Is the rising education level a driving factor of economic growth in cities?

A panel data analysis on cities in the Netherlands

ERASMUS UNIVERSITEIT ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS

BSc. Thesis Urban, Port and Transport economics

Baus Verbeek
422849
Supervisor: Dr. Erik Braun

Abstract

In earlier literature there is a strong implication for education as a driver for economic growth in cities. Broadly categorized, cities offer advantages on the production and consumption side of the economy. Cities in general can lead to higher productivity, an increased adaptability of the workforce and knowledge spillovers. In theory, a higher educated population accelerates these positive externalities. The main goal of this study is to understand (economic) city growth and its driving factors. Using panel data on the ten largest cities in the Netherlands, I aim to quantify the relationship between higher education and economic development of cities and differentiate results on population size. I conclude that there is a strong correlation between higher education and economic city growth, even when implementing various control variables. However, it is not clear what impact the difference in population size has on this relationship.

Content

1. Introduction 4
2. Literature Review 8
    2.1 City growth 8
        2.1.1 Early authors 8
        2.1.2 Recent authors 10
        2.1.3 Empirical studies 12
    2.2 Education 14
    2.3 Combined Studies 16
3. Data 19
4. Methodology 21
5. Results 25
6. Conclusion & discussion 32
7. References 34
8. Appendix 36

1. Introduction

According to a recent report by the United Nations (2018), currently about 55% of the world’s population lives in cities. They estimate that by 2050 this figure will have risen to 68% (United Nations, 2018), taking into account the world population growth in the coming decades. In other words, there will be another 2.5 billion people added to urban areas. It has proven to be very difficult to give a comprehensive explanation to the urbanization trend. However, economic growth in cities is evidently a factor. Throughout history, people have been moving to cities in search of employment and economic welfare. On average, cities perform better economically than rural areas (McKinsey Global Institute, 2011). However, to this day, there has been a long ongoing debate about the driving factors behind the economic growth of cities and the urbanization trend.

Looking at the United States, the last few decades have shown some interesting examples of city development. In July 2013, the city of Detroit filed for bankruptcy. After years of decline in the private sector there was no other option. The city’s economy developed quickly in the beginning of the 19th century because of the industrial activities. It became home to the producers of engines, railway carts and eventually cars. Henry Ford established the automated assembly line in Detroit in the beginning of the 20th century, making the automobile affordable for the American people. However, the economy of Detroit took a turn after 1950. According to Glaeser (2013), due to mismanagement of policymakers, like investing unnecessarily in infrastructure instead of education, not taking into account the declining population, severely underfunding pensions and the retiree health system, the city’s debt grew to 18 billion USD. At that point, bankruptcy was inevitable. Meanwhile, the population plummeted from 1.8 million in 1950 to less than 700.000 in 2016.

Boston, Massachusetts suffered from similar economic decline in the 1950-1980 period. The climate in Boston, a significant factor in economic development (Glaeser, Kolko, & Saiz, 2001; Glaeser & Shapiro, 2003) , is similar to that of Detroit, and like Detroit, Boston was mainly a manufacturing city. However, unlike Detroit and other ‘rustbelt’ cities, Boston had universities and a relatively well-educated population. In the 1980s, the education level rose quickly. With the revolution of the computer, the demand for an educated workforce increased and many inhabitants returned to school. Because of the relatively high education level, Boston performed extremely well economically. Other manufacturing cities benefited far less from the technology revolution because of the lower levels of education. When the manufacturing economy was still thriving in the 1950s, the universities in Boston may have seemed frivolous. However, because of those universities Boston was able to make the transformation from a manufacturing economy to an information economy. Once upon a time, Boston and Detroit were very similar. A factor in this example that cannot be neglected is the abundance of human capital. Because of the presence of a highly-skilled workforce, Boston adapted to a changing economy and avoided the same fate as Detroit (Glaeser, 2005).

In the recent decades, the bigger cities of the Netherlands have seen a significant increase in economic growth and population. Half a century ago however, Amsterdam and other large cities were struggling with squatters, a deurbanization trend and urban decay. People working in the city would much rather live in the suburbs. Just before the turn of the century, there was a shift in urban development. In the last two decades, city economies have thrived in the Netherlands. Together with steep population growth, Amsterdam has continuously performed above the national average economically (de Groot, Ossokina, & Teulings, 2016).

The urban revival in the Netherlands coincided with a strong increase in the relative share of the population that has received higher education (HBO and WO level). On a national level, it rose from 10% in 1970 to around 30% in 2010. However, in Amsterdam these figures were even more impressing. In some areas of the city, 70% of the inhabitants have received higher education (de Groot et al., 2016). As research shows in the United States, education is one of the significant factors contributing to urban growth, and the education level of the population can make the difference in the fate of a city. With 68% of the world's population predicted to live in cities by 2050 and cities delivering over 80% of the world's Gross Domestic Product (GDP) (McKinsey Global Institute, 2011; United Nations, 2018), understanding the effect of education on city growth is vital for national and international economies.

Research question

In this paper I will zoom in on one of the factors that have been proven to contribute greatly to (economic) growth in cities: education. First, an evaluation of the relevant literature will give insights into the driving factors of growth in cities. To understand the influence of education on city development, one must start with the theories on development of cities in general. There is a wide selection of literature available on the many factors that influence city development outside the Netherlands, mainly in the United States. Second, I will review the literature on education as a driver for economic growth in general and in cities, followed by various case studies that incorporate both education and economic growth. Finally, by performing a case study in the Netherlands, I aim to gain insight into the impact of education on economic development in the ten largest Dutch cities in terms of population: Amsterdam, Rotterdam, the Hague, Utrecht, Eindhoven, Groningen, Tilburg, Almere, Breda and Nijmegen. By comparing the impact of education in these cities, an assumption can be made on the importance of education for economic growth and differences per city. In my research I will differentiate on population size, as differently sized cities have different characteristics. For instance, a large city in the Netherlands will have one or two universities and several applied universities, which is not the case for many medium-sized cities. Furthermore, the available employment opportunities will differ greatly depending on city size. Therefore, this differentiation will help to give a better understanding of the effect of education on economic growth in cities in general. My main hypotheses are:

Hypothesis 1: Higher education is positively related to economic growth in cities.

Hypothesis 2: Higher educated citizens have a greater positive effect on economic development in larger cities compared to smaller cities.

Academic relevance

Rapidly growing cities worldwide present challenges and opportunities for policymakers. With nearly three quarters of the world's population living in cities by 2050 (United Nations, 2018), insight into the many factors that contribute to city growth is vital. In other research it has already been proven that there is an undeniable link between education and city growth, in terms of population and economic development. With this paper, I aim to compare those results with my findings in the Netherlands. If the same relation holds, I can differentiate the results on city size and compare large and medium-sized cities. Hopefully, this will improve the understanding of city growth in general and particularly in the Netherlands.

2. Literature Review

2.1 City growth

2.1.1 Early authors

Input sharing

In *Principles of Economics* (Marshall, 1920), Alfred Marshall incorporated Adam Smith's line of thought on division of labor (Smith, 1775) in his theory of localization. The central argument in Marshall's theory is reducing the role of different forms of transport costs on the production side, by locating near other producers (agglomeration economies). Transport costs are divided into three categories. From the perspective of a firm, locating near other firms helps to reduce the costs of obtaining inputs, like raw materials and the shipping costs to customers. A firm compares the costs of transporting materials needed for production and transportation of finished goods to customers. If the costs of transport and production in a cut-rate production location combined, are less than the costs of producing near the customer, the firm will move the production. When firms locate together, transportation costs benefit from economies of scale. Locating near other firms lowers production costs and increases profit.

Labor market pooling

The second argument for agglomeration comes from the labor side of production. Firms can benefit from agglomeration by utilizing the scale economies that derive from a large labor pool. When a firm becomes more productive and reduces the amount of labor available, workers can switch between employers. Therefore, firms can share the risk of a large labor pool. This maximizes firm productivity and stabilizes workers' wages. However, this is only possible when firms share the same type of workers.

