Does the Finnish healthcare sector show signs of Baumol's cost disease?

Åbo Akademi University

Kristian Gästgifvars
Master's Thesis in Economics
Supervisor: Edvard Johansson
Åbo Akademi University
Faculty of Social Sciences, Business and Economics
2020

Åbo Akademi University – Faculty of Social Sciences, Business and Economics

Subject: Economics	
Writer: Kristian Gästgifvars	
Title: Does the Finnish healthcare sector show signs of Baumol's cost disease?	
Supervisor: Edvard Johansson	

Abstract:

Health expenditures are growing globally and the causes are somewhat unclear. Baumol's cost disease could potentially act as a viable explanation for rising healthcare costs. The formalisation and method constructed by Atanda, Menclova and Reed (2018) allow the estimation of two of Baumol's (1967) original propositions of the cost disease, which are seldom tested. The models measure how the healthcare sector's share of the labour force is affected by a) the relative productivity between health and non-health sector, and b) the total productivity in the economy. In this thesis, the method is applied to measure if the Finnish healthcare sector displays any signs of the cost disease.

The data are gathered from 70 sub-regions in Finland between 2000 and 2016. Sub-regions are by regional size between municipals and hospital districts. Primary care is provided by municipalities and specialised care is provided by hospital regions. The Finnish health system is exceptionally decentralised and the cross-regional variation is rather large.

The empirical impact of the cost disease is tested with three different fixed effects models, similar to the ones used in most of the contributions in the same field of study. The results suggest that relative and total productivity are positively related to healthcare share of the labour force, which are both anticipated theoretically as well. The effect is statistically significant but economically affects only 600-850 employees annually in the health sector, which employs around 400 000 people. Therefore, the cost disease is perhaps noticeable after a decade-long intersectoral transition in the Finnish labour market.

This thesis strengthens the results of Ministry of Finance (2013) study of the healthcare sector which suggests that Finnish healthcare is suffering from Baumol's cost disease but decreases the economical severity of the impact.

Key words: Baumol's cost disease, the Finnish healthcare system, productivity, labour market, panel data, fixed effects.

Date: 15.5.2020	**Number of pages:** 58

Table of Contents

Tables

Figures

1. Introduction

Health expenditures have grown rapidly in the last decades in the world and they account for a considerable amount of the national income in most developed countries. The proportion of health expenditures as a percentage of gross domestic product have increased from roughly four percent up to one tenth of the national income in most EU and OECD countries in the last 40 years (Medeiros & Schwierz, 2013; OECD, 2019). The Finnish healthcare expenditures have doubled in the same period and the costs continues to put increasing pressure on the public budget and the fiscal stability. Finland provides universal health coverage with 90 percent coverage rate and public health spending is expected to grow 2 % annually in the next decade (Dieleman, et.al., 2018). Today, managing healthcare expenditures is a challenge of a global scale, which will be even more difficult in the future.

The field of health economics has seen many detailed studies about the costs of the health sector. One main determinant of rising costs is argued to be the income of the population. In fact, no other variables have yielded as clear effects on health expenditure as GDP per capita (Gerdtham & Jönsson, 2000; Baltagi & Moscone, 2010; Benavides, 2018). A key metric that has naturally derived from the countless estimations is the income elasticity. Growth of income increases the seeking of care. However, there is no consensus of whether one should consider health services as necessities or luxury goods. Thus, there are no clear estimations in this particular field of the future of health expenditures either; will their growth continue to increase or eventually slow down. The income elasticity has been shown to vary across different levels of income and, somewhat surprisingly, richer countries even show a decreasing income elasticity of demand (Di Matteo, 2003).

Given that health expenditures have grown more than income in developed countries, implies that there are some omitted factors which have yet to be tested. In recent years, many authors have tried to explain this effect with technological differences and Baumol's cost disease (Hartwig, 2008; Colombier, 2012 & 2017; Atanda, Menclova & Reed, 2018).

This would explain rising costs if healthcare would be relatively unproductive compared to the rest of the economy. Baumol’s cost disease is a theory of the evolution of market structures, more specifically how different productivity growth rates can cause higher costs in some sectors over time. A non-productive sector would see less cost-decreasing innovations and be expected to make less profit and increase selling prices more than a productive sector.

