupations: 'First-Line Supervisors of Transportation and Material-Movement'; and the 'Heavy and Tractor-Trailer Truck Drivers'.

Table B.5: Differences in abilities clusters 0612

Ability	Cluster 1	Cluster 2	Cluster 3
Oral.Comprehension	4.01	3.98	0.03
Written.Comprehension	3.91	4.00	-0.09
Oral.Expression	4.05	4.0	-0.0
Written.Expression	3.73	3.85	-0.12
Fluency.of.Ideas	3.05428571428571	2.55428571428571	0.5
Originality	2.91142857142857	2.5	0.411428571428571
Problem.Sensitivity	4.03571428571429	3.48142857142857	0.554285714285714
Deductive.Reasoning	3.89428571428571	3.51571428571429	0.378571428571429
Inductive.Reasoning	3.98142857142857	3.49714285714286	0.484285714285714
Information.Ordering	3.60714285714286	3.49857142857143	0.108571428571429
Category.Flexibility	3.26714285714286	3.03571428571429	0.231428571428572
Mathematical.Reasoning	2.53571428571429	2.51857142857143	0.0171428571428569
Number.Facility	2.60714285714286	2.42857142857143	0.178571428571429
Memorization	2.76571428571429	2.32142857142857	0.444285714285714
Speed.of.Closure	3.01571428571429	2.33857142857143	0.677142857142857
Flexibility.of.Closure	3.35571428571429	2.62285714285714	0.732857142857143
Perceptual.Speed	3.25	2.44714285714286	0.802857142857143
Spatial.Orientation	2.32285714285714	1	1.32285714285714
Visualization	2.87571428571429	2.12714285714286	0.748571428571429
Selective.Attention	3.30428571428571	2.82428571428571	0.48
Time.Sharing	2.91	2.41142857142857	0.498571428571429
Arm.Hand.Steadiness	2.76714285714286	1.60428571428571	1.16285714285714
Manual.Dexterity	2.60857142857143	1.71428571428571	0.894285714285714
Finger.Dexterity	2.77	2.39285714285714	0.377142857142857
Control.Precision	2.64285714285714	1.62571428571429	1.01714285714286
Multilimb.Coordination	2.66142857142857	1.05428571428571	1.60714285714286
Response.Orientation	2.44571428571429	1.07142857142857	1.37428571428571
Rate.Control	2.16285714285714	1.05428571428571	1.10857142857143
Reaction.Time	2.48285714285714	1.05428571428571	1.42857142857143
Wrist.Finger.Speed	1.59	1.28285714285714	0.307142857142857
Speed.of.Limb.Movement	2.02	1	1.02
Static.Strength	2.46285714285714	1.01714285714286	1.44571428571429
Explosive.Strength	1.92857142857143	1	0.928571428571429
Dynamic.Strength	2.12571428571429	1.07142857142857	1.05428571428571
Trunk.Strength	2.39285714285714	1.33714285714286	1.05571428571429
Stamina	2.33857142857143	1	1.33857142857143
Extent.Flexibility	2.28428571428571	1.01714285714286	1.26714285714286
Dynamic.Flexibility	1.03428571428571	1	0.0342857142857143
Gross.Body.Coordination	2.25142857142857	1	1.25142857142857
Gross.Body.Equilibrium	2.08857142857143	1	1.08857142857143
Near.Vision	3.80428571428571	3.58857142857143	0.215714285714286
Far.Vision	3.35571428571429	2.33714285714286	1.01857142857143
Visual.Color.Discrimination	2.83714285714286	1.96428571428571	0.872857142857143
Night.Vision	1.98142857142857	1	0.981428571428572
Peripheral.Vision	2.01714285714286	1	1.01714285714286
Depth.Perception	2.77	1.25	1.52
Glare.Sensitivity	2.09	1.03571428571429	1.05428571428571
Hearing.Sensitivity	2.57	1.73428571428571	0.835714285714286

Table B.5: Differences in abilities clusters 0612

Ability	Cluster 1	Cluster 2	Cluster 3
Auditory.Attention	2.82428571428571	2.12428571428571	0.7
Sound.Localization	2.03714285714286	1.05285714285714	0.984285714285714
Speech.Recognition	3.76857142857143	3.68	0.0885714285714285
Speech.Clarity	3.80571428571429	3.64285714285714	0.162857142857143

Table B.6: Differences in abilities clusters 1311

Ability	Cluster 1	Cluster 2	Difference
Oral.Comprehension	4.25	3.08333333333333	1.16666666666667
Written.Comprehension	4.125	2.70833333333333	1.41666666666667
Oral.Expression	4.19	3.02333333333333	1.16666666666667
Written.Expression	4	2.48	1.52
Fluency.of.Ideas	3.75	2.27333333333333	1.47666666666667
Originality	3.75	2.17	1.58
Problem.Sensitivity	3.94	3.085	0.855
Deductive.Reasoning	4.06	3.02333333333333	1.03666666666667
Inductive.Reasoning	4	2.835	1.165
Information.Ordering	3.56	2.98	0.58
Category.Flexibility	3.44	2.93833333333333	0.501666666666666
Mathematical.Reasoning	3.13	1.795	1.335
Number.Facility	3	1.89833333333333	1.10166666666667
Memorization	2.625	2.14666666666667	0.478333333333333
Speed.of.Closure	2.75	2.39833333333333	0.351666666666667
Flexibility.of.Closure	3.19	2.77166666666667	0.418333333333333
Perceptual.Speed	2.75	2.64833333333333	0.101666666666667
Spatial.Orientation	1.185	2.39666666666667	-1.21166666666667
Visualization	2.935	2.625	0.31
Selective.Attention	2.94	2.81333333333333	0.126666666666666
Time.Sharing	2.81	2.355	0.455
Arm.Hand.Steadiness	1	3.315	-2.315
Manual.Dexterity	1	3.27166666666667	-2.27166666666667
Finger.Dexterity	2.25	3.04166666666667	-0.791666666666667
Control.Precision	1.75	3.25166666666667	-1.50166666666667
Multilimb.Coordination	1.685	3.39666666666667	-1.71166666666667
Response.Orientation	1.06	2.585	-1.525
Rate.Control	1.06	2.71	-1.65
Reaction.Time	1.125	2.835	-1.71
Wrist.Finger.Speed	1.06	2.20833333333333	-1.14833333333333
Speed.of.Limb.Movement	1	2.29333333333333	-1.29333333333333
Static.Strength	1	3.39666666666667	-2.39666666666667
Explosive.Strength	1	1.645	-0.645
Dynamic.Strength	1	2.85666666666667	-1.85666666666667
Trunk.Strength	1.06	3.23	-2.17
Stamina	1	2.93666666666667	-1.93666666666667
Extent.Flexibility	1	2.99833333333333	-1.99833333333333
Dynamic.Flexibility	1	1.565	-0.565
Gross.Body.Coordination	1	2.68833333333333	-1.68833333333333
Gross.Body.Equilibrium	1	2.27166666666667	-1.27166666666667
Near.Vision	3.63	3.16833333333333	0.461666666666666
Far.Vision	2.81	3.085	-0.275
Visual.Color.Discrimination	1.94	2.58333333333333	-0.643333333333334

Table B.6: Differences in abilities clusters 1311

Ability	Cluster 1	Cluster 2	Difference
Night.Vision	1	1.85666666666667	-0.856666666666667
Peripheral.Vision	1	2.045	-1.045
Depth.Perception	1.875	2.81333333333333	-0.938333333333333
Glare.Sensitivity	1	2.085	-1.085
Hearing.Sensitivity	1.875	2.48166666666667	-0.606666666666666
Auditory.Attention	2	2.295	-0.295
Sound.Localization	1.06	1.87666666666667	-0.816666666666667
Speech.Recognition	4.06	3.04333333333333	1.01666666666667
Speech.Clarity	4.06	2.85333333333333	1.20666666666667

B.5 Skills

This section presents the results [12] of the aggregation analysis of the skills from the 1012 O*Net occupations to their corresponding 113 BRC occupations.

B.5.1 Results Cluster Analysis Skills

Table B.7 shows the results of the cluster analysis of the skills of the O*Net occupations within their corresponding BRC-code. As can be seen, Only the BRC-code 1311 ('Beroepsgroep overig" or 'All others") has more than 1 cluster of O*Net occupations within it –namely, two clusters. Figure B.3 shows which clusters are found within the code 1311. Looking at the O*Net occupations of the clusters, we see that the first cluster correspondent with more managing type of occupations whereas the second cluster considers the more operational type of occupations. This can be further explained by looking at the actual differences in skills between the two clusters (see table B.8). Here one can see that the managing cluster scores higher at most skills than the cluster with the operational jobs except at the technical skills. Technical skills are considered to be more important for the second (operational) cluster than for the cluster with the managing occupations.

Table B.7: Results cluster analysis skills

Job	ratio	method	ac	nc
0611 Overheidsbestuurders	14	ward	0.55	1
1311 Beroepsgroep overig	8	ward	0.82	2
0511 Algemeen directeuren	3	ward	0.23	1
0521 Managers zakelijke en administra.	19	ward	0.59	1
0522 Managers verkoop en marketing	9	ward	0.56	1
0531 Managers productie	19	ward	0.61	1
0532 Managers logistiek	3	complete	0.16	1
0533 Managers ICT	1	NAP	NAP	1
0534 Managers zorginstellingen	4	complete	0.13	1
0535 Managers onderwijs	4	complete	0.39	1
0536 Managers gespecialiseerde dienst.	12	ward	0.47	1
0541 Managers horeca	2	n=2	n=2	n=2
0542 Managers detail- en groothandel	1	NAP	NAP	1
0543 Managers commerciële en persoonl.	10	ward	0.40	1
0711 Biologen en natuurwetenschappers	27	ward	0.67	1
0712 Ingenieurs (geen elektrotechniek)	23	ward	0.65	1
0713 Elektrotechnisch ingenieurs	3	ward	0.11	1
0714 Architecten	6	ward	0.47	1
0221 Grafisch vormgevers en producton...	2	n=2	n=2	n=2
1011 Artsen	21	ward	0.64	1
1012 Gespecialiseerd verpleegkundigen	6	complete	0.44	1
1013 Fysiotherapeuten	11	complete	0.48	1
0111 Docenten hoger onderwijs en hoog...	35	ward	0.73	1
0112 Docenten beroepsgerichte vakken ...	2	n=2	n=2	n=2
0113 Docenten algemene vakken secunda...	1	NAP	NAP	1
0114 Leerkrachten basisonderwijs	7	ward	0.33	1
0115 Onderwijskundigen en overige doc...	5	ward	0.56	1
0411 Accountants	4	complete	0.33	1
0412 Financieel specialisten en economen	6	complete	0.33	1
0413 Bedrijfskundigen en organisatiea.	4	complete	0.19	1
0414 Beleidsadviseurs	6	ward	0.31	1
0415 Specialisten personeels- en loop.	2	n=2	n=2	n=2

[12] r-scripts used for this analysis: *skillsclustering.R*, *clusterfunctionSkills.R* and *finalframeSkills.R*

Table B.7: Results cluster analysis skills

Job	ratio	method	ac	nc
0311 Adviseurs marketing, public rela...	5	ward	0.27	1
0811 Software- en applicatieontwikkel.	18	ward	0.62	1
0812 Databank- en netwerkspecialisten	15	ward	0.57	1
0621 Juristen	4	ward	0.09	1
0211 Bibliothecarissen en conservatoren	4	ward	0.27	1
1022 Psychologen en sociologen	13	ward	0.56	1
1021 Maatschappelijk werkers	4	ward	0.28	1
0212 Auteurs en taalkundigen	4	ward	0.20	1
0213 Journalisten	1	NAP	NAP	1
0214 Beeldend kunstenaars	2	n=2	n=2	n=2
0215 Uitvoerend kunstenaars	11	ward	0.69	1
0721 Technici bouwkunde en natuur	45	ward	0.76	1
0722 Productieleiders industrie en bouw	2	n=2	n=2	n=2
0723 Procesoperators	13	ward	0.54	1
1211 Dekofficieren en piloten	10	ward	0.63	1
1031 Laboranten	9	ward	0.41	1
1032 Apothekersassistenten	1	NAP	NAP	1
1035 Medisch vakspecialisten	20	complete	0.62	1
1033 Verpleegkundigen (mbo)	2	n=2	n=2	n=2
1034 Medisch praktijkassistenten	6	complete	0.36	1
0421 Boekhouders	14	complete	0.59	1
0321 Vertegenwoordigers en inkopers	9	ward	0.55	1
0422 Zakelijke dienstverleners	18	ward	0.63	1
0423 Directiesecretaresses	5	complete	0.36	1
0612 Overheidsambtenaren	14	ward	0.63	1
0631 Politie-inspecteurs	6	ward	0.45	1
1041 Sociaal werkers, groeps- en woon.	6	ward	0.40	1
0121 Beroepsgroep sportinstructeurs	5	ward	0.51	1
0222 Fotografen en interieurontwerpers	4	ward	0.39	1
1112 Koks	6	ward	0.55	1
0821 Gebruikersondersteuning ICT	2	n=2	n=2	n=2
0822 Radio- en televisietechnici	4	ward	0.27	1
0431 Administratief medewerkers	17	ward	0.68	1
0432 Secretaresses	3	complete	0.20	1
0433 Receptionisten en telefonisten	8	complete	0.62	1
0434 Boekhoudkundig medewerkers	6	ward	0.43	1
0435 Transportplanners en logistiek m.	2	n=2	n=2	n=2
1111 Reisbegeleiders	5	complete	0.49	1
1113 Kelners en barpersoneel	1	NAP	NAP	1
1114 Kappers en schoonheidsspecialisten	6	ward	0.51	1
1115 Conciërges en teamleiders schoon.	1	NAP	NAP	1
1116 Verleners van overige persoonlij.	7	ward	0.62	1
0334 Callcentermedewerkers outbound e...	7	complete	0.59	1
0331 Winkeliers en teamleiders detail...	2	n=2	n=2	n=2
0332 Verkoopmedewerkers detailhandel	2	n=2	n=2	n=2
0333 Kassamedewerkers	1	NAP	NAP	1
0131 Leidsters kinderopvang en onderw...	2	n=2	n=2	n=2
1051 Verzorgenden	6	ward	0.34	1
0632 Politie en brandweer	8	ward	0.59	1
0633 Beveiligingspersoneel	10	ward	0.61	1
0911 Land- en bosbouwers	10	ward	0.77	1
0912 Hoveniers, tuinders en kwekers	2	n=2	n=2	n=2
0913 Veetelers	7	ward	0.72	1

Table B.7: Results cluster analysis skills

Job	ratio	method	ac	nc
0731 Bouwarbeiders ruwbouw	13	ward	0.65	1
0732 Timmerlieden	2	n=2	n=2	n=2
0733 Bouwarbeiders afbouw	11	ward	0.69	1
0734 Loodgieters en pijpfitters	3	ward	0.14	1
0735 Schilders en metaalspuiters	3	complete	0.07	1
0741 Metaalbewerkers en constructiewe.	22	ward	0.66	1
0742 Lassers en plaatwerkers	5	complete	0.47	1
0743 Automonteurs	8	ward	0.63	1
0744 Machinemonteurs	7	complete	0.54	1
0755 Medewerkers drukkerij en kunstni.	19	ward	0.74	1
0761 Elektriciens en elektronicamonteurs	23	ward	0.67	1
0751 Slagers	3	complete	0.29	1
0752 Bakkers	1	NAP	NAP	1
0753 Productcontroleurs	11	ward	0.63	1
0754 Meubelmakers, kleermakers en sto.	14	ward	0.74	1
0771 Productiemachinebedieners	53	ward	0.82	1
0772 Assemblagemedewerkers	7	complete	0.38	1
1215 Bedieners mobiele machines	20	ward	0.76	1
1212 Chauffeurs auto's, taxi's en bes.	5	ward	0.46	1
1213 Buschauffeurs en trambestuurders	4	ward	0.38	1
1214 Vrachtwagenchauffeurs	2	n=2	n=2	n=2
1121 Schoonmakers	5	ward	0.44	1
0921 Hulpkrachten landbouw	5	ward	0.49	1
0781 Hulpkrachten bouw en industrie	15	ward	0.60	1
1221 Laders, lossers en vakkenvullers	6	ward	0.59	1
1122 Keukenhulpen	4	complete	0.26	1
1222 Vuilnisophalers en dagbladenbezo.	11	ward	0.64	1

According to the Gap Statistic method, only the BRC-code 1311 has an optimal number of clusters that is higher than 1. Table B.3 shows the actual differences of the relevant –e.g. higher than two– skills between these clusters between. According to equation B.1, the sum of the squared differences is 1.653. Since there is only one BRC with two optimal clusters, we use this number as the benchmark for the aggregation analysis of the BRC-codes with exactly two O*Net occupations. Analysis of the BRC-codes with a ratio of 2 found that none of these BRC-codes should be split. We could therefore thus aggregate all the O*Net occupations within the BRC-codes with a ratio of 2.

Table B.8: Difference in skills between clusters 1311

Skill	Cluster 1	Cluster 2	Difference
Reading Comprehension	4.06	2.62666666666667	1.43333333333333
Active.Listening	4	2.95833333333333	1.04166666666667
Writing	3.94	2.335	1.605
Speaking	4.19	2.89333333333333	1.29666666666667
Mathematics	3	1.815	1.185
Science	1.815	1.605	0.21
Critical.Thinking	4.19	3.02166666666667	1.16833333333333
Active.Learning	3.75	2.45833333333333	1.29166666666667
Learning.Strategies	3.25	2.19	1.06
Monitoring	3.87	2.83333333333333	1.03666666666667
Social Perceptiveness	4	2.66666666666667	1.33333333333333
Coordination	3.935	2.85666666666667	1.07833333333333

Table B.8: Difference in skills between clusters 1311

Skill	Cluster 1	Cluster 2	Difference
Persuasion	3.87	2.20833333333333	1.66166666666667
Negotiation	3.5	2.18666666666667	1.31333333333333
Instructing	3.185	2.27166666666667	0.913333333333334
Service.Orientation	3.185	2.31333333333333	0.871666666666667
Complex Problem Solving	4.19	2.64666666666667	1.54333333333333
Operations.Analysis	2.87	1.46	1.41
Technology.Design	1.685	1.23	0.455
Equipment.Selection	1.12	2.39666666666667	-1.27666666666667
Installation	1	1.20833333333333	-0.208333333333333
Programming	1.62	1	0.62
Operation.Monitoring	2.065	2.83333333333333	-0.768333333333334
Operation.and.Control	1.685	2.875	-1.19
Equipment.Maintenance	1	2.54333333333333	-1.54333333333333
Troubleshooting	1.06	2.45833333333333	-1.39833333333333
Repairing	1	2.48	-1.48
Quality Control Analysis	2	2.39666666666667	-0.396666666666667
Judgment and Decision Making	4.125	2.73	1.395
Systems.Analysis	3.87	1.79166666666667	2.07833333333333
Systems.Evaluation	3.87	1.77	2.1
Time.Management	3.69	2.645	1.045
Management of Financial.Resources	3.37	1.52166666666667	1.84833333333333
Management of Material.Resources	3.13	1.66666666666667	1.46333333333333
Management of Personnel.Resources	3.815	2.35666666666667	1.45833333333333

B.6 Knowledge

This section presents the results [13] of the aggregation analysis of the types of knowledge from the 1012 O*Net occupations to their corresponding 113 BRC occupations.

B.6.1 Results Cluster Analysis knowledge

Table B.9 shows the results of the cluster analysis of the type of knowledge needed. As can be seen, no clusters within a BRC-code with a ratio higher than 2 are found. Because of these results, no benchmark for knowledge can be calculated. We, therefore, use the minimal value of the mean-squares of the skills and abilities cluster analyses. This minimum is the mean square of 0612 of the abilities analysis: 0.75.

Using this benchmark we found that the BRC-code 0221 ('Grafische vormgevers" or 'Graphical designers') have to be split according to its two O*Net occupations: 'Commercial and Industrial design'; and 'Graphic designers'.

Table B.9: Results cluster analysis knowledge

BRC	ratio	method	ac	nc
0611 Overheidsbestuurders	14	ward	0.58	1
1311 Beroepsgroep overig	8	ward	0.58	1

[13] r-scripts used for this analysis: *knowledgeclustering.R*, *clusterfunctionKnowledge.R* and *finalframeKnowledge.R*

Table B.9: Results cluster analysis knowledge

BRC	ratio	method	ac	nc
0511 Algemeen directeuren	3	complete	0.16	1
0521 Managers zakelijke en administra.	19	ward	0.63	1
0522 Managers verkoop en marketing	9	ward	0.59	1
0531 Managers productie	19	ward	0.63	1
0532 Managers logistiek	3	complete	0.07	1
0533 Managers ICT	1	NAP	NAP	1
0534 Managers zorginstellingen	4	ward	0.20	1
0535 Managers onderwijs	4	complete	0.25	1
0536 Managers gespecialiseerde dienst.	12	ward	0.61	1
0541 Managers horeca	2	n=2	n=2	n=2
0542 Managers detail- en groothandel	1	NAP	NAP	1
0543 Managers commerciÃ«le en persoonl.	10	ward	0.50	1
0711 Biologen en natuurwetenschappers	27	ward	0.69	1
0712 Ingenieurs (geen elektrotechniek)	23	ward	0.69	1
0713 Elektrotechnisch ingenieurs	3	ward	0.23	1
0714 Architecten	6	ward	0.51305566124123	1
0221 Grafisch vormgevers en producton...	2	n=2	n=2	n=2
1011 Artsen	21	ward	0.62	1
1012 Gespecialiseerd verpleegkundigen	6	complete	0.36	1
1013 Fysiotherapeuten	11	ward	0.44	1
0111 Docenten hoger onderwijs en hoog...	35	ward	0.75	1
0112 Docenten beroepsgerichte vakken ...	2	n=2	n=2	n=2
0113 Docenten algemene vakken secunda...	1	NAP	NAP	1
0114 Leerkrachten basisonderwijs	7	ward	0.42	1
0115 Onderwijskundigen en overige doc...	5	ward	0.29	1
0411 Accountants	4	ward	0.19	1
0412 Financieel specialisten en economen	6	ward	0.38	1
0413 Bedrijfskundigen en organisatiea.	4	complete	0.32	1
0414 Beleidsadviseurs	6	ward	0.41	1
0415 Specialisten personeels- en loop.	2	n=2	n=2	n=2
0311 Adviseurs marketing	5	ward	0.35	1
0811 Software- en applicatieontwikkel.	18	ward	0.63	1
0812 Databank- en netwerkspecialisten	15	ward	0.61	1
0621 Juristen	4	complete	0.19272451750642	1
0211 Bibliothecarissen en conservatoren	4	ward	0.34	1
1022 Psychologen en sociologen	13	ward	0.59	1
1021 Maatschappelijk werkers	4	ward	0.22	1
0212 Auteurs en taalkundigen	4	complete	0.24	1
0213 Journalisten	1	NAP	NAP	1
0214 Beeldend kunstenaars	2	n=2	n=2	n=2
0215 Uitvoerend kunstenaars	11	ward	0.54	1
0721 Technici bouwkunde en natuur	45	ward	0.80	1
0722 Productieleiders industrie en bouw	2	n=2	n=2	n=2
0723 Procesoperators	13	ward	0.469	1
1211 Dekofficieren en piloten	10	ward	0.55	1
1031 Laboranten	9	ward	0.69	1
1032 Apothekersassistenten	1	NAP	NAP	1
1035 Medisch vakspecialisten	20	ward	0.69	1
1033 Verpleegkundigen (mbo)	2	n=2	n=2	n=2
1034 Medisch praktijkassistenten	6	complete	0.29	1
0421 Boekhouders	14	ward	0.50	1
0321 Vertegenwoordigers en inkopers	9	ward	0.52	1
0422 Zakelijke dienstverleners	18	ward	0.60	1

Table B.9: Results cluster analysis knowledge

BRC	ratio	method	ac	nc
0423 Directiesecretaresses	5	ward	0.40	1
0612 Overheidsambtenaren	14	ward	0.61	1
0631 Politie-inspecteurs	6	ward	0.499	1
1041 Sociaal werkers	6	ward	0.40	1
0121 Beroepsgroep sportinstructeurs	5	complete	0.36	1
0222 Fotografen en interieurontwerpers	4	complete	0.23	1
1112 Koks	6	complete	0.56	1
0821 Gebruikersondersteuning ICT	2	n=2	n=2	n=2
0822 Radio- en televisietechnici	4	ward	0.40	1
0431 Administratief medewerkers	17	ward	0.61	1
0432 Secretaresses	3	complete	0.15	1
0433 Receptionisten en telefonisten	8	complete	0.49	1
0434 Boekhoudkundig medewerkers	6	ward	0.37	1
0435 Transportplanners en logistiek m.	2	n=2	n=2	n=2
1111 Reisbegeleiders	5	complete	0.18	1
1113 Kelners en barpersoneel	1	NAP	NAP	1
1114 Kappers en schoonheidsspecialisten	6	ward	0.3	1
1115 ConciÃ«rges en teamleiders schoon.	1	NAP	NAP	1
1116 Verleners van overige persoonlij.	7	ward	0.45	1
0334 Callcentermedewerkers outbound e...	7	complete	0.51	1
0331 Winkeliers en teamleiders detail...	2	n=2	n=2	n=2
0332 Verkoopmedewerkers detailhandel	2	n=2	n=2	n=2
0333 Kassamedewerkers	1	NAP	NAP	1
0131 Leidsters kinderopvang en onderw...	2	n=2	n=2	n=2
1051 Verzorgenden	6	ward	0.35	1
0632 Politie en brandweer	8	ward	0.57	1
0633 Beveiligingspersoneel	10	ward	0.60	1
0911 Land- en bosbouwers	10	ward	0.54	1
0912 Hoveniers, tuinders en kwekers	2	n=2	n=2	n=2
0913 Veetelers	7	ward	0.42	1
0731 Bouwarbeiders ruwbouw	13	ward	0.64	1
0732 Timmerlieden	2	n=2	n=2	n=2
0733 Bouwarbeiders afbouw	11	complete	0.67	1
0734 Loodgieters en pijpfitters	3	ward	0.10	1
0735 Schilders en metaalspuiters	3	complete	0.22	1
0741 Metaalbewerkers en constructiewe.	22	ward	0.74	1
0742 Lassers en plaatwerkers	5	complete	0.49	1
0743 Automonteurs	8	ward	0.55	1
0744 Machinemonteurs	7	complete	0.468	1
0755 Medewerkers drukkerij en kunstni.	19	complete	0.64	1
0761 Elektriciens en elektronicamonteurs	23	ward	0.65	1
0751 Slagers	3	ward	0.31	1
0752 Bakkers	1	NAP	NAP	1
0753 Productcontroleurs	11	ward	0.50	1
0754 Meubelmakers	14	ward	0.55	1
0771 Productiemachinebedieners	53	ward	0.80	1
0772 Assemblagemedewerkers	7	ward	0.54	1
1215 Bedieners mobiele machines	20	ward	0.67	1
1212 Chauffeurs auto's	5	ward	0.42	1
1213 Buschauffeurs en trambestuurders	4	complete	0.2	1
1214 Vrachtwagenchauffeurs	2	n=2	n=2	n=2
1121 Schoonmakers	5	ward	0.44	1
0921 Hulpkrachten landbouw	5	ward	0.24	1

Table B.9: Results cluster analysis knowledge

BRC	ratio	method	ac	nc
0781 Hulpkrachten bouw en industrie	15	ward	0.67	1
1221 Laders	6	ward	0.58	1
1122 Keukenhulpen	4	ward	0.30	1
1222 Vuilnisophalers en dagbladenbezo.	11	ward	0.55	1

B.7 Results descriptive analysis

This section presents the results[14] of the descriptive analysis of the competencies of the Dutch workers. According to CBS, there are 8587.000 people in the Netherlands with a job. Table B.10 shows the top 5 of the most common jobs in the Netherlands. Here we see that job 0332 ('Verkoopmedewerkers Detailhandel' or 'Shop-assistant') is the most common job in the Netherlands. Moreover, we see that job 1211 ('Laders, Lossers en Vakkenvullers' or 'Loaders, Unloaders and Compartment Fillers') is the most common job for the men in the Netherlands whereas job 0332 ('Verkoopmedewerkers Detailhandel' or 'Shop-assistant') is the most common job for the women in the Netherlands.

Table B.10: Top 5 jobs in the Netherlands

Job	Total.men.and.women.x.1000	men	women
0332 Verkoopmedewerkers detailhandel	335.00	107.00	228.00
0431 Administratief medewerkers	278.00	82.00	196.00
1121 Schoonmakers	265.00	61.00	204.00
1051 Verzorgenden	250.00	16.00	234.00
1221 Laders, lossers en vakkenvullers	235.00	173.00	62.00

Moreover, as stated before, the CBS aggregates professions according to their BRC-codes. The following tables provides the number of workers per 3-digit BRC-code and 2-digit BRC-code:

Table B.11: Number of dutch workers per 2-digit occupations

job	Total.x1000	men	women
04 Bedrijfseconomische en administrat...	1621	742	879
07 Technische beroepen	1244	1075	168
10 Zorg en welzijn beroepen	1210	247	962
03 Commerciële beroepen	984	431	553
11 Dienstverlenende beroepen	810	295	514
12 Transport en logistiek beroepen	638	528	109
01 Pedagogische beroepen	617	178	440
05 Managers	428	321	108
08 ICT beroepen	368	316	53
06 Openbaar bestuur, veiligheid en ju.	300	187	114
02 Creatieve en taalkundige beroepen	223	122	101
09 Agrarische beroepen	188	148	40
13 Beroepsklasse overig	143	95	48

[14] r-scripts used for this analysis: *descriptiveclustering.R*, and *descriptive.R*

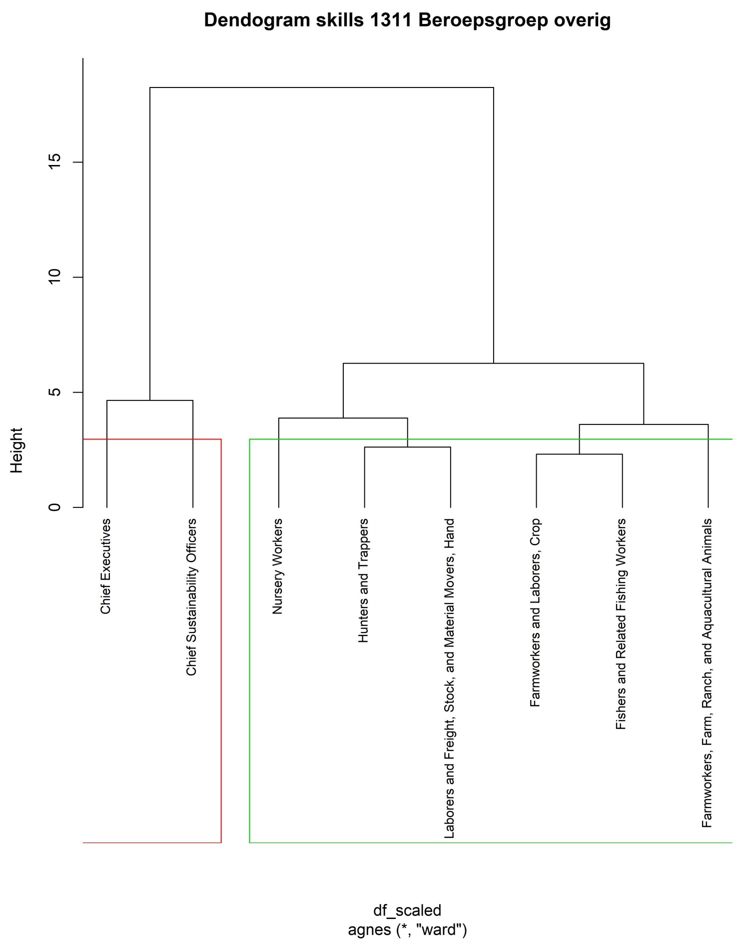


Figure B.3: Clustering skills 1311

B.7.1 Distribution of skills in the Netherlands

Table B.13 shows the median and mean for each skill in the Netherlands. The skills/competencies in this table are ordered based on their median –the competence with the highest median on top. In this table, we see that, in general, the basic skills –especially in the category 'Content'—, have the highest median. This means that when looking at the number of people for each job, and the level of importance for skills for each job, these basic skills are the most important for the most people in the Netherlands. As stated earlier, according to O*Net, skills that score lower than 2 are not considered to be important for that job. This means that skills that have a median lower than 2, are not considered not important for the greater part of workers in the Netherlands. looking at the skill headers, we see that these not important-skills consists of almost only cross-functional skills. Almost all basic-skills are considered to be important for more than 50% of the Dutch workers. The only basic-skill that is considered to be not important for most of the Dutch workers is the basic skill 'Science'.