Knowledge spillovers

Finally, the concentration of an industry in a city speeds the flow of ideas (knowledge spillovers) between firms, which causes that industry to grow. Marshall argued that historically, when an industry had chosen a location, it was likely to remain there for a long period. For workers in that industry, it thus made sense to live in close proximity. The skills needed for the trade would be passed on to neighbors, family and future generations, living in the same settlement. Employers in search of workers with a specific skill would localize their business near any place with a suitable workforce, while workers seeking employment would go to places where their skill is needed or where they can acquire new skills.

Finally, Marshall describes a localization argument from another dimension. Besides the reduction in transport costs for the production side, there is also a localization argument for the consumer. From a consumer's perspective, for a small, reoccurring and common purchase, the nearest shop will suffice. However, for an uncommon, expensive and important purchase, the consumer is prepared to travel further and incur greater costs (besides the purchase) to satisfy specific needs. Furthermore, a larger settlement generally accommodated a few sizeable and established industries. A district that depends for the most part on one industry is more liable to shocks in demand and supply. These shocks can greatly affect the workers, who depend on the wages. The presence of a few larger industries significantly diminishes this threat and leaves the population of a settlement less vulnerable to economic fluctuations.

Marshall-Arrow-Romer

Arrow (1962) focused on the concept of knowledge as a driver for economic growth. He described the important role of experience in increasing productivity and technological advancement. Romer (1986) later extended this theory. In Arrow's model the rate of growth of per capita output is limited by the rate of growth of the labor force, since he assumed that the marginal product of capital is diminishing given a fixed supply of labor. Romer (1986) disagreed with this implication. These theories combined would later be referred to as the Marshall-Arrow-Romer theory on externalities (MAR). The research on knowledge externalities by Arrow and Romer, applied to cities by Marshall, forms a theory on knowledge spillovers between firms of the same industry in cities. In essence, this theory promotes industrial specialization of cities. If an industry is relatively large in a city, knowledge spillovers between neighboring firms will help speed up innovation and thus industry growth (Glaeser, Kallal, Scheinkman, & Schleifer, 1991). A popular example of this theory was the computer chip industry in Silicon Valley (Arthur, 1990). Furthermore, the MAR theory favors local monopoly to stimulate economic developments over local competition. A local monopoly is assumed to restrict the flow of ideas to other firms allowing the monopolist to internalize externalities. This in turn would stimulate innovation and growth (Glaeser et al., 1991).

2.1.2 Recent authors

Urban vs. rural

An author that cannot be excluded in research on city growth is Jane Jacobs. Jacobs (1969) emphasized the importance of cities in economic development. Contrary to common believe, Jacobs stated that agriculture originated and developed out of cities and not the other way around, a common misunderstanding. In hunter-gatherer age, people lived in small, self-sustaining groups and settled in villages before agriculture emerged. When these groups learned to cultivate grain and raise livestock, they settled and stable villages emerged. This would be the beginning of organized social systems with more complex division of labor and large economic projects. Jacobs argues that this common theory on the origin of cities fails to address simple observations we see in cities today. Cities are usually surrounded by highly productive agricultural areas, whereas agriculture is not very productive if it does not incorporate many goods and services produced in cities or transplanted from cities. Early villages, 'cities', were needed for agriculture to develop. When a parent city, supporting multiple farming villages, would cease to exist due to some sort of misfortune, the orphaned farming villages would not develop further. There would be no new technological advancements coming from the city economy. The farming village would have to focus on being self-sufficient and could not focus on development. If the agriculture-first-theory holds, then developments in agriculture and of rural resources in general is basic and cities, depending on rural developments, are secondary. However, this is not the case. Modern agriculture is the result of countless innovations that were exported from the cities to the countryside.

In cities new goods and services are first created, even inventions directly for farming. According to Jacobs it is impossible to increase rural productivity first and city productivity later. Jacobs agreed with Adam Smith (1775), who stated that the most highly developed agricultural nations of that time, were the nations in which commerce and industry were developed the furthest. On the other hand, the countries with the poorest developed agriculture were those depending the most on agriculture. Smith already observed that the most developed, prosperous and innovative agriculture was to be found near cities. However, Smith did not go as far as to make the assumption that agriculture arose from cities. He justified the advancement of cities by stating that industry must inherently be better suited for division of labor than agriculture. Jacobs, at a later point in time, was convinced that early cities laid the grounds for agriculture.

"Rural production is literally the creation of city consumption. That is to say, city economies invent the things that are to become city imports from the rural world, so it can supply those imports. This, as far as I can see, is the only way in which rural economies develop at all, the dogma of agricultural primacy notwithstanding." (Jacobs, 1969, p. 40)

Competitive advantage

Porter (1990) agreed with the MAR theory on externalities in a sense that geographically-concentrated, specialized industries can cause knowledge spillovers, which stimulate growth. Contrary to MAR, Porter argued that local competition cultivates economic development in a city. Competition fuels the need for companies to innovate and rapidly adopt. If a company fails to keep up, it will be pushed out by the competition. A monopoly on the other hand reduces the necessity to innovate. Porter substantiated his theory with an example of the Italian ceramic tile industry. With rapidly growing domestic and foreign demand many new (family) businesses emerged in the 1950s. In the 1960s, the tremendous number of ceramic tile producers in the Sassuolo area in Italy led to heavy competition and rivalry. Individual firms were forced to innovate and decrease production costs. New innovations drastically decreased labor intensity and production speed, which spread rapidly amongst the competition. Furthermore, the export of the new machinery grew considerably. In essence, Porter's theory predicted that externalities are optimally utilized in industrially specialized cities with local competition.

2.1.3 Empirical studies

In later research by Glaeser, Kallal, Scheinkman, & Schleifer (1991) these three main theories of city growth, the MAR theory on externalities, Porter and Jacobs were empirically tested. In summary, the MAR theory focused on knowledge spillovers that occur between firms in the same industry. Furthermore, the MAR theory favored local monopoly over local competition to stimulate growth. Porter (1990) agreed with the MAR theory that knowledge spillovers occur predominantly within the same industry. However, Porter emphasized that local competition is a better driver for growth than local monopoly. Finally, Jacobs's theory disagreed with MAR and Porter by stating that most important knowledge spillovers come from outside the core industry. According to Jacobs, an industrially diverse geographically concentrated environment stimulates innovation and growth the most. A schematic overview of the theories is presented in figure 1 below.

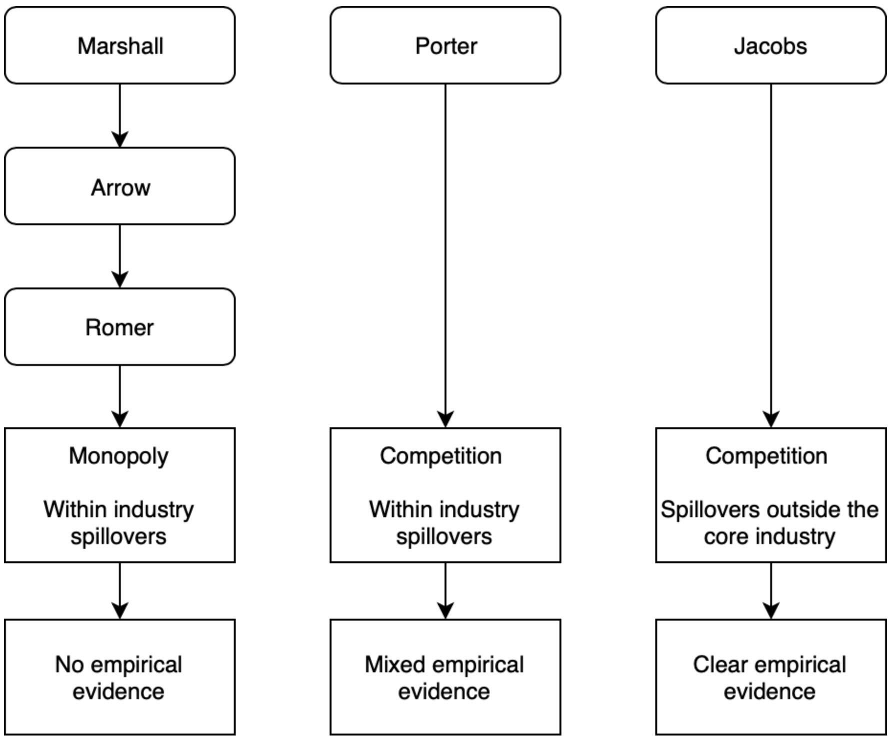

Figure 1: Overview of theories on city growth

Empirically testing these theories gave support for Jacobs' theory (Glaeser et al., 1991). Industries grow slower in cities where they are heavily overrepresented. Furthermore, industries were found to grow faster in cities where the average size of the firm was smaller than the national average size of firms in that industry. Assuming that an increasing number of firms depending on the same labor pool will increase local competition, Glaeser et al.