Baumol’s (1967) defining characteristics of the cost disease, however, are quite difficult to apply in an econometric study. The sufficient data is rarely gathered and, if accessible, it might be incomparable with other countries. The idea of using cost disease as an explanation for globally rising expenditures might shed light on the ongoing discussion in health economics. Even though the theory was born half a century ago, the application in empirical research has only begun in the last decade. One difficulty is that there are yet no explicit explanatory variables explaining the cost disease. Baumol’s theoretical work provided only propositions of suggestible side effects, which are difficult to measure and newer studies are constantly trying to construct new models to resolve the problem.

1.1. Research objective

The aim of this thesis is to study if the Finnish healthcare market suffers from Baumol’s cost disease, by focusing on theory built by Baumol (1967) and following the approach that Atanda, Menclova and Reed (2018) have reconstructed. The new approach follows Baumol’s (1967) propositions more closely than prior studies of the cost disease in the healthcare sector. This thesis will contribute to the discussion by applying the new model for smaller regions in one country, instead of using it to measure effects between various heterogeneous countries. The data are collected from Finnish sub-regions between 2000 and 2016. Finland is divided into 70 sub-regions, which each contain 2-7 municipalities, and every sub-region belongs to a larger hospital district. Each hospital district has the sole responsibility to provide its own specialised care and each municipality is responsible for the provision of primary care, to the extent permitted by the national legal framework. The Finnish health system is quite decentralised, which gives the opportunity to study varying productivity measures under the same fundamental health structure.

1.2. Disposition

Chapter 2 will introduce the characteristics of the Finnish health system. The chapter discusses the organisational structure and the problems with efficiency it faces. Next, in chapter 3, the theory behind Baumol's cost disease is explained and how theory can be modified to test empirically. Chapter 3 includes a review of prior studies as from global data sets and one study from Finnish healthcare as well. The econometrical method and the data of Finland is presented in chapter 4. Chapter 5 presents the results of the main estimations in the thesis as well as a discussion of the findings from Finland. Chapter 6 concludes the thesis. The appendix provides additional definitions, calculations and estimation results, which might be beneficial for the reader.

2. The Finnish healthcare system

Chapter 2.1 briefly explains the characteristics of the health system in Finland; how it is governed and who is responsible for provision of care. Chapter 2.2 is an assessment of the health system and discusses the main problems with efficiency today.

2.1. Overview

Finland spends a relatively small proportion of its income on healthcare. Compared to its Nordic neighbours and other similarly governed countries, Finnish healthcare expenditures per capita and as a percentage of GDP have generally been less in recent years. In 2017, the per capita health expenditures were 4 100 €, which can be translated to a total of 9.2 % of the GDP. Health expenditures in Norway, Denmark and Sweden have exceeded 10 % of GDP. Approximately 20 % of the total health expenditures in Finland are paid with out-of-pocket payments, 70 % is financed with municipal and state taxes, and the rest is from other private sources. A third of the health expenditures come from private service providers in Finland, and the private sector has grown by 10 percentage points since 2000. In general, health expenditures have grown drastically in the better part of the last sixty years, which much like the rest of the developed countries, displays a concerning picture of the future. (Keskimäki et al, 2019; Dieleman, et.al., 2018).

The provision of care in Finland is quite decentralised. There are three levels of administration: The Ministry of Social Affairs and Health (MSAH), hospital districts and municipalities. MSAH is responsible for the legislation and direction of general health policy. MSAH prepares key reforms in healthcare and guides the implementation. The municipalities are solely responsible for organising the primary healthcare for their local residents. The municipalities may cooperate with public health provision, which is common in low-populated neighbouring regions in Finland. In 2020, the number of municipalities is 311, of which 70 % have under 10 000 inhabitants. Hospital districts are federations of municipalities and they are responsible for specialised medical care provision. Specialised medical care includes examinations and treatments in hospitals, which are performed by medical specialists. There are 20 different hospital districts and every municipality must be a member of one. Three quarters of the hospital districts are inhabited by more than 100 000 people, one of which has 30 % of the Finnish population. Furthermore, the care of most demeaning and rare diseases in Finland is centralised into five university hospitals, which are located in larger cities around the country. The five university hospitals do not have any explicit decision-making powers, but they plan the care and treatment of the most severe diseases. The flow of information and finances is thus quite complicated. The regional differences between Finnish municipalities are substantial as well, since the specialised care expenditures as of total health expenditures are almost uniformly distributed between 20 and 50 percent. (Keskimäki et al, 2019; Statistics Finland, 2019, Ministry of Social Affairs and Health, 2016a)