Figure B.4 (a) shows the distribution of basic-skills and cross-functional skills in more detail. Here we see that most workers in the Netherlands have a job where the requirement of basic-skills is important –most of the mass under the PDF of basic-skills lies above 2. The level of importance for cross-functional skills is more evenly distributed across dutch workers.

If we look closer to the distribution of skills based on their category, we see that for most workers in the Netherlands it is not important to have technical skills (see Table B.13 and Figure B.4 (b)). On the other hand, the possession of social skills, process skills, content skills and system skills is of importance for most of the Dutch workers.

Figure B.5 and Figure B.6 show the density functions of the skills per category. If we look at the distribution of technical skills we see that skill installation' has a very skewed distribution around the value 1. According to table B.13 the median of skill installation is also one. However, the mean of installation is slightly higher, 1.15. This indicates that for almost no workers in the Netherlands, the skill 'Installation' is needed in order to exercise their jobs. Only a few workers need the skill 'Installation' (see Table B.14).

Table B.13: Importance of skills in the Netherlands

Competence	Header	Category	Median	Mean
Active Listening	Basic Skill	Content	3.75	3.64
Speaking	Basic Skill	Content	3.56	3.56
Reading Comprehension	Basic Skill	Content	3.48	3.44
Critical Thinking	Basic Skill	Process	3.43	3.47
Monitoring	Basic Skill	Process	3.23	3.29
Social Perceptiveness	Cross-Functional Skills	Social Skills	3.23	3.25
Judgment and Decision Making	Cross-Functional Skills	Systems Skills	3.21	3.21
Time Management	Cross-Functional Skills	Resource Management Skills	3.13	3.14
Coordination	Cross-Functional Skills	Social Skills	3.12	3.18
Complex Problem Solving	Cross-Functional Skills	Technical Skills	3.12	3.14
Writing	Basic Skill	Content	3.08	3.14
Service Orientation	Cross-Functional Skills	Social Skills	3.03	3.08
Active Learning	Basic Skill	Process	3.00	3.06
Mathematics	Basic Skill	Content	2.96	3.00
Persuasion	Cross-Functional Skills	Social Skills	2.88	2.89
Instructing	Cross-Functional Skills	Social Skills	2.83	2.81
Learning Strategies	Basic Skill	Process	2.75	2.73
Negotiation	Cross-Functional Skills	Social Skills	2.69	2.74
Management of Personnel Resources	Cross-Functional Skills	Resource Management Skills	2.63	2.64
Systems Evaluation	Cross-Functional Skills	Systems Skills	2.56	2.56
Systems Analysis	Cross-Functional Skills	Systems Skills	2.50	2.62
Operation Monitoring	Cross-Functional Skills	Technical Skills	2.27	2.34
Quality Control Analysis	Cross-Functional Skills	Technical Skills	2.19	2.22
Management of Material Resources	Cross-Functional Skills	Resource Management Skills	1.94	1.99
Operations Analysis	Cross-Functional Skills	Technical Skills	1.84	1.94
Management of Financial Resources	Cross-Functional Skills	Resource Management Skills	1.83	1.91
Troubleshooting	Cross-Functional Skills	Technical Skills	1.82	1.82
Operation and Control	Cross-Functional Skills	Technical Skills	1.78	1.97
Technology Design	Cross-Functional Skills	Technical Skills	1.62	1.66
Science	Basic Skill	Content	1.50	1.66
Equipment Selection	Cross-Functional Skills	Technical Skills	1.44	1.53
Programming	Cross-Functional Skills	Technical Skills	1.43	1.52
Equipment Maintenance	Cross-Functional Skills	Technical Skills	1.22	1.48
Repairing	Cross-Functional Skills	Technical Skills	1.16	1.46
Installation	Cross-Functional Skills	Technical Skills	1.00	1.15

B.7.2 Distribution of Abilities in the Netherlands

Table B.15 shows the median and mean for the importance of the abilities for Dutch workers. Looking at the header of abilities, we see that, in general, abilities that fall under the header ‘cognitive abilities’ have the highest median. Abilities of the header Sensory abilities are at the bottom. The verbal abilities ‘Oral Expression’ and ‘Oral Comprehension’ are at the top with a median of 3.90 and 3.88. This means that for over 50 per cent of the Dutch worker these abilities are of high importance.

Table B.15: Importance of abilities in the Netherlands

Competence	Header	Category	Median	Mean
Oral Comprehension	Cognitive Abilities	Verbal Abilities	3.90	3.79
Oral Expression	Cognitive Abilities	Verbal Abilities	3.88	3.75
Problem Sensitivity	Cognitive Abilities	Idea Generation and Reasoning Abilities	3.56	3.55
Near Vision	Sensory Abilities	Visual Abilities	3.56	3.53
Written Comprehension	Cognitive Abilities	Verbal Abilities	3.53	3.50
Speech Recognition	Sensory Abilities	Auditory and Speech Abilities	3.50	3.46
Speech Clarity	Sensory Abilities	Auditory and Speech Abilities	3.44	3.45
Deductive Reasoning	Cognitive Abilities	Idea Generation and Reasoning Abilities	3.41	3.44
Written Expression	Cognitive Abilities	Verbal Abilities	3.37	3.30
Information Ordering	Cognitive Abilities	Idea Generation and Reasoning Abilities	3.27	3.32
Inductive Reasoning	Cognitive Abilities	Idea Generation and Reasoning Abilities	3.25	3.34
Category Flexibility	Cognitive Abilities	Idea Generation and Reasoning Abilities	3.05	3.07
Selective Attention	Cognitive Abilities	Attentiveness	3.00	3.02
Fluency of Ideas	Cognitive Abilities	Idea Generation and Reasoning Abilities	2.88	2.86
Originality	Cognitive Abilities	Idea Generation and Reasoning Abilities	2.88	2.79
Flexibility of Closure	Cognitive Abilities	Perceptual Abilities	2.87	2.83
Far Vision	Sensory Abilities	Visual Abilities	2.81	2.83
Perceptual Speed	Cognitive Abilities	Perceptual Abilities	2.75	2.75
Time Sharing	Cognitive Abilities	Attentiveness	2.65	2.65
Visualization	Cognitive Abilities	Spatial Abilities	2.62	2.62
Finger Dexterity	Psychomotor Abilities	Fine Manipulative Abilities	2.62	2.63
Number Facility	Cognitive Abilities	Quantitative Abilities	2.55	2.54
Mathematical Reasoning	Cognitive Abilities	Quantitative Abilities	2.54	2.57
Arm Hand Steadiness	Psychomotor Abilities	Fine Manipulative Abilities s	2.50	2.40
Speed of Closure	Cognitive Abilities	Perceptual Abilities	2.44	2.45
Memorization	Cognitive Abilities	Perceptual Abilities	2.41	2.41
Visual Color Discrimination	Sensory Abilities	Visual Abilities	2.39	2.41
Manual Dexterity	Psychomotor Abilities	Fine Manipulative Abilities	2.38	2.33
Trunk Strength	Physical Abilities	Physical Strength Abilities	2.31	2.25
Auditory Attention	Sensory Abilities	Auditory and Speech Abilities	2.25	2.32
Control Precision	Psychomotor Abilities	Control Movement Abilities	2.19	2.23
Multilimb Coordination	Psychomotor Abilities	Control Movement Abilities	2.13	2.15
Hearing Sensitivity	Sensory Abilities	Auditory and Speech Abilities	2.07	2.18
Static Strength	Physical Abilities	Physical Strength Abilities	2.00	1.96
Depth Perception	Sensory Abilities	Visual Abilities	1.98	2.05
Stamina	Physical Abilities	Endurance	1.91	1.85
Extent Flexibility	Physical Abilities	Flexibility, Balance, and Coordination	1.88	1.90
Gross Body Coordination	Physical Abilities	F Flexibility, Balance, and Coordination	1.84	1.78
Dynamic Strength	Physical Abilities	Physical Strength Abilities	1.73	1.74
Gross Body Equilibrium	Physical Abilities	Flexibility, Balance, and Coordination	1.68	1.66
Reaction Time	Psychomotor Abilities	Reaction Time and Speed Abilities	1.65	1.78

Wrist Finger Speed	Psychomotor Abilities	Reaction Time and Speed Abilities	1.57	1.65
Response Orientation	Psychomotor Abilities	Control Movement Abilities s	1.50	1.68
Rate Control	Psychomotor Abilities	Control Movement Abilities	1.50	1.65
Speed of Limb Movement	Psychomotor Abilities	Reaction Time and Speed Abilities	1.44	1.54
Spatial Orientation	Cognitive Abilities	Spatial Abilities	1.31	1.47
Explosive Strength	Physical Abilities	Physical Strength Abilities	1.29	1.31
Glare Sensitivity	Sensory Abilities	Visual Abilities	1.19	1.36
Peripheral Vision	Sensory Abilities	Visual Abilities	1.18	1.33
Sound Localization	Sensory Abilities	Auditory and Speech Abilities	1.18	1.33
Night Vision	Sensory Abilities	Visual Abilities	1.13	1.28

Looking at the density function of the abilities by header we see indeed that for most workers the cognitive abilities are of high importance –most of the mass of this PDF lies between 2.5 and 4 (see Figure B.7. The probability function of the abilities under the header 'Psychical abilities' shows that the physical abilities are of least importance for most workers. Figure B.8, Figure B.9, Figure B.10 and figure B.11 show the density plots of the abilities grouped by category. Here we see that the all the abilities of category: 'Attractiveness's'; 'Idea generation abilities' ; and 'Verbal abilities' are important (score higher than 2) for all the workers in the Netherlands (see Figure B.8 (a), B.9 (c) and Figure B.11 (a)). Moreover, the abilities: speech recognition and speech clarity of category 'Auditory and Speech recognition' have also a level higher than 2 for all the Dutch workers. The ability 'Dynamic Flexibility' is for no worker in the Netherlands important.

Analysis towards the distribution of abilities across the Dutch working population, founded that the sensory abilities 'Night Vision', 'Peripheral Vision', 'Sound Localization' and 'Glare Sensitivity'; and the physical ability 'Explosive Strength are the least needed for Dutch workers in order to exercise their profession. The following table shows the jobs and number of workers who at least require one of these abilities.

B.7.3 Distribution of Knowledge in the Netherlands

Table B.9 present the median and mean for the level of importance of each type of knowledge of the Dutch workers. Here we see that the type of Knowledge 'Customer and Personal Service' has the highest median and the knowledge type 'Fine Arts' the lowest. In general, types of knowledge of the category 'Arts and Humanities' are of less importance for most of the Dutch workers than types of knowledge of category 'Business and Management'.

Table B.17: Importance of types of knowledge in the Netherlands

Competence	Category	Median	Mean
Customer and Personal Service	Business and Management	3.74	3.65
Mathematics	Mathematics and Science	2.96	3.00
Administration and Management	Business and Management	2.95	3.06
Computers and Electronics	Engineering and Technology	2.92	2.89
Clerical	Business and Management	2.79	2.77
Education and Training	Education and Training	2.76	2.89
Public Safety and Security	Law and Public Safety	2.75	2.70
Law and Government	Law and Public Safety	2.44	2.52
Personnel and Human Resources	Business and Management	2.34	2.45
Sales and Marketing	Business and Management	2.29	2.39
Production and Processing	Manufacturing and Production	2.28	2.33
Psychology	Mathematics and Science	2.27	2.43
Communications and Media	Communications	2.27	2.31

Economics and Accounting	Business and Management	2.13	2.28
Telecommunications	Communications	2.07	2.08
Transportation	Transportation	2.03	2.16
Mechanical	Engineering and Technology	1.94	2.13
Engineering and Technology	Engineering and Technology	1.82	2.05
Chemistry	Mathematics and Science	1.81	1.82
Geography	Mathematics and Science	1.79	1.85
Sociology and Anthropology	Mathematics and Science	1.72	1.91
Design	Engineering and Technology	1.70	1.95
Foreign Language	Arts and Humanities	1.61	1.65
Physics	Mathematics and Science	1.58	1.70
Medicine and Dentistry	Health Services	1.49	1.73
Therapy and Counseling	Health Services	1.46	1.76
Building and Construction	Engineering and Technology	1.44	1.69
Philosophy and Theology	Arts and Humanities	1.43	1.58
Biology	Mathematics and Science	1.40	1.62
History and Archeology	Arts and Humanities	1.35	1.42
Food Production	Manufacturing and Production	1.31	1.44
Fine Arts	Arts and Humanities	1.21	1.33

The density plots of the categories of the types of knowledge show that the types of knowledge 'Education and Training', 'National language', 'Administration and Management', 'Clerical', 'Customer and Personal Service' are important to have for all dutch workers (see Figure B.14, Figure B.13 and Figure B.15). The density plots of the types of knowledge with category 'Health and Services' shows that these types of knowledge score lower than 2 for most of the Dutch workers.

B.7.4 Results cluster-analysis Dutch occupations clustered by competencies

To indicate which occupations are similar in their competencies, a hierarchical cluster analysis is performed on all the Dutch occupations. Since there are 116 distinct occupations, a dendrogram of all these occupations would be too chaotic. We, therefore, have chosen to present results of the hierarchical cluster analysis –the clusters– in a table (See Table B.18). The column '41 Clusters' shows to which cluster each job belongs when we chose 41 as the number of clusters. The column '13 Clusters' shows to which cluster each job belongs when we have chosen the number of clusters to be similar to the number of 2-digit groups according to the CBS. Note, however, that these cluster numbers only represent the cluster of a job. It has no normative value. For example, the job 0111 has a value 1 for the factor '41 Clusters'. This only means that it belongs to the same cluster as job 1022 –which also has value 1 at factor 41 Clusters.

Table B.18: Clusters of Dutch occupations by their competencies

Job	41 Clusters	13 Clusters
0111 Docenten hoger onderwijs en hoog...	1	1
1022 Psychologen en sociologen	1	1
0112 Docenten beroepsgerichte vakken ...	2	1
0113 Docenten algemene vakken secunda...	2	1
0114 Leerkrachten basisonderwijs	2	1
0115 Onderwijskundigen en overige doc...	2	1
0121 Beroepsgroep sportinstructeurs	3	2
0131 Leidsters kinderopvang en onderw...	3	2
0211 Bibliothecarissen en conservatoren	4	3
0212 Auteurs en taalkundigen	5	3
0213 Journalisten	5	3
0214 Beeldend kunstenaars	6	4
0215 Uitvoerend kunstenaars	7	3
0222 Fotografen en interieurontwerpers	8	4
0221 Grafisch vormgevers en producton... -Graphic Designers	8	4
0311 Adviseurs marketing, public rela...	9	5
0413 Bedrijfskundigen en organisatiea.	9	5
0414 Beleidsadviseurs	9	5
0422 Zakelijke dienstverleners	9	5
0522 Managers verkoop en marketing	9	5
0321 Vertegenwoordigers en inkopers	10	6
0411 Accountants	10	6
0412 Financieel specialisten en economen	10	6
0421 Boekhouders	10	6
0331 Winkeliers en teamleiders detail...	11	5
0541 Managers horeca	11	5
0542 Managers detail- en groothandel	11	5
0332 Verkoopmedewerkers detailhandel	12	2
1051 Verzorgenden	12	2
1111 Reisbegeleiders	12	2
1116 Verleners van overige persoonlij.	12	2
0333 Kassamedewerkers	13	7
0431 Administratief medewerkers	13	7
0334 Callcentermedewerkers outbound e...	14	7
1114 Kappers en schoonheidsspecialisten	14	7
0415 Specialisten personeels- en loop.	15	6
0433 Receptionisten en telefonisten	15	6
0434 Boekhoudkundig medewerkers	15	6
1041 Sociaal werkers, groeps- en woon.	15	6

Table B.18: Clusters of Dutch occupations by their competencies

Job	41 Clusters	13 Clusters
0612 Overheidsambtenaren- office	15	6
0423 Directiesecretaresses	16	6
0432 Secretaresses	16	6
0435 Transportplanners en logistiek m.	17	2
1214 Vrachtwagenchauffeurs - First-Line Supervisors of Transportation and Material-Moving Machine and Vehicle Operators	17	2
0511 Algemeen directeuren	18	5
0521 Managers zakelijke en administra.	18	5
0531 Managers productie	18	5
0532 Managers logistiek	18	5
0534 Managers zorginstellingen	18	5
0535 Managers onderwijs	18	5
0536 Managers gespecialiseerde dienst.	18	5
0543 Managers commerciÃĪle en persoonl.	18	5
0611 Overheidsbestuurders	18	5
0533 Managers ICT	19	8
0811 Software- en applicatieontwikkel.	19	8
0812 Databank- en netwerkspecialisten	19	8
0621 Juristen	20	1
1021 Maatschappelijk werkers	20	1
0631 Politie-inspecteurs	21	9
0633 Beveiligingspersoneel	21	9
0612 Overheidsambtenaren- operational	21	9
0632 Politie en brandweer	22	9
0711 Biologen en natuurwetenschappers	23	8
0712 Ingenieurs (geen elektrotechniek)	23	8
0713 Elektrotechnisch ingenieurs	23	8
0714 Architecten	24	8
0721 Technici bouwkunde en natuur	25	10
0723 Procesoperators	25	10
0822 Radio- en televisietechnici	25	10
0722 Productieleiders industrie en bouw	26	10
1211 Dekofficieren en piloten	26	10
0731 Bouwarbeiders ruwbouw	27	10
0732 Timmerlieden	27	10
0734 Loodgieters en pijpfitters	27	10
0733 Bouwarbeiders afbouw	28	11
0735 Schilders en metaalspuiters	28	11
0781 Hulpkrachten bouw en industrie	28	11
0921 Hulpkrachten landbouw	28	11
1311 Beroepsgroep overig - operational	28	11
0741 Metaalbewerkers en constructiewe.	29	11
0742 Lassers en plaatwerkers	29	11
0753 Productcontroleurs	29	11
0754 Meubelmakers, kleermakers en sto.	29	11
0755 Medewerkers drukkerij en kunstni.	29	11
0771 Productiemachinebedieners	29	11
0772 Assemblagemedewerkers	29	11
0743 Automonteurs	30	10
0744 Machinemonteurs	30	10
0761 Elektriciens en elektronicamonteurs	30	10
0751 Slagers	31	7
0752 Bakkers	31	7

Table B.18: Clusters of Dutch occupations by their competencies

Job	41 Clusters	13 Clusters
1122 Keukenhulpen	31	7
0821 Gebruikersondersteuning ICT	32	8
0911 Land- en bosbouwers	33	12
0912 Hoveniers, tuinders en kwekers	33	12
0913 Veetelers	33	12
1011 Artsen	34	13
1012 Gespecialiseerd verpleegkundigen	34	13
1013 Fysiotherapeuten	34	13
1033 Verpleegkundigen (mbo)	34	13
1031 Laboranten	35	2
1032 Apothekersassistenten	35	2
1034 Medisch praktijkassistenten	35	2
1035 Medisch vakspecialisten	35	2
1112 Koks	36	2
1113 Kelners en barpersoneel	36	2
1115 ConciÃĨrges en teamleiders schoon.	37	7
1121 Schoonmakers	37	7
1212 Chauffeurs auto's, taxi's en bes.	38	12
1221 Laders, lossers en vakkenvullers	38	12
1222 Vuilnisophalers en dagbladenbezo.	38	12
1213 Buschauffeurs en trambestuurders	39	12
1215 Bedieners mobiele machines	39	12
1214 Vrachtwagenchauffeurs - Heavy and Tractor-Trailer Truck Drivers	39	12
1311 Beroepsgroep overig - managing	40	5
0221 Grafisch vormgevers en producton. -Commercial and Industrial Designers	41	4

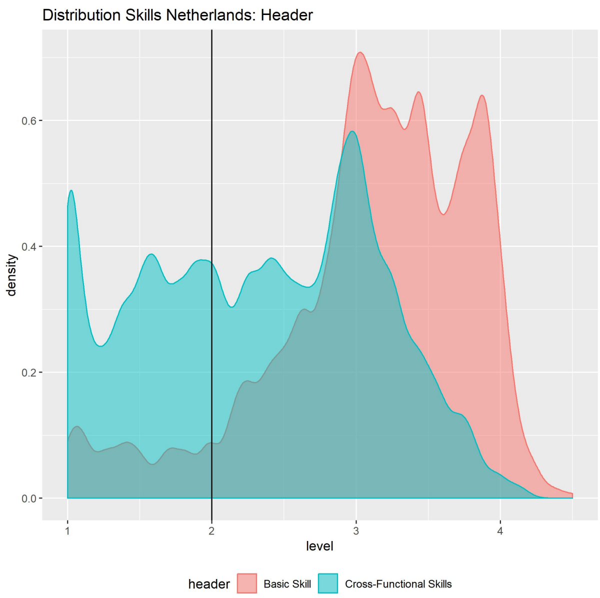
Looking at the clusters of Dutch occupations according to their competencies, we see that the clusters follow more or less the aggregation of the CBS. For example, the CBS groups all managers together with the 2-digit code 05. However, when we look at the cluster analysis of the competencies, we see that these managers do not belong to the same group. When divided into 13 clusters, managers can be found in cluster 5 and cluster 8. This is also true for the 2-digit group 10 ('Zorg en welzijn beroepen' or 'Care and welfare professions') and 03 ('Commerciele beroepen' or 'Commercial professions').

Table B.12: Number of dutch workers per 3-digit occupations

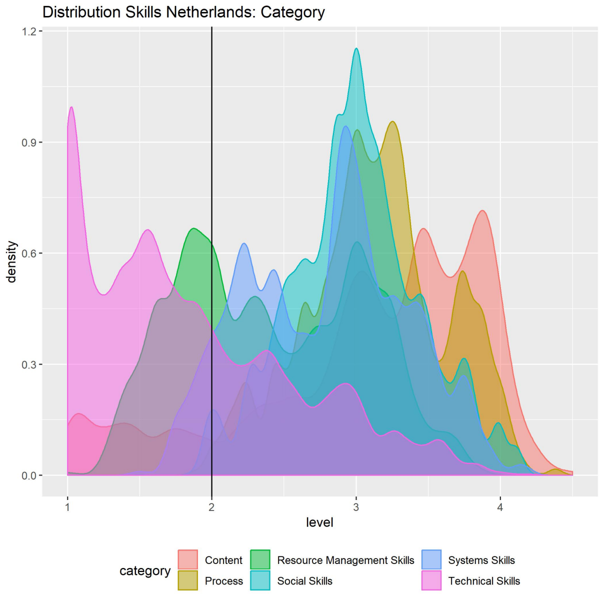
job	Total.x1000	men	women
Totaal	8774	4686	4088
043 Administratief personeel	821	324	497
033 Verkopers	672	239	433
111 Medewerkers persoonlijke dienstve...	480	187	292
041 Specialisten bedrijfsbeheer en ad.	478	277	201
011 Docenten	429	140	289
122 Hulpkrachten transport en logistiek	336	246	90
112 Schoonmakers en keukenhulpen	330	108	222
101 Artsen, therapeuten en gespeciali...	328	90	238
042 Vakspecialisten bedrijfsbeheer en...	322	141	181
081 Specialisten ICT	305	262	43
121 Bestuurders voertuigen en bediene...	302	283	19
073 Bouwarbeiders	257	252	5
103 Vakspecialisten gezondheidszorg	252	49	203
105 Verzorgenden	247	20	227
104 Sociaal werkers, groeps- en woonb...	239	56	184
071 Ingenieurs en onderzoekers wis.	212	171	41
072 Vakspecialisten natuur en techniek	185	163	21
074 Metaalarbeiders, machinemonteurs	183	179	4
031 Adviseurs marketing, public relat.	163	86	77
091 Tuinders, akkerbouwers en veetelers	160	130	29
032 Vertegenwoordigers en inkopers	149	106	43
053 Managers productie en gespecialis...	147	113	33
102 Specialisten op maatschappelijk g...	144	33	110
131 Beroepssegment overig	143	95	48
013 Leidsters kinderopvang en onderwi.	142	14	128
063 Beveiligingswerkers	140	108	32
021 Auteurs en kunstenaars	138	74	64
075 Voedselverwerkende beroepen en ov...	128	86	42
052 Managers op administratief en com...	123	88	35
077 Productiemachinebedieners en asse...	123	101	23
061 Overheidsambtenaren en -bestuurders	89	44	45
022 Vakspecialisten op artistiek en c.	85	48	37
076 Elektriciens en elektronicamonteurs	85	84	2
054 Managers horeca, detailhandel en.	75	53	22
051 Algemeen directeuren	74	60	14
062 Juristen	71	35	36
078 Hulpkrachten bouw en industrie	70	39	31
082 Vakspecialisten ICT	64	54	10
012 Sportinstructeurs	47	24	23
092 Hulpkrachten landbouw	28	18	10
055 Managers z.n.d.	10	7	3

Table B.14: Jobs with importance of skill Installation

Job	Installation	Total men and women x1000	men	women
0761 Elektriciens en elektronicamonteurs	2.82	85.00	84.00	1.00
0734 Loodgieters en pijpfitters	2.42	38.00	38.00	0.00
0744 Machinemonteurs	2.27	47.00	47.00	0.00
0743 Automonteurs	2.09	63.00	62.00	1.00
0731 Bouwarbeiders ruwbouw	1.90	62.00	60.00	2.00

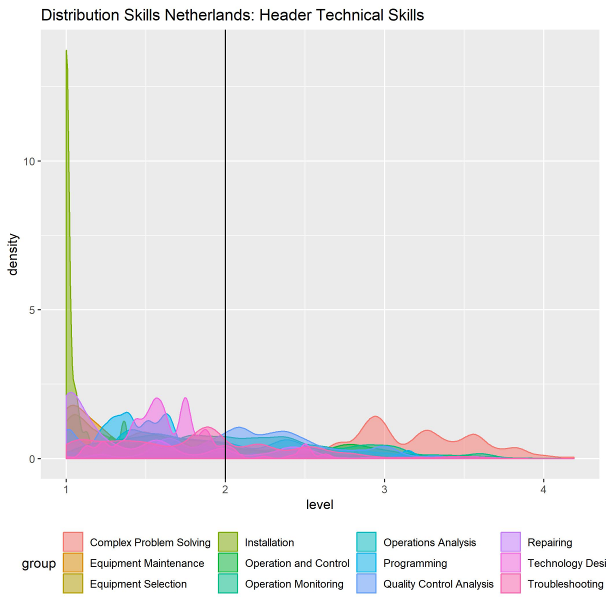


(a) Header

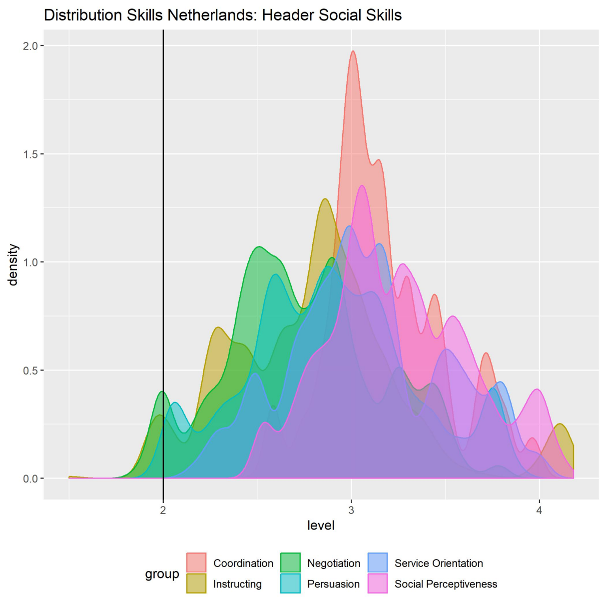


(b) Category

Figure B.4: Density plots of skills header and categories

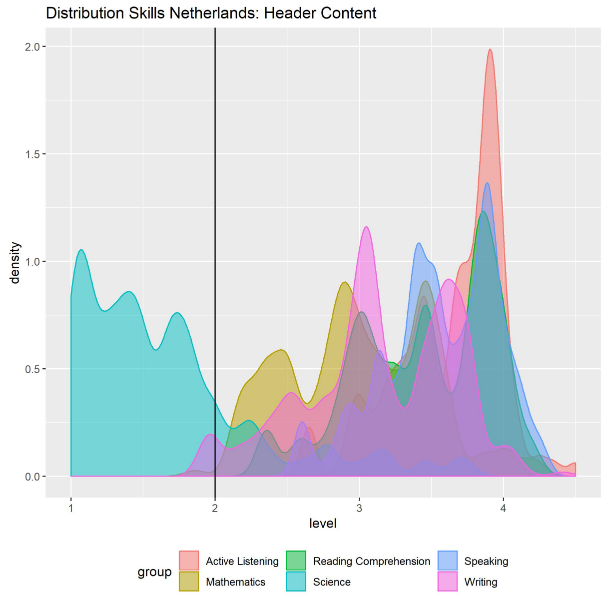


(a) Technical Skills

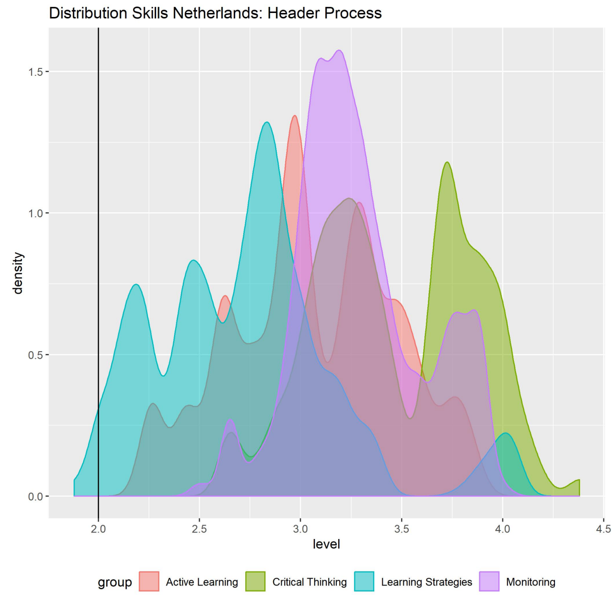


(b) Social Skills

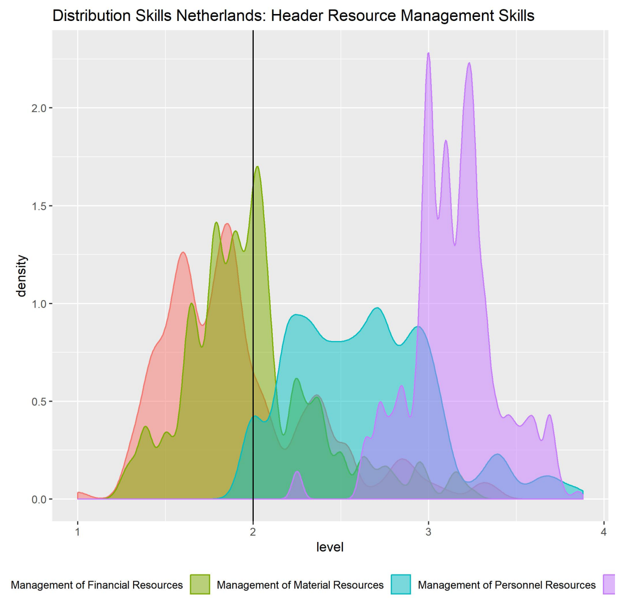
Figure B.5: Density plots of social and technical skills