(1991) concluded that local competition stimulates economic growth in cities, in line with Porter and Jacobs' theories. Finally, results showed that city-industries grow faster when the city is less specialized, as Jacobs predicted. There is therefore no evidence in this research on the MAR theory on externalities. The evidence is mixed on Porter and overall in line with Jacobs' theory. In essence, local competition and urban variety appear to encourage employment growth in industries and therefore stimulate economic growth.

Further research on this topic uncovered some of the key factors influencing population growth in cities. Glaeser, Scheinkman, & Shleifer (1995) found, looking at urban characteristics in the US in 1960 and comparing those with urban growth from 1960 to 1990, that income and population have similar growth rates. Furthermore, population and income are positively related to initial schooling of the population. Education level is consequently a predictor of population growth. Population and income are also negatively related to initial share of unemployment and initial share of employment in manufacturing. These results show similarities with the case study of Boston (Glaeser, 2005). An initially higher level of schooling gives a significant economic advantage. Furthermore, although Boston was a manufacturing city in the 1950s, the level of education made the transformation to an information economy possible. The absence of an educated population in Detroit is further proof for this theory, as Detroit tragically failed as a city after the industrial heydays. Some other factors, which on the surface would appear to influence urban development, like racial composition and segregation are uncorrelated in this model.

Consumer amenities

Thus far, many of the theories discussed have been based on the concept that the benefits of the city mostly arise from the production side. Many earlier authors incorporate the central assumption that people live and work in the city because of the increased productivity and therefore higher wages, but see the consumer side of living in the city as a disadvantage, offsetting the productivity premium. Living in the city often means paying higher rent and dealing with negative externalities that impact the quality of life, like air pollution, congestion and higher crime rates. However, more recent literature focuses on the urban amenities that are becoming increasingly important in studying urban growth. Broadly categorized, there are four different categories of urban amenities. First, the presence of a rich variety of goods and services. Second, the aesthetics of living in a city. Third, the better public services available and fourth, the speed of which goods and people can move around (Glaeser et al., 2001). The first amenity, a rich variety of goods and services also applies to local services like museums, restaurants, theatres. Cities in the US and France with more of these amenities have grown more quickly over the last two decades. Glaeser et al. implicate that while a rise in the level of education has caused a rise in productivity, the quality of life in cities has risen even faster in places with more educated workers.

2.2 Education

To properly understand the effect of education on the economic growth of cities, it is vital to understand the effect of education on society as a whole. Many economists have already claimed that there are more benefits to a higher level of education than the individual benefits of higher education. The social benefits exceed the private benefits. Rauch (1991) laid the groundwork in this aspect by comparing different levels of human capital across cities. His assumption was that differences in human capital level will translate into differences in productivity, as human capital causes knowledge spillovers. Findings indicated that there was indeed a positive relation between human capital and productivity. Each additional year of average education would give an increase of total factor productivity of 0.8% in his dataset. However, a limitation of this approach was that initial schooling was treated as historically predetermined.

Moretti (2004a) extended this research by analyzing the social economic effect of education on different income groups. It is evident that in general, individuals with a higher level of education receive higher wages. However, Moretti researched the social return of education. The results show that when the level of college graduates, a higher educated group, rises in a city, the wages of high-school dropouts, the lowest educated group, also rises. The increase in wages for high-school dropouts is around four times the increase of the rise in wages for college graduates. Naturally, in researching a phenomenon like this, one cannot assume causality. It is in many individual cases entirely plausible that higher wages attract higher educated people, indicating reversed causality. Similarly, there could be other unobservable characteristics in cities that can cause biased results. However, Moretti made efforts to correct for both individual characteristics, like individual ability and city characteristics, like industrial structure. Nonetheless, it remains to be extremely difficult to fully account for the heterogeneity of cities.

In later research, Moretti (2004b) investigated the spillovers at plant level. His argument was that, if spillovers exist, production facilities in areas with a high level of human capital should be more productive than similar facilities in areas with a lower level of human capital. His findings indicate that in areas with a higher growth rate of college graduates, production facilities reported higher increases in productivity on average, compared to other areas with lower college graduate's growth rates. Interestingly, manufacturing wages showed a similar relation when comparing human capital in different areas (again, city heterogeneity cannot be fully accounted for).

Besides direct economic growth, society benefits from education on other aspects as well. According to Krueger & Lindahl, (1999), these benefits can differ per socioeconomic group. For instance, the increased payoff of an extra year of schooling for disadvantaged youth is higher than for advantaged youth. Furthermore, the increase of education for disadvantaged children also helps to reduce crime and welfare dependence, besides raising incomes.

The positive externalities of education were further studied by Moretti (2003). Although there is not enough comprehensive empirical evidence in some areas, there are many theories in which education has positive externalities. Apart from the production side of the economy, as has been discussed in previous articles, society as a whole can benefit from a higher education level. A popular example is that higher education is associated with a reduction in crime levels, which impacts the overall quality of life. Having received higher education expands the options for an individual and can reduce the probability of engaging in criminal activities. Not all studies found a significant link between crime and education, but there is enough empirical evidence to prove that there is a relation between the two. Another example is the influence of education on engagement in politics. According to Moretti (2003), there are two types of possible externalities in this respect. First, educated citizens may be more able to select proper candidates. Second, educated citizens may desire to participate more in civic issues. Although very plausible, it is hard to empirically study this theory. For instance, it is impossible to objectively differentiate candidates in good and bad, taking every aspect into account. Furthermore, the increased civic participation of higher educated citizens is difficult to measure.

Another view by Glaeser (1999) emphasizes on the thought that cities exist to facilitate learning. In this research and in earlier research it is confirmed that there is a wage premium in urban areas. This premium exists also when controlling for race, demographics, occupation and education. However, workers who move to the city are not paid higher wages instantly, they experience higher wage growth. This is in line with the model where workers come to the city to learn. Glaeser (1999) built on the theory of Alfred Marshall (1920), who argued that individuals acquire skills. His model predicts that cities attract more talented workers because of the higher mean of skills and knowledge in cities. Furthermore, people learn through imitation. The more contact a person has per period, the more opportunities he or she has to learn. Living in a city drastically increases the amount of contact for an individual. Therefore, learning and the exchange of skills happens more quickly in cities, which attracts new residents. Young and not too risk-averse individuals that will benefit the most from learning will come to the cities to increase their knowledge and as a result their wage. Glaeser states that urbanization will rise when the return to skills, the ability to learn by imitation and the general health of the economy also rises.

2.3 Combined Studies

Combining the theories so far, there are three main arguments as to why the presence of an educated population is vital to a city's success (Glaeser & Saiz, 2004). The first view depicts the knowledge spillovers that occur within cities. Jacobs (1969) gives a comprehensive explanation of the city as a facilitator of the flow of ideas. Knowledge spillovers are a central concept in urban economics and are interlinked with education. Second, the theory based around the adaptability of a city as a key survival element as proposed by Glaeser (2005). Glaeser argues that in the transformation from a manufacturing to an information economy, the relatively high level of average education played a vital role in Boston, when comparing Boston to Detroit. This view in essence values human capital as means to adapt to change. Third, (Glaeser et al., (2001) took a different approach from earlier research by analyzing the consumer side of cities and looking at city amenities. Most of the theories on urban growth in general focus on the productivity gains cities can offer, but lack the implementation of the potential benefits cities can offer to consumers. Nowadays, many choose to live in the city for the variety of the goods and services available in close proximity. Furthermore, the aesthetics of the city, public services and the speed of transferring goods, people and ideas, add to the joy of living in the city.