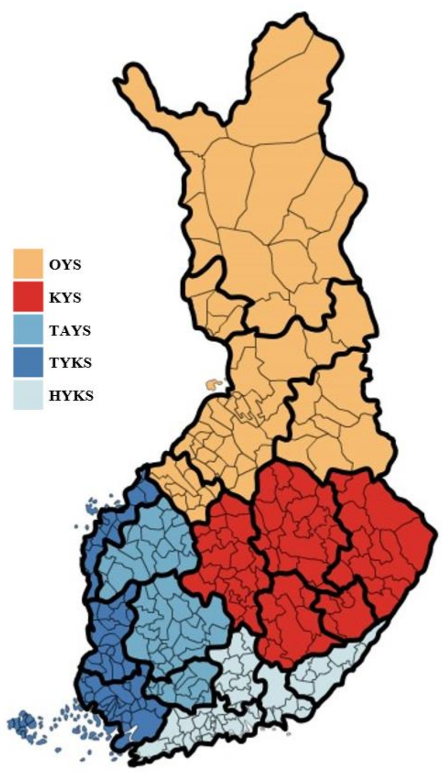

Figure 2.1 *Finnish municipalities, hospital districts and university hospital responsibility regions*. **Note:** Municipalities within thin lines, hospital districts within thick lines. OYS: Oulu University Hospital. KYS: Kuopio University Hospital. TAYS: Tampere University Hospital. TYKS: Turku University hospital. HYKS: Helsinki University Hospital.

Every Finnish resident is covered by national health insurance and one in six of the population have an additional voluntary health insurance. Out-of-pocket expenses have defined price caps in each individual medical case, but the expenses are higher than in the rest of the Nordic regions (Ministry of Social Affairs and Health, 2016b). There are three channels for first-contact care: the municipal system, national health insurance system and occupationally provided care. The municipal system provides the majority of all care and acts as gatekeepers for specialised care in public healthcare. The national health coverage system reimburses partially costs of prescribed medicines and health visits to private facilities. Additionally, many employees are covered by occupational health coverage, which is quite well developed in Finland. The private sector has a substantial stake in the occupational healthcare system and private care is often obtained much faster than publicly provided care. Occupational care is completely free of charge for the patients as well. (Teperi et al, 2009; Keskimäki et al, 2019).

2.2. Efficiency and productivity

Healthcare system in Finland performs quite efficiently compared to other countries around the world. Health outcomes have improved since 2000 and preventable mortality rates are below EU average. According to Keskimäki et al (2019), technical efficiency and productivity have grown considerably in specialised care and the pharmaceutical sector. Other sectors in healthcare perform relatively well compared to other countries, however, the allocation of resources between specialised and primary care in Finland is somewhat inefficient and unbalanced. Resources for development are provided more toward specialised care, which has been visible in the flow of medical professionals between sectors as well. Nationally, between 2000 and 2016 the number of patients per physicians decreased in health centres by 30 %, whereas it increased in hospitalised care more than 50 %. The decentralised nature of the health system makes it difficult to tackle this problem, since municipalities and hospital regions are on two different administrative levels and no clear cooperating procedure is applied. Any acts of reforms or renewals in healthcare are thus difficult to achieve and the interaction between administrative levels is considered complex by the decision makers. Public officials in healthcare and executive members of NGOs have different perceptions of the organisational structure of the healthcare system. The majority would prefer a more centralised system for healthcare provision and financing. (Ministry of Social Affairs and Health, 2016a; Keskimäki et al, 2019)