(a) Content Skills



(b) Process Skills



(c) Resource Management Skills

Figure B.6: Density plots of content, process and resource management skills

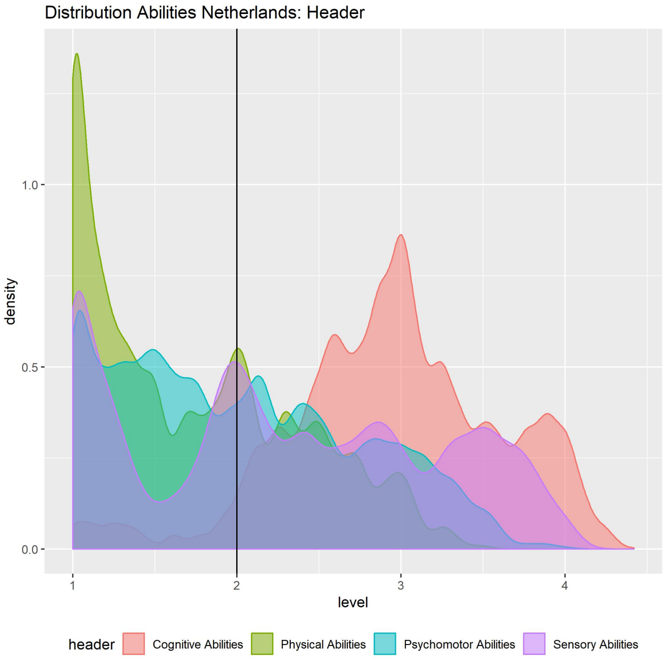
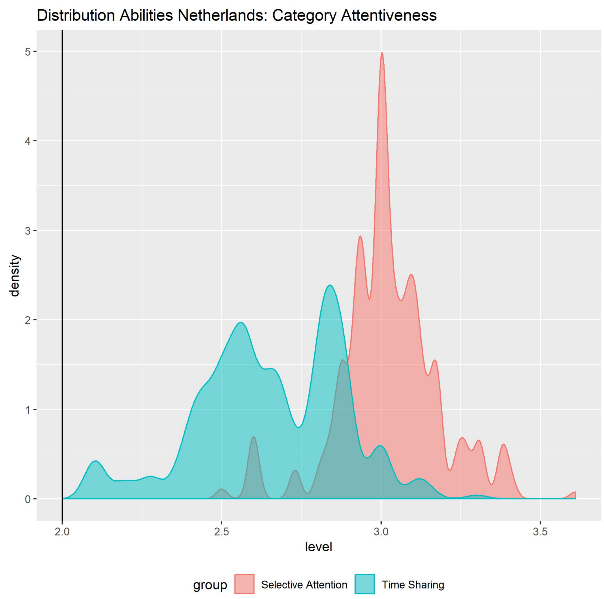


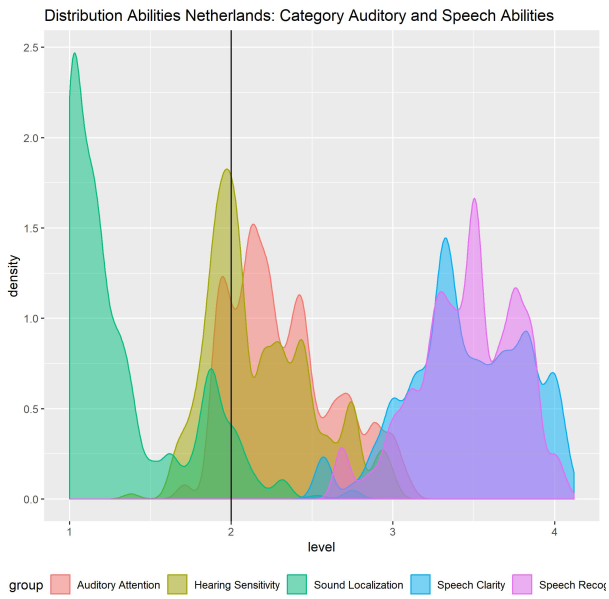
Figure B.7: Density plot of abilities by header

Table B.16: Occupations whith importance of: Explosive Strength,Night vision, Glare sensitivity, Periheral Vision and Sound Localization

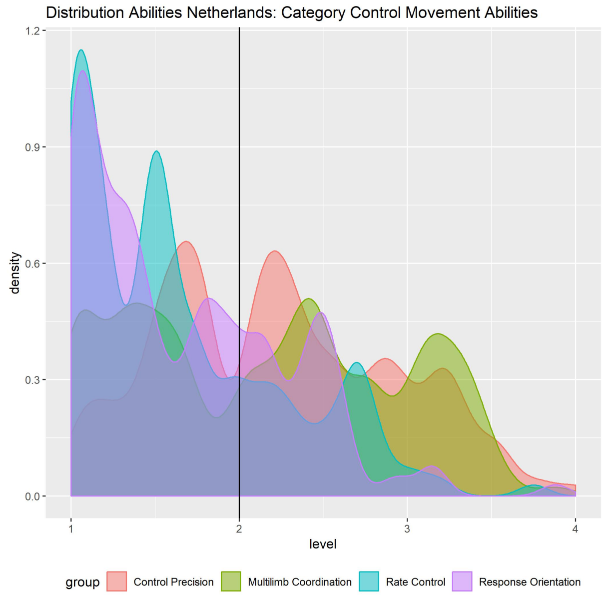
job	Explosive Strength	Night Vision	Peripheral Vision	Glare Sensitivity	Sound Localization	Total x 1000	men x 1000	Women x 1000
0121 Beroepsgroep sportinstructeurs	2.25	1.22	1.73	1.43	1.30	47.00	24.00	23.00
0632 Politie en brandweer	2.44	2.31	2.41	2.58	2.31	41.00	29.00	12.00
0633 Beveiligingspersoneel	2.09	1.66	1.85	1.85	1.75	62.00	47.00	15.00
0722 Productieleiders industrie en bouw	1.25	1.94	1.94	2.06	1.94	55.00	52.00	3.00
0723 Procesoperators	1.55	1.68	1.77	2.04	1.86	26.00	24.00	2.00
0731 Bouwarbeiders ruwbouw	1.67	1.86	1.97	2.20	1.89	63.00	62.00	1.00
0732 Timmerlieden	1.82	1.88	1.94	2.06	1.88	85.00	84.00	1.00
0734 Loodgieters en pijpfitters	1.59	1.92	2.00	2.00	1.88	38.00	38.00	1.00
0743 Automonteurs	1.59	1.77	1.85	2.03	2.17	58.00	57.00	1.00
0744 Machinemonteurs	1.52	1.88	1.97	2.25	2.11	47.00	45.00	1.00
0761 Elektriciens en elektronicamonteurs	1.45	1.84	1.89	2.04	2.01	85.00	84.00	2.00
0781 Hulpkrachten bouw en industrie	1.67	1.78	1.94	2.06	1.83	70.00	39.00	31.00
0911 Land- en bosbouwers	1.49	1.94	2.13	2.01	2.03	29.00	25.00	4.00
0912 Hoveniers, tuinders en kwekers	1.44	1.94	2.06	2.06	2.00	74.00	64.00	9.00
0913 Veetelers	1.48	1.90	2.04	2.12	1.98	57.00	41.00	16.00
0921 Hulpkrachten landbouw	1.65	1.85	2.08	2.13	1.85	28.00	18.00	10.00
1211 Dekofficieren en piloten	1.34	2.21	2.30	2.44	2.07	30.00	27.00	3.00
1212 Chauffeurs auto's, taxi's en bes.	1.35	2.22	2.37	2.37	2.08	67.00	57.00	10.00
1213 Buschauffeurs en trambestuurders	1.40	2.62	2.88	2.63	2.53	23.00	20.00	3.00
1215 Bedieners mobiele machines	1.30	2.19	2.51	2.58	2.32	74.00	72.00	2.00
1221 Laders, lossers en vakkenvullers	1.48	1.88	2.02	2.15	1.90	253.00	180.00	73.00
1311 Beroepsgroep overig - operational	1.64	1.86	2.04	2.08	1.88	143.00	95.00	48.00
0612 Overheidsambtenaren- operational	1.93	1.98	2.02	2.09	2.04	67.00	30.00	37.00
1214 Vrachtwagenchauffeurs - First-Line Supervisors	1.00	1.75	2.00	1.75	2.00	108.00	106.00	3.00
1214 Vrachtwagenchauffeurs -Heavy [..] Truck Drivers	1.50	3.00	3.00	2.88	2.75	108.00	106.00	3.00



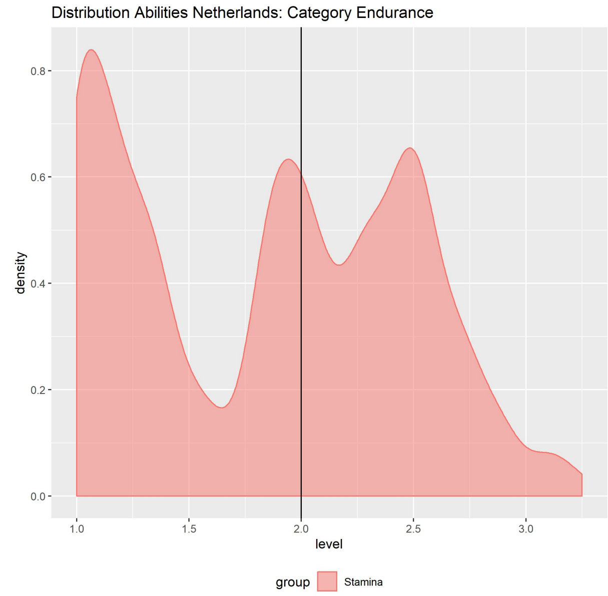
(a) Attentiveness



(b) Auditory and Speech

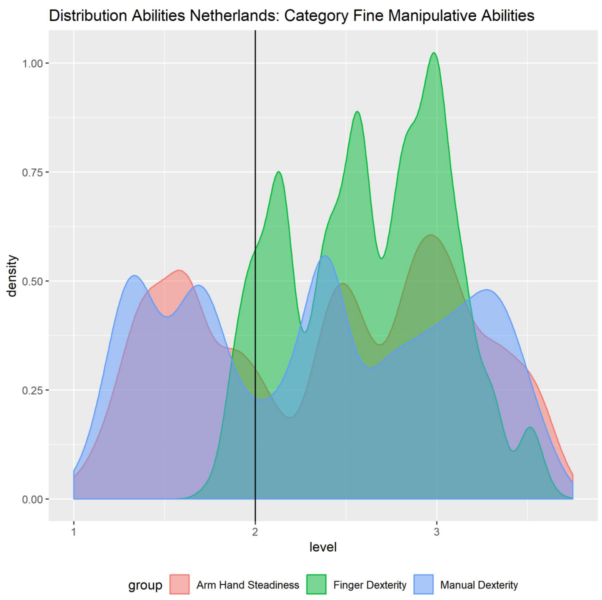


(c) Control movement abilities

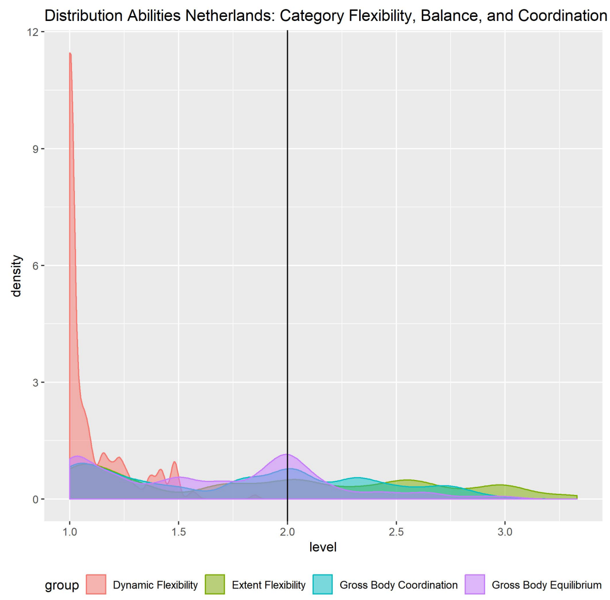


(d) Endurance

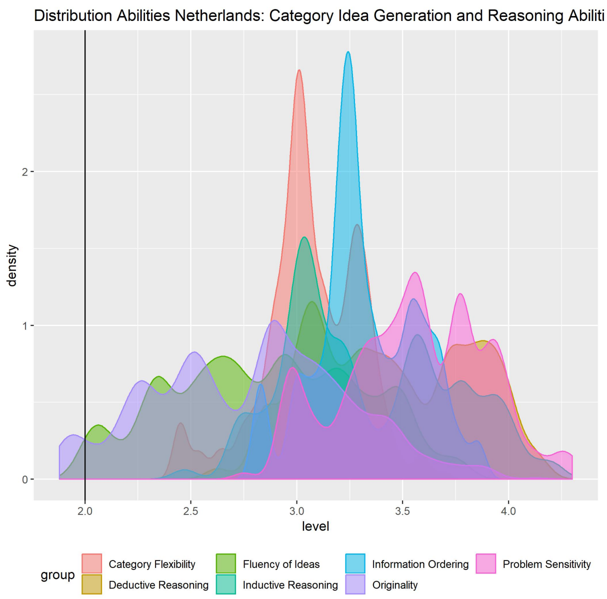
Figure B.8: Density plots of abilities from category: Attentiveness; Auditory and Speech; Control movement abilities; and Endurance



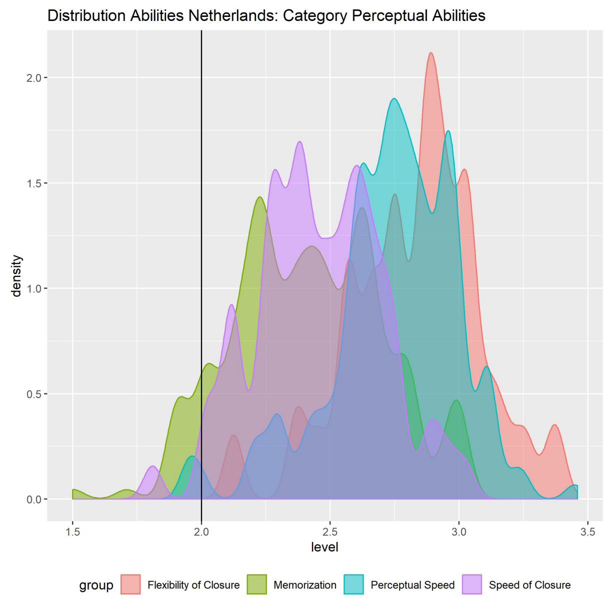
(a) Fine Manipulative abilities



(b) Flexibility and balance abilities

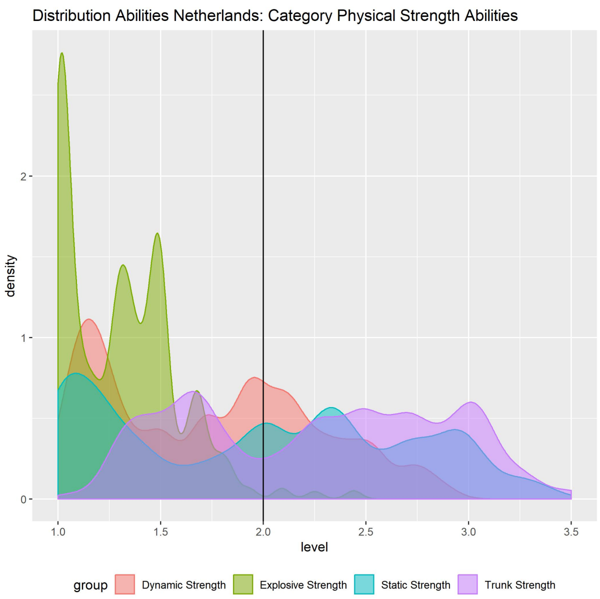


(c) Idea generation abilities

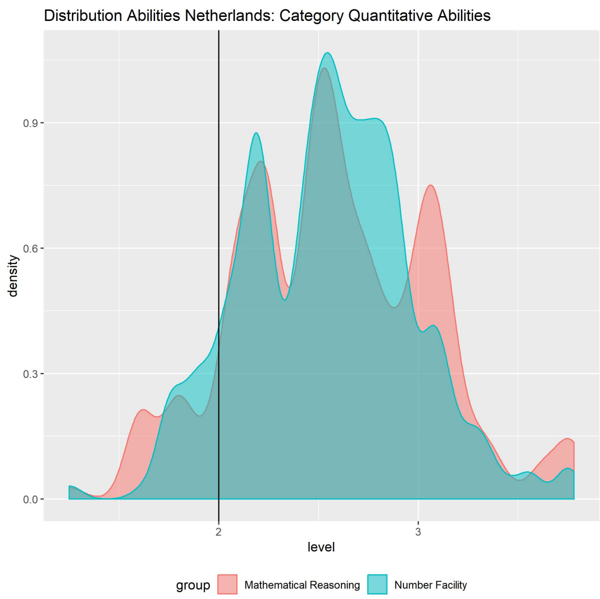


(d) perceptual abilities

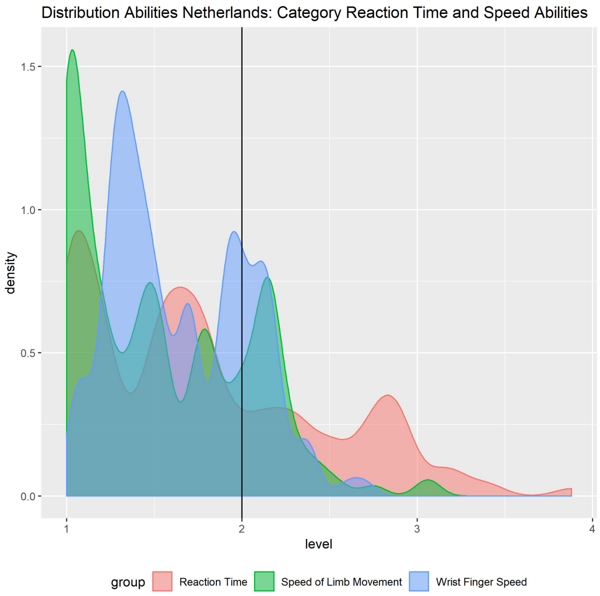
Figure B.9: Density plots of abilities from category: Fine Manipulative abilities; Flexibility and balance abilities; Idea generation abilities; and perceptual abilities



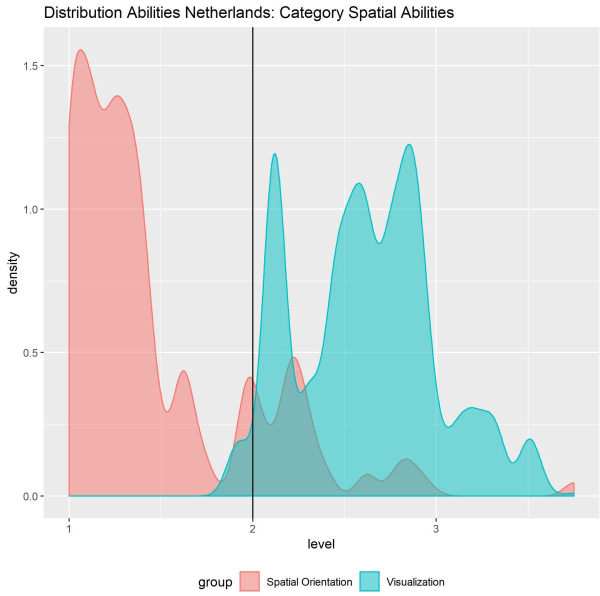
(a) Physical Strength Abilities



(b) Quantitative Abilities

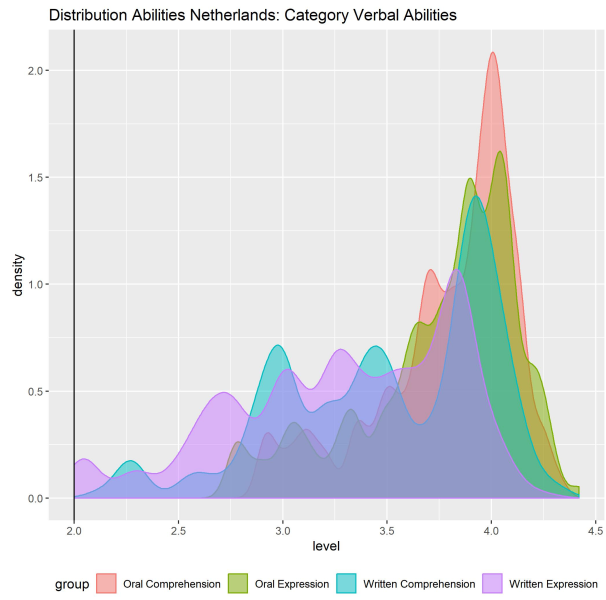


(c) Reaction time and speed abilities

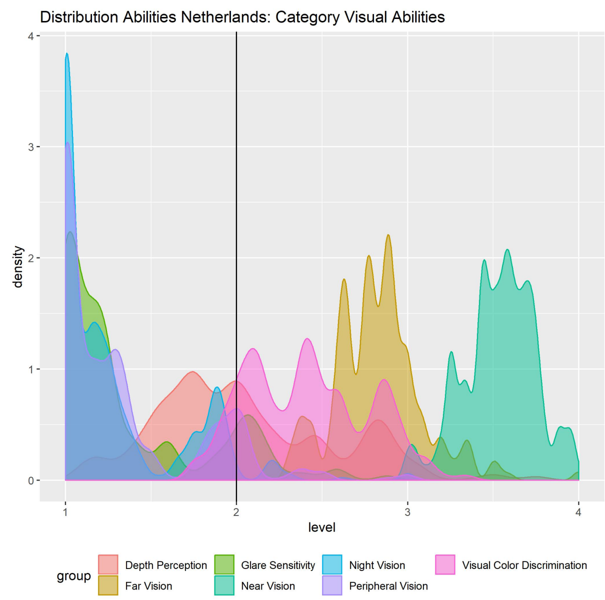


(d) Spatial abilities

Figure B.10: Density plots of abilities from category: Physical strength abilities; Quantitative Abilities; Reaction time and speed abilities; and spatial abilities



(a) Verbal Abilities



(b) Visual Abilities

Figure B.11: Density plots of abilities from category verbal abilities and visual abilities

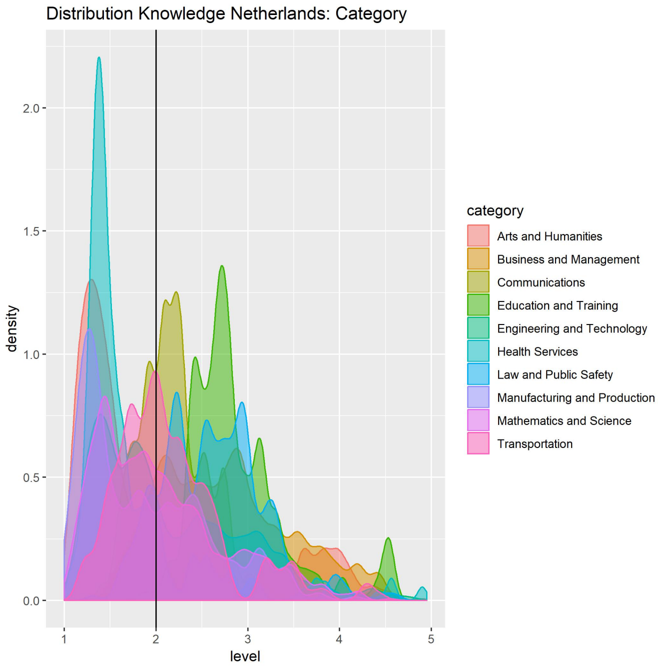
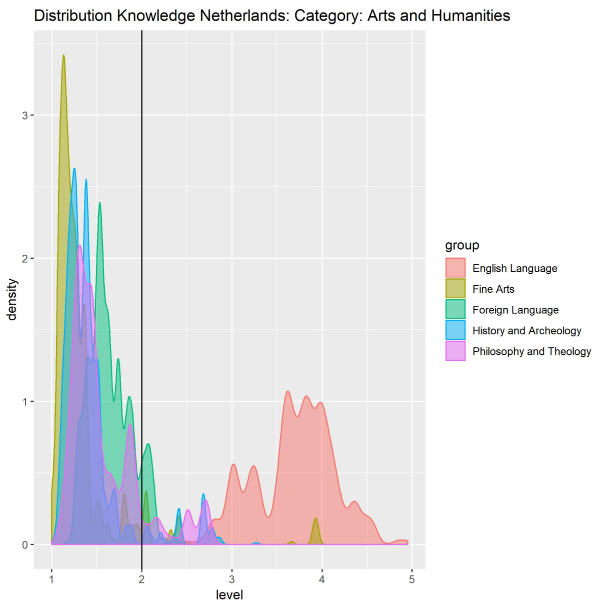
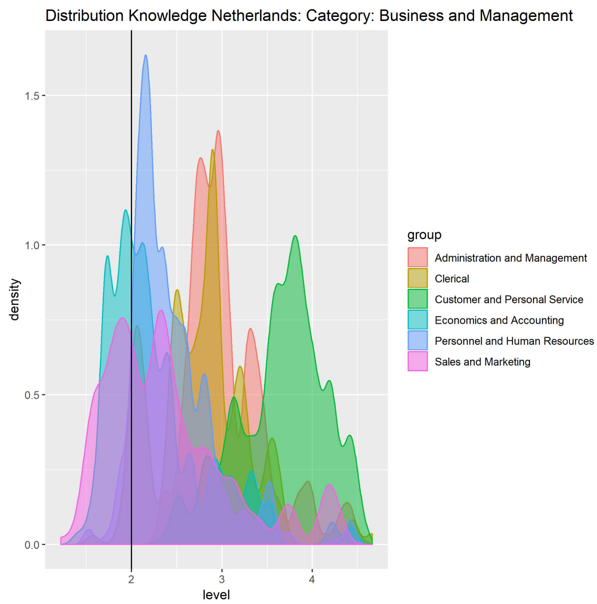


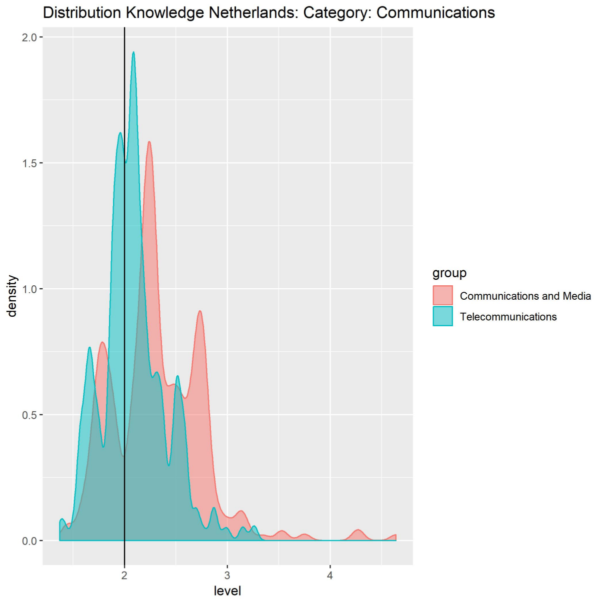
Figure B.12: Density plot of knowledge by category



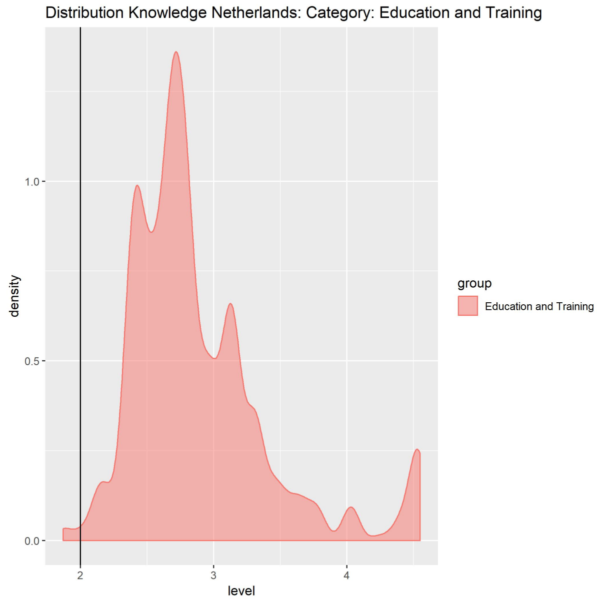
(a) Arts and Humanities



(b) Business and Management

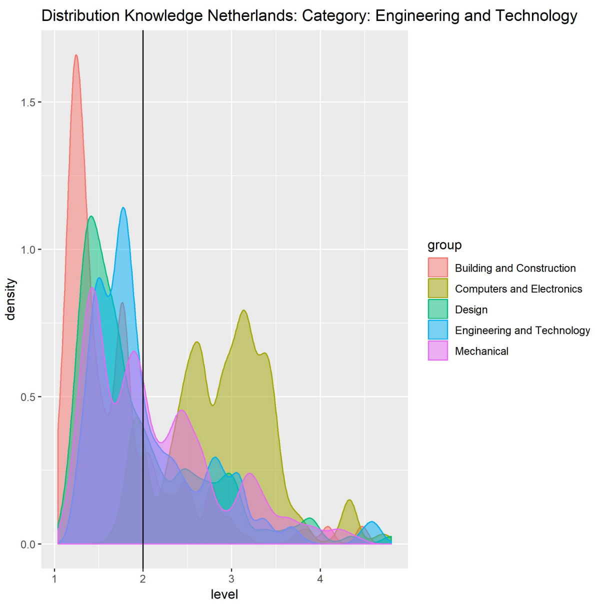


(c) Communications

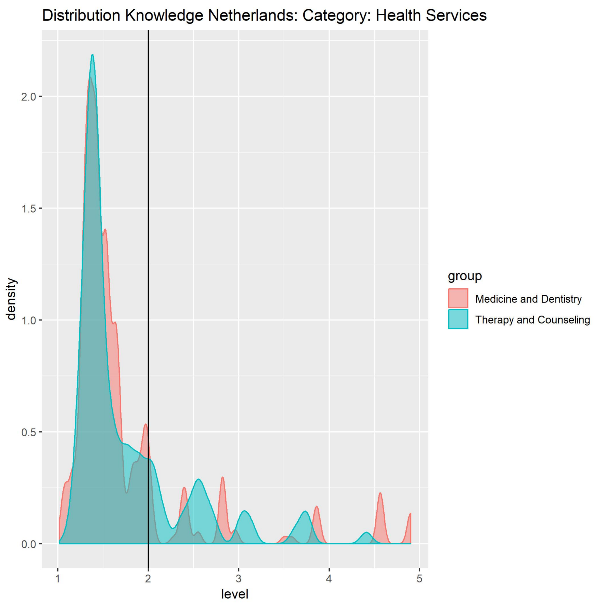


(d) Education and Training

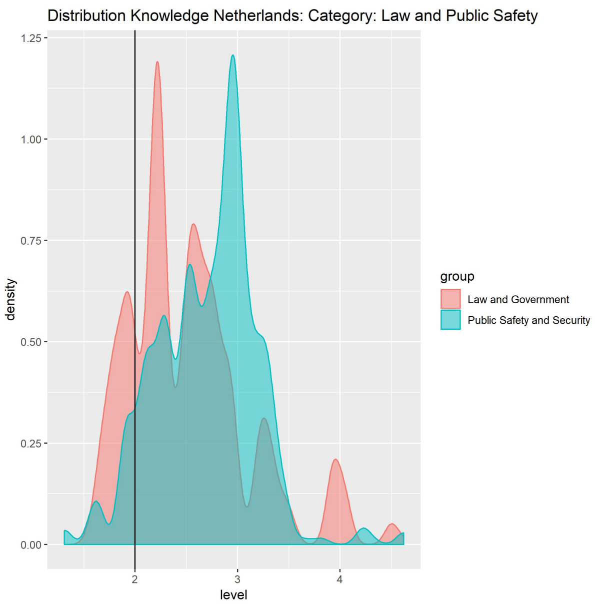
Figure B.13: Density plots of knowledge from category: Arts and Humanities; Business and Management; Communications; and Education and Training



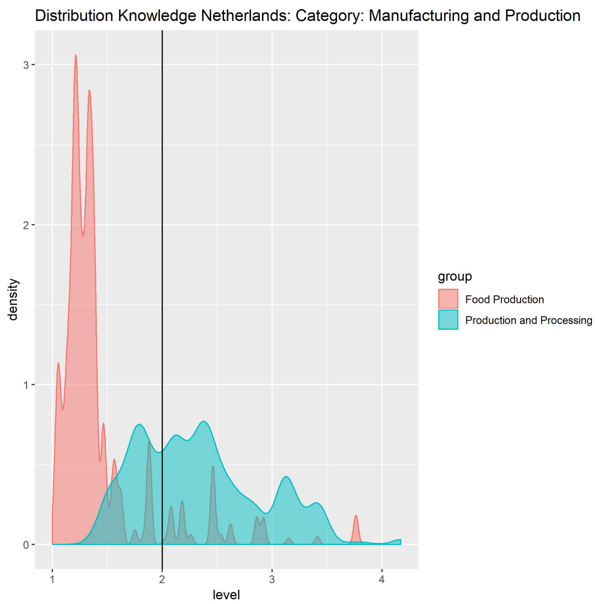
(a) Engineering and Technology



(b) Health Services

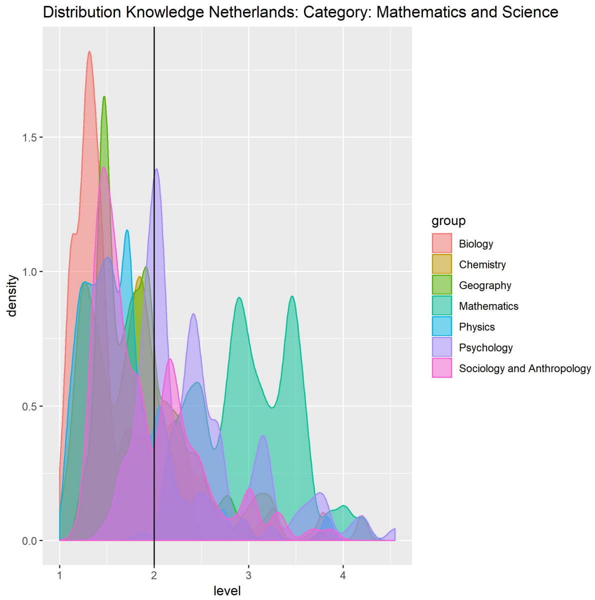


(c) Law and Public Safety

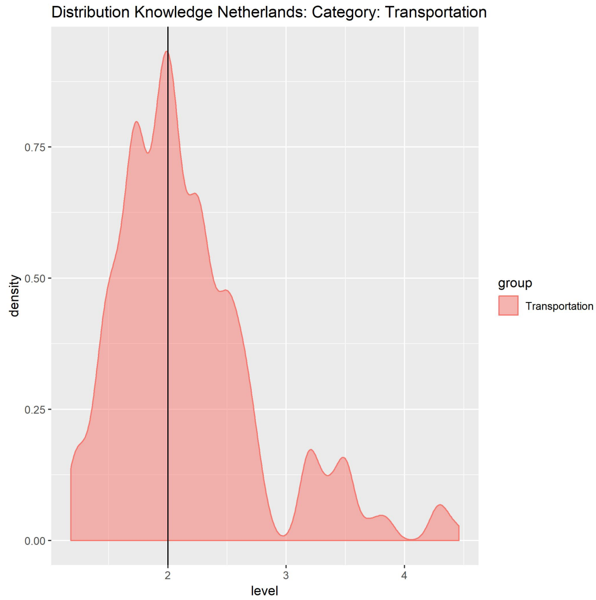


(d) Manufacturing and Production

Figure B.14: Density plots of knowledge from category: Engineering and Technology; Health Services; Law and Public Safety; and Manufacturing and Production



(a) Mathematics and Science



(b) Transportation

Figure B.15: Density plots of knowledge from category: Mathematics and Science; and Transportation

Appendix C

The risk of automation

C.1 Classification method Frey & Osborne

This section provides some more detail about the method used by FO and which is applied in this analysis on Dutch data. Since this analysis is fully based on the methodology of FO, for the validation of the methodology, this analysis relies on the results of the FO analysis. This means we did not test the best classification method for our analysis but used the GP approach –with a squared exponential covariance matrix– as validated by FO. Nevertheless, this section will discuss the approach and validation as done by FO in more detail. It starts with a brief discussion of the Gaussian process used for machine-learning, followed by the validation of the different classification methods done by FO.