Glaeser & Saiz (2004) use various measures to test these three theories. Using extensive data on large US cities and Metropolitan Statistical Areas (MSA), Glaeser found that in the last century, educated cities have grown more quickly in population than comparable cities with less human capital. Even when implementing various control variables, the relation remains significant and the possibility for reversed causality was thoroughly investigated. It is very plausible that economically developed and populated areas attract a more educated population. However, according to this research, skilled cities are growing because they are becoming more productive economically, not because they are increasingly attractive as a place to live.

By using patent data, more evidence was found for the view of the information city, that is driven by knowledge spillovers. The share of the population with a bachelor's degree was found to be an important predictor for technological growth. Furthermore, there was a significant relationship between an increase in the share of college graduates and the number of patents filed. The reinvention view of the city, as earlier described (Glaeser, 2005; Jacobs, 1969), was also further extended by using data on immigrants. Assuming the reinvention city view to be correct, logically the presence of human capital should not matter in cities with large supplies of immigrants. In cities without immigrants, human capital should be an important predictor of the adaptability of the city and therefore economic growth, which is supported by the results in this study. Nonetheless, the authors conclude that the reinvention hypothesis needs further research to fully understand the relation. Finally, Glaeser & Saiz (2004) found that the consumer city view, where the increased quality of life is what attracts the higher educated to cities, is also proven to be a factor. Looking at housing prices and population growth, Glaeser & Saiz find that human capital is strongly associated with rising consumer amenity levels. However, results differ at

MSA and city level. Especially at city level, consumer amenities seem to have a strong relation, whereas at MSA level, the growth comes mainly from the productivity increase of human capital. According to Glaeser & Saiz, human capital is the most powerful predictor of growth, besides climate which was already proven in previous research (Glaeser et al., 2001; Glaeser & Shapiro, 2003) and controlled for in this study.

A similar study was conducted by Shapiro (2005). According to Shapiro, there are three main explanations for the relation between human capital and growth in cities. First, the increase in productivity. Second, the increase in the quality of life (consumer amenities) and finally, the presence of omitted variable bias. The results align with Glaeser & Saiz (2004) and show support for a causal interpretation of the relationship between human capital and increase in the quality of life. Two possible explanations are given for this finding. On the one hand, higher educated residents may encourage the growth of consumer amenities, like restaurants, bars and museums. On the other hand, a higher educated population can have more influence, privately or politically, which they can use to increase the quality of living by minimizing negative externalities of cities, like crime and air pollution. Testing for these hypotheses, a significant effect for consumer amenities and no substantial evidence for private or political influencing was found. Furthermore, like Glaeser & Saiz (2004), research indicated a causal effect for concentrations of college graduates and employment growth. Roughly one-third of the economic growth comes from improvement in quality of life. The rest is the result of the increased productivity of high human capital areas. The findings of these more recent studies are important to urban economics, because they give a comprehensive explanation to for the relationship between human capital in cities and economic growth. Predominantly, results suggest that urban growth can be promoted by stimulating the increase of human capital in cities. In the next section, I will present the data used in this analysis.

3. Data

There are extensive datasets available, with various kinds of growth measures and demographics on the 10 largest cities of the Netherlands. The 'Centraal Bureau voor Statistiek' (CBS) collects and publishes this data, dating back to around 2000. In this study I use a macro-panel data set consisting of 15 years for model 1 and a macro-panel data set containing 8 years in model 2 (2010-2017). The usage of a panel dataset can offer several relevant advantages, one of which is the possibility to incorporate unobservable characteristics. Furthermore, panel data generally consists of more observations and allows for more degrees of freedom. Incorporating more observations can reduce collinearity among the explanatory variables in the model. However, there is a risk of heterogeneity, which can skew the results. Given the advantages of a balanced dataset, I will limit my research to the 2003-2017 and 2010-2017 timeframe, which is continuously documented by the CBS.

The 10 cities I limit my research to obviously differ in size. By categorizing them based on population size, I can distinguish the effect of education on economic growth in differently sized cities. Cities that differ greatly in size will have different characteristics. I use an ordinal variable for the categories of city size, with 7 categories as can be seen in table 1 of the appendix. Population size can impact several factors of a city, which in turn can influence the effect of education on economic development. For economic growth, I will use the Gross Regional Product per capita (GRP) measure as the dependent variable. However, the CBS data on economic development is not available at municipality-level. Therefore, I will use the GRP per capita for the area surrounding the city (COROP and COROP-Plus area). A more detailed description can be found in table 2 of the appendix. Although monitoring at municipality-level would be more precise, this data is sufficiently detailed and can be used as a proxy for economic growth in the city. Furthermore, the GRP of a city will also incorporate people working in the city and living outside the municipality, which can cause biased results. This issue will be eliminated with the use of the COROP data.

The independent variable will be the share of the city's population per education level. Education will therefore be an ordinal variable, divided into three different categories: lower education, medium education and higher education. Each category gives the absolute value of the working and non-working population of that city, aged 15-65 and divided by 1000. An individual's education level is determined by the highest completed level of education. Table 3 of the appendix gives an overview of the education levels. Based on the discussed literature, I will include the following control variables on the city level in model 1: unemployment, population and the interaction effects between population and higher education. In model 2 I will incorporate control variables for consumer amenities and economic sectors. Although the climate has proven to be one of the most important drivers for city growth, I will not take this into account in my research. The cities in this study are located in close proximity and therefore the climate is largely similar.

As discussed earlier, Glaeser et al. (2001) mentions the consumer aspect of city growth. Most of the theories focus mainly on the advantages cities can offer for the production side of the economy. However, Glaeser et al. (2001) show that there are several factors that contribute to people's desire to live in cities apart from employment opportunities. I will include several control variables to incorporate this effect into the model. As proxies to measure consumer amenities, I will use the following four variables: the number of café's, restaurants, cinemas and train stations within a 5km radius. With these variables I aim to improve the accuracy of the model and refine the relationship between higher education and GRP per capita.

Glaeser et al., (1995) found that a relative share of the population in manufacturing is negatively correlated with economic growth. Furthermore, the presence of a specialized business services industry naturally needs educated workers. This has an effect on economic growth and the level of education. Therefore, I will include both these variables in the model. For the data on economic sectors as well as consumer amenities, I will use CBS data. The sector data is categorized according to the Standaard Bedrijfs Indeling (SBI) 2008 classification and is available on municipality level. The consumer amenity data is also available on municipality level.

4. Methodology

Panel data, which I will use in this model, is essentially a combination of time-series data and cross-section data. This combination makes it possible to compare variables for multiple cities, in multiple years. The aim of this study is to quantify the impact of an educated population on the economic growth of cities. Therefore, the dependent variable will be the GRP per capita of the COROP area. The effect to be measured is education, which will be represented in an ordinal variable with three different categories, as the independent variable.

An important distinction to be made when using panel data, is how to account for possible omitted variables. In this study, the possibility of omitted variables is very likely. Besides education, there are countless factors that affect economic development of cities. Obviously, it is impossible to implement all these factors in a fully comprehensive model of city growth. Therefore, I must assume there are (at least a few) omitted variables in this model. The usefulness of the model depends on how these omitted variables are accounted for. Panel data is most useful when there are unobservable explanatory variables that are correlated with the variables in the model.

If there are omitted variables, they have an omitted effect on the dependent variable. If this effect is constant over time (time-invariant), this effect can be consistently measured using panel data estimators. The two main models for controlling for omitted variables in panel data are the fixed effects, also referred to as the Least Squares Dummy Variables (LSDV) model and the random effects model. In some datasets, there are no panel effects because the observations are independent of each other and there is no within-panel correlation. In that case, a pooled OLS model is the most suitable.