The access to care is quite unbalanced in Finland and unmet medical needs are considerably higher than EU average and the Nordic neighbours. (European Commission, 2020; Ministry if Social Affairs and Health, 2016b). Within Finland, access to primary and specialised care varies drastically depending on the geographical region. This is mostly due to small numbers of health centres in low populated municipalities. Waiting times is one factor of unmet medical needs, which also is an indicator of productivity. In economics, patients' waiting time can be seen as a function of number of healthcare professionals and number of patients. However, the population density is an insignificant factor in waiting times in Finland. On average, the proportion of the patients that waited over one week for primary care visits was 45 % in 2015 and it increased to 55 % in 2017. In fifteen municipalities, the same statistic was less than 10 %, and in five municipalities, every patient waited more than one week in 2017 (Finnish institute of health and welfare, 2020). There have been attempts to reach patients faster in recent years, which have turned out to be mostly unsuccessful. One explanation for this might be that a change in waiting times usually affects the behaviour of patients. The demand is expected to increase by shortening queues, which, in turn, eventually increases the waiting time. Often, it is rather difficult to determine the cause and effect of patients' waiting times in healthcare. However, in Atella et al (2019), the results suggest that lowering waiting times increases productivity in healthcare more than cost containment and regulation. (Iversen & Lurås, 2002; Ministry of Social Affairs and Health, 2016a; Keskimäki et al, 2019)