C.1.1 Gaussian Process for machine-learning

In order to determine the automation probabilities of US jobs, FO applied a *supervised machine learning* technique: which is the problem of learning input-output mappings from empirical data (the training data-set) (Rasmussen, 2004). "Depending on the characteristics of the output, this problem is known as either regression, for continuous outputs, or classification, when outputs are discrete" (p.2). In their analysis, FO used the classification technique. A simple example of the classification technique in supervised machine-learning is given by Rasmussen (2004):

> A well-known example is the classification of images of handwritten digits. digit classification. The training set consists of small digitized images, together with a classification from 0, . . . , 9, normally provided by a human. The goal is to learn a mapping from image to classification label, which can then be used on new, unseen images. Supervised learning is an attractive way to attempt to tackle this problem since it is not easy to specify accurately the characteristics of, say, the handwritten digit 4. (p.2)

In general, the input is denoted as x[1] and the output (target) is denoted as y. In the FO model, the target y is discrete –which is always the case in a classification problem– and represents the occupational labels of automatability: $y \in \{0,1\}^{70}$[2]. In the FO model, the training data-set, D, is denoted as $D = (X, y)$ where $X \in \mathbb{R}^{70\times 9}$ is a matrix of variables and $y \in \{0,1\}^{70}$ gives the associated labels. Given this training data-set, the goal of FO was to make predictions of new inputs x_*. For this, they used a probabilistic classification algorithm which exploits the patterns existent in the training-data, D, to return the probability $P(y_* = 1|(x_*, X, y))$ of a new, unlabelled, *test-datum* which features x_*, having class label $y_* = 1$

[1] The input y is usually represented as a vector since there are often more input variables. In our application of the FO model, the input is denoted as a vector $x\mathbb{R}^9$ where $X \in \mathbb{R}^{19\times 9}$ is the matrix od O*Net variables for the subset of 19 Dutch training occupations.

[2] In our application of the FO analysis, we have 19 occupational labels of automatibility. In our analysis, the output y is this denoted as: $y \in \{0,1\}^{19}$

(Frey & Osborne, 2013).

For this, FO, needed to move from a finite data-set, D, to a function f that could make predictions for all possible input values x_*. In order to do this, FO had to make assumptions about the characteristics of the underlying function, as otherwise, any function which is consistent with the training data would be equally valid (Rasmussen, 2004). In the literature, a wide variety of methods are proposed in order to overcome this supervised machine-learning problem. In their article, Rasmussen (2004) explain the two most common approaches:

> The first [approach] is to restrict the class of functions that we consider, for example by only considering linear functions of the input. The second approach is to give a prior probability to every possible function, where higher probabilities are given to functions that we consider to be more likely. The first approach has an obvious problem in that we have to decide upon the richness of the class of functions considered; if we are using a model based on a certain class of functions and the target function is not well modelled by this class, then the predictions will be poor. [..] The second approach appears to have a serious problem, in that surely there are an uncountably infinite set of possible functions, and how are we to compute with this set in finite time? This is where the *Gaussian process* comes to our rescue. (p.3)

In their approach, FO used *Gaussian Process (GP) classifiers*. These classifiers model the underlying latent function, f with a Gaussian process: a non-parametric probability distribution over functions (Frey & Osborne, 2013). In other words, one can think of a GP as a generalization of the Gaussian probability distribution. However, rather than describing random variables which are scalars or vectors –what probability distributions do–, a stochastic process governs the characteristics and properties of *functions*. We can, therefore, think of a function as a "very long vector, each entry in the vector specifying the function value $f(x)$ at a particular input x"(Rasmussen, 2004, p. 4).

C.1.2 Validation of various classifiers

To summarize, a GP is defined as a distribution over the functions $f : X \rightarrow \mathbb{R}$ where the distribution over the possible function values is multivariate Gaussian. According to Frey and Osborne (2013), for a function $f(x)$, the prior distributing over its values $\underline{f}$ on a subset $\underline{x} \subset X$ are completely specified by the covariance matrix K. This covariance matrix is generated by the function: $\kappa : X \times X \mapsto \mathbb{R}$. Since a GP model is expressed by the covariance function, the model depends thus on the choice of κ. In their analysis, FO have considered three choices of κ:

- The exponential quadratic (squared exponential)
- Rational quadratic
- Linear covariances[3]

For their analysis, FO validated these three functions of κ by randomly selecting a reduced training set of half the available data D and a test-set which is the remaining data. On this test-set, they evaluated how closely the classification' algorithms matched the hand labels according to the metrics: UAC[4] and the log-likelihood. Table C.1 shows the performance of the three classification models that FO tested. According to Frey and Osborne (2013), the model with the best performance is the GP process model with an exponential quadratic covariance function. They used this model for their research. For our analysis, we rely on this validation and also perform the analysis with a GP exponential quadratic model.

[3] This choice is equivalent to logistic regression with a Gaussian prior taken on the weights $\underline{w}$

[4] Metric to assess the performance of a classifier where Area Under the Curve (AUC) is equal to one for a perfect classifier and one half for a completely random classifier

Table C.1: Performance of various classifiers; best performance in bold.

Classifier Model	AUC	Log-likelihood
Exponential quadratic	**0.894**	**-163.3**
Rational quadratic	0.393	-163.7
Linear (logit-regression)	0.827	-205.9

C.2 Application FO model on Dutch data

This appendix presents the results of the application of the FO model on Dutch data.

C.2.1 Mapping Dutch training-data

In order to formulate a training-set for our classification model, we used the training-set as provided by FO. This training set, however, needed to be mapped to Dutch occupations. Table C.3 represent the mapping of the FO training data to the Dutch data-set. It shows how the O*NET-SOC code used by FO is mapped to its corresponding BRC-code through the ISCO-2008 code. This table also shows for each BRC-code the ratio of labelled O*Net-SOC occupations to unlabelled O*Net-SOC occupations. For example, within BRC-code 1011, one-third of all the underlying O*NET-SOC codes are labelled by FO.

For our training set, define a BRC-code as labelled when the ratio is equal or higher than 0.5. Note, however, that this ratio is set arbitrary and that the number of labelled BRC-codes is highly sensitive to this ratio (see Table C.2)

Table C.2: Sensitivity of number of labels to ratio threshold

ratio	n labelled BRC-codes
0.33	26
0.5	19
0.66	17

Table C.3: Mapping training dataset

Training data set codification	Frey and Osborne	ISCO-08 equivalent occupation	BRC equivalent occupation	Ratio	Use label/ ISCO
0	Physicians and Surgeons	Generalist medical practitioners; Specialist medical practitioners	1011 Artsen	1/3	N
0	Dentists, General	Dentist	1011 Artsen	1/3	N
0	Social and Community Service Managers	Social welfare managers	0534 Managers zorginstellingen	1/5	N
0	Preschool Teachers, Except Special Education	Early childhood educators	0114 Leerkrachten basisonderwijs	1/5	N
0	Clergy	Religious professionals; Religious associate professionals	1022 Psychologen en sociologen	2/3	Y
0	Registered Nurses	Nursing and midwifery professionals; Nursing associate professionals	1033 Verpleegkundigen (mbo)	1/3	N
0	Marriage and Family Therapists	Psychologists	1022 Psychologen en sociologen	2/3	Y

Table C.3: Mapping training dataset

Training data set codification	Frey and Osborne	ISCO-08 equivalent occupation	BRC equivalent occupation	Ratio label/ ISCO	Use
0	Chief Executives	Directors and chief executives	0511 Algemeen directeuren	1	Y
0	Education Administrators Preschool and Childcare Centre/Program	Education managers	0535 Managers onderwijs	1	Y
0	Civil Engineers	Civil Engineers	0712 Ingenieurs (geen elektrotechniek)	1/2	Y
0	Fashion Designers	Product and garment designers	0221 Grafisch vormgevers en producton...	1	Y
0	Substance Abuse and Behavioral Disorder Counselors	Social work and counselling professionals	1021 Maatschappelijk werkers	1/2	Y
0	Lawyers	Lawyers	0621 Juristen	1/2	Y
0	Meeting, Convention, and Event Planners	Conference and event planners	0422 Zakelijke dienstverleners	1/6	N
0	Landscape Architects	Landscape Architects	0714 Architecten	2/5	N
0	Healthcare Practicioners and Technical Workers, All	Health professionals, other; Health associate	1013 Fysiotherapeuten	1	Y
1	Bus Drivers, Transit and Intercity	Bus and tram drivers	1213 Buschauffeurs en trambestuurders	1	Y
1	Light Truck or Delivery Services Drivers	Heavy-truck and lorry drivers	1214 Vrachtwagenchauffeurs - Heavy and Tractor- Trailer Truck Drivers	1/4	N
0	Maids and Housekeeping Cleaners	Domestic housekeepers; Domestic helpers and clearners	1115 Conciërges en teamleiders schoon.	1/4	N
1	Civil Engineering Technicians	Civil engineering technicians	0721 Technici bouwkunde en natuur	1/6	N
1	Dishwashers	Kitchen helpers	1122 Keukenhulpen	2/3	Y
0	Hunters and Trappers	Hunters and Trappers	1311 Beroepsgroep overig	1/15	N
1	Cooks, Fast Food	Cooks	1122 Keukenhulpen	2/3	Y
1	Electrical and Electronics Drafters	Electrical engineering technicians;	0721 Technici bouwkunde en natuur	1/6	N
1	Sheet Metal Workers	Sheet metal workers	0742 Lassers en plaatwerkers	1/2	Y
1	Meter Readers, Utilities	s Meter readers	1222 Vuilnisophalers en dagbladenbezo.	1/5	N
1	Computer-Controlled Machine Tool Operators, Metal and Plastic	Stationary plant and machine operators, other	0771 Productiemachinebedieners	2/34	N
1	Parking Lot Attendants				N
1	Medical Transcriptionists	Medical assistants; Medical secretaries	1034 Medisch praktijkassistenten	1/3	N
1	Technical Writers				N
1	Sewing Machine Operators	Sewing Machine Operators	0771 Productiemachinebedieners	2/34	N
1	Taxi Drivers and Chauffeurs	Car, taxi and van drivers	1212 Chauffeurs auto's, taxi's en bes.	1/3	N
0	Compliance Officers	Process control technicians, other	0723 Procesoperators	1/7	N

Table C.3: Mapping training dataset

Training data set codification	Frey and Osborne	ISCO-08 equivalent occupation	BRC equivalent occupation	Ratio label/ ISCO	Use
0	Childcare Workers	Child-care workers	0131 Leidsters kinderopvang en onderw...	1/3	N
0	Chefs and Head Cooks	Chefs	1112 Koks	1/2	Y
0	Electrical Engineers	Electrical engineers	0713 Elektrotechnisch ingenieurs	1/4	N
0	Physicists	Physicists and astronomers	0711 Biologen en natuurwetenschappers	1/10	N
0	Hairdressers, Hairstylists, and Cosmetologists	Hairdressers; Beauticians and related workers	1114 Kappers en schoonheidsspecialisten	1	Y
0	Concierges	Hotel receptionists	0433 Receptionisten en telefonisten	2/9	N
0	Athletes and Sports Competitors	Athletes, sportspersons and related associate professionals	0121 Beroepsgroep sportinstructeurs	1/4	N
0	Zoologists and Wildlife Biologists	Biologists	0711 Biologen en natuurwetenschappers	1/10	N
0	Plimbers, Pipefitters, and Steamfitters	Plumbers and pipe fitters	0734 Loodgieters en pijpfitters	1	Y
0	Flight Attendants	Travel attendants and travel stewards	1111 Reisbegeleiders	1/9	N
1	Surveyors	Cartographers and surveyors	0714 Architecten	2/5	N
0	Judges, Magistrate Judges, and Magistrates	Judges	0621 Juristen	1/2	Y
1	Judicial Law Clerks	Legal secretaries Legal and related associate professionals	0423 Directiesecretaresses	1/4	N
0	Economists	Economists	0412 Financieel specialisten en economenen	1/2	N
1	Cost Estimators	Valuers and loss assessors	0421 Boekhouders	1	Y
0	Transportation, Storage, and Distribution Managers	Supply, distribution and related managers	0532 Managers logistiek	1	Y
1	Market research Analyst and Marketing Specialists	Advertising and marketing professionals	0311 Adviseurs marketing, public rela...	1/5	N
1	Motorboat Operators S	Ships' deck officers and pilots	1211 Dekofficieren en piloten	1/6	N
1	Human Resources Assistants, Except Payroll and Timekeeping	Personnel clerks	0431 Administratief medewerkers	1/8	N
1	Tax Examiners and Collectors, and Revenue Agents	Government tax and excise officials	0612 Overheidsambtenaren	1/6	N
1	Industrial Truck and Tractor Operators	Mobile farm and forestry plant operators	1215 Bedieners mobiele machines	1/8	N
1	Accountants and Auditors	Accountants	0411 Accountants	1	Y
0	Waiters and Waitresses	Waiters	1113 Kelners en barpersoneel	1/3	N

Table C.3: Mapping training dataset

Training data set codification	Frey and Osborne	ISCO-08 equivalent occupation	BRC equivalent occupation	Ratio label/ ISCO	Use
1	Couriers and Messengers	Messengers, Package Deliverer and Luggage Porters	1222 Vuilnisophalers en dagbladenbezo.	1/5	N
1	Paralegals and Legal Assistants	-	-	-	N
1	Electrical and Electronic Equipment Assemblers	Electrical and electronic equipment assemblers	0772 Assemblagemedewerkers	1/5	N
1	Switchboard Operators, Including Answering Servic	Telephone switchboard operators	0433 Receptionisten en telefonisten	2/9	N
1	Gaming Dealers				N
1	Farm Labour Contractors				N
1	Cashiers	Cashiers and ticket clerks	0333 Kassamedewerkers	1	Y
1	File Clerks	Filing and copying clerks	0431 Administratief medewerkers	1/8	N
1	Credit Authorizers, Checkers, and Clerks	Credit and loans officers	0421 Boekhouders	1	Y
1	Claims Adjusters, Examiners, and Investigators		0421 Boekhouders	1	Y
1	Credit Analysts	Financial analyst	0412 Financieel specialisten en economen	1/2	N
1	Loan Officers	Credit and loan officers	0421 Boekhouders	1	Y
1	Data Entry Keyers	Data entry clerks	0432 Secretaresses	1/4	N
1	Insurance Underwriters	Insurance representatives	0321 Vertegenwoordigers en inkopers	1/5	N

C.2.2 Probability of automation of Dutch jobs

After constructing our training set, D, we used a GP classification approach in order to estimate the automation probability for all the 116 BRC-codes occupations. Table C.4 shows the results of this analysis. For each BRC-code, the variable *p auto* represents the probability of automation for that job. The variable *Risk* shows whether this probability is classified as 'high risk', 'medium risk' or 'low risk' with a cut-off at 0.3 and 0.7.

Table C.4: Probability of becoming automated of Dutch BRC-code occupations

Job	Total x1000	Women x1000	Men x1000	Pauto	Risk
0434 Boekhoudkundig medewerkers	138	100	38	0,816	High
0431 Administratief medewerkers	285	193	92	0,814	High
0333 Kassamedewerkers	92	83	9	0,795	High
1122 Keukenhulpen	75	31	44	0,779	High
0423 Directiesecretaresses	110	92	19	0,759	High
0751 Slagers	19	5	14	0,754	High
0432 Secretaresses	53	50	3	0,751	High
1121 Schoonmakers	255	191	64	0,722	High
1115 ConciÃ«rges en teamleiders schoon.	43	16	28	0,704	High
0921 Hulpkrachten landbouw	28	10	18	0,698	Medium
1222 Vuilnisophalers en dagbladenbezo.	83	17	66	0,694	Medium
1041 Sociaal werkers groeps- en woon	239	184	56	0,691	Medium
0334 Callcentermedewerkers outbound e...	123	75	48	0,69	Medium

Table C.4: Probability of becoming automated of Dutch BRC-code occupations

Job	Total x1000	Women x1000	Men x1000	Pauto	Risk
0781 Hulpkrachten bouw en industrie	70	31	39	0,65	Medium
1311 Beroepsgroep overig - operational	143	48	95	0,649	Medium
0411 Accountants	99	30	68	0,648	Medium
0771 Productiemachinebedieners	81	11	70	0,636	Medium
1212 Chauffeurs auto's	67	10	57	0,626	Medium
1213 Buschauffeurs en trambestuurders	23	3	20	0,616	Medium
0742 Lassers en plaatwerkers	37	1	36	0,616	Medium
1215 Bedieners mobiele machines	74	2	72	0,614	Medium
0772 Assemblagemedewerkers	42	11	31	0,611	Medium
0741 Metaalbewerkers en constructiewe.	42	1	41	0,609	Medium
0412 Financieel specialisten en economen	82	25	57	0,595	Medium
0435 Transportplanners en logistiek m.	183	39	143	0,586	Medium
0612 Overheidsambtenaren- office	67	37	30	0,584	Medium
0415 Specialisten personeels- en loop.	98	60	38	0,583	Medium
1032 Apothekersassistenten	25	23	1	0,576	Medium
0735 Schilders en metaalspuiters	33	1	32	0,575	Medium
0633 Beveiligingspersoneel	62	15	47	0,564	Medium
0911 Land- en bosbouwers	29	4	25	0,549	Medium
1221 Laders lossers en vakkenvullers	253	73	180	0,548	Medium
0421 Boekhouders	112	49	64	0,544	Medium
0754 Meubelmakers kleermakers en sto.	40	10	29	0,543	Medium
0752 Bakkers	20	9	11	0,538	Medium
1214 Vrachtwagenchauffeurs -Heavy and Tractor-Trailer Truck Drivers	108	3	106	0,52	Medium
0753 Productcontroleurs	20	8	12	0,517	Medium
0812 Databank- en netwerkspecialisten	70	12	58	0,501	Medium
0912 Hoveniers tuinders en kwekers	74	9	64	0,5	Medium
0723 Procesoperators	26	2	24	0,497	Medium
0811 Software- en applicatieontwikkel.	235	31	204	0,478	Medium
0821 Gebruikersondersteuning ICT	45	8	37	0,468	Medium
0111 Docenten hoger onderwijs en hoog...	56	30	27	0,448	Medium
0913 Veetelers	57	16	41	0,441	Medium
0755 Medewerkers drukkerij en kunstni.	29	9	19	0,429	Medium
0733 Bouwarbeiders afbouw	38	0	38	0,417	Medium
1031 Laboranten	26	19	7	0,415	Medium
0711 Biologen en natuurwetenschappers	39	11	27	0,398	Medium
1116 Verleners van overige persoonlij.	37	19	19	0,395	Medium
0713 Elektrotechnisch ingenieurs	17	1	16	0,382	Medium
0414 Beleidsadviseurs	72	41	31	0,38	Medium
0721 Technici bouwkunde en natuur	105	17	88	0,374	Medium
1214 Vrachtwagenchauffeurs -First-Line Supervisors of Transportation and Material-Moving Machine and Vehicle Operators	108	3	106	0,374	Medium
0731 Bouwarbeiders ruwbouw	63	1	62	0,368	Medium
1051 Verzorgenden	247	227	20	0,367	Medium
0433 Receptionisten en telefonisten	162	115	47	0,366	Medium
1034 Medisch praktijkassistenten	77	72	5	0,362	Medium
0743 Automonteurs	58	1	57	0,345	Medium
1035 Medisch vakspecialisten	52	24	28	0,34	Medium
1111 Reisbegeleiders	21	15	6	0,338	Medium
0413 Bedrijfskundigen en organisatiea.	127	45	82	0,337	Medium
0212 Auteurs en taalkundigen	32	20	12	0,329	Medium
1112 Koks	71	19	52	0,327	Medium

Table C.4: Probability of becoming automated of Dutch BRC-code occupations

Job	Total x1000	Women x1000	Men x1000	Pauto	Risk
0712 Ingenieurs (geen elektrotechniek)	117	15	102	0,31	Medium
0761 Elektriciens en elektronicamonteurs	85	2	84	0,31	Medium
0422 Zakelijke dienstverleners	99	41	59	0,304	Medium
1114 Kappers en schoonheidsspecialisten	76	70	6	0,296	Low
0321 Vertegenwoordigers en inkopers	149	43	106	0,287	Low
0121 Beroepsgroep sportinstructeurs	47	23	24	0,274	Low
0734 Loodgieters en pijpfitters	38	1	38	0,262	Low
0732 Timmerlieden	85	1	84	0,252	Low
0211 Bibliothecarissen en conservatoren	8	4	4	0,25	Low
0533 Managers ICT	19	1	18	0,25	Low
0822 Radio- en televisietechnici	19	2	17	0,246	Low
1033 Verpleegkundigen (mbo)	71	65	7	0,246	Low
0115 Onderwijskundigen en overige doc...	70	47	23	0,245	Low
0744 Machinemonteurs	47	1	45	0,242	Low
1211 Dekofficieren en piloten	30	3	27	0,236	Low
0612 Overheidsambtenaren- operational	67	37	30	0,233	Low
0214 Beeldend kunstenaars	14	7	7	0,226	Low
0714 Architecten	40	14	26	0,223	Low
1113 Kelners en barpersoneel	231	154	77	0,213	Low
0531 Managers productie	44	4	40	0,209	Low
0332 Verkoopmedewerkers detailhandel	339	225	114	0,191	Low
0536 Managers gespecialiseerde dienst.	28	4	24	0,189	Low
0632 Politie en brandweer	41	12	29	0,187	Low
0532 Managers logistiek	18	1	17	0,182	Low
0112 Docenten beroepsgerichte vakken ...	33	18	15	0,179	Low
0631 Politie-inspecteurs	11	4	7	0,178	Low
0215 Uitvoerend kunstenaars	53	20	34	0,175	Low
0543 Managers commerciÃ«le en persoonl.	12	4	8	0,167	Low
1012 Gespecialiseerd verpleegkundigen	129	110	19	0,164	Low
0222 Fotografen en interieurontwerpers	25	12	13	0,153	Low
0722 Productieleiders industrie en bouw	55	3	52	0,15	Low
0131 Leidsters kinderopvang en onderw...	142	128	14	0,146	Low
1011 Artsen	121	71	50	0,139	Low
0213 Journalisten	31	14	17	0,138	Low
0221 Grafisch vormgevers en producton... - Graphic Designers	60	25	35	0,138	Low
0113 Docenten algemene vakken secunda...	109	61	48	0,131	Low
1013 Fysiotherapeuten	78	57	21	0,129	Low
0331 Winkeliers en teamleiders detail...	117	49	68	0,128	Low
0521 Managers zakelijke en administra.	77	24	52	0,126	Low
1022 Psychologen en sociologen	75	57	18	0,122	Low
0621 Juristen	71	36	35	0,121	Low
0311 Adviseurs marketing public rela...	163	77	86	0,118	Low
0522 Managers verkoop en marketing	47	11	36	0,109	Low
0541 Managers horeca	24	9	15	0,099	Low
1311 Beroepsgroep overig - managing	143	48	95	0,097	Low
0542 Managers detail- en groothandel	39	9	30	0,096	Low
0114 Leerkrachten basisonderwijs	161	134	27	0,095	Low
0611 Overheidsbestuurders	23	9	14	0,088	Low
1021 Maatschappelijk werkers	69	53	16	0,077	Low
0221 Grafisch vormgevers en producton..- Commercial and Industrial Designers	60	25	35	0,068	Low

Table C.4: Probability of becoming automated of Dutch BRC-code occupations

Job	Total x1000	Women x1000	Men x1000	Pauto	Risk
0511 Algemeen directeuren	74	14	60	0,053	Low
0534 Managers zorginstellingen	21	14	6	0,034	Low
0535 Managers onderwijs	17	9	9	0,024	Low

Appendix D

Specification of the model

D.1 Agents

This section explains the specification of the agents. For each type of agents –workers and firms– their attributes and rules are presented.

D.1.1 Workers

Table D.1 shows the attributes of the workers in the agent-based model. It also shows it numeric specification and sometimes an additional explanation. Here, each job j is a member of all 112 jobs J. Furthermore, each industry i is a member of all 24 industries I. Table D.2 show the specification and the assumptions of the workers.

Algorithm 5 shows the the rules/algorithm of workers within in each tick. As can be seen, this algorithm is structured through the earlier defined expect-plan-act-evaluate process.

```
begin tick;
if unemployed = True then
    Expect;
    Set trust; Determine expected p-auto Determine best retrain job;
    Calculate NCB retrain;
    Plan;
    if if NCB-retrain > 0 then
        Determine p-retrain;
        Determine retrain?;
        if Retrain = True then
            Retrain
        else
            No retrain
        end
    else
        No retrain
    end
    evaluate;
    if Employed = true then
        Update Experiences + 1
    else
        Update Experiences + 0
    end
else
    do nothing
end
```

Algorithm 5: Algorithm Workers

Table D.1: Attributes of workers

Attribute	Specification	Description
employed	[True, False]	
job	$j \in J$	
initial-job	$j \in J$	
companyid	Integer	If employed, which firm
uncertainty	Float [0, 1]	
age	Integer [18, 67]	
experience-job	Integer [0, 43]	Number years experience in job
experience-firm	Integer [0,43]	Number years experience in firm
skills	Array [Competencies]	
wage	1	
wealth	Integer	
productivity	1	
unit-costs	1	
replaced-by	List [FirmID]	List of firms that have replaced/fired worker
automated	[True, False]	
p-auto	Float [0, 1]	FO automation probabiliy worker
expected-p-auto	Float [0, 1]	automation probabiliy worker as believed by worker
to-be-replaced	[true, False]	Replaced in next tick?
foresight	3	
p-retrain	Float [0, 1]	Retrain probability of worker
retrain?	[True, False]	
retrain-plans	$j \in J$	
firm-retrain-plans	$j \in J$	
retrain-costs	Integer	
train-history	List	Previous trainings worker
retrained-by	list [FirmID]	
last-wage	Integer	
unemployed-years	Integer	if unemployed, years unemployed
contract	Boolean [0,1]	
network	list [workers]	colleagues and previous colleagues of worker
industry	$i \in I$	

D.1.2 Firms

Table D.3 shows the attributes of the firms in the agent-based model. It also shows it numeric specification and sometimes an additional explanation. Here, each job j is a member of all 112 jobs J. Furthermore, each industry i is a member of all 24 industries I. The complete specification of the firms is given by table D.4.

Algorithm 6 shows the the rules/algorithm of firms within in each tick. As can be seen, this algorithm is structured through the earlier defined expect-plan-act-evaluate process.

D.2 Specification globals

Table D.5 shows the specification of our global variables.

D.3 Initial set-up

This appendix discusses the initial set-up of the model. This ‘set-up’ is done before each model-run and will be the same for each model run executed in this analysis

Table D.2: Specification workers

Assumption	Specification
Trust	$T_c(t+\Delta t) = Tc(t) + (X_c(t) - (T_c(t)) * \Delta$
automation probability	$Pauto_c^j(t+\Delta t) = P_{t+\Delta t}^j(Y=1) * T_c(t+\Delta t)$
Wage	1
Productivity	1
Unit costs	1
Retrain possibility	$p_{retrain} = p_{jobexperience} * p_{age}$
$p_{jobexperience}$	$p_{jobexperience} = 1 - \frac{1}{49} Y_{jobexperience}$
p_{age}	$p_{age} = 1 - \frac{1}{49}(Y_{age} - 18)$
Retrain costs	$C_{retrain} = \sum_{i=1}(\lvert S_i^{required} - S_i \rvert) * C_{unit} * (1 - subsidy)$
NCB retrain	$NCB = (p_{auto} * profit_{retrain}) - ((1 - p_{auto}) * c_{retrain})$
Retrain profitable?	$NCB > 0$ and $wealth > c_{retrain}$
Retrain?	$p_{retrain} = 1$
Wealth	If employed $wealth_t = wealth_{t-1} + wage_t$

D.3.1 Distribution of workers within sectors

This appendix shows the workforce distributions for firms in each sector. In our model, we assume that every sector has 10 firms and every firms has an initial workforce distribution. During the run of the model, firms can expand or decrease their workforce only with jobs were part of the initial distribution. This assumption controls for extreme workforce changes within firms. In other words, a firms from the financial sector cannot suddenly become a firm with only teachers.