Fixed effects model

In the fixed effects model, individual-specific (city-specific) characteristics are allowed to be correlated with the regressors. The coefficients of these effects are assumed to be similar, but the intercepts vary between cities. The basic fixed effects model is stated below.

$$Y_{it} = \beta_1 X_{1,it} + \cdots + \beta_k X_{k,it} + \alpha_i + u_{it}$$

Where

- a_i $(i = 1 \dots n)$ is the unknown intercept for each entity (n entity-specific intercepts)
- Y_{it} is the dependent variable with i = entity (city) and t = time.
- $X_{k,it}$ $(k = 1 \dots n)$ represents one independent variable
- β_1 is the coefficient for independent variable $X_{1,it}$
- u_{it} is the error term

Random effects model

In the random effects model, the correlation between entities is assumed to be random and uncorrelated with the independent variables in the model. In other words, the unobserved individual effects are expected to show no correlation with the regressors. If there are reasons to believe that differences across entities have some influence on the dependent variable, then the random effects model should be applied.

$$Y_{it} = \beta X_{it} + \alpha + u_{it} + \varepsilon_{it}$$

Where

- α is the intercept
- u_{it} is the between-entity error term
- ε_{it} is the between-entity error term

Pooled OLS model

In the pooled OLS model, it is expected that there are no unique attributes of entities (cities) within the data and no universal effects across time.

$$Y_{it} = \alpha_0 + \beta_1 X_{1,it} \dots + \beta_k X_{k,it} + u_{it}$$

To choose between these models, I will make use of the Breusch-Pagan Lagrange Multiplier (LM) and the Hausman test (Hausman, 1978), which are commonly used to decide between pooled OLS, fixed effects and random effects. The Breusch-Pagan test tests for variances across entities. If these variances are zero, there is no significant difference across units, indicating that there is no panel effect. In this case pooled OLS will suffice. If variances across entities are not zero and the null hypothesis is rejected, pooled OLS is not suitable and the random or fixed effects model is applicable. The Hausman test decides between these two. In this statistic test the null hypothesis is the random effects model and thus the alternative hypothesis is linked to fixed effects. In essence, it tests whether the unique errors (u_i) are correlated with the regressors. The null hypothesis states that they are not. These results will indicate which model is best suited for the data.

The model for the regressions in this study is depicted below.

Model 1

$$GRP_{i,t} = a_0 + \beta \cdot higher\ education_{i,t} + \beta \cdot medium\ education_{i,t} + \beta \cdot lower\ education_{i,t} + \beta \cdot unemployment\ rate_{i,t} + \beta \cdot population\ size\ (category)_{i,t} + \mathcal{E}_{i,t}$$

Model 2

$$GRP_{i,t} = a_0 + \beta \cdot higher\ education_{i,t} + \beta \cdot medium\ education_{i,t} + \beta \cdot lower\ education_{i,t} + \beta \cdot consumer\ amenities_{i,t} + \beta \cdot relative\ sector\ share_{i,t} + \mathcal{E}_{i,t}$$

These models will help determine the relation between GRP and education in the data. To properly analyze the results, I will divide my research into two separate models, with different datasets. The first model contains data from 2003-2017 and incorporates the unemployment rate and population size. In the second model, I will use a dataset with data on 2010-2017 for the relative sector share and consumer amenities. This division will help to understand the relative effects of the different variables. In the first model, I will also incorporate interaction effects between population and higher education. Interaction effects will help to gain a better understanding of the impact of city size on the relationship of higher education and GRP per capita. It is entirely probable that this relationship is stronger in cities with a bigger population, as hypothesis 2 suggests. I will add two interaction terms, one linear and one quadratic and regress these both separately and combined to properly analyze these effects. Together with the ordinal variable and the absolute variable of population size, I expect to extrapolate the population size difference from the data, should this exist.

Furthermore, it is possible that the relationship central in this study is not linear and is better explained using percentage. For instance, a rise in education level may have a more significant impact in a city that has a relatively low education level to begin with. Besides, the data is not necessarily normally distributed. To account for both of these problems, I will include two regressions with natural logarithmic variable values.

Finally, I will perform various statistical tests to ensure a proper interpretation of the results. Data collected over time can be subject to serial correlation, heteroskedasticity and cross-sectional dependence. Serial correlation effectively means that the error terms from different time periods are correlated. Therefore, the errors associated with a given time period are carried over into future time periods. Serial correlation can be tested using a Lagram-Multiplier test. Heteroskedasticity implies that the variability of a specific variable is unequal across the range of values of a second variable that predicts it. Heteroskedasticity is tested for with the Wald test for groupwise heteroskedasticity. Cross-sectional dependence occurs when there is 'between-group' dependence in the data, in this case between cities. This phenomenon is mostly found in social studies and can bias results. To test for cross-sectional dependence, I will use the Breusch-Pagan LM test of independence.

5. Results

Model 1

A summary of the first dataset is presented in table 1. As mentioned before, in the first model I use a balanced panel data set, so for every variable there are 150 observations. First, the differences among cities in terms of education level and GRP per capita are assessed. Looking at the GRP per capita in figure 2, it becomes clear that there has been a rising trend in all cities. In the majority of the cities GRP per capita is relatively similar. However, it is also clear that the GRP is higher in the four largest cities of the Netherlands, especially in Amsterdam. Furthermore, the GRP of Groningen fluctuates significantly in comparison to the other cities. This is most likely due to the severe decline in natural gas production, which had a significant impact on GRP (NRC, 2018).

Variable	N	Mean	Std. Dev.	Min	Max
GRP per Capita	150	42700.35	13554.2	22490	89137
Higher Education	150	87.87	63.93	25	315
Medium Education	150	88.06	51.84	39	210
Lower Education	150	76.83	50.75	30	207
Unemployed (total)	150	18.30	14.17	4.52	58.97
Population Category (1-7)	150	2.75	2.06	1	7
Population (total)	150				

Table 1: Descriptive statistics model 1

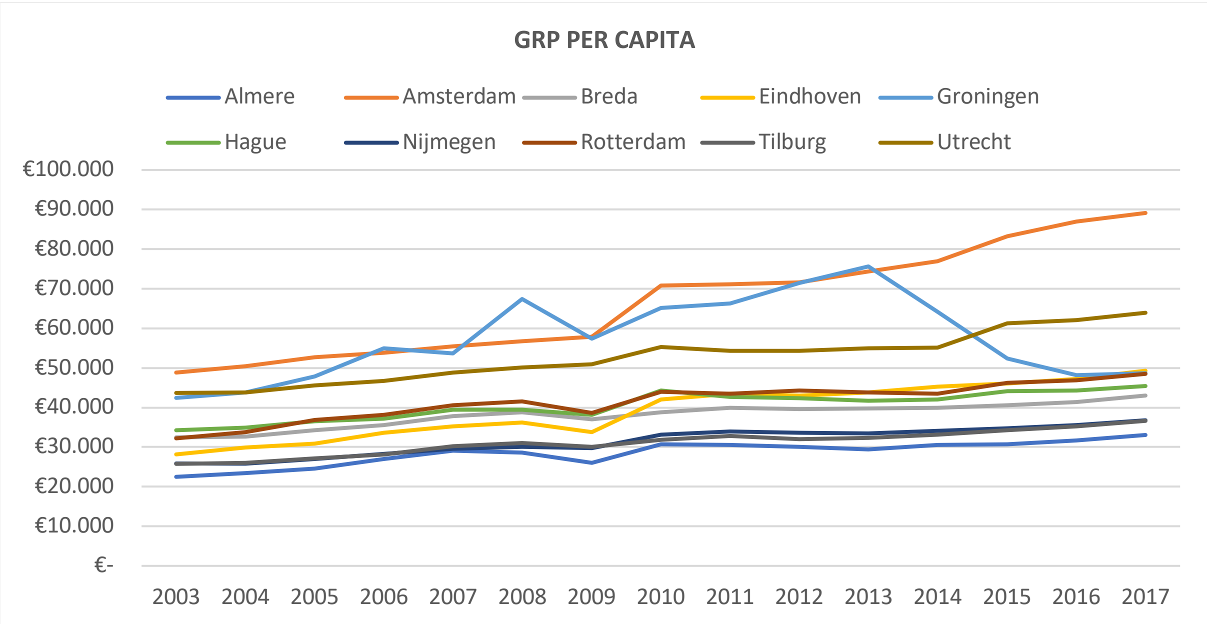

Figure 2: GRP per capita per city (2003-2017)

Looking at education in figure 3, we see that the percentage of higher education is the highest in Utrecht, where more than half of the population has received a higher education. The lowest percentage of higher education is in Almere and Rotterdam, although Rotterdam has seen a significant increase over the past years. This is also visible in figure 4, presenting the share of lower education per city. All cities in this study have seen an increase in higher education and a decrease in lower education.