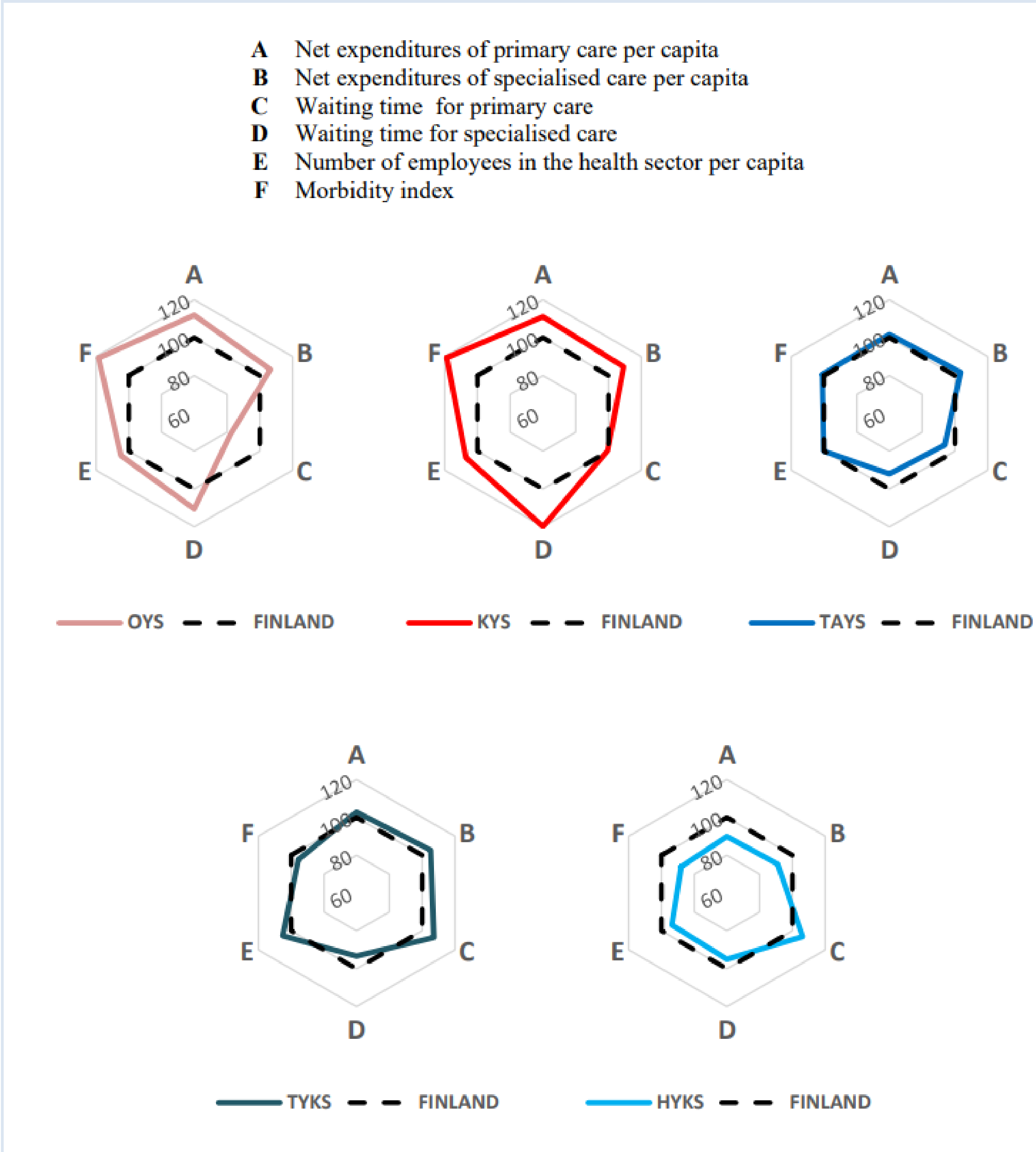

Figure 2.2 *Efficiency in university hospital responsibility regions 2016.* **Note:** Values indexed to the average values in Finland. OYS: Oulu University Hospital. KYS: Kuopio University Hospital. TAYS: Tampere University Hospital. TYKS: Turku University hospital. HYKS: Helsinki University Hospital.

Figure 2.2 displays the variation of different efficiency variables in Finnish university hospital regions (Finnish institute of health and welfare, 2020). A large area in the coloured hexagons indicates less efficient or unproductive regions. Generally, the Helsinki University hospital region (HYKS) would rank as the most efficient healthcare region in

Finland according to Table 2.2, where waiting time in primary care is the only indicator above the Finnish average. Expenditures per capita are relatively high in OYS and KYS, which could be explained by the regions' low population density. Waiting time for primary care is highest in TYKS and HYKS, whereas the waiting time for specialised care is shorter than average in these regions. The indicator for health care employees per capita has perhaps the lowest regional variation and the morbidity index has the highest.

After the general election in 2015, the government actively tried to improve the health and social care system with an extensive reform. The main objective was to centralise the system by integrating the provision of primary, specialised and social care under the same administrative organ within larger regions. The reform would have, presumably, decreased the disparity in healthcare access and helped to control costs. Alongside the structural changes, a crucial aim was to give the patients a greater freedom of choice, with the intention of increasing the competition in the health sector. However, the size of the reform and the partially conflicting objectives concluded with no actions taken and the government's early resignation. The reform was postponed and the current government have written a similar structural reform in their Government programme. (Sinervo, Tynkkynen & Vehko, 2016; Keskimäki, et al, 2019, Ministry of Social Affairs and Health, 2016b).

3. Theoretical Background and Empirical Review

This chapter presents thoroughly the theory and empirical research of Baumol's cost disease. The following sub-chapter explains how different elements in the market affect each other and how they are expected to change. Chapter 3.2 is a review of the known contributions for the application of the cost disease in the health sector. Lastly in chapter 3.3, the new approach to measure the cost disease is explained in great detail. The new model is used in the empirical part of this thesis.

3.1. Baumol's cost disease

The cost disease, commonly known as the Baumol's cost disease, is a theoretical framework to show why some sectors of the economy face expanding costs, even though other sectors' costs would remain stable. The key premise is that the output per man-hour grows at different rates in every sector of the economy. Some sectors can apply technological changes and innovations more naturally than others, which inevitably will enlarge the gap in productivity between various sectors over time. The second assumption is that the growth in wages is heavily correlated across all sectors. Wages tend to increase at the same rate in the whole economy, regardless of the differences in productivity levels. Growth in productivity implies that the same output volume can be produced with less labour and, therefore, average costs (per input unit) in slowly growing sectors are expected to increase as well.