Workforce distributions within firms are derived from aggregated workforce data of the UWV. This data shows the number of workers per job within a sector. For our model, we apply this sector distribution on firm level. For instance, if in a sector work 10 cleaners and 30 teachers, we assume that the ratio of teachers versus cleaners within a firm is 1:30. Table **??** presents the distribution of jobs within the 24 sectors. Due to the length of this table, we use only the BRC code of a job and a number for each industry. The numbering of sectors and their corresponding English translation is shown in table D.6.

Table D.3: Attributes of firms

Attribute	Specification	Description
id	Integer	
workforce	List [WorkersID]	List of employees
trust	Integer [0, 1]	
experience-rate	List [0, 1]	Investment experiences
costs	Integer	
revenue	Integer	
output	Integer	
profit	Integer	
previous-profit	Integer	Previous profit
invest-money	Integer	Money left of profit to invest
vacancies	List {[Job, demand]}	
my-robots	List [Robots]	
expected-robot-productivity	Float	Expected productivity of robots
experiences	List [0,1]	
my-plans	List [{To hire:Workers},	{To-replace:Workers}, {Robots-to-buy:Robots},
{Retrain:Workers}]		
class	String [”First”, ”Second”]	Confidence class
my-expectations	List [{Jobs:Expected-p-auto}]	
forsight	3	Years for which p-auto is calculated
initial-jobs	$j \in J$	Initial distribution of workforce
not-invested-in	List [Robots]	Because of believe, not invested in
industry	String $i \in I$	

Table D.4: Specification firms

Assumption	Specification
Trust	$T_c(t+\Delta t) = Tc(t) + (X_c(t) - (T_c(t)) * \Delta t$
automation probability	$Pauto_c^j(t+\Delta t) = P_{t+\Delta t}^j(Y=1) * T_c(t+\Delta t)$
No automation probability	$NoPauto_c^j(t+\Delta t) = 1 - Pauto_c^j(t+\Delta t)$
First movers	Firms with $T > 0.5$
Second movers	Firms with $T < 0.5$
Expected unit-costs robot	$UC_{j,r} = Pauto_c^j(t+\Delta t) * UC_{j,p=1} + NoPauto_c^j(t+\Delta t) * UC_{j,p=0}$
Replace with robot	If $UC_{j,r} < UC_{j,w}$
Retrain worker	If $RetrainCosts < FireCosts$
Fire worker	If $RetrainCosts > FireCosts$
Fire Costs	If Contract = flex $FireCosts = 0$ If Contract = fixed $FirceCosts = wage$
Costs	$costs_t = \sum wage_w + \sum costs_r$
Output	$output_t = \sum Productivity_{w,t} + \sum Productivity_{r,t}$
Revenue	$revenue_t = output_t * R_{ad}$
Profit	$profit_t = revenue_t - costs_t$
Previous profit	$profit_{t-1}$
Profit increase	$R_p = profit_t - profit_{t-1}$
Positive experience	$R_p > 0$
Negative experience	$R_p < 0$

```
Begin Tick;
Update Trust;
Expect;
for j in J do
    calculate expected p-auto;
    calculate expected robot-productivity;
    calculate expected unit-costs workers;
    calculate expected unit-costs robots;
end
Plan;
check replace;
for Worker in workforce do
    if unit-cost worker > expected-unit-costs robots then
        set my plan replace worker;
    else
        do nothing
    end
end
while Money-to-invest > 0 do
    if unit-costs new-worker > expected-unit-costs robots then
        set my plan buy robot
    else
        if Any unemployed workers with job = True then
            hire worker;
        else
            ;determine retrain costs;
            if retrain-costs workers < money-to invest then
                Set my plain retrain worker
            else
                Do nothing
            end
        end
    end
end
Act;
if Class = "First" then
    Act;
else
    Do nothing;
end
Evaluate ; calculate profit if Previous profit < profit then
    Update experiences + 1
else
    Update experiences + 0
end
if p-auto = expected-p-auto then
    Update experiences + 1
else
    Update experiences + 0
end
if plan workers to hire = workers hired then
    Update experiences + 1
else
    Update experiences + 0
end
```

Algorithm 6: Algorithm Firms

Table D.5: Specification global assumptions

Assumption	Specification
Cumulative automation probability	$p_j^{cum} = p_a uto_j^{FO}$
Conditional automation probability	$p_{j,t}^{con} = \frac{p_j^{cum}}{1+e^{(-t-foresight)-p_j^{cum}}} - \frac{p_j^{cum}}{1+e^{-t-p_j^{cum}}}$
Gap between demand and supply	$G_j = D_j - S_j$
Labour demand	$D_j = \sum_{i=1} v_{j,i} + h_{j,i} - f_{j,i}$
Labour supply	$S_j = sum_{i=1} w_j$
Aggregate demand growth	$r_{ad} = 1 - \frac{ad_{t-1} - ad_t}{ad_{t-1}}$

Table D.6: Labour Sectors in the Netherlands, Dutch and English

	Sector Dutch	Sector English
1	Land- en tuinbouw, visserij, winning van delfstoffen	Agriculture and horticulture, fishing, mining and quarrying
2	Procesindustrie Voedings- en genotsmiddelen Chemie	Food and Process Industry
3	Grafimedia	Graphics media
4	Metalektro en metaalnijverheid	Metallurgy and metal industry
5	Overige industrie, inclusief energie en nutsbedrijven milieu en recycling	Other industry including energy and utilities environment and recycling
6	Bouwnijverheid en bouwinstallatie	Construction industry and construction installation
7	Mobiliteitsbranche (Handel in en reparatie van motorvoertuigen, tweewielers en caravans)	Mobility sector (Trade in and repair of motor vehicles, two-wheelers and caravans)
8	Groothandel	Wholesale
9	Detailhandel	Retail
10	Vervoer en opslag	Transport and storage
11	Horeca, catering en verblijfsrecreatie	Hospitality, catering and accommodation recreation
12	Informatie en communicatie	Information and communication
13	Financiële dienstverlening	Financial services
14	Arbeidsbemiddeling, uitzendbureaus en personeelsbeheer	Employment services, employment agencies and personnel management
15	Schoonmaak	Cleaning
16	Overige zakelijke dienstverlening	Other business services
17	Openbaar bestuur, inclusief overheidsdiensten	Public administration, including government services
18	Onderwijs	Education
19	Zorg	Health
20	Welzijn, Kinderopvang Jeugd, Maatschappelijke opvang, Sociaal-cultureel werk	Well-being, Childcare Youth, Social relief, Socio-cultural work
21	Cultuur, sport en recreatie	Culture, sport and recreation
22	Overige dienstverlening, huishoudens en extraterritoriale organisaties	Other services, households and extraterritorial organizations
23	Installatiebranche	Installation industry
24	Bouwinstallatie	Construction installation

Appendix E

Verification

This appendix shows the results of the verification of the model.

E.1 Single-agent testing

In this section, the implemented rules of the agents are tested. For this, we followed the behavior and variable changes for both firms and workers in during each tick. Table E.1 shows the parameter settings that are used for this single-agent model verification.

Table E.1: Default parameter settings verification

Variable	Value
Number firms	5
Forsight	3
Information	Medium
Flexibility	Medium
Subsidy	None
Tick end	20

E.1.1 Firms

This section shows the results of the singe-agent verification of firms. As can be seen, the implementation of the rules of the firms correspond with the earlier formulated rules in the model specification (see Chapter 8).

Expect

- In the expect phase, firms make expectations about the automatizing probabilities of jobs according to their own confidence and FO probabilities. **Confirmed**
- The trust level of firms is determined every expect phase and is derived from its previous experiences. **Confirmed**
- Firms with a trust-level lower than 0,5 will be 'waiters' and firms with a trust-level above 0.5 will be 'first-movers'.
- Every expect phase, the difference between the initial working distribution and current working distribution is calculated. **Confirmed**
- Firms only have jobs in their workforce that are consistent with their initial workforce. **Confirmed**

Plan

- Planning is made only for workers who are considered to be first movers. **Confirmed**
- The planning will be in line with the initial workforce distribution of firms. **Confirmed**
- The planning will not exceed the money available to invest of firms. **Confirmed**
- When no workers are available to hire, the firm makes vacancies for that job. **Confirmed**
- More than one firm can target a possible worker in his planning. **Confirmed**
- When firing costs are lower than retrain costs, when the firm replaces a worker for a robot, the firm will fire the worker instead of retraining. **Confirmed**
- When unit-costs of a robot is less than that off a new worker, the firm will plan to buy a robot. **Confirmed**.

Act

- Firms retrain the workers the workers that are planned to retrain. **Confirmed**
- Firms buy the robots that they planned to buy. **Confirmed**
- When no other firm has hired the worker already, firms hire the workers they planned to hire. **Confirmed**
- When a firm planned to hire a worker but this worker is already taken, the firm will increase its vacancies for that job. **Confirmed**.

Evaluate

- The costs of a firm is the sum of the wages of its workforce and the costs of its robots. **Confirmed**
- The output of a firm is the sum of the productivity of its workforce and robots. **Confirmed**
- The revenue of a firm is its output times the aggregate demand. **Confirmed**
- Firms who have a higher profit increase than last tick, have a positive experience. **Confirmed**
- Firms who have a lower profit increase than last tick, have a negative experience. **Confirmed**
- Firms who have waited and who could have make a higher profit by acting have a negative experience. **Confirmed**
- Firms who have waited and who made a higher profit than when acting, have a positive experience. **Confirmed**

E.1.2 Worker

This section shows the results of the singe-agent verification of workers. As can be seen, the implementation of the rules of the workers correspond with the earlier formulated rules in the model specification (see Chapter 8).

Expect

- If the workers is unemployed, expectations have to be made about the demand for his job. **Confirmed**
- Since we assume a negative relation between age and retraining probability, workers with a higher age but the same working experience must have a lower retrain probability. **Confirmed**

- Since we assume a negative relation between working experience and retraining probability, workers with a higher but same age as must have a lower retrain probability. **Confirmed**
- The sum of the contributing weights of government information and network information in the workers expectations should be 1. **Confirmed**

Plan

- The actual decision of retraining is determined in the plan phase and is the result of the retrain probability function. **Confirmed**
- Workers that do not have enough money to retrain themselves, do not retrain themselves. **Confirmed**

Act

- Workers that have a plan to retrain themselves retrain themselves. **Confirmed**.
- The wealth of the retrained workers decreased by its training costs. **Confirmed**

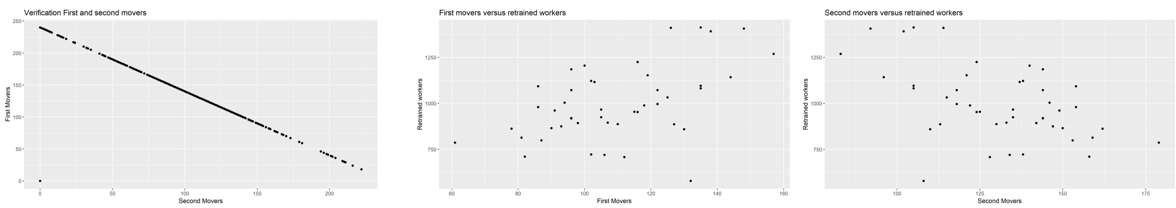
Evaluate

- Workers that are hired because of retraining in last tick, have a positive experience. **Confirmed**

E.2 Model testing

This section presents the results of the verification of the whole model. Figure E.1 shows the verification of first and second movers in our model. As discussed in chapter 9, less investments are made when there are more second movers. Furthermore, since firms can either be a first or second mover, there is an inversely proportional relationship between the number first and second movers (see figure E.1 a).

For this verification of first and second movers, we have executed 45 replications with the parameter setting as specified in Table E.1.



(a) First and second movers (b) First movers (c) Second movers

Figure E.1: Verification First and second movers

Appendix F

Sensitivity Analysis

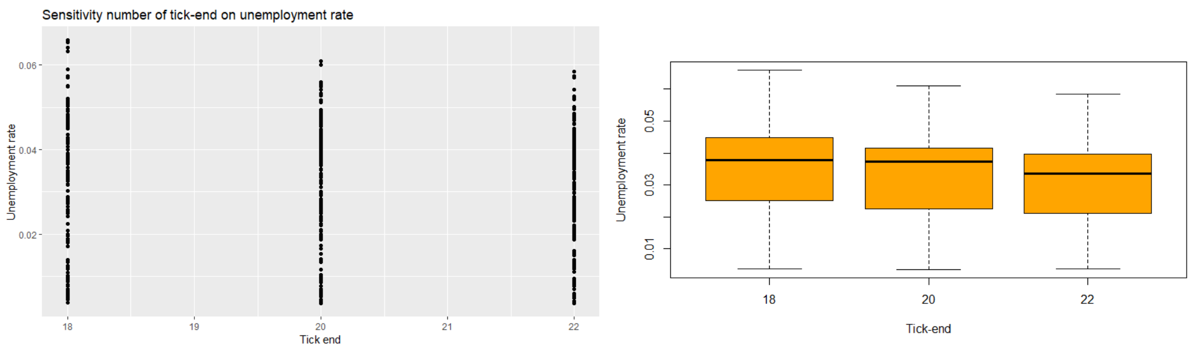
This appendix shows the results of the sensitivity analysis. Table F.1 shows the parameter settings that are used for the sensitivity analyses.

Table F.1: Parameter settings for the sensitivity analysis. Variance depicts the different values for which changes in the unemployment rate is compared.

Parameter	Default	Variance	Repetitions
Number of firms	5	4,5,6	10
Foresight variable	3	2,3,4,10	10
Tick-end	20	18,20,22	10
Logistic growth rate	5	1,2,3,4,5	10

F.1 Tick end

We have chosen to run our agent-based model for 20 years. To test how this choice may influence the outcomes of the model, we have performed a sensitivity analysis of the unemployment rate to run-time. Figure F.1 shows the results of this analysis. As can be seen from the figure, the model is not sensitive to small changes in the run-time.



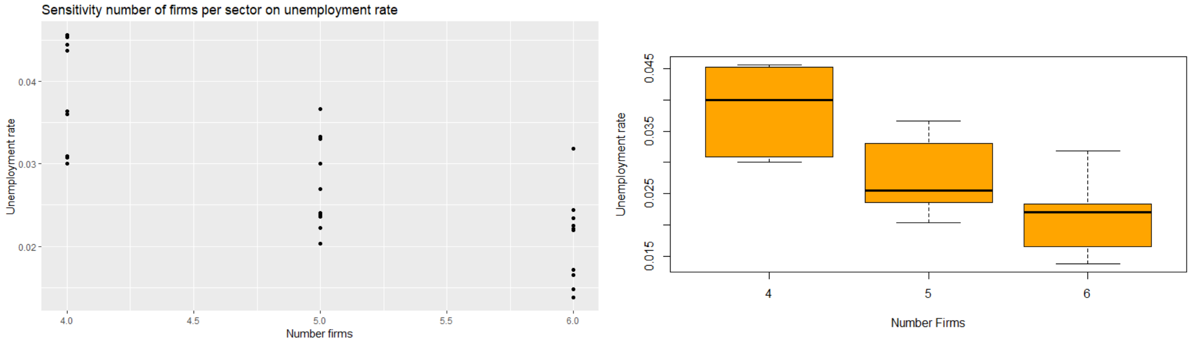
(a) Scatter plot (b) Box plot

Figure F.1: Sensitivity of the unemployment rate on the length of the runtime

F.2 Number of firms

Figure F.2 shows the results of the sensitivity the unemployment rate on the number of firms. As one, can see, the number of firms has a negative correlation with the unemployment rate:

the higher the number of firms, the lower the unemployment rate.

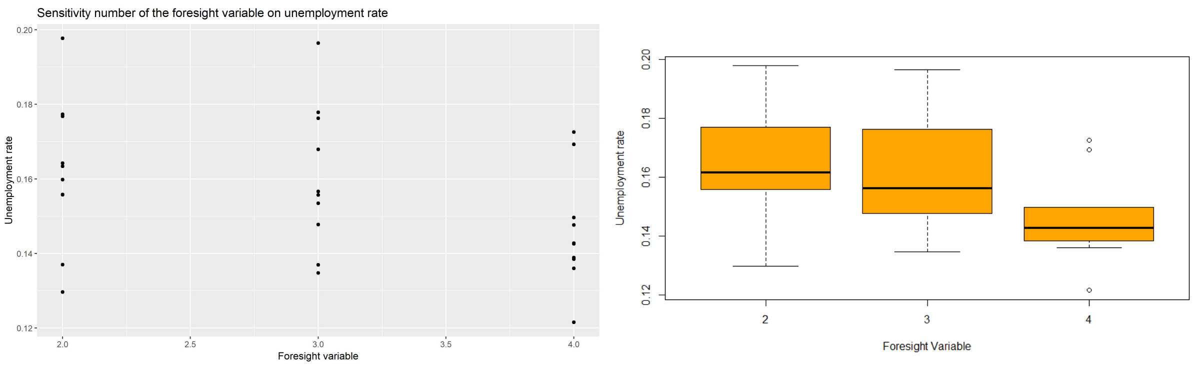


(a) Scatter plot (b) Box plot

Figure F.2: Sensitivity of the unemployment rate on the number of firms.

F.3 Foresight variable

Figure F.3 shows the results of the sensitivity analysis of the foresight variable. This figure shows that the unemployment rate is sensitive for small changes in the foresight variable. The higher the foresight variable, the lower the unemployment rate.

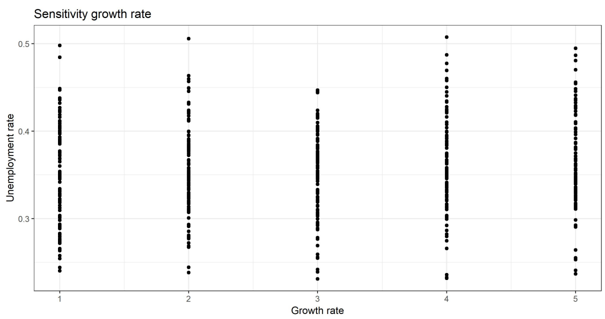


(a) Scatter plot (b) Box plot

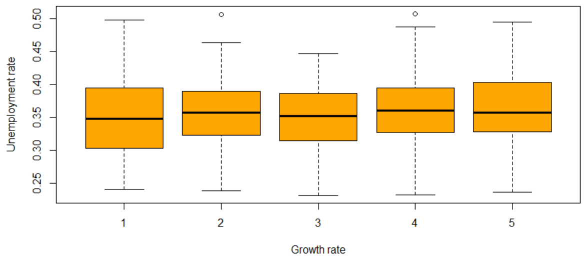
Figure F.3: Sensitivity of the unemployment rate on the foresight variable.

F.4 Logistic growth rate

Figure F.4 shows the results of the sensitivity analysis of the growth rate. As one can see, the unemployment rate is not sensitive to changes in the growth rate.



(a) Scatter plot



(b) Box plot

Figure F.4: Sensitivity of the unemployment rate on growth rate.

Appendix G

Model Configurations

This Appendix will present the model implementation in NetLogo of both configurations. In our model, running the model consists of two steps: the *setup procedure* and the *go procedure*(See Figure G.1) . The setup procedure is called before actually running the model and sets up all the needed agent, parameters and variables of the model. After this, the go procedure can be executed. Since we have two model configurations, we have two separate go procedures. The setup procedure, however, is the same for both model configurations.

This appendix starts with the implementation of the setup procedure in NetLogo. Here, the script of the this procedure is presented. Next, the implementation of both model configurations is presented. Note however, that due to the fact that our NetLogo scripts has approximately 2000 lines, only an abstracted version is presented.

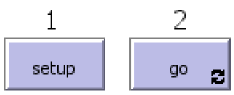


Figure G.1: Go and Setup buttons in NetLogo model

G.1 Model setup

This section presents the setup procedure of the model. This setup procedure is the same for both model configuration.

G.1.1 Script

When the setup procedure is called, the script of listings G.1 is executed. As mentioned, due to the length of the original script, this listings only present an abstracted version of the setup procedure.

```

to setup
  clear-all
  reset-ticks
  setup-workers
  setup-firms
  setup-unemployment
  setup-robots
  make-tables-globals
end

to setup-firms;
  foreach all-industries [x ->
    create-firms nr-firms [set size 2 set color blue setxy random-xcor random-ycor
```

```
      set industry x
      set id who + 10000
      set vacancies table:make
      set my-robots no-turtles
      set trust random-float 2
      set experiences (list random(2))
      set my-plans table:make
      set my-expectations table:make
      set bought-robots-for-belief no-turtles
      set workforce-all no-turtles]

    let workers-now workers with [industry = x]
    let workers-count count workers-now
    let workers-per-firm floor (workers-count / 10)
    ask firms with [industry = x][
      ask n-of workers-per-firm workers-now with [employed = false][
        set employed true
        set color green
        set companyid [id] of myself
      ]
      set workforce workers with [companyid = [id] of myself]]
    ask workers-now with [employed = false][
      set employed true
      set companyid [id] of one-of firms with [industry = x ]]]

  ask firms [
    set workforce workers with [companyid = [id] of myself]
    set workforce-all workforce
    foreach all-jobs [x ->
      table:put vacancies x 0]
    set initial-jobs table:make
    let c-all count workforce-all
    foreach remove-duplicates [job] of workforce-all [x ->
      let n count workforce-all with [job = x]
      let per n / c-all
      table:put initial-jobs x per ]
    set diff-distribution table:make
    foreach table:keys initial-jobs [x ->
      table:put diff-distribution x 0]]

  ask firms [
    set strategy-list remove-duplicates [job] of workforce
    ask workforce [
      set network [workforce] of myself]]

to setup-workers
  if debug = true [print "Begin setup-workers"]
  file-open "NetlogoWorkers2.csv" // open the file with the turtle data
  // We'll read all the data in a single loop
  while [ not file-at-end? ] [
    // here the CSV extension grabs a single line and puts the read data in a list
    let data csv:from-row file-read-line
    let n-18-25 item 131 data
    let n-25-35 item 132 data
    let n-35-45 item 133 data
    let n-45-55 item 134 data
    let n-55-65 item 135 data
    let n-65-70 item 136 data ;;
    if is-number? n-18-25 [
     create-workers ceiling n-18-25 [
        set age 18 + random(8)
        set-rest-worker data
    ]]
    if is-number? n-25-35 [
     create-workers ceiling n-25-35 [
        set age 25 + random(11)
        set-rest-worker data
    ]]
    if is-number? n-35-45 [
     create-workers ceiling n-35-45 [
        set age 35 + random(11)
```

```
        set-rest-worker data
    ]]
    if is-number?  n-45-55 [
     create-workers ceiling n-45-55 [
        set age 45 + random(11)
        set-rest-worker data
    ]]
    if is-number?  n-55-65 [
     create-workers ceiling n-55-65 [
        set age 55 + random(11)
        set-rest-worker data
    ]]
    if is-number?  n-65-70 [
     create-workers ceiling n-65-70 [
        set age 65 + random(3)
        set-rest-worker data
    ]]

  ]
  file-close
  set all-jobs sort remove-duplicates [job] of workers
  set all-industries remove-duplicates [industry] of workers

end


to setup-unemployment
  let count-employed-workers-25 count workers with [employed = true and age < 25]
  let total-25 count-employed-workers-25 / 0.928
  let n-25-unemployed total-25 * 0.072
  set n-25-unemployed round n-25-unemployed
  create-workers n-25-unemployed [
   make-unemployed workers
  ]

  let count-employed-workers-45 count workers with [employed = true and age >= 25 and age < 45]
  let total-45 count-employed-workers-45 / 0.972
  let n-45-unemployed total-45 * 0.028
  set n-45-unemployed round n-45-unemployed
  create-workers n-45-unemployed [
    make-unemployed workers]

  let count-employed-workers-67 count workers with [employed = true and age >= 45 and age <= 67]
  let total-67 count-employed-workers-67 / 0.964
  let n-67-unemployed total-67 * 0.036
  set n-67-unemployed round n-67-unemployed
  create-workers n-67-unemployed [
    make-unemployed workers]
end

to setup-robots
  create-robots (count workers)[set job-replace false]
  ask workers [
    let robots-left robots with [job-replace = false]
    let my-robot one-of robots-left
    ask my-robot [
      set job-replace [job] of myself
      set my-firms no-turtles
      set productivity [productivity] of myself
      set costs 5
      set expected-productivity 6.5
      set expected-unit-costs costs / expected-productivity
      set unit-costs costs / productivity
      set buying-by no-turtles
      set firms-expectations table:make]]
end

to make-tables-globals
  set auto-table table:make
  set endogenous-auto table:make
  foreach all-jobs [x ->
```

```
    let auto-c [p-auto] of one-of workers with [job = x]
    table:put auto-table x auto-c
    if endogenous-TC = true [table:put endogenous-auto x auto-c] ]

  set demand-jobs table:make
  set unfilled-demand table:make
  set unfilled-demand-previous table:make
  set demand-difference table:make
  set demand-jobs-rate table:make

  set unit-training-costs 0.01
  set ad-increase 1

  let unemployed count workers with [employed = false]
  set unemployment-rate unemployed / count workers

  let flex count workers with [contract = 0]
  set percentage-flexible flex / count workers
end
```

Listing G.1: Script setup procedure

G.2 The Frey & Osborne model configuration

This section will present the model specification of the Frey & Osborne model configuration (FO configuration). For this, an (abstracted) version of the model script of the FO Model Configuration is shown. This section also presents the initial parameters –or so-called 'set-up parameters'– that are used in the FO Model Configuration.

G.2.1 Script

Listings G.2 shows the script as used for the FO Model Configuration. Note, that this is the real model script and basic knowledge about the programming NetLogo is needed (for more information about the programming language of NetLogo, visit: https://ccl.northwestern.edu/netlogo/docs Moreover, since all important algorithms are already discussed in chapter 8, this listings shows only the important elements of the code.

```
to go
  ifelse fo-conf = true [go-fo][go-normal]
  tick
  if ticks >= tick-end + 1 [stop]
end

to go-fo
  determine-strategy-firms-fo
  act-fo
  technological-progress-event
  end-tick-fo
end

to determine-strategy-firms-fo
  ask firms [
    table:put my-plans ”replace” no-turtles
    if length jobs-automated > 0 [
    foreach jobs-automated [x ->
        let replacers workforce with [job = x]
        let already-replace table:get my-plans ”replace”
        table:put my-plans ”replace” (turtle-set replacers already-replace)
        ask replacers [buying-robot myself x]
      ]
    ]
  ]
  if debug = true [print ”determine-strategy fo completed”]
end
```

```
to act-fo
  ask firms [
    set fired no-turtles
    let workers-to-fire table:get my-plans "replace"
    fire workers-to-fire self
    table:put my-plans "replace" no-turtles
    set workforce workers with [companyid = [id] of myself]

    let robots-to-buy table:get my-plans "robots-to-buy"
    set my-robots (turtle-set my-robots robots-to-buy)
    ask robots-to-buy [
      set my-firms (turtle-set self my-firms)]
    table:put my-plans "robots-to-buy" no-turtles ]
end

to technological-progress-event
  let all-jobs-na jobs-not-automated
  set jobs-automated (list)
  let change 0
  foreach all-jobs-na [x ->
    ifelse endogenous-TC = true
        [set change table:get endogenous-auto x]
        [set change table:get auto-table x]
    let conditional-change conditional-tech change ticks
    let happened probability conditional-change
    if happened = true [
      if debug = true [output-print x]
      automatize x]]
end

to end-tick-fo
  set total-workforce count workers
  let unemployed count workers with [employed = false]
  set unemployment-rate unemployed / total-workforce

  set retrained-workers count workers with [initial-job != job]
  set replaced-workers count workers with [count replaced-by > 0]
  set retrained-by-firm count workers with [count retrained-by > 0]
  set firm-robots sum [count my-robots] of firms
  set first-movers count firms with [class = "first"]
  set second-movers count firms with [class = "second"]
  set most-replaced-jobs modes [job] of workers with [count replaced-by > 0]
  set most-robot-jobs modes [job-replace] of robots with [count my-firms > 0]
end
```

Listing G.2: Script FO Model Configuration

G.2.2 Setup parameters

Table G.1 shows the value of the parameters used in the experiments with the FO Model Configuration. For this experiment, only the parameters 'tick-end', 'nr-firms', 'fo-config' and 'ex-growth-rate' (exponential growth rate) are important.

G.3 The Extended Model Configuration

In this section, the script and setup parameters of the E-MC are presented. Note that just as in the setup and FO-MC scripts, this script is an abstracted version of the original script.

G.3.1 Script

```
to go
  ifelse fo-conf = true [go-fo][go-normal]
  tick
  if ticks >= tick-end + 1 [stop]
```

Table G.1: Parameter setup: FO configuration

Parameter	Value
Tick end	20
nr-firms	5
fo-config	True
Flexibility	”none”
Information Policy	”none”
retrain-subsidy	0
endogenous-TC	False
scenario	”none”
foresight variable	”none”
ex-growth-rate	2

```
end

to go-normal
  begin-tick
  determine-strategy-firms
  determine-strategy-workers
  act
  technological-progress-event
  evaluate
  end-tick
end

to begin-tick // set parameters at begin tics
  if debug = true [print ”begin tick”]
  set-unfilled-demand //set the unfilled demand
  set-firms // set all parameters for firms right at beginning tick
  set-workers // set all parameters for workers right at beginning tick
  determine-AD // determine aggregate demand
end

to determine-strategy-workers
  let unemployed workers with [employed = false]
  let employed-workers workers with [employed = true]
  // which job is best to retrain to, calculate costs
  if ticks != 0 [
    ask unemployed [
      let random-jobs [job] of employed-workers
      let filtered-random-job filter [i -> not member? i all-jobs-automated] random-jobs
      let random-job one-of filtered-random-job
      let believe-demand 0
      ifelse highest-demand-increase != ”nothing” [
        let belief-demand highest-demand-increase
        if network != no-turtles [
          let network-jobs [job] of network
          let network-jobs-filtered filter [i -> not member? i all-jobs-automated] network-jobs
          let network-bestjob random-job
          if length network-jobs-filtered > 0
            [set network-bestjob one-of modes network-jobs-filtered]
          let ip table:get policy-values Information-policy
          let network? probability (1 - ip)
          if network? = true [set believe-demand network-bestjob]]
      ]
      [
        ifelse network != no-turtles [
          let network-jobs [job] of network
          let network-bestjob one-of modes network-jobs
          set believe-demand network-bestjob][
          set believe-demand random-job]]

      if believe-demand = 0 [set believe-demand max-unfilled-demand]
      if believe-demand = ”nothing” [set believe-demand random-job]
      let wanted-skills sum table:values table:get skill-table believe-demand
```

```
      let sum-skills sum table:values skills
      let skills-gap abs (wanted-skills - sum-skills)
      let train-costs skills-gap * unit-training-costs
      set train-costs train-costs * (1 - retrain-subsidy)
      set p-retrain my-p-retrain experience-job age self
      let no-auto (1 - expected-p-auto)
      let profit-of-retrain wage * forsight
      let ncb-retrain (expected-p-auto * profit-of-retrain) - (no-auto * train-costs)
      if ncb-retrain > 0 [set retrain-plans believe-demand set retrain-costs train-costs]
  ]]
end


to determine-strategy-firms
  check-replace  // checks if worker can be replaced by bot
  make-invest-decision // what to do with profit? buy bot or worker?
end

to act
  act-firms // firms executes their strategy: buying bot, replace or do nothing
  act-workers // workers executes their strategy: retrain? or no retrain?
end

to technological-progress-event
  let all-jobs-na jobs-not-automated
  set jobs-automated (list)
  let change 0
  foreach all-jobs-na [x ->
    ifelse endogenous-TC = true
        [set change table:get endogenous-auto x]
        [set change table:get auto-table x]
    let conditional-change conditional-tech change ticks
    let happened probability conditional-change
    if happened = true [
      if debug = true [output-print x]
      automatize x]]
  if debug = true [print "technological-process completed"]
end

to evaluate
  evaluate-firms // determination of first and second movers
  if endogenous-TC = true [calculate-new-probability]
end
```

Listing G.3: Script Extended Model Configuration

G.3.2 Set-up parameters

Table G.2 shows the parameter setup as used for the experiments with the Extented model configuration. As can be seen from the table, multiple parameters has several values. Experiments are ran for each combination of these parameters values. Hence the large total number of runs for this experiment (20,000).