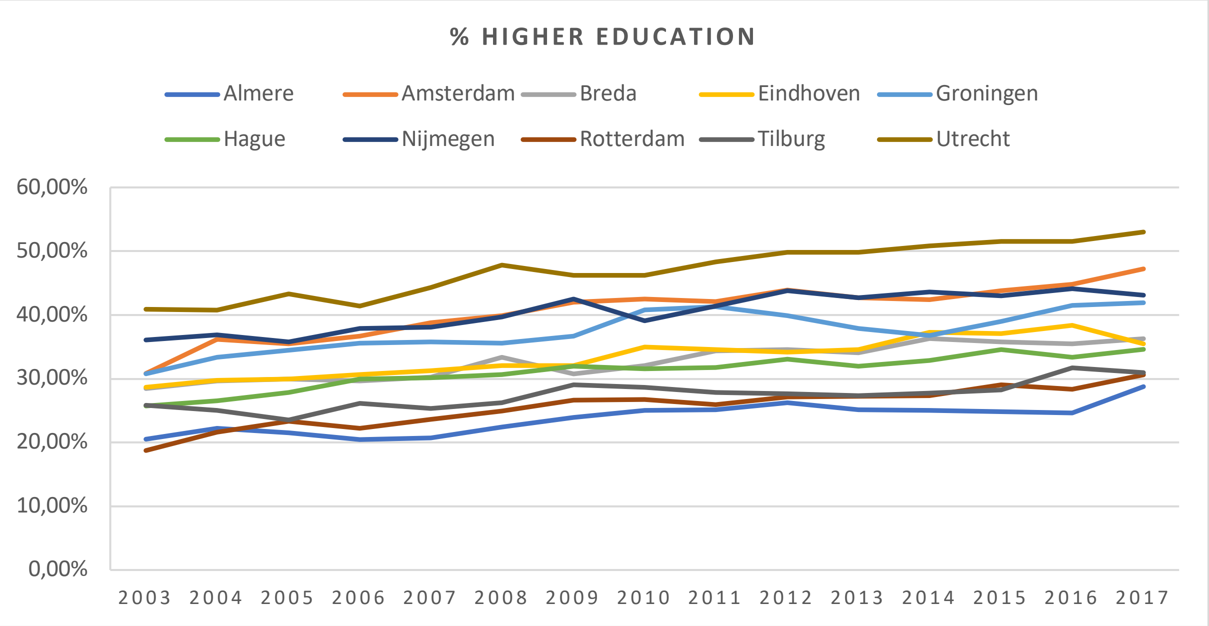

Figure 3: Higher education - % of total population (2003-2017)

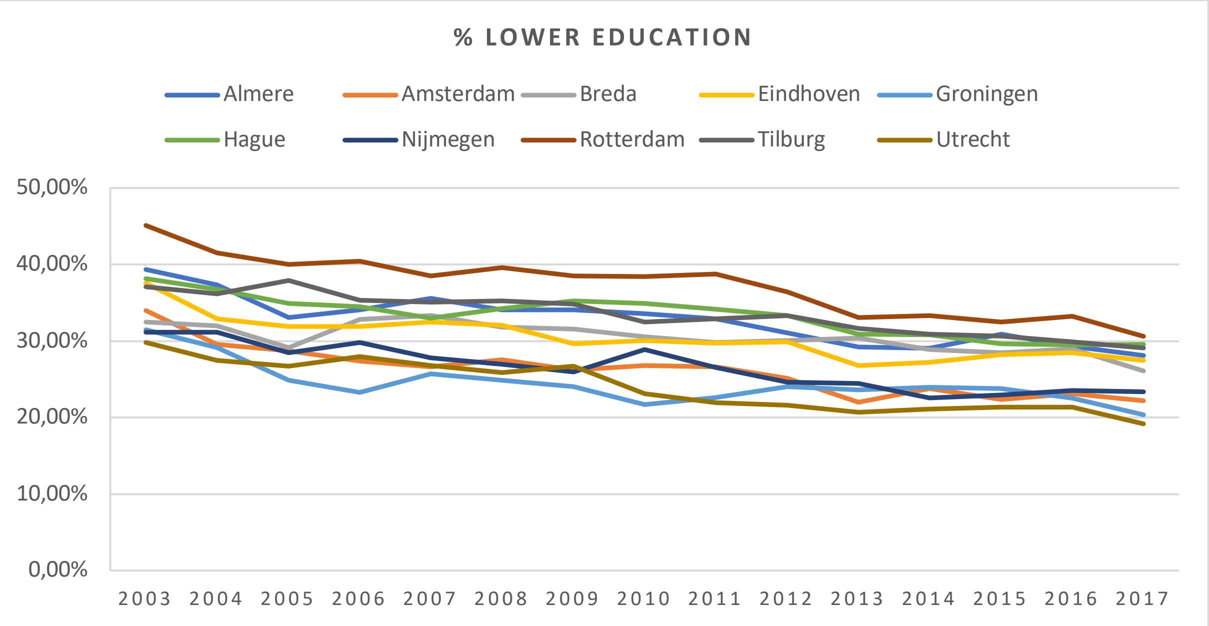

Figure 4: Lower education - % of total population (2003-2017)

A simple test of the data in dataset 1 shows that there is indeed serial correlation (p-value = 0.0). Furthermore, the null hypothesis of the test for heteroskedasticity is also rejected, implicating heteroskedasticity (p-value 0.0). To control for both these effects, I will use robust standard errors and cluster at city level.

To run the regressions, a choice has to be made between a pooled OLS, random effects or a fixed effects model. As mentioned before, there are two tests that help to make this decision. The Breusch-Pagan Lagrange multiplier (LM) tests for model 1 is significant (p-value 0.0), meaning that the null hypothesis is rejected and there are panel effects. Pooled OLS is therefore not suitable. The Hausman test gives a significant result (p-value 0.0), indicating that the fixed effects model is best suited for this dataset.

Table 2 presents the results of the regressions. First, a regression without the control variables is run. There is a significant relation between GRP per capita and higher education at the 99% level ($p<0.01$). Including the other education levels, medium and lower education, changes the constant to negative. The coefficients for medium and lower education are lower than for higher education, as can be expected. However, only high and medium education are statistically significant.

When including the other variables: unemployment and population, higher and lower education levels have significant coefficients. However, the coefficient for unemployment does not. Surprisingly, the coefficient of population category is significant, but negative. This would implicate that a larger population category has a negative effect on the GRP. This is unexpected, as the average GRP is the highest in the four largest cities. Unemployment has an expected negative effect on GRP but is not statistically significant. Finally, when replacing population category with the absolute population value, the coefficients change slightly.

Regressions 5 and 6 show the results of the ln regressions. In these two regressions, the natural logarithmic values are taken of all variables. This allows for easy percentage interpretation and enhances results for datasets that are not fully normally distributed. An analysis of GRP per capita shows that the data is slightly positively skewed, which may cause biased results. The results in regression 5 indicate that a 1% increase in higher education is associated with a 0.76% increase in GRP per capita. In regression 6, a 1% increase in higher education is associated with a 0.66% increase in GRP per capita, but this coefficient is only significant at the 95% level. Furthermore, unemployment has a negative coefficient that is significant in regression 6, similar to regression 4.