Baumol (1967) introduced the formalised version of the cost disease and derived four propositions for the economy that would be visible. His analysis uses a simplified version of an economy with two sectors: a progressive and a non-progressive sector (henceforth PS and NPS), in terms of growth in productivity. Firstly, suppose that the output (Y), in both sectors is given by

$$Y_{NPS}(t) = aL_{NPS}(t) \qquad (3.1)$$

$$Y_{PS}(t) = bL_{PS}(t)e^{rt}, \qquad (3.2)$$

where L is the employment in each sector and a and b are arbitrary positive constants. The productivity in the NPS is constant over time, while it continuously grows at the rate of r in the PS.

Secondly, Baumol (1967) assumed that all wages in the economy grow at the same rate r as the rapidly growing sector and remains equal in both sectors for all t. The equation is shown as

$$W(t) = W_0 e^{rt} \ . \qquad (3.3)$$

The constant W_0 is an insignificant start value for the wage equation. Furthermore, given equations (3.1), (3.2) and (3.3), we can derive the unit costs in both sectors as

$$C_{NPS}(t) = \frac{W(t)L_{NPS}(t)}{Y_{NPS}(t)} = \frac{W_0 e^{rt} L_{NPS}(t)}{aL_{NPS}(t)} = \frac{W_0 e^{rt}}{a} \qquad (3.4)$$

and

$$C_{PS}(t) = \frac{W(t)L_{PS}(t)}{Y_{PS}(t)} = \frac{W_0 e^{rt} L_{PS}(t)}{bL_{PS}(t)e^{rt}} = \frac{W_0}{b} \ . \qquad (3.5)$$

By using (3.4) and (3.5) to show the relative unit cost of both sectors, we have

$$C_{NPS|PS}(t) = \frac{C_{NPS}(t)}{C_{PS}(t)} = \frac{W_0 e^{rt}/a}{W_0/b} = \frac{be^{rt}}{a} \qquad (3.6)$$

and a change in t can be expressed as

$$\frac{\partial\left(C_{NPS|PS}(t)\right)}{\partial t} > 0 \ . \qquad (3.7)$$

P1. **The first proposition**: Unit costs in the NPS will grow unrestrictedly over time, while the costs in the PS will remain constant. The demand for jobs and the labour market equilibrium will eventually produce a same rate of growth in wages in both sectors. Assuming that the wages in both sectors grow at the same rate as the productivity in the PS, the burden of labour wage growth will be increasingly difficult for the NPS. The relative costs between sectors increase, which is the key problem with Baumol's cost disease. Prices in each sector are in proportion to their costs, therefore, a positive effect of prices in the NPS is expected as well.

Following the issue of increasing relative costs between the NPS and PS, we can continue the analysis by examining the expected demand for NPS goods. Suppose that prices in the NPS market would be in an equilibrium and the price elasticity of demand for its goods would be negative[1]. The relative output on goods in the economy would thus be shown as

$$\frac{Y_{NPS}(t)}{Y_{PS}(t)} = \frac{aL_{NPS}(t)}{bL_{PS}(t)e^{rt}} \ . \tag{3.8}$$

Additionally, we can see that

$$\lim_{t\to\infty}\left(\frac{Y_{NPS}(t)}{Y_{PS}(t)}\right) = \frac{a}{b}\lim_{t\to\infty}\left(\frac{L_{NPS}(t)}{L_{PS}(t)}\cdot\frac{1}{e^{rt}}\right) = 0 \tag{3.9}$$

if the growth rate in $L_{NPS}(t)$ is less than the sum of growth rates in both PS labour and in PS productivity, which is certainly a reasonable assumption. In other words, the demand for NPS goods will approach zero over time.

P2. **The second proposition**: In an economy suffering from unbalanced growth, the NPS will produce less and less goods and the demand will eventually vanish. Only markets with highly inelastic price elasticity of demand will survive increasing

[1] Goods that cause a negative change in demanded quantity by a price increase.

costs. However, the government can keep the ratio of production between these sectors constant with transfers. Sometimes, the desired outcome is to maintain a constant level of demand and, thus, subsidise the production. This will, consequently, lead to higher transfer costs as well, as the difference in the rate of growth between the PS and NPS increases.