Table G.2: Parameter setup: Extended model configuration

Parameter	Value
Tick end	20
nr-firms	5
fo-config	false
Flexibiliy	”none”, ”low”, ”medium”, ”high”, ”all”
Information Policy	”none”, ”low”, ”medium”, ”high”, ”all”
retrain-subsidy	0, 0.25, 0.5, 0.75, 1
endogenous-TC	True, False
scenario	”high”, ”low”
foresight variable	1, 3, 6, 11
ex-growth-rate	2

Appendix H

Results FO configuration

In this appendix the results of the FO Model Configuration will be presented.

H.1 Differences between FO probabilities and model outcomes

Table H.1 shows for each job the difference between the FO probabilities ('P.Auto') and the percentage of that job being automated through all runs.

Table H.1: Differences model outcomes and FO probabilities

Job	Total men and women x.1000	P(automated)	Risk	Proportion automated to all runs	Difference
1115 ConciĂŤrges en teamleiders schoon.	43	0.7	High	0.46	0.24
0431 Administratief medewerkers	285	0.81	High	0.58	0.23
1032 Apothekersassistenten	25	0.58	Medium	0.36	0.22
1215 Bedieners mobiele machines	74	0.61	Medium	0.39	0.22
0333 Kassamedewerkers	92	0.79	High	0.57	0.22
0921 Hulpkrachten landbouw	28	0.7	Medium	0.5	0.2
0423 Directiesecretaresses	110	0.76	High	0.56	0.2
0771 Productiemachinebedieners	81	0.64	Medium	0.45	0.19
0435 Transportplanners en logistiek m.	183	0.59	Medium	0.4	0.19
0612 Overheidsambtenaren-office	67	0.58	Medium	0.39	0.19
0751 Slagers	19	0.75	High	0.56	0.19
1311 Beroepsgroep overig - operational	143	0.65	Medium	0.47	0.18
0742 Lassers en plaatwerkers	37	0.62	Medium	0.45	0.17
1213 Buschauffeurs en trambestuurders	23	0.62	Medium	0.45	0.17
0754 Meubelmakers, kleermakers en sto.	40	0.54	Medium	0.37	0.17
1212 Chauffeurs auto's, taxi's en bes.	67	0.63	Medium	0.47	0.16
0434 Boekhoudkundig medewerkers	138	0.82	High	0.66	0.16
0753 Productcontroleurs	20	0.52	Medium	0.36	0.16
0911 Land- en bosbouwers	29	0.55	Medium	0.39	0.16
0411 Accountants	99	0.65	Medium	0.5	0.15

Table H.1: Differences model outcomes and FO probabilities

Job	Total men and women x.1000	P(automated)	Risk	Proportion automated to all runs	Difference
1116 Verleners van overige persoonlij.	37	0.4	Medium	0.26	0.14
0735 Schilders en metaal-spuiters	33	0.58	Medium	0.44	0.14
1122 Keukenhulpen	75	0.78	High	0.64	0.14
1214 Vrachtwagenchauffeurs-Heavy and Tractor-Trailer Truck Drivers	108	0.52	Medium	0.38	0.14
0432 Secretaresses	53	0.75	High	0.61	0.14
0772 Assemblagemedewerkers	42	0.61	Medium	0.47	0.14
1041 Sociaal werkers, groeps- en woon.	239	0.69	Medium	0.55	0.14
0821 Gebruikersondersteuning ICT	45	0.47	Medium	0.34	0.13
0741 Metaalbewerkers en constructiewe.	42	0.61	Medium	0.48	0.13
0734 Loodgieters en pijpfitters	38	0.26	Low	0.14	0.12
0111 Docenten hoger onderwijs en hoog...	56	0.45	Medium	0.34	0.11
0752 Bakkers	20	0.54	Medium	0.43	0.11
0711 Biologen en natuurwetenschappers	39	0.4	Medium	0.29	0.11
0781 Hulpkrachten bouw en industrie	70	0.65	Medium	0.54	0.11
0415 Specialisten personeels- en loop.	98	0.58	Medium	0.47	0.11
0214 Beeldend kunstenaars	14	0.23	Low	0.13	0.1
0723 Procesoperators	26	0.5	Medium	0.4	0.1
0733 Bouwarbeiders afbouw	38	0.42	Medium	0.32	0.1
0913 Veetelers	57	0.44	Medium	0.34	0.1
1051 Verzorgenden	247	0.37	Medium	0.28	0.09
1221 Laders, lossers en vakkenvullers	253	0.55	Medium	0.46	0.09
0334 Callcentermedewerkers outbound e...	123	0.69	Medium	0.6	0.09
0712 Ingenieurs (geen elektrotechniek)	117	0.31	Medium	0.22	0.09
0543 Managers commerciĂŤle en persoonl.	12	0.17	Low	0.09	0.08
0632 Politie en brandweer	41	0.19	Low	0.11	0.08
0533 Managers ICT	19	0.25	Low	0.17	0.08
0713 Elektrotechnisch ingenieurs	17	0.38	Medium	0.3	0.08
0115 Onderwijskundigen en overige doc...	70	0.25	Low	0.18	0.07
0531 Managers productie	44	0.21	Low	0.14	0.07
0121 Beroepsgroep sportinstructeurs	47	0.27	Low	0.2	0.07
0721 Technici bouwkunde en natuur	105	0.37	Medium	0.3	0.07
0421 Boekhouders	112	0.54	Medium	0.47	0.07
0633 Beveiligingspersoneel	62	0.56	Medium	0.49	0.07
1031 Laboranten	26	0.42	Medium	0.36	0.06
1114 Kappers en schoonheidsspecialisten	76	0.3	Low	0.24	0.06

Table H.1: Differences model outcomes and FO probabilities

Job	Total men and women x.1000	P(automated)	Risk	Proportion automated to all runs	Difference
0413 Bedrijfskundigen en organisatiea.	127	0.34	Medium	0.28	0.06
0321 Vertegenwoordigers en inkopers	149	0.29	Low	0.23	0.06
0912 Hoveniers, tuinders en kwekers	74	0.5	Medium	0.44	0.06
1222 Vuilnisophalers en dagbladenbezo.	83	0.69	Medium	0.63	0.06
0131 Leidsters kinderopvang en onderw...	142	0.15	Low	0.1	0.05
0631 Politie-inspecteurs	11	0.18	Low	0.13	0.05
0755 Medewerkers drukkerij en kunstni.	29	0.43	Medium	0.38	0.05
0212 Auteurs en taalkundigen	32	0.33	Medium	0.28	0.05
0722 Productieleiders industrie en bouw	55	0.15	Low	0.1	0.05
0414 Beleidsadviseurs	72	0.38	Medium	0.33	0.05
0332 Verkoopmedewerkers detailhandel	339	0.19	Low	0.14	0.05
0215 Uitvoerend kunstenaars	53	0.17	Low	0.12	0.05
0221 Grafisch vormgevers en producton- Graphic Designers	60	0.14	Low	0.1	0.04
1111 Reisbegeleiders	21	0.34	Medium	0.3	0.04
0112 Docenten beroepsgerichte vakken ...	33	0.18	Low	0.14	0.04
0211 Bibliothecarissen en conservatoren	8	0.25	Low	0.21	0.04
0621 Juristen	71	0.12	Low	0.08	0.04
1113 Kelners en barpersoneel	231	0.21	Low	0.17	0.04
0114 Leerkrachten basisonderwijs	161	0.1	Low	0.07	0.03
0521 Managers zakelijke en administra.	77	0.13	Low	0.1	0.03
0822 Radio- en televisietechnici	19	0.25	Low	0.22	0.03
0731 Bouwarbeiders ruwbouw	63	0.37	Medium	0.34	0.03
0221 Grafisch vormgevers en producton -Commercial and Industrial Designers	60	0.07	Low	0.04	0.03
1013 Fysiotherapeuten	78	0.13	Low	0.1	0.03
0113 Docenten algemene vakken secunda...	109	0.13	Low	0.1	0.03
0714 Architecten	40	0.22	Low	0.19	0.03
0412 Financieel specialisten en economen	82	0.59	Medium	0.56	0.03
0433 Receptionisten en telefonisten	162	0.37	Medium	0.35	0.02
1033 Verpleegkundigen (mbo)	71	0.25	Low	0.23	0.02
0611 Overheidsbestuurders	23	0.09	Low	0.07	0.02
0811 Software- en applicatieontwikkel.	235	0.48	Medium	0.46	0.02
0536 Managers gespecialiseerde dienst.	28	0.19	Low	0.17	0.02
1022 Psychologen en sociologen	75	0.12	Low	0.1	0.02
0732 Timmerlieden	85	0.25	Low	0.23	0.02
1121 Schoonmakers	255	0.72	High	0.7	0.02

Table H.1: Differences model outcomes and FO probabilities

Job	Total men and women x.1000	P(automated)	Risk	Proportion automated to all runs	Difference
0612 Overheidsambtenaren-operational	67	0.23	Low	0.21	0.02
1214 Vrachtwagenchauffeurs -First-Line Supervisors of Transportation and Material-Moving Machine and Vehicle Operators	108	0.37	Medium	0.35	0.02
0311 Adviseurs marketing, public rela...	163	0.12	Low	0.11	0.01
1011 Artsen	121	0.14	Low	0.13	0.01
1035 Medisch vakspecialisten	52	0.34	Medium	0.33	0.01
0532 Managers logistiek	18	0.18	Low	0.17	0.01
0744 Machinemonteurs	47	0.24	Low	0.23	0.01
0534 Managers zorginstellingen	21	0.03	Low	0.02	0.01
0743 Automonteurs	58	0.34	Medium	0.33	0.01
1112 Koks	71	0.33	Medium	0.33	0
1311 Beroepsgroep overig - managing	143	0.1	Low	0.1	0
0213 Journalisten	31	0.14	Low	0.14	0
0761 Elektriciens en elektronicamonteurs	85	0.31	Medium	0.31	0
1034 Medisch praktijkassistenten	77	0.36	Medium	0.37	-0.01
0422 Zakelijke dienstverleners	99	0.3	Medium	0.31	-0.01
0542 Managers detail- en groothandel	39	0.1	Low	0.12	-0.02
1211 Dekofficieren en piloten	30	0.24	Low	0.26	-0.02
0511 Algemeen directeuren	74	0.05	Low	0.07	-0.02
0535 Managers onderwijs	17	0.02	Low	0.04	-0.02
1012 Gespecialiseerd verpleegkundigen	129	0.16	Low	0.18	-0.02
1021 Maatschappelijk werkers	69	0.08	Low	0.11	-0.03
0331 Winkeliers en teamleiders detail...	117	0.13	Low	0.16	-0.03
0522 Managers verkoop en marketing	47	0.11	Low	0.14	-0.03
0812 Databank- en netwerkspecialisten	70	0.5	Medium	0.53	-0.03
0222 Fotografen en interieurontwerpers	25	0.15	Low	0.19	-0.04
0541 Managers horeca	24	0.1	Low	0.18	-0.08

Appendix I

T-SNE results

This appendix presents the results of the t-SNE analysis of the policy levers, flexibility, information policy and retrain subsidy. It starts with presenting the results for flexibility, followed by the t-SNE results of information policy and retrain subsidy.

I.1 Method

Before analyzing the effect of different policy levers and scenario's, we first start with a general model exploration that reveals the behaviour of the model. This model exploration starts with an analysis of the different states of the model. For this, we conduct a t-Distributed Stochastic Network Embedding (t-SNE) analysis on the outcomes of the model. By using the t-SNE analysis technique, we are able to do a multidimensional reduction[1] and visualize the different 'states' of the model (van der Maaten & Hinton, 2008).

T-SNE analysis is a method that is used for data dimension reduction and is an extension on the commonly used Stochastic Neighbor Embedding Technique (SNE). Here, we will explain the procedure of t-SNE and its advantages over the general SNE technique briefly. More details about this technique and its applications can be found in the article *Visualizing data using t-SNE* by van der Maaten and Hinton (2008).

The original SNE analysis –for which the t-SNE analysis is variance– starts with the conversion of the high-dimensional Euclidean distances between data points (see also chapter 4) into conditional probabilities that represent similarities (van der Maaten & Hinton, 2008). This conditional probability is given by:

$$p_{j|i} = \frac{exp(-||x_i - x_j||^2/2\sigma_i^2)}{\sum_{k\neq i} exp(-||x_i - x_j||^2/2\sigma_i^2)} \tag{I.1}$$

Here, σ_i is the variance of the Gaussian centred on datapoint x_i[2]. This conditional probabiliy represents thus "the probability that x_i would pick x_j as its neighbor if neighbors were picked in proportion to their probability density under a Gaussian centered at x_i (p. 2581)". If the variance of the Gaussian is set to $\frac{1}{\sqrt{2}}$, the similarity of map point y_i to map point y_j is given by:

$$q_{j|i} = \frac{exp(-||x_i - x_j||^2)}{\sum_{k\neq i} exp(-||x_i - x_j||^2/)} \tag{I.2}$$

Accordingly, if the map points y_i and y_j correctly model the similarity between the high-dimensional datapoints x_i and x_j , the conditional probabilities $p_{j|i}$ and $q_{j|i}$ will be equal. The SNE technique aims to find a data representation that minimized the mismatch between $p_{j|i}$ and $q_{j|i}$.

[1]Dimensionality reduction methods are able to convert a high dimensional data set, $X = \{x_1, x_2 ... x_n\}$, into two-dimensional data set. $Y = \{y_1, y_2 ... y_n\}$, that can be easily visualized –for example in a scatterplot (van der Maaten & Hinton, 2008)

[2]The method for determining σ_i is presented in the article by van der Maaten and Hinton (2008)

This SNE technique is an often used method in order to do dimension reduction and to visualize high dimensional data. However, one disadvantage of this technique is the fact that it is hampered by a cost function that is difficult to optimize and a problem that is referred to as the 'crowding problem'[3]. In order to overcome these limitations, van der Maaten and Hinton (2008) altered the general SNE technique (which was originally developed by Hinton and Roweis (2002)):

> The cost function used by t-SNE differs from the one used by SNE in two ways: (1) it uses a symmetrized version of the SNE cost function with simpler gradients that was briefly introduced by Cook et al. (2007) and (2) it uses a Student-t distribution rather than a Gaussian to compute the similarity between two points in the low-dimensional space. t-SNE employs a heavy-tailed distribution in the low-dimensional space to alleviate both the crowding problem and the optimization problem of SNE. (van der Maaten & Hinton, 2008, p. 2583)

Without going into too much detail, experiments with the t-SNE technique proved that this technique is much better to optimize and produces significantly better visualizations than the SNE technique (van der Maaten & Hinton, 2008). This is especially true for high dimensional data that lie on different but related low-dimensional manifolds –such as the data of our experiments with the Extended Model Configuration. Hence, our choice for t-SNE.

[3]Explantion of these problems can be found in van der Maaten and Hinton (2008) page 2583

I.2 Flexibility

As figure I.1 shows, we can clearly distinct the states of the model based on the level of flexibility. For example, we see that large unemployment (a high circumference of the circles) correlates with a high level of flexibility. On the other hand, these states (with large unemployment) do not have the highest percentage of retrained workers by firms. Model states with a high percentage of retrained workers by firms often have a low unemployment rate and low flexibility (represented by the color black).

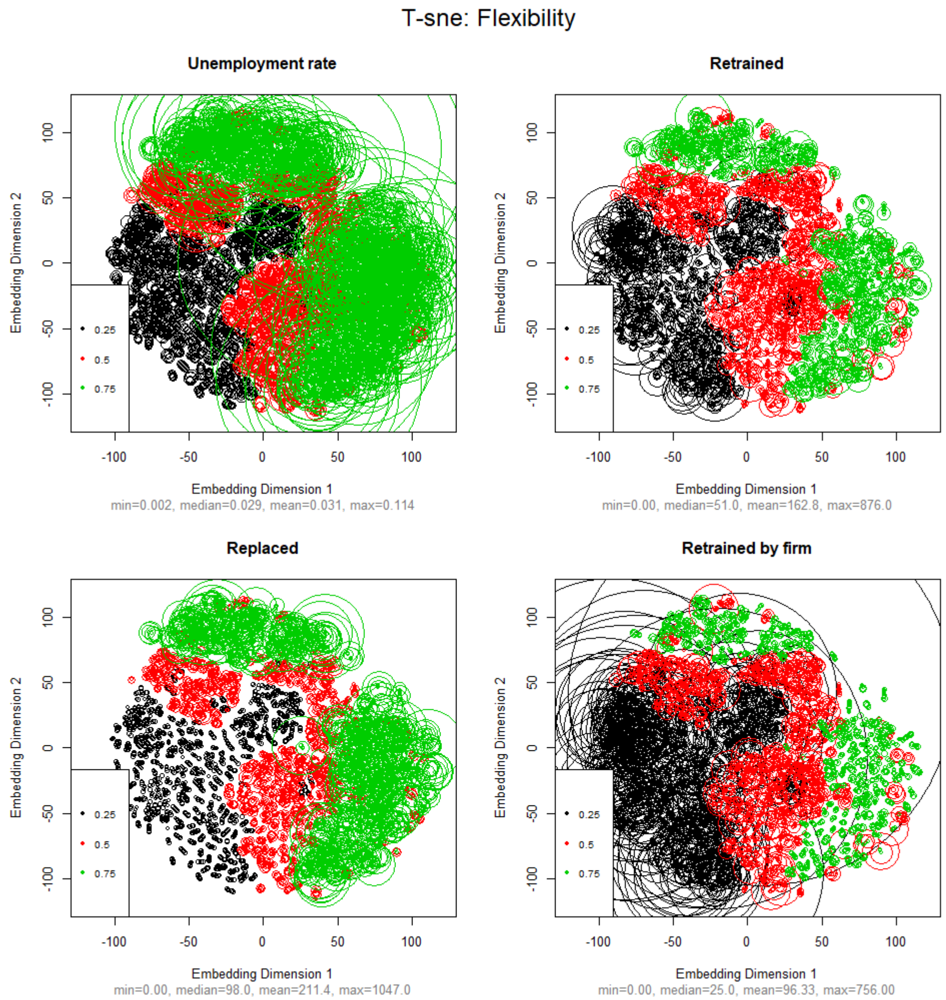


Figure I.1: t-SNE: States of the model and Flexibility

I.3 Information Policy

Whereas the level of flexibility clearly showed distinct states, no patterns can be found in the results of the t-SNE analysis of the information policy (see Figure I.2). While the size of the circles do show some patterns, this cannot be correlated with the level of information policy.

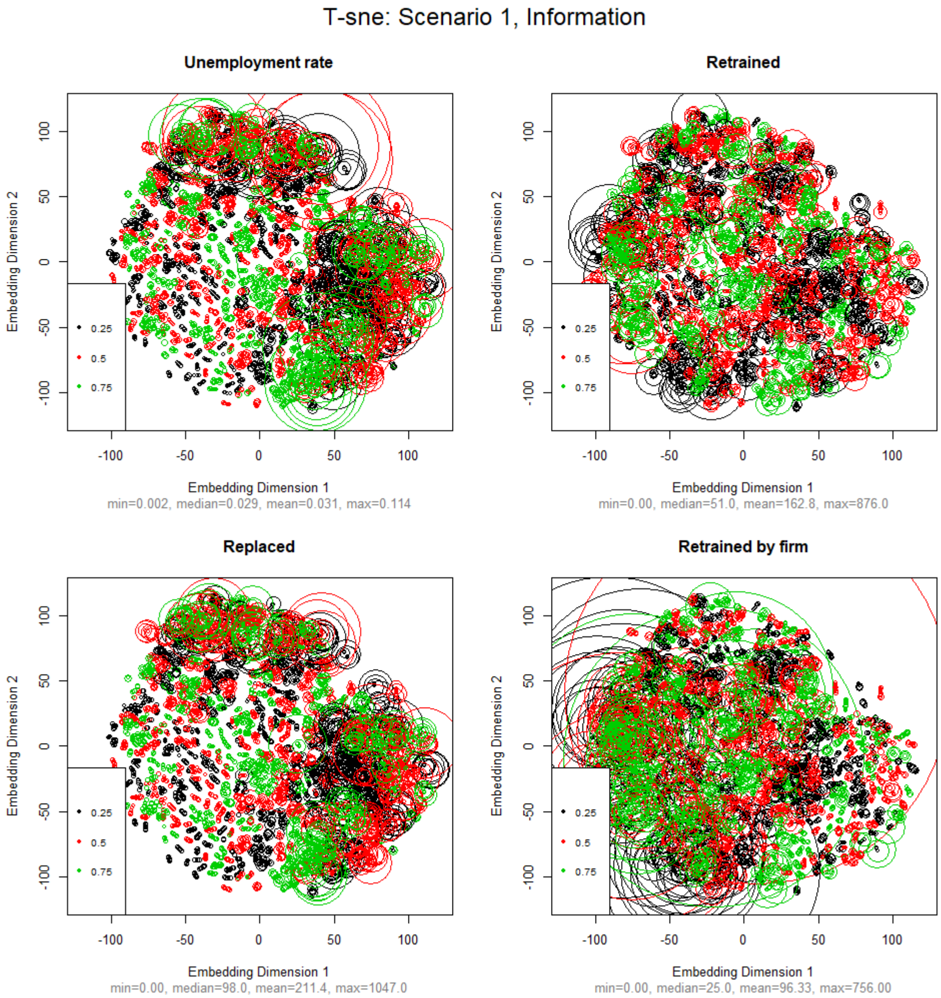


Figure I.2: t-SNE: States of the model and Information Policy

I.4 Retrain Subsidy

Just as with the t-SNE analysis of the information policy, no specific patterns can be found in the results of the t-SNE of the retrain subsidy.

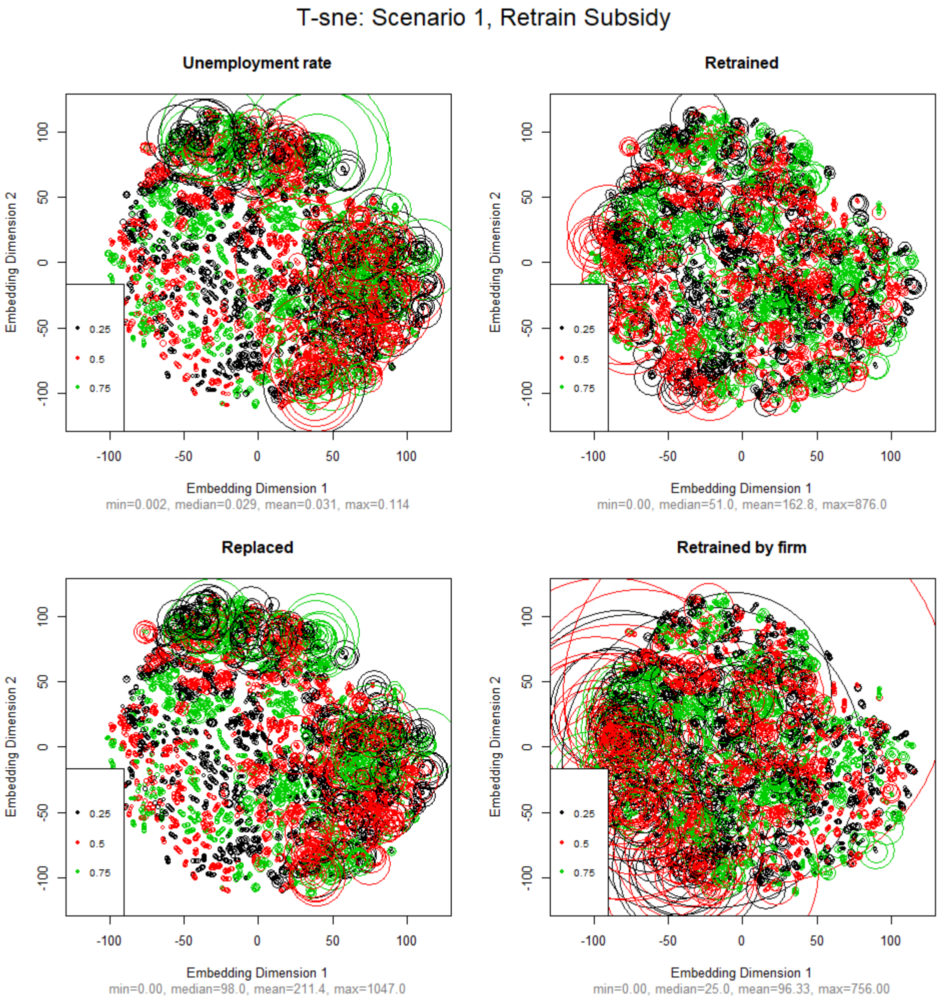


Figure I.3: t-SNE: States of the model and Retrain Subsidy

I.5 Correlation

Table I.1 shows the found correlations between the unemployment rate, retrained workers, replaced workers and retrained workers by firms. Here, we see that the highest correlation is between the unemployment rate and the number of replaced workers. The lowest correlation is between the percentage of retrained workers by firms and the percentage of replaced workers.

Table I.1: Correlation matrix: outcomes EMC

Outcome 1	Outcome 2	cor	p
retrained workers	unemployment rate	0.124	0
retrained workers	retrained by firm	0.775	0
unemployment rate	retrained by firm	-0.083	0
retrained workers	replaced workers	0.341	0
unemployment rate	replaced workers	0.875	0
retrained by firm	replaced workers	-0.044	0

Appendix J

Mediation Analysis

In order to see how the percentage of uncertain firms indirectly influences the model, we performed a causal mediation analysis. This chapter briefly explains this method and presents the results of this analysis.

J.1 Method

In order to see how the variable 'percentage' uncertain act as a mediator, we performed an mediation analysis in R. For this, we defined two linear regression models: fit_direct and fit_indirect. The fit_direct model depicts the effect of the model parameters on the mediator (here: percentage uncertain). The fit_indirect model depicts the whole effect of all parameters and the mediator on the model outcome (here: the unemployment rate). Furthermore, classified our data-set in two groups: one group of data where the percentage of uncertain firms is equal to zero; and one group were the percentage of uncertain firms is higher than zero. The latter can bee seen as the 'treatment group' whereas the former can bee seen as the 'control group'. By dividing our data set into these two groups we can determine the intermediary effects of the variable 'percentage uncertain'.

This is done as followed:

> Given these model objects [fit_direct and fit_indirect], the estimation proceeds by simulating the model parameters based on their approximate asymptotic distribution (i.e., the multivariate normal distribution with the mean equal to the parameter estimates and the variance equal to the asymptotic variance estimate), and then computing causal mediation effects of interest for each parameter draw. This method of inference can be viewed as an approximation to the Bayesian posterior distribution due to the Bernstein-Von Mises Theorem (King et al., 2000). (Imai et al., 2009, p. 3)

This method is implemented in R as a package which we have used for our analysis. Script J.1 shows the script used for the mediation analysis.

```
results_n$group <- ifelse(results_n$percentageuncertain == 0, 0, 1)

fit_direct <- lm(percentageuncertain ~ group + Flexibility +
    Information.Policy + retrain.subsidy, data=results_n)
fit_indirect <- lm(unemployment.rate ~ group + percentageuncertain
    + Flexibility + Information.Policy + retrain.subsidy, data=results_n)

med.out <- mediate(fit_direct, fit_indirect, mediator = "percentageuncertain", treat = "group",
                   control.value = 0, robustSE = TRUE, sims=100)
```

Listing J.1: Script Mediation analysis

J.2 Results

Table J.1 show the results of the mediation analysis. Here, ACME (Average Causal Mediated Effect) depicts the effect of the mediator –e.g. the effect of the mediator alone. The ADE (Average Direct Effect) depicts the unmediated effect and the Total Effect is the sum of ADE and ACME. The table shows us that the ACME is much higher than the ADE (0.011 vs 0.002). This means that a large proportion of the total effect of the policy levers on the unemployment rate arose from through the percentage of uncertain firms. In other words: the effect of the policy levers on the unemployment rate is for 87.8% determined by the indirect effect of the percentage uncertain firms.

Table J.1: Causal Mediation Analysis

	Estimate	95 %CI Lower	95% CI Upper	p-value
ACME	0.011	0.011	0.01	<2e-16
ADE	0.002	0.001	0.00	<2e-16
Total Effect	0.013	0.013	0.01	<2e-16
Prop. Mediated	0.877636	0.833	0.92	<2e-16

Appendix K

Policy identification

This appendix will present the results of the policy identification. Moreover, it also presents the results towards the parameter 'foresight variable'.

K.1 Flexibility

This section presents the results of the analysis towards the level of flexibility. As can be seen from K.1, there is a clear difference in the median of the unemployment rate between the levels of flexibility. Moreover, K.2 shows that the residuals for the unemployment rate are normally distributed.

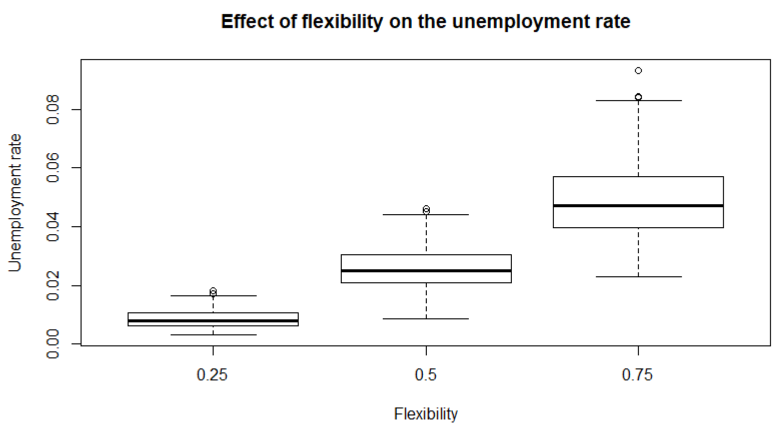


Figure K.1: Box plot: flexibility and unemployment rate

K.1.1 Generalized Linear model

This section present the results of the generalized linear model (with a 'Gaussian family') for the levels of flexibility. For this, the following model is estimated:

$$U = \beta_0 + \beta_1 Flex \times \beta_2 p_{replaced} \times \beta_3 p_{retrained} \times \beta_4 p_{retrainedfirm} \tag{K.1}$$

As can be seen from table K.1, the residual deviance is low. This means that the model is a good fit. Note, however, that for this analysis we only use regression to visualize patterns in our data and not to fit and predict data.

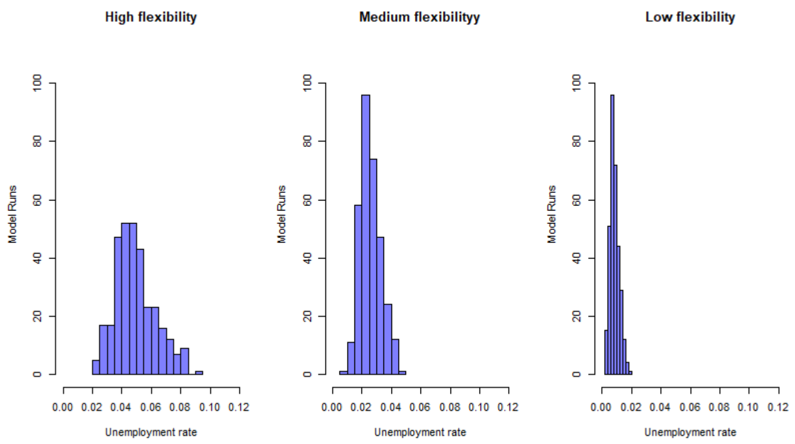


Figure K.2: Histogram: flexibility and unemployment rate

Table K.1: General Linear Model: Flexibility

	Intercept	Flexibility	percentage retrained workers	percentage replaced workers	percentage retrained by firms
Coefficients	0.010	-0.036	-0.002	0.010	0.002
Degrees of Freedom	971				
Residual Deviance	0.01572				

K.2 Information Policy

This section shows the results of the policy analysis of the policy parameter 'information policy'. As shown by figure K.6, variation of the information policy does not have a significant effect on the unemployment rate. Moreover, as can be seen from figure K.7, the values for the unemployment rate have a more or less Poisson distribution. In other words: it is not normally distributed.