Regressions 7, 8 and 9 incorporate the interaction effects between population and higher education. The coefficient of higher education remains significant in these three regressions and in regression 9, both the interaction effects are significant. However, when adding the interaction effects (Population*Higher Education) and $(\text{Population*Higher Education})^2$ separately, the coefficients are not significant. This is unexpected, as it is probable that if there are notable interaction effects, a single term would also have significant effect. Furthermore, it is remarkable that one of the interaction terms is negative and the other is positive. This further obstructs compelling evidence for interaction effects.

Overall, the explanatory power of the model (R^2) is the highest in regression 5. Adding the control variables to the model in regression 6 decreases overall R^2 and hence causes the model to lose explanatory power. However, when looking at the within R^2, the explanatory power for the model when comparing the grand mean of the variables, is the highest when including the control variables. The within R^2 does not take into account individual characteristics of the cities. Finally, the between R^2, which disregards the time element and considers only the mean of each city individually, is the highest in regression 5. The regressions with interaction effects (7, 8 and 9) also have relatively high explanatory power, especially for the within R^2. Judging by explanatory power, regression 5 with ln variables is best suited for this dataset. However, this does not necessarily mean that regression 5 gives the best depiction of the relationship between the variables.

GRP per capita	*Linear*				*Ln*		*Interaction effect*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(constant)	11664.98** (1776.05)	-20066.11 (14725.24)	-22073.39 (10168.59)	-24148.46 (15419.94)	7.98** (0.87)	3.42 (8.44)	-20193.65 (18131.44)	-18305.25 (16514.14)	-16064.97 (16077.82)
Higher Education	353.18** (20.21)	383.89** (29.13)	417.93** (30.23)	404.87** (141.93)	0.76** (0.11)	0.66* (0.25)	512.27* (185.25)	467.09* (170.67)	609.27** (136.31)
Medium Education		204.81* (83.27)	355.85** (99.08)	290.28 (141.93)	0.04 (0.09)	0.05 (0.31)	298.58* (124.64)	311.30 (149.61)	-5.61 (123.97)
Lower Education		143.15 (74.25)	130.68 (72.78)	152.31* (65.31)	-0.20 (0.11)	-0.28 (0.20)	211.16* (66.03)	210.59* (85.06)	-131.93 (132.34)
Unemployed (Total)			-43.15 (62.68)	-105.26* (43.54)		-0.07* (0.02)			
Population Category (1-7)			-4564.16 (3211.83)						
Population				-0.01 (0.10)		0.44 (0.91)	-0.07 (0.11)	-0.07 (0.13)	0.16* (0.06)
Higher Education * Population							-0.00 (0.00)		-0.00* (0.00)
(Higher Education * Population)^2								0.00 (0.00)	0.00* (0.00)
R^2 Within	0.65	0.67	0.71	0.67	0.69	0.70	0.67	0.67	0.70
R^2 Between	0.59	0.46	0.47	0.45	0.66	0.60	0.48	0.47	0.55
R^2 Overall	0.55	0.41	0.42	0.40	0.64	0.57	0.43	0.42	0.54

* significant at the 95% level
** significant at the 99% level

Table 2: Regression analysis model 1 (robust/clustered standard errors)

Model 2

A summary of the second dataset is presented in table 3. The results of the Breusch-Pagan Lagrange Multiplier test show that the variance across entities is not zero (p-value 0.0). This indicates that there are panel effects and the pooled OLS model is not appropriate. Interestingly, when performing the Hausman test on the second dataset, containing fewer years, the results show that a random effect model is the most appropriate (p-value 0.15). Furthermore, the test for heteroskedasticity, the test for serial correlation and the test for cross-sectional dependence all give a p-value of 0.0. This implies that in this dataset heteroskedasticity, serial correlation and cross-sectional dependence should be taken into account.

The results of model 2 in table 4 are ambiguous and give no clear explanation of the effect of the variables. In the first regression the relative effects of the different education levels are significantly lower than in model 1. However, only high education remains significant. Adding the different control variables in regression 2 and 3 distorts the model. Coefficients for consumer amenities and economic sectors give unexpected results and are insignificant. Including both consumer amenities and economic sector distribution gives some significant coefficients but unexpected implications. Cinemas seem to have an extraordinarily large positive effect on GRP per capita, whereas the number of cafés in a 5km radius shows a negative correlation. Looking at the explanatory power of the model we see that the highest overall R^2 is found in the model that includes consumer amenities as well as relative sector share (regression 4). However, the within R^2 for that regression is the lowest, indicating that when the individual characteristics of the cities are disregarded, there is little correlation over time. On the other hand, the between R^2 is high, hence when the time component is neglected, there is more correlation between cities. Although the explanatory power for regression 6 is high, the coefficients for education are insignificant and the coefficients for consumer amenities, although some significant, give ambiguous results. When using panel data, the R^2 is not the only indicator of a proper model and coefficients have to be taken into account as well.

Variable	N	Mean	Std. Dev.	Min	Max
GRP per Capita	80	46985.14	14670	29420	89137
Higher Education	80	97.04	70.07	35	315
Medium Education	80	89.86	52.12	41	210
Lower Education	80	73.75	47.97	30	178
Cafés (<5km)	80	166.91	126.84	17.2	496.3
Restaurants (<5km)	80	267.58	244.93	51.6	984.7
Train stations (<5km)	80	2.49	0.48	1.7	3.6
Cinemas (<5km)	80	2.86	1.68	0.6	7
Share in manufacturing	69	18.19	11.19	5.4	45.7
Share in specialized business services	80	47.51	36.35	10.6	153.6

Table 3: Descriptive statistics model 2 (2010-2017)

GRP per capita		(1)	(2)	(3)	(4)
(constant)		34231.63** (5358.52)	35608.14** (9960.67)	31379.34** 2772.17	23236.73* (11623.21)
Higher Education		253.57** (29.58)	204.41 (173.82)	214.55** (40.27)	-348.76 (290.08)
Medium Education		-83.48* (35.26)	-123.29 (78.23)	-104.42** (29.90)	156.42 (199.20)
Lower Education		-58.99 (36.06)	-84.16 (138.90)	-102.31 (79.20)	-239.09 (173.26)
Cafés (<5km)			5.41 (86.44)		-99.93** (36.03)
Restaurants (<5km)			30.62 (48.68)		173.45* (82.36)
Cinemas (5km)			-1240.64 (3014.75)		5103.71* (2135.04)
Train stations (<5km)			1315.05 (1562.52)		2419.16 (4237.16)
Share in manufacturing				369.14 (389.80)	1020.53 (565.62)
Share in specialized business services				106.10 (2772.17)	-163.21 (213.10)
R^2	Within	0.24	0.26	0.40	0.19
	Between	0.75	0.75	0.73	0.89
	Overall	0.71	0.71	0.78	0.86

* significant at the 95% level
** significant at the 99% level

Table 4: Regression analysis model 2 (robust/clustered standard errors)

6. Conclusion & discussion

In the existing literature many authors agree on the link between higher education and growth in cities. However, the extent of that relation differs per study and so do the factors impacting this relation. In general, the presence of a higher educated population is found to have a positive impact on economic development in cities. Broadly categorized, cities offer advantages on the production side and the consumption side of the economy. Educated residents help to speed up this process and have other positive externalities. While a city in general can lead to a higher productivity, increased adaptability of the workforce and knowledge spilllovers, higher educated residents have a stronger impact on these factors. Higher educated workers offer a higher productivity in general, more easily adapt to changes in career paths and possess more knowledge in their field.

On the other hand, cities can become centers of consumption that offer a high degree of amenities and likeminded communities, to a point where people desire to live in cities despite the negative externalities. The findings in the literature of relative higher increase in rents compared to wages and the rise of reversed commuting support this theory. In theory, higher educated residents are expected to have an increased demand for consumer amenities and higher educated neighbors. However, the relation between skilled residents and the consumer city is far less evident in empirical studies. In general, it can be concluded that the benefits of skilled residents to a city are reflected primarily on the production side.

Overall, as in the literature, the results of the empirical study indicate that higher educated citizens and economic growth are strongly correlated. All the regressions in the tested models give a significant coefficient for higher education (except regressions 2 and 4 in model 2). Even when including several control variables and interaction effects, the relationship remains significant. Therefore, I can conclude that hypothesis 1 holds for the ten largest cities of the Netherlands. The presence of a higher educated population is positively related to economic growth in the ten largest Dutch cities. However, apart from higher education, the tested models often give unexpected and volatile results. The results of the other variables are ambiguous and do not give a distinct answer for hypothesis 2. The peculiarities of the coefficients and the differences in explanatory power are an indication that in this case study other factors are excluded, which may have significant impact on the results. Therefore, there is no compelling evidence in this model for a relation between city size and the impact of a higher educated population in the Netherlands. Hypothesis 2 cannot be confirmed based on this empirical study.

In the field of urban economics and in particular in city development, there is a high risk of omitted variable bias, given the many factors that influence city growth. As most of the factors that are expected to be of importance in earlier research were included, the potentially omitted variables may be exclusive to Dutch cities. However, there are also other limitations to my research. This model only considers city development in a single country. Although results may be similar in other developed countries, external validity of the model cannot be assumed. Furthermore, the data used in this study is limited in terms of years and detail. The data on economic performance of the cities is available only at COROP level, which may differ slightly with actual city economies. The 2003-2017 and 2010-2017 timeframes may be too short to capture change over time and the relation between education and economic development, or other factors that impact this relation. As can be seen in the previous century, the economy can evolve rapidly and the stage of an economy may significantly influence the need for education. Finally, it is also very likely that reversed causality causes biased results. Higher education and city growth may have a self-reinforcing effect, since higher education drives economic development, which in turn draws more higher-educated citizens to the city. It can be argued that many of the factors researched in this study but also in earlier research are interlinked and have to be considered as such.

Although crucial steps have been made in the field of urban economics and one can assume that many of the key factors influencing city development have been identified, the individual impact of these factors and especially how they relate to one another remains largely unknown. To properly interpret this relation, future studies would benefit from including a larger time period in their research as well as providing a comprehensive understanding of reversed causality and interaction effects. Furthermore, comparing different countries with different economic stages of development and correctly modelling individual characteristics of cities and regions could widen our perception of the relation between higher education and economic development in cities.

7. References

Arrow, K. J. (1962). The economic implications of learning by doing. *The Review of Economic Studies*, *29*(3), 155.

Arthur, W. B. (1990). ‘Silicon Valley’ locational clusters: when do increasing returns imply monopoly? *Mathematical Social Sciences*, *19*(3), 235–251.

de Groot, H., Ossokina, I., & Teulings, C. (2016). “Hoogopgeleiden zijn een motor voor stedelijke groei.” *Me Judice*, (27 juni 2016).

Glaeser, E. L. (1999). Learning in Cities. *Journal of Urban Economics*, *46*, 254–277.

Glaeser, E. L. (2005). Reinventing Boston: 1630-2003. *Journal of Economic Geography*, *5*(2), 119–153.

Glaeser, E. L. (2013). In Detroit, bad policies bear bitter fruit. *Boston Globe*.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Schleifer, A. (1991). Growth in Cities, (3607), 1963–1987.

Glaeser, E. L., Kolko, J., & Saiz, A. (2001). The consumer city. *Journal of Economic Geography*, *1*, 27–50. https://doi.org/10.4337/9781784717445.00009

Glaeser, E. L., & Saiz, A. (2004). The Rise of the Skilled City. *SSRN*, *1*, 47–105. https://doi.org/10.2139/ssrn.569867

Glaeser, E. L., Scheinkman, J. A., & Shleifer, A. (1995). Economic growth in a cross-section of cities. *Journal of Monetary Economics*, *36*(1), 117–143. https://doi.org/10.1016/0304-3932(95)01206-2

Glaeser, E. L., & Shapiro, J. M. (2003). Urban Growth in the 1990s: Is City Living Back? *Journal of Regional Science*, *43*(1), 139–165. https://doi.org/10.1111/1467-9787.00293

Hausman, J. A. (1978). Specification Tests in Econometrics, *46*(6), 1251–1271.

Jacobs, J. (1969). *Economies of Cities*. Vintage.

Krueger, A. B., & Lindahl, M. (1999). *Education for growth in Sweden and the world*. *ICY REVIEW* (Vol. 6).

Marshall, A. (1920). *Principles of Economics*.

McKinsey Global Institute. (2011). Urban world : Mapping the economic power of cities. *Journal of Monetary Economics*, *36*(March), 49. https://doi.org/10.3152/147154400781777368

Moretti, E. (2003). Human Capital Externalities in Cities. *Handbook of Regional and Urban*

Economics, *4*(April). https://doi.org/10.2139/ssrn.242708

Moretti, E. (2004a). Estimating the social return to higher education: Evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics*, *121*, 175–212. https://doi.org/10.1016/j.jeconom.2003.10.015

Moretti, E. (2004b). Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions. *The American Economic Review*, *94*(3), 656–690.

NRC. (2018). Grootste economische groei in Flevoland en Noord-Brabant. *24 April 2018*. Retrieved from https://www.nrc.nl/nieuws/2018/04/23/cbs-grootste-economische-groei-in-flevoland-en-noord-brabant-a1600485

Porter, M. (1990). The Competitive Advantage of Nations. *Administrative Science Quarterly*, *37*(3), 507–510. https://doi.org/10.2307/2393460

Rauch, J. E. (1991). Productivity Gains From Geographic Concentration of human Capital: Evidence From the Cities. *Journal of Urban Economics*, *34*, 380–400.

Romer, P. M. (1986). Increasing Returns and Long-Run Growth. *American Economic Review*, *94*(5), 1002–1037. https://doi.org/10.1086/261420

Shapiro, J. M. (2005). Smart Cities: Quality of Life, Productivity, and the Growth Effects of Human Capital. *NBER Working Paper No. 11615*, *88*(May), 324–335. Retrieved from http://www.nber.org/papers/w11615

Smith, A. (1775). An Inquiry into the Nature and Causes of the Wealth of Nations. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9089453

United Nations. (2018). *World Urbanization Prospects: The 2018 Revision*.

8. Appendix

Category	Population size
1	0 – 200.000
2	200.001 – 300.000
3	300.001 – 400.000
4	400.001 – 500.000
5	500.001 – 600.000
6	600.001 – 700.000
7	700.000 +

Table 1: Population categories

City	COROP area
Almere	Almere (CP)
Amsterdam	Amsterdam (CP)
Breda	West-Noord-Brabant (CR)
Eindhoven	Zuidoost-Noord-Brabant (CR)
The Hague	Aggl.'s-Gravenhage excl. Zoetermeer (CP)
Groningen	Overig Groningen (CR)
Nijmegen	Arnhem/Nijmegen (CR)
Rotterdam	Rijnmond (CP)
Tilburg	Midden-Noord-Brabant (CR)
Utrecht	Stadsgewest Utrecht (CP)

Table 2: COROP classification (Source: CBS)

Education level	Definition
Lower education	High-school (VMBO level completed)
Medium education	Senior-level (grade 3+) of Havo/VWO high-school level or MBO 2+
Higher education	Bachelor’s degree on HBO or University level

Table 3: Education levels (Source: CBS)

Variable	Model 1	Model 2
Breusch – Pagan Lagrange Multiplier	0.0	0.0
Hausman test	0.0	0.15
Serial correlation	0.0	0.0
Heteroskedasticity	0.0	0.0
Cross-sectional dependence (Pasaran CD)	0.0	0.0

Table 4: P-values statistical tests