K.3 Generalized Linear Model

This section presents the results of the regression analysis (with a 'Gaussian family') of the information policy. Here, the following regression model is estimated:

$$U = \beta_0 + \beta_1 IP \times \beta_2 p_{replaced} \times \beta_3 p_{retrained} \times \beta_4 p_{retrainedfirm} \tag{K.2}$$

Note, however, that the residuals are not normally distributed –hence, the assumption of normality is violated (see Figure K.7). However, since we do not use this model as a predictive data fitting model, we do not consider this as a problem. Table K.2 shows the summary statistics of the estimated model. As can be seen, the coefficients are very low indicating that the combination of these parameters do not contribute much to the explanation of variance of the unemployment rate.

K.4 Retrain Subsidy

In this section, we present the results of the policy analysis of the policy 'retrain subsidy'. In this analysis, we have found that this parameter does not influences the unemployment rate (see Figure K.11). Moreover, as figure K.12 shows, the values of the unemployment follow a Poisson distribution.

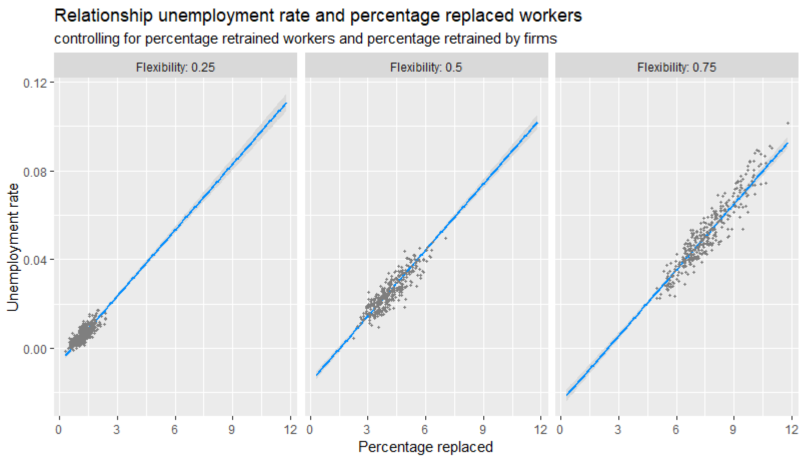


Figure K.3: GML: Percentage replaced

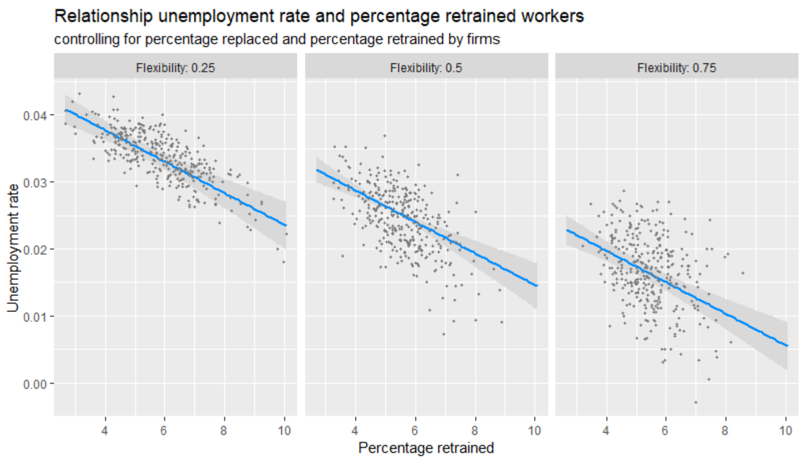


Figure K.4: GML: Percentage retrained workers

K.5 Generalized Linear Model

This section presents the results of the generalized linear model when taking the retrain subsidy into account. For this, the following regression model is estimated:

$$U = \beta_0 + \beta_1 RS \times \beta_2 p_{replaced} \times \beta_3 p_{retrained} \times \beta_4 p_{retrainedfirm} \tag{K.3}$$

As can be seen from table K.3, the coefficients for this estimated regression model are very low. Nevertheless, the low residual deviance indicates to a 'good fit' of the model.

K.6 Foresight variable

While not a policy parameter, this section presents the results of the analysis of the foresight variable. As can be seen from figure K.16, changes in the value of the foresight variable have an influence on the unemployment rate: the higher the foresight variable, the lower the unemployment rate. Moreover, as can be seen from figure K.17, the values of the unemployment rate, separated per foresight variable, follow a Poisson distribution.

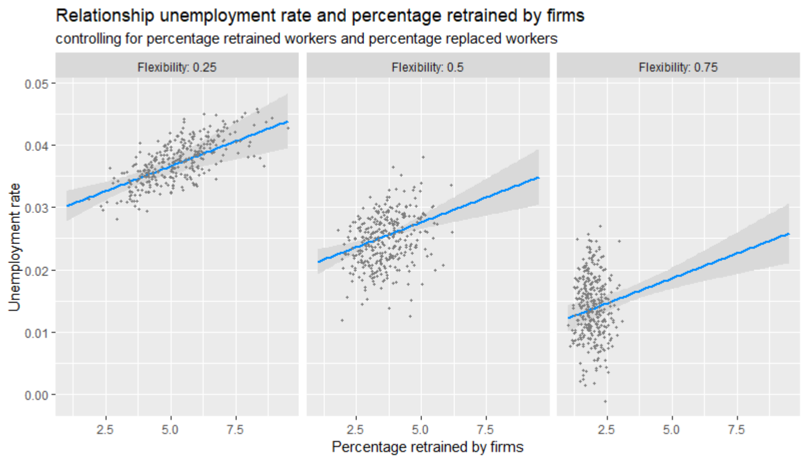


Figure K.5: GML: Percentage retrained workers by firms

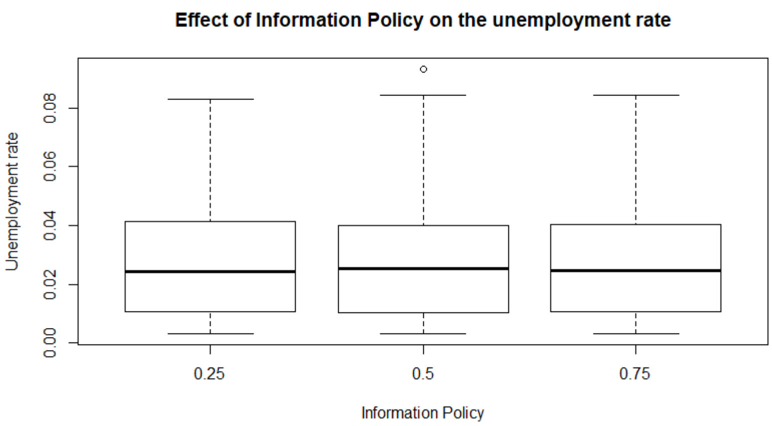


Figure K.6: Box plot: Information Policy and unemployment rate

K.6.1 Generalized Linear Model

This section presents the results of the regression analsyis of the foresight variable. For this, the following regression model is estimated:

$$U = \beta_0 + \beta_1 FS \times \beta_2 p_{replaced} \times \beta_3 p_{retrained} \times \beta_4 p_{retrainedfirm} \tag{K.4}$$

As can be seen from table K.4, the coefficients are rather low. Nevertheless, the low value of the residual deviance indicates that the model is a good fit.

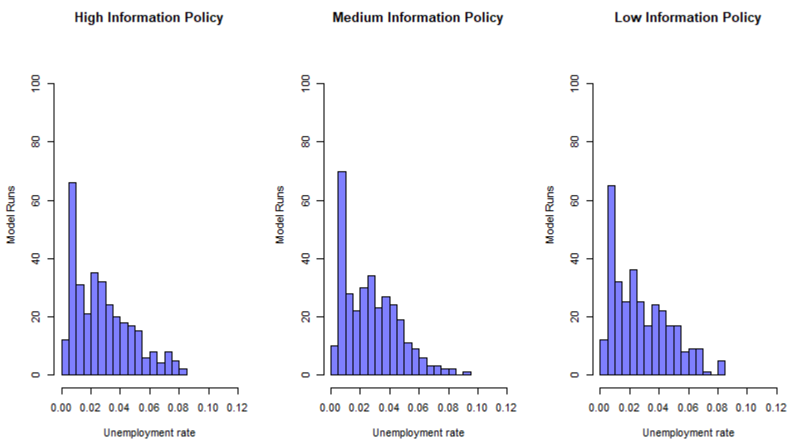


Figure K.7: Histogram: Information Policy and unemployment rate

Table K.2: General Linear Model: Information Policy

	Intercept	Information Policy	percentage retrained workers	percentage replaced workers	percentage retrained by firms
Coefficients	-0.003	-0.002	-0.003	0.008	0.002
Degrees of Freedom	971				
Residual Deviance	0.01995				

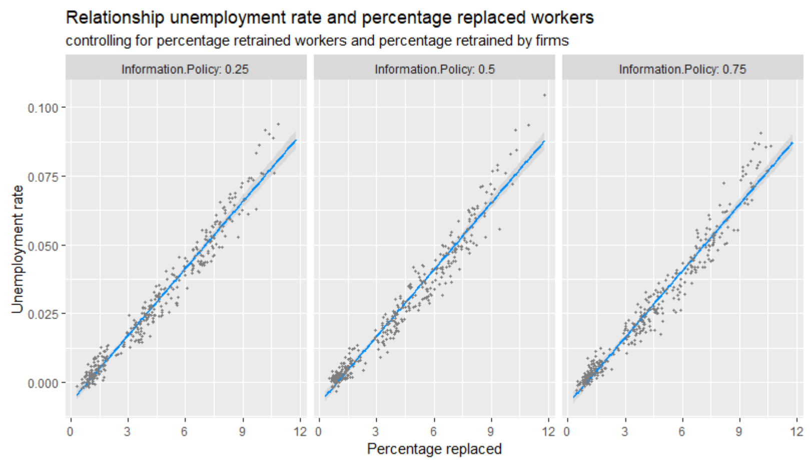


Figure K.8: GML:Information Policy & Percentage replaced

Table K.3: General Linear Model: Retrain Subsidy

	Intercept	Retrain subsidy	percentage retrained workers	percentage replaced workers	percentage retrained by firms
Coefficients	-0.004	-0.001	-0.002	0.008	0.002
Degrees of Freedom	971				
Residual Deviance	0.02004				

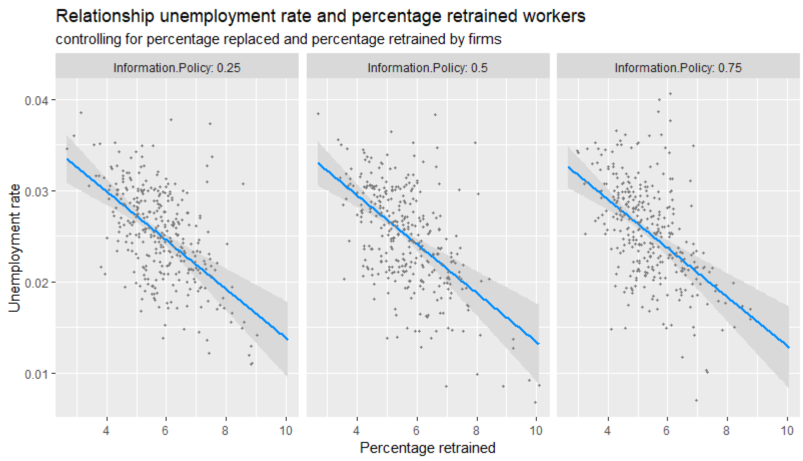


Figure K.9: GML: Information Policy & Percentage retrained workers

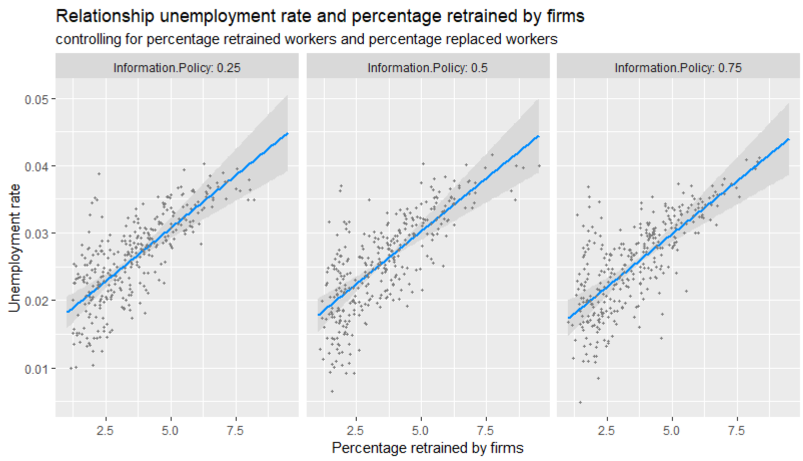


Figure K.10: GML: Information policy & Percentage retrained workers by firms

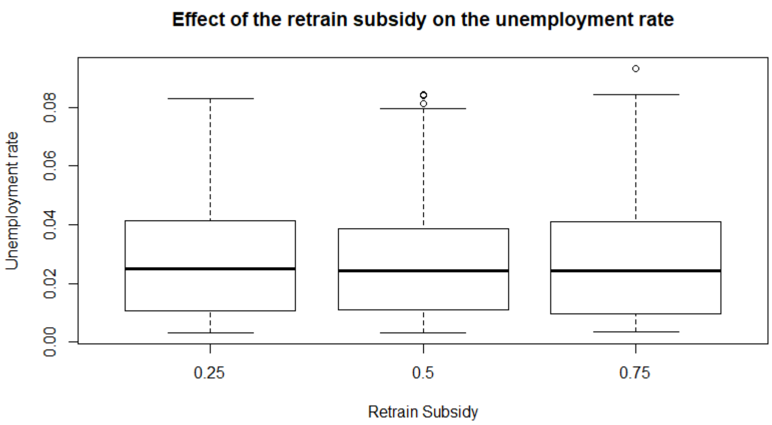


Figure K.11: Box plot: Retrain Subsidy and unemployment rate

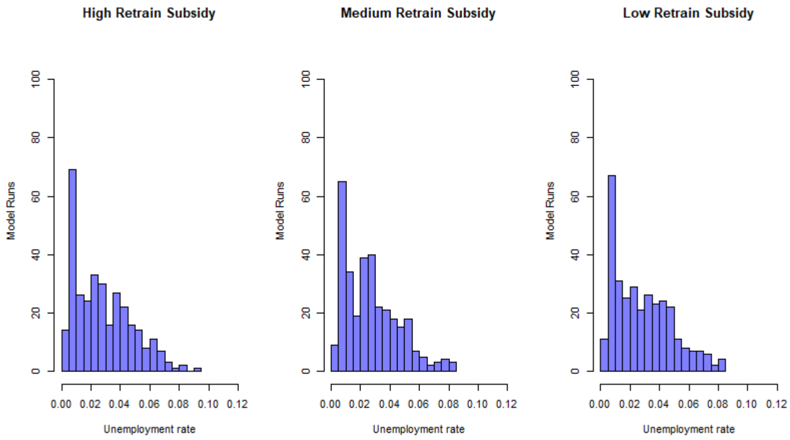


Figure K.12: Histogram: Retrain Subsidy and unemployment rate

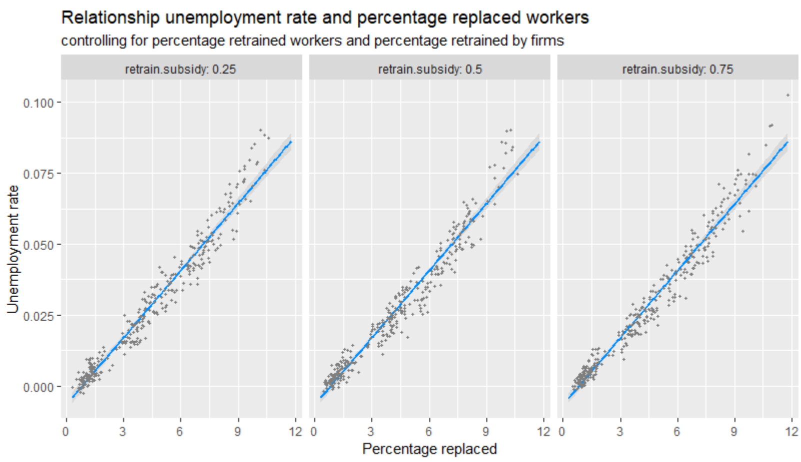


Figure K.13: GML: Retrain subsidy & Percentage replaced

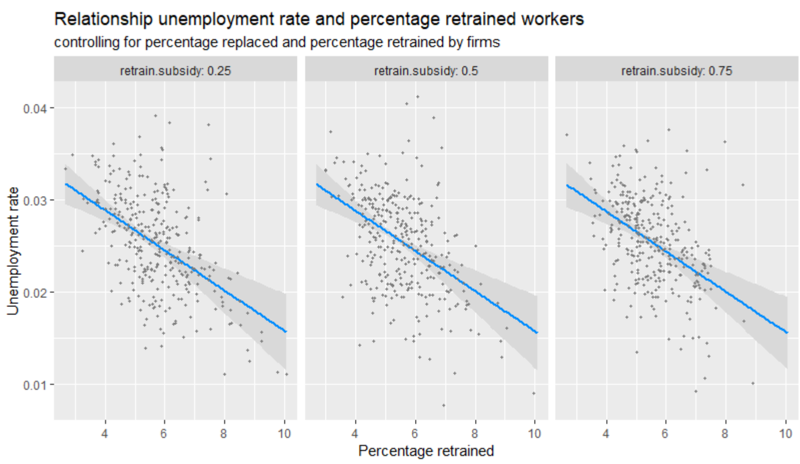


Figure K.14: GML: Retrain subsidy & Percentage retrained workers

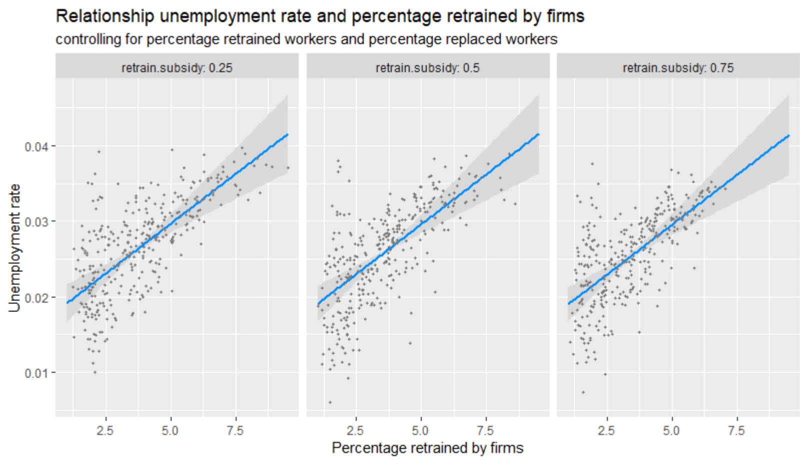


Figure K.15: GML: Retrain subsidy & Percentage retrained workers by firms

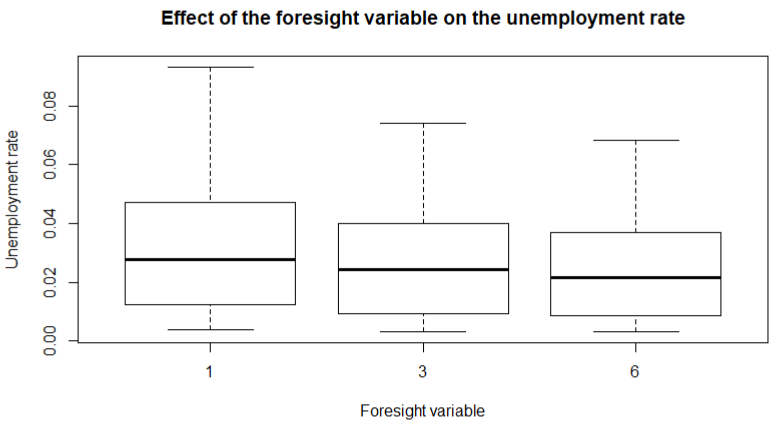


Figure K.16: Box plot: Foresight variable and unemployment rate

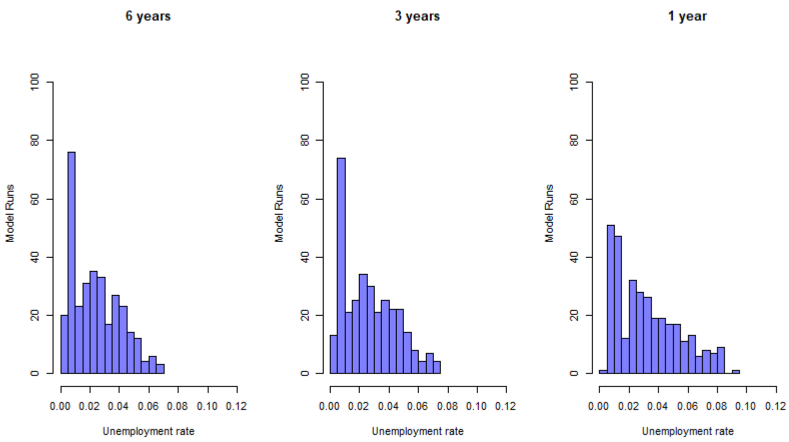


Figure K.17: Histogram: Foresight variable and unemployment rate

Table K.4: General Linear Model: Foresight Variable

	Intercept	Foresight Variable	percentage retrained workers	percentage replaced workers	percentage retrained by firms
Coefficients	-0.003	-0.001	-0.002	0.008	0.003
Degrees of Freedom	971				
Residual Deviance	0.01276				

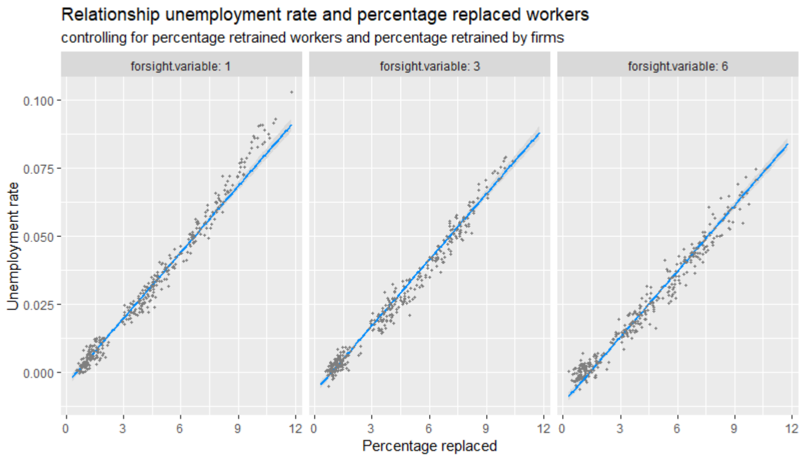


Figure K.18: GML: Foresight variabley & Percentage replaced

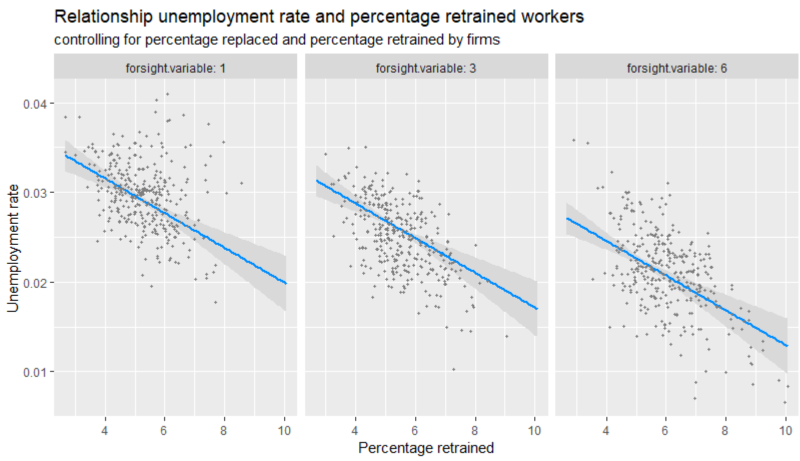


Figure K.19: GML: Foresight variable & Percentage retrained workers

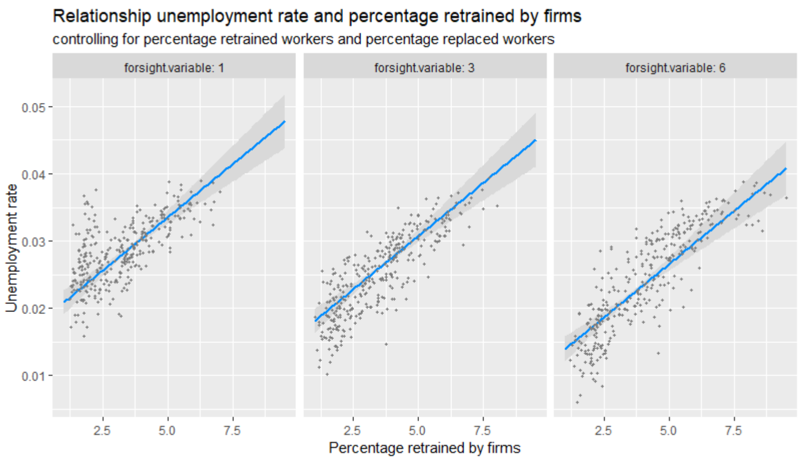


Figure K.20: GML: Foresight variable & Percentage retrained workers by firms

Appendix L

Scenario Analysis

This appendix will present the results of the scenario analysis. This appendix starts with presenting the results of the scenario analysis of economic growth. Next, it presents the results of the two different technological change scenario's.

L.1 Economic growth

In this section, the results of the scenario analysis of economic growth will be presented. As can be seen from figure L.1, a different economic growth does not affect the unemployment rate. It does, however, slightly affects the percentage of uncertain firms (see figure L.2).

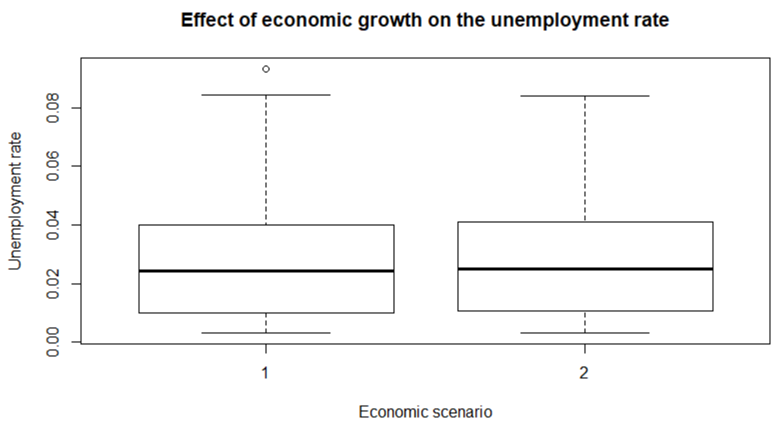


Figure L.1: Box plot: Unemployment rate per scenario

L.2 Endogenous or Endogenous technological change

This section presents the result of the scenario analysis of the two different scenario's for technological change. As can be seen from table L.1, there is a slight difference in the mean unemployment rate between the two different scenario's . Moreover, there is a slight difference in the percentage of automated jobs between both scenario's where more jobs are automated in the endogenous-TC scenario (see Table L.2). Table L.3 shows the frequency (in %) of being automated for each job.

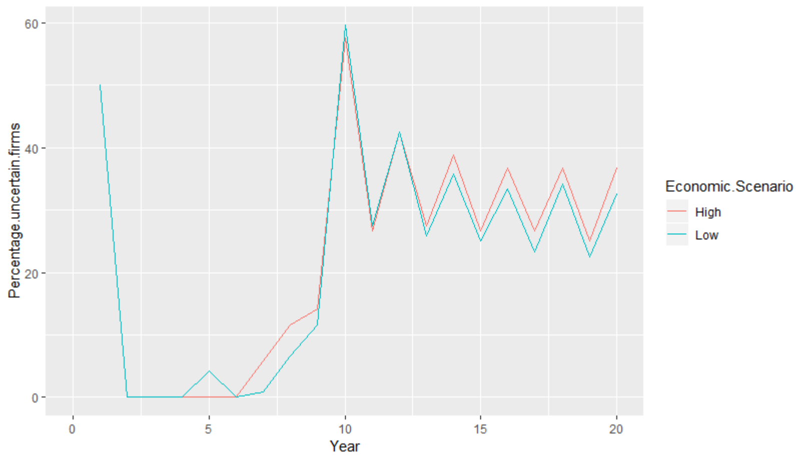


Figure L.2: Plot: Median percentage uncertain firms per scenario

Table L.1: Statistics: Unemployment rate endogenous and exogenous TC

	Min	1st Qu.	Median	Mean	3rd Qu.	Max
Exogenous TC	0.003	0.010	0.024	0.027	0.041	0.084
Endogenous TC	0.003	0.010	0.024	0.028	0.040	0.093

Table L.3: Percentage job automated: Endogenous TC and Exogenous TC

Job	Endogenous TC % automated	Exogenous TC % automated
0111 Docenten hoger onderwijs en hoog...	0.18	0.19
0112 Docenten beroepsgerichte vakken ...	0.08	0.09
0113 Docenten algemene vakken secunda...	0.06	0.06
0114 Leerkrachten basisonderwijs	0.05	0.06
0115 Onderwijskundigen en overige doc...	0.1	0.11
0121 Beroepsgroep sportinstructeurs	0.13	0.1
0131 Leidsters kinderopvang en onderw...	0.08	0.07
0211 Bibliothecarissen en conservatoren	0.09	0.09
0212 Auteurs en taalkundigen	0.14	0.13
0213 Journalisten	0.06	0.06
0214 Beeldend kunstenaars	0.09	0.09
0215 Uitvoerend kunstenaars	0.08	0.07
0221 Grafisch vormgevers en producton -Commercial and Industrial Designers	0.03	0.02
0221 Grafisch vormgevers en producton. -Graphic Designers	0.05	0.04
0222 Fotografen en interieurontwerpers	0.08	0.08
0311 Adviseurs marketing, public rela...	0.05	0.05
0321 Vertegenwoordigers en inkopers	0.13	0.13
0331 Winkeliers en teamleiders detail...	0.06	0.07
0332 Verkoopmedewerkers detailhandel	0.07	0.08
0333 Kassamedewerkers	0.3	0.31
0334 Callcentermedewerkers outbound e...	0.26	0.26
0411 Accountants	0.26	0.25
0412 Financieel specialisten en economen	0.24	0.23
0413 Bedrijfskundigen en organisatiea.	0.15	0.14
0414 Beleidsadviseurs	0.17	0.15
0415 Specialisten personeels- en loop.	0.24	0.22
0421 Boekhouders	0.23	0.2

Table L.3: Percentage job automated: Endogenous TC and Exogenous TC

Job	Endogenous TC % automated	Exogenous TC % automated
0422 Zakelijke dienstverleners	0.14	0.16
0423 Directiesecretaresses	0.29	0.27
0431 Administratief medewerkers	0.31	0.27
0432 Secretaresses	0.29	0.26
0433 Receptionisten en telefonisten	0.15	0.15
0434 Boekhoudkundig medewerkers	0.3	0.3
0435 Transportplanners en logistiek m.	0.26	0.24
0511 Algemeen directeuren	0.03	0.03
0521 Managers zakelijke en administra.	0.06	0.06
0522 Managers verkoop en marketing	0.04	0.05
0531 Managers productie	0.09	0.08
0532 Managers logistiek	0.07	0.08
0533 Managers ICT	0.11	0.09
0534 Managers zorginstellingen	0.02	0.01
0535 Managers onderwijs	0.01	0.01
0536 Managers gespecialiseerde dienst.	0.06	0.09
0541 Managers horeca	0.04	0.05
0542 Managers detail- en groothandel	0.06	0.03
0543 Managers commerciĂŤle en persoonl.	0.06	0.07
0611 Overheidsbestuurders	0.04	0.04
0612 Overheidsambtenaren- office	0.18	0.23
0612 Overheidsambtenaren- operational	0.12	0.1
0621 Juristen	0.05	0.05
0631 Politie-inspecteurs	0.07	0.07
0632 Politie en brandweer	0.08	0.07
0633 Beveiligingspersoneel	0.24	0.21
0711 Biologen en natuurwetenschappers	0.15	0.18
0712 Ingenieurs (geen elektrotechniek)	0.12	0.13
0713 Elektrotechnisch ingenieurs	0.14	0.17
0714 Architecten	0.09	0.1
0721 Technici bouwkunde en natuur	0.14	0.14
0722 Productieleiders industrie en bouw	0.08	0.07
0723 Procesoperators	0.18	0.18
0731 Bouwarbeiders ruwbouw	0.17	0.14
0732 Timmerlieden	0.13	0.11
0733 Bouwarbeiders afbouw	0.15	0.16
0734 Loodgieters en pijpfitters	0.12	0.11
0735 Schilders en metaalspuiters	0.22	0.22
0741 Metaalbewerkers en constructiewe.	0.24	0.22
0742 Lassers en plaatwerkers	0.23	0.22
0743 Automonteurs	0.13	0.13
0744 Machinemonteurs	0.1	0.1
0751 Slagers	0.26	0.26
0752 Bakkers	0.19	0.19
0753 Productcontroleurs	0.19	0.19
0754 Meubelmakers, kleermakers en sto.	0.23	0.21
0755 Medewerkers drukkerij en kunstni.	0.15	0.16
0761 Elektriciens en elektronicamonteurs	0.15	0.16
0771 Productiemachinebedieners	0.24	0.24
0772 Assemblagemedewerkers	0.23	0.26
0781 Hulpkrachten bouw en industrie	0.23	0.26
0811 Software- en applicatieontwikkel.	0.18	0.21

Table L.3: Percentage job automated: Endogenous TC and Exogenous TC

Job	Endogenous TC % automated	Exogenous TC % automated
0812 Databank- en netwerkspecialisten	0.22	0.2
0821 Gebruikersondersteuning ICT	0.2	0.19
0822 Radio- en televisietechnici	0.11	0.1
0911 Land- en bosbouwers	0.2	0.23
0912 Hoveniers, tuinders en kwekers	0.21	0.2
0913 Veetelers	0.17	0.19
0921 Hulpkrachten landbouw	0.24	0.26
1011 Artsen	0.05	0.06
1012 Gespecialiseerd verpleegkundigen	0.08	0.08
1013 Fysiotherapeuten	0.06	0.05
1021 Maatschappelijk werkers	0.05	0.04
1022 Psychologen en sociologen	0.05	0.06
1031 Laboranten	0.16	0.16
1032 Apothekersassistenten	0.22	0.2
1033 Verpleegkundigen (mbo)	0.11	0.11
1034 Medisch praktijkassistenten	0.17	0.16
1035 Medisch vakspecialisten	0.13	0.15
1041 Sociaal werkers, groeps- en woon.	0.29	0.25
1051 Verzorgenden	0.15	0.18
1111 Reisbegeleiders	0.14	0.14
1112 Koks	0.15	0.14
1113 Kelners en barpersoneel	0.1	0.1
1114 Kappers en schoonheidsspecialisten	0.13	0.12
1115 ConciĂĪrges en teamleiders schoon.	0.27	0.26
1116 Verleners van overige persoonlij.	0.17	0.15
1121 Schoonmakers	0.32	0.27
1122 Keukenhulpen	0.28	0.28
1211 Dekofficieren en piloten	0.1	0.08
1212 Chauffeurs auto’s, taxi’s en bes.	0.25	0.25
1213 Buschauffeurs en trambestuurders	0.23	0.22
1214 Vrachtwagenchauffeurs - First-Line Supervisors of Transportation and Material-Moving Machine and Vehicle Operators	0.14	0.14
1214 Vrachtwagenchauffeurs -Heavy and Tractor-Trailer Truck Drivers	0.2	0.21
1215 Bedieners mobiele machines	0.25	0.26
1221 Laders, lossers en vakkenvullers	0.26	0.26
1222 Vuilnisophalers en dagbladenbezo.	0.26	0.26
1311 Beroepsgroep overig - managing	0.03	0.05
1311 Beroepsgroep overig - operational	0.23	0.24

Table L.2: Statistics: percentage automated jobs endogenous and exogenous TC

	Min	1st Qu.	Median	Mean	3rd Qu.	Max
Exogenous TC	0.0000	0.0700	0.1400	0.1437	0.2125	0.3100
Endogenous TC	0.0000	0.0775	0.1400	0.1454	0.2225	0.3200

Appendix M

Unfilled demand

This appendix presents the analysis towards the total labour demand and unfilled demand for jobs. As can be seen from figure M.1, both the total demand as its proportion of unfillled demand increases over time. Table M.1 shows the total demand for each job and table M.2 shows the proportion of total unfilled demand for each job.

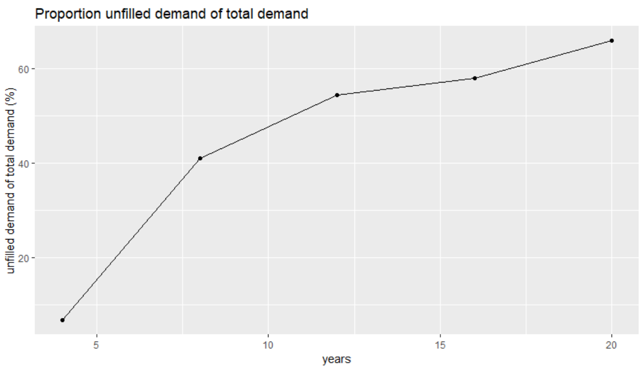


Figure M.1: Proportion of unfilled demand of total demand

Table M.1: Proportion of labour demand to total labour demand of each job

Job	4 Years	8 Years	12 Years	16 Years	20 Years	difference 20-4 years
0511 Algemeen directeuren	0.968	1.2	1.53	2.035	2.302	1.334
0311 Adviseurs marketing, public rela...	1.99	2.133	2.601	2.99	3.027	1.037
0321 Vertegenwoordigers en inkopers	1.801	1.989	2.428	2.597	2.618	0.817
0712 Ingenieurs (geen elektrotechniek)	1.432	1.562	1.944	2.117	2.237	0.805
0521 Managers zakelijke en administra.	0.932	0.96	1.215	1.538	1.726	0.794
0722 Productieleiders industrie en bouw	0.678	0.769	0.991	1.275	1.442	0.764
0721 Technici bouwkunde en natuur	1.289	1.409	1.743	1.911	2.034	0.745
0522 Managers verkoop en marketing	0.581	0.635	0.827	1.099	1.276	0.695
0413 Bedrijfskundigen en organisatiea.	1.54	1.644	2.021	2.141	2.215	0.675
0531 Managers productie	0.558	0.658	0.847	1.071	1.216	0.658
1311 Beroepsgroep overig - managing	0.463	0.514	0.682	0.93	1.092	0.629
0761 Elektriciens en elektronicamonteurs	1.031	1.172	1.443	1.56	1.652	0.621
1021 Maatschappelijk werkers	0.827	0.82	1.024	1.278	1.429	0.602
0744 Machinemonteurs	0.572	0.651	0.843	1.034	1.166	0.594
0433 Receptionisten en telefonisten	1.984	2.112	2.588	2.613	2.555	0.571
0422 Zakelijke dienstverleners	1.188	1.234	1.51	1.587	1.674	0.486

0215 Uitvoerend kunstenaars	0.627	0.652	0.813	1.001	1.11	0.483
0732 Timmerlieden	1.003	0.986	1.206	1.349	1.434	0.431
0734 Loodgieters en pijpfitters	0.476	0.525	0.662	0.8	0.891	0.415
0812 Databank- en netwerkspecialisten	0.861	0.959	1.208	1.181	1.234	0.373
0621 Juristen	0.807	0.688	0.845	1.058	1.179	0.372
0414 Beleidsadviseurs	0.861	0.868	1.077	1.148	1.209	0.348
0821 Gebruikersondersteuning ICT	0.563	0.593	0.775	0.801	0.87	0.307
0115 Onderwijskundigen en overige doc...	0.795	0.682	0.852	0.976	1.049	0.254
1116 Verleners van overige persoonlij.	0.432	0.443	0.575	0.618	0.67	0.238
0536 Managers gespecialiseerde dienst.	0.33	0.291	0.372	0.482	0.567	0.237
0912 Hoveniers, tuinders en kwekers	0.856	0.873	1.052	1.025	1.062	0.206
1022 Psychologen en sociologen	0.843	0.684	0.83	0.969	1.04	0.197
0212 Auteurs en taalkundigen	0.389	0.335	0.429	0.507	0.572	0.183
0781 Hulpkrachten bouw en industrie	0.869	1	1.123	1.039	1.034	0.165
0733 Bouwarbeiders afbouw	0.433	0.386	0.483	0.544	0.597	0.164
0822 Radio- en televisietechnici	0.239	0.208	0.269	0.324	0.371	0.132
1311 Beroepsgroep overig - operational	0.46	0.526	0.603	0.578	0.586	0.126
0714 Architecten	0.449	0.334	0.414	0.504	0.57	0.121
0415 Specialisten personeels- en loop.	1.19	1.207	1.389	1.313	1.294	0.104
0432 Secretaresses	0.686	0.748	0.863	0.798	0.781	0.095
1222 Vuilnisophalers en dagbladenbezo.	1.022	1.15	1.311	1.132	1.114	0.092
0331 Winkeliers en teamleiders detail...	1.325	1.033	1.207	1.364	1.412	0.087
0754 Meubelmakers, kleermakers en sto.	0.459	0.455	0.534	0.536	0.545	0.086
0412 Financieel specialisten en economen	0.974	1.009	1.148	1.062	1.056	0.082
0633 Beveiligingspersoneel	0.729	0.716	0.83	0.799	0.806	0.077
0771 Productiemachinebedieners	0.952	1.011	1.133	1.046	1.029	0.077
0742 Lassers en plaatwerkers	0.45	0.445	0.516	0.516	0.525	0.075
0752 Bakkers	0.255	0.234	0.295	0.3	0.325	0.07
0743 Automonteurs	0.68	0.551	0.653	0.716	0.747	0.067
1115 ConciĂŤrges en teamleiders schoon.	0.534	0.572	0.663	0.604	0.586	0.052
0221 Grafisch vormgevers en producton... - Graphic Designers	0.203	0.147	0.18	0.225	0.254	0.051
1215 Bedieners mobiele machines	0.919	0.999	1.125	1.003	0.968	0.049
1214 Vrachtwagenchauffeurs -First-Line Supervisors of Transportation and Material-Moving Machine and Vehicle Operators	0.324	0.241	0.297	0.337	0.37	0.046
0421 Boekhouders	1.334	1.369	1.515	1.398	1.379	0.045
0411 Accountants	1.192	1.288	1.422	1.268	1.232	0.04
0213 Journalisten	0.35	0.261	0.306	0.365	0.389	0.039
0753 Productcontroleurs	0.237	0.191	0.241	0.251	0.265	0.028
0611 Overheidsbestuurders	0.243	0.162	0.192	0.237	0.264	0.021
0543 Managers commerciĂŤle en persoonl.	0.145	0.102	0.124	0.145	0.157	0.012
0532 Managers logistiek	0.195	0.125	0.149	0.178	0.202	0.007
0755 Medewerkers drukkerij en kunstni.	0.334	0.251	0.285	0.321	0.341	0.007
0214 Beeldend kunstenaars	0.14	0.099	0.113	0.131	0.139	-0.001
0533 Managers ICT	0.218	0.155	0.181	0.203	0.216	-0.002
1214 Vrachtwagenchauffeurs -Heavy and Tractor-Trailer Truck Drivers	0.328	0.247	0.305	0.311	0.325	-0.003
0423 Directiesecretaresses	1.348	1.482	1.655	1.43	1.343	-0.005
0221 Grafisch vormgevers en producton... - Commercial and Industrial Designers	0.061	0.039	0.044	0.048	0.046	-0.015
0713 Elektrotechnisch ingenieurs	0.198	0.131	0.156	0.169	0.179	-0.019
0211 Bibliothecarissen en conservatoren	0.079	0.047	0.055	0.06	0.059	-0.02
0751 Slagers	0.229	0.184	0.199	0.195	0.196	-0.033
0631 Politie-inspecteurs	0.125	0.074	0.083	0.089	0.088	-0.037
1111 Reisbegeleiders	0.237	0.161	0.188	0.195	0.2	-0.037

1035 Medisch vakspecialisten	0.607	0.426	0.508	0.544	0.564	-0.043
0921 Hulpkrachten landbouw	0.315	0.253	0.278	0.267	0.264	-0.051
0612 Overheidsambtenaren- operational	0.212	0.129	0.145	0.155	0.159	-0.053
0535 Managers onderwijs	0.192	0.109	0.124	0.137	0.135	-0.057
1112 Koks	0.812	0.629	0.741	0.75	0.751	-0.061
0222 Fotografen en interieurontwerpers	0.27	0.159	0.178	0.194	0.203	-0.067
0534 Managers zorginstellingen	0.229	0.143	0.161	0.168	0.162	-0.067
0735 Schilders en metaalspuiters	0.371	0.273	0.305	0.305	0.304	-0.067
0121 Beroepsgroep sportinstructeurs	0.528	0.35	0.406	0.443	0.455	-0.073
0131 Leidsters kinderopvang en onderw...	1.611	1.191	1.366	1.51	1.538	-0.073
0541 Managers horeca	0.268	0.163	0.177	0.185	0.186	-0.082
0723 Procesoperators	0.295	0.187	0.199	0.212	0.213	-0.082
0542 Managers detail- en groothandel	0.42	0.256	0.276	0.324	0.334	-0.086
0612 Overheidsambtenaren- office	0.214	0.132	0.142	0.136	0.127	-0.087
0911 Land- en bosbouwers	0.316	0.207	0.222	0.227	0.22	-0.096
1211 Dekofficieren en piloten	0.32	0.191	0.206	0.216	0.217	-0.103
1213 Buschauffeurs en trambestuurders	0.245	0.147	0.151	0.147	0.136	-0.109
0435 Transportplanners en logistiek m.	2.239	2.431	2.712	2.304	2.126	-0.113
0112 Docenten beroepsgerichte vakken ...	0.366	0.214	0.242	0.255	0.251	-0.115
0434 Boekhoudkundig medewerkers	1.709	1.894	2.054	1.749	1.584	-0.125
1032 Apothekersassistenten	0.284	0.174	0.192	0.178	0.157	-0.127
0811 Software- en applicatieontwikkel.	2.768	2.548	2.99	2.755	2.636	-0.132
0731 Bouwarbeiders ruwbouw	0.701	0.475	0.534	0.559	0.56	-0.141
0711 Biologen en natuurwetenschappers	0.429	0.258	0.283	0.284	0.28	-0.149
0334 Callcentermedewerkers outbound e...	1.473	1.478	1.617	1.379	1.314	-0.159
0772 Assemblagemedewerkers	0.478	0.337	0.347	0.333	0.318	-0.16
0632 Politie en brandweer	0.469	0.274	0.29	0.31	0.306	-0.163
1031 Laboranten	0.299	0.154	0.159	0.148	0.123	-0.176
0741 Metaalbewerkers en constructiewe.	0.466	0.307	0.294	0.288	0.27	-0.196
1212 Chauffeurs auto's, taxi's en bes.	0.723	0.616	0.614	0.535	0.489	-0.234
0913 Veetelers	0.626	0.378	0.409	0.417	0.389	-0.237
1114 Kappers en schoonheidsspecialisten	0.844	0.536	0.587	0.606	0.589	-0.255
0111 Docenten hoger onderwijs en hoog...	0.636	0.372	0.405	0.383	0.357	-0.279
1113 Kelners en barpersoneel	2.686	2.293	2.656	2.48	2.345	-0.341
1013 Fysiotherapeuten	0.862	0.465	0.496	0.476	0.428	-0.434
1033 Verpleegkundigen (mbo)	0.783	0.407	0.418	0.376	0.317	-0.466
1034 Medisch praktijkassistenten	0.851	0.45	0.472	0.432	0.382	-0.469
0113 Docenten algemene vakken secunda...	1.193	0.669	0.709	0.728	0.677	-0.516
1122 Keukenhulpen	0.821	0.49	0.361	0.331	0.292	-0.529
1011 Artsen	1.327	0.748	0.8	0.764	0.68	-0.647
0333 Kassamedewerkers	1.023	0.612	0.398	0.357	0.319	-0.704
1012 Gespecialiseerd verpleegkundigen	1.411	0.749	0.769	0.698	0.592	-0.819
0332 Verkoopmedewerkers detailhandel	3.895	3.066	3.484	3.237	3.033	-0.862
0114 Leerkrachten basisonderwijs	1.761	0.93	0.962	0.938	0.824	-0.937
0431 Administratief medewerkers	3.377	3.33	3.412	2.73	2.36	-1.017
1121 Schoonmakers	3.026	2.911	2.889	2.254	1.913	-1.113
1041 Sociaal werkers, groeps- en woon.	2.789	2.351	1.992	1.547	1.295	-1.494
1221 Laders, lossers en vakkenvullers	2.953	2.541	2.258	1.647	1.32	-1.633
1051 Verzorgenden	2.681	1.409	1.427	1.183	0.97	-1.711

Table M.2: Proportion unfilled demand to total labour demand per Job

Job	4 Years	8 Years	12 Years	16 Years	20 Years	difference 20-4 Years
0311 Adviseurs marketing, public rela...	0.215	1.25	1.711	2.131	2.311	2.096
0511 Algemeen directeuren	0.148	0.789	1.117	1.648	1.989	1.841
0321 Vertegenwoordigers en inkopers	0.183	1.184	1.618	1.83	1.989	1.806
0433 Receptionisten en telefonisten	0.223	1.237	1.713	1.802	1.893	1.67
0712 Ingenieurs (geen elektrotechniek)	0.152	0.922	1.299	1.519	1.755	1.603
0413 Bedrijfskundigen en organisatiea.	0.154	0.952	1.329	1.499	1.697	1.543
0811 Software- en applicatieontwikkel.	0.19	1.269	1.717	1.615	1.708	1.518
0721 Technici bouwkunde en natuur	0.135	0.832	1.165	1.376	1.603	1.468
0332 Verkoopmedewerkers detailhandel	0.208	1.241	1.651	1.529	1.622	1.414
0521 Managers zakelijke en administra.	0.094	0.537	0.791	1.141	1.405	1.311
1113 Kelners en barpersoneel	0.16	1.038	1.393	1.309	1.389	1.229
0761 Elektriciens en elektronicamonteurs	0.089	0.7	0.966	1.113	1.294	1.205
0422 Zakelijke dienstverleners	0.109	0.694	0.968	1.089	1.276	1.167
0435 Transportplanners en logistiek m.	0.244	1.441	1.734	1.405	1.411	1.167
0722 Productieleiders industrie en bouw	0.079	0.463	0.684	0.988	1.212	1.133
1021 Maatschappelijk werkers	0.066	0.434	0.638	0.919	1.14	1.074
0522 Managers verkoop en marketing	0.055	0.366	0.556	0.847	1.074	1.019
0732 Timmerlieden	0.075	0.519	0.735	0.913	1.082	1.007
0431 Administratief medewerkers	0.291	1.805	1.982	1.393	1.297	1.006
0531 Managers productie	0.065	0.404	0.591	0.83	1.021	0.956
0744 Machinemonteurs	0.059	0.387	0.579	0.787	0.966	0.907
0131 Leidsters kinderopvang en onderw...	0.053	0.413	0.589	0.794	0.957	0.904
0331 Winkeliers en teamleiders detail...	0.05	0.393	0.57	0.78	0.934	0.884
1311 Beroepsgroep overig - managing	0.054	0.301	0.468	0.729	0.929	0.875
0812 Databank- en netwerkspecialisten	0.092	0.571	0.82	0.828	0.954	0.862
0434 Boekhoudkundig medewerkers	0.218	1.153	1.325	1.088	1.063	0.845
0414 Beleidsadviseurs	0.071	0.469	0.677	0.779	0.914	0.843
0621 Juristen	0.042	0.299	0.454	0.69	0.879	0.837
0421 Boekhouders	0.112	0.759	0.934	0.855	0.947	0.835
0215 Uitvoerend kunstenaars	0.047	0.355	0.513	0.717	0.88	0.833
0415 Specialisten personeels- en loop.	0.12	0.67	0.86	0.837	0.919	0.799
0423 Directiesecretaresses	0.139	0.879	1.067	0.901	0.926	0.787
0411 Accountants	0.119	0.751	0.906	0.796	0.858	0.739
0334 Callcentermedewerkers outbound e...	0.128	0.806	0.98	0.81	0.864	0.736
1121 Schoonmakers	0.26	1.545	1.726	1.123	0.988	0.728
0115 Onderwijskundigen en overige doc...	0.046	0.302	0.47	0.624	0.766	0.72
0912 Hoveniers, tuinders en kwekers	0.054	0.467	0.647	0.656	0.766	0.712
1022 Psychologen en sociologen	0.038	0.275	0.42	0.594	0.738	0.7
1222 Vuilnisophalers en dagbladenbezo.	0.123	0.698	0.867	0.756	0.823	0.7
0734 Loodgieters en pijpfitters	0.043	0.3	0.436	0.587	0.72	0.677
0412 Financieel specialisten en economen	0.076	0.557	0.706	0.656	0.734	0.658
0781 Hulpkrachten bouw en industrie	0.104	0.613	0.743	0.695	0.762	0.658
0771 Productiemachinebedieners	0.068	0.566	0.694	0.634	0.7	0.632
0821 Gebruikersondersteuning ICT	0.063	0.337	0.517	0.568	0.686	0.623
1215 Bedieners mobiele machines	0.095	0.584	0.724	0.641	0.682	0.587
0633 Beveiligingspersoneel	0.06	0.376	0.494	0.498	0.569	0.509
1116 Verleners van overige persoonlij.	0.028	0.231	0.361	0.421	0.512	0.484
0432 Secretaresses	0.088	0.444	0.562	0.527	0.568	0.48
0743 Automonteurs	0.028	0.218	0.318	0.407	0.499	0.471
1112 Koks	0.034	0.232	0.345	0.392	0.461	0.427
0536 Managers gespecialiseerde dienst.	0.021	0.127	0.207	0.326	0.441	0.42

1221 Laders, lossers en vakkenvullers	0.183	1.166	1.211	0.734	0.597	0.414
0733 Bouwarbeiders afbouw	0.026	0.171	0.267	0.344	0.436	0.41
0212 Auteurs en taalkundigen	0.029	0.146	0.238	0.33	0.43	0.401
1041 Sociaal werkers, groeps- en woon.	0.169	1.053	1.174	0.678	0.558	0.389
1311 Beroepsgroep overig - operational	0.052	0.315	0.397	0.393	0.44	0.388
0714 Architecten	0.025	0.114	0.191	0.298	0.404	0.379
1115 ConciĂŤrges en teamleiders schoon.	0.05	0.324	0.42	0.385	0.412	0.362
0754 Meubelmakers, kleermakers en sto.	0.031	0.233	0.312	0.329	0.379	0.348
0742 Lassers en plaatwerkers	0.031	0.226	0.3	0.317	0.366	0.335
1035 Medisch vakspecialisten	0.03	0.132	0.213	0.274	0.347	0.317
0731 Bouwarbeiders ruwbouw	0.014	0.124	0.181	0.234	0.295	0.281
1114 Kappers en schoonheidsspecialisten	0.011	0.114	0.165	0.218	0.274	0.263
0822 Radio- en televisietechnici	0.017	0.088	0.146	0.21	0.279	0.262
0121 Beroepsgroep sportinstructeurs	0.009	0.079	0.132	0.189	0.251	0.242
0213 Journalisten	0.009	0.077	0.121	0.19	0.245	0.236
1214 Vrachtwagenchauffeurs -First-Line Supervisors of Transportation and Material-Moving Machine and Vehicle Operators	0.014	0.078	0.131	0.181	0.242	0.228
0113 Docenten algemene vakken secunda...	0.012	0.072	0.113	0.183	0.236	0.224
1212 Chauffeurs auto's, taxi's en bes.	0.028	0.262	0.313	0.243	0.248	0.22
0752 Bakkers	0.018	0.106	0.165	0.183	0.232	0.214
0755 Medewerkers drukkerij en kunstni.	0.011	0.081	0.114	0.156	0.208	0.197
1011 Artsen	0.018	0.09	0.145	0.176	0.206	0.188
1214 Vrachtwagenchauffeurs -Heavy and Tractor-Trailer Truck Drivers	0.018	0.083	0.139	0.156	0.2	0.182
0114 Leerkrachten basisonderwijs	0.007	0.054	0.089	0.136	0.172	0.165
0221 Grafisch vormgevers en producton. - Graphic Designers	0.012	0.046	0.075	0.126	0.173	0.161
0753 Productcontroleurs	0.016	0.072	0.119	0.139	0.176	0.16
0611 Overheidsbestuurders	0.009	0.035	0.063	0.113	0.163	0.154
0542 Managers detail- en groothandel	0.006	0.038	0.056	0.11	0.158	0.152
0735 Schilders en metaalspuiters	0.011	0.081	0.116	0.129	0.162	0.151
0913 Veetelers	0.008	0.062	0.092	0.117	0.146	0.138
0921 Hulpkrachten landbouw	0.01	0.087	0.113	0.115	0.143	0.133
0772 Assemblagemedewerkers	0.009	0.091	0.114	0.11	0.138	0.129
0533 Managers ICT	0.007	0.039	0.063	0.093	0.127	0.12
0111 Docenten hoger onderwijs en hoog...	0.01	0.05	0.084	0.096	0.126	0.116
0532 Managers logistiek	0.006	0.024	0.044	0.078	0.12	0.114
0632 Politie en brandweer	0.005	0.033	0.048	0.078	0.115	0.11
0711 Biologen en natuurwetenschappers	0.009	0.04	0.063	0.083	0.118	0.109
1013 Fysiotherapeuten	0.014	0.039	0.069	0.091	0.119	0.105
0112 Docenten beroepsgerichte vakken ...	0.008	0.028	0.053	0.079	0.11	0.102
0751 Slagers	0.011	0.063	0.079	0.084	0.106	0.095
1111 Reisbegeleiders	0.009	0.037	0.061	0.075	0.102	0.093
0713 Elektrotechnisch ingenieurs	0.009	0.029	0.051	0.071	0.1	0.091
0741 Metaalbewerkers en constructiewe.	0.008	0.068	0.083	0.08	0.099	0.091
0222 Fotografen en interieurontwerpers	0.01	0.024	0.039	0.066	0.099	0.089
0543 Managers commerciĂŤle en persoonl.	0.009	0.028	0.047	0.072	0.098	0.089
1012 Gespecialiseerd verpleegkundigen	0.008	0.044	0.07	0.081	0.094	0.086
0723 Procesoperators	0.005	0.031	0.043	0.061	0.09	0.085
0911 Land- en bosbouwers	0.007	0.04	0.055	0.068	0.091	0.084
1211 Dekofficieren en piloten	0.006	0.024	0.037	0.056	0.084	0.078
1034 Medisch praktijkassistenten	0.012	0.029	0.051	0.064	0.087	0.075
0541 Managers horeca	0.006	0.021	0.033	0.051	0.076	0.07
0214 Beeldend kunstenaars	0.007	0.023	0.035	0.056	0.076	0.069
0333 Kassamedewerkers	0.015	0.1	0.117	0.071	0.083	0.068

0612 Overheidsambtenaren- operational	0.007	0.019	0.032	0.049	0.073	0.066
0534 Managers zorginstellingen	0.008	0.027	0.043	0.057	0.073	0.065
1122 Keukenhulpen	0.01	0.075	0.091	0.062	0.068	0.058
1051 Verzorgenden	0.014	0.085	0.122	0.084	0.066	0.052
0535 Managers onderwijs	0.005	0.012	0.021	0.04	0.056	0.051
1032 Apothekersassistenten	0.009	0.03	0.046	0.046	0.052	0.043
0612 Overheidsambtenaren- office	0.009	0.022	0.031	0.034	0.045	0.036
1213 Buschauffeurs en trambestuurders	0.005	0.019	0.025	0.029	0.04	0.035
1033 Verpleegkundigen (mbo)	0.007	0.016	0.027	0.032	0.041	0.034
0631 Politie-inspecteurs	0.004	0.009	0.015	0.023	0.034	0.03
0211 Bibliothecarissen en conservatoren	0.004	0.008	0.012	0.019	0.025	0.021
0221 Grafisch vormgevers en producton..- Commercial and Industrial Designers	0.004	0.008	0.011	0.015	0.018	0.014
1031 Laboranten	0.004	0.006	0.009	0.01	0.013	0.009

M.1 Unfilled demand per 2-digits job group

In this section, the results of the analysis of labour demand per 2-digits BRC code will be presented. Here, figure M.2, shows the total labour demand per 2-digits job group and figure M.2 shows its proportion of unfilled demand (open vacancies).

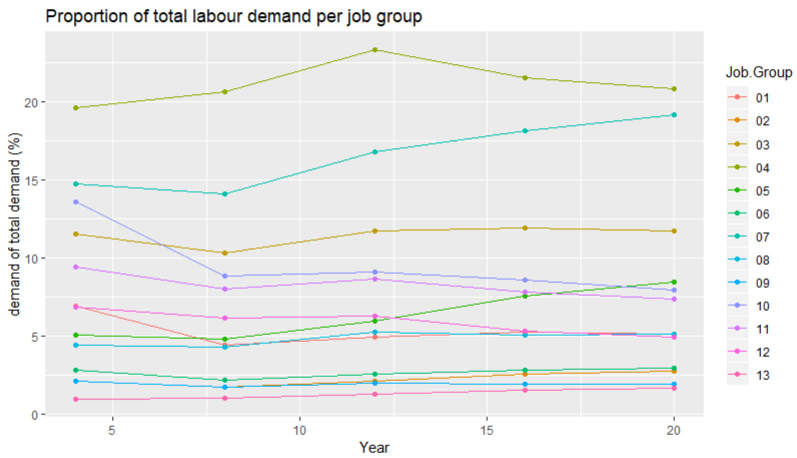


Figure M.2: Proportion of labour demand to total demand per CBS two-digit job code

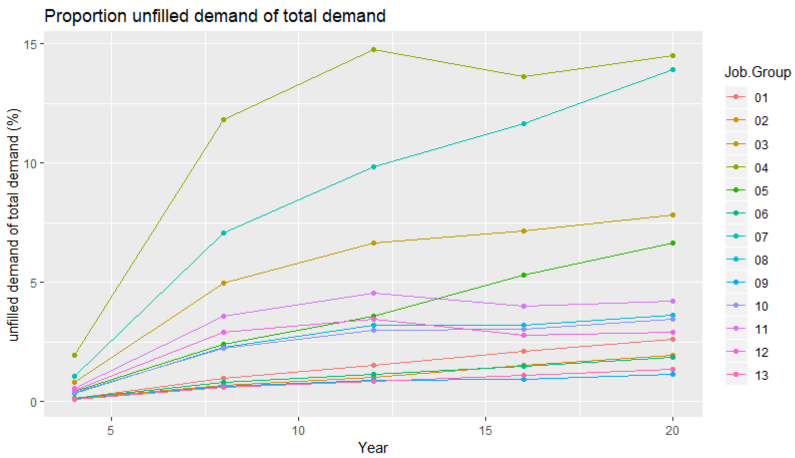


Figure M.3: Proportion of unfilled demand to total demand per CBS two-digit job code