The next effect of Baumol's (1967) cost disease is derived from the assumption that an external part, such as the government, is subsidising the production. Given the total labour supply

$$L(t) = L_{NPS}(t) + L_{PS}(t) \tag{3.10}$$

and by multiplying equation (3.8) with an arbitrary positive constant $\frac{b}{a}$ we have

$$\frac{b}{a}\frac{Y_{NPS}(t)}{Y_{PS}(t)} = \frac{L_{NPS}(t)}{L_{PS}(t)e^{rt}} \ . \tag{3.11}$$

By keeping the production ratio between both sectors stable, we can write the right-hand side of the equation (3.11) equal to a constant γ. Now if we combine equations (3.10) and (3.11) we can express the required labour supply in both sectors with the following steps:

$$\frac{L_{NPS}(t)}{L_{PS}(t)e^{rt}} = \gamma \Leftrightarrow L_{NPS}(t) = \gamma L_{PS}(t)e^{rt}$$

$$L_{PS}(t) = L(t) - L_{NPS}(t)$$

$$L_{NPS}(t) = \gamma\big(L(t) - L_{NPS}(t)\big)e^{rt}$$

$$\Leftrightarrow L_{NPS}(t) = \frac{\gamma L(t)e^{rt}}{1 + \gamma e^{rt}} \ . \tag{3.12}$$

The labour function in the progressive sector can be derived in the same way and is expressed as

$$L_{PS}(t) = \frac{L(t)}{1 + \gamma e^{rt}} \ . \tag{3.13}$$

Finally, if we let the value t approach infinity in equation (3.12), the analysis is quite clear:

$$\lim_{t\to\infty} L_{NPS}(t) = \lim_{t\to\infty}\left(\frac{\gamma L(t) e^{rt}}{1 + \gamma e^{rt}}\right)$$
$$= \lim_{t\to\infty}\left(\frac{e^{rt}}{e^{rt}}\left[\frac{\gamma}{1/_{e^{rt}} + \gamma} L(t)\right]\right) = \lim_{t\to\infty} L(t) \tag{3.14}$$

The required labour supply in the NPS increases by time and approaches the limit of the total labour supply.

P3. **The third proposition**: If we hold the produced amount of goods constant in all sectors, the labour force of the PS will continue to shift towards the NPS until the employment in the PS is zero. Technological changes allow the PS firms to operate with fewer employees and they must move elsewhere. There seems to be no clear evidence that long-term unemployment has changed over time or has drastically been affected by technological growth (Lucas & Rapping, 1969; Feldmann, 2013; Khraief et.al., 2018). Thus, the change in the NPS employees per total labour force ought to be positive. It is highly unlikely that a market could run without human interaction, however, it is far more important to realise the shift in the employment distribution. Suppose that the rate of production is constant, by nature or by transfers: now the economy will face a situation where a growing proportion of the labour force is entering into a non-productive field of business, making it increasingly difficult to enhance the labour productivity.

Baumol's (1967) final proposition explores the effects of the whole economy if the ratio of production quantities between sectors is kept constant. If the relative costs are always rising and the total employment is moving to the NPS, the growth of the national income is going to slow down. Given the properties of the equations above, the rate of growth in total production is approaching zero.[2]

P4. **The fourth Proposition**: The resulting effects of the cost disease, if the rate of production between the two sectors is kept constant, is that the transfers puts pressure on the whole economy. The transfers required to retain the same ratio of output increases, which will force the growth of the whole economy to slow down. This implies that if the three prior effects of the cost disease hold, the economy would grow to a finite limit, after which the rate of change in growth would be zero.

Even though this formalised theory has been known for more than fifty years, it has proven to be problematic to estimate econometrically. One concern is that it is questionable to use productivity, or its growth, as an identical measurement between two sectors. Baumol (1996; 2012) argues in later discussion pieces that prices are a great indicator when analysing different markets, seemingly affected by the cost disease. The most important effect, according to Baumol, is increasing costs, which naturally influences the prices of the end-users. The first proposition (*P1*) is therefore an essential determinant of disparity in prices in the economy.

[2] See calculations in Appendix B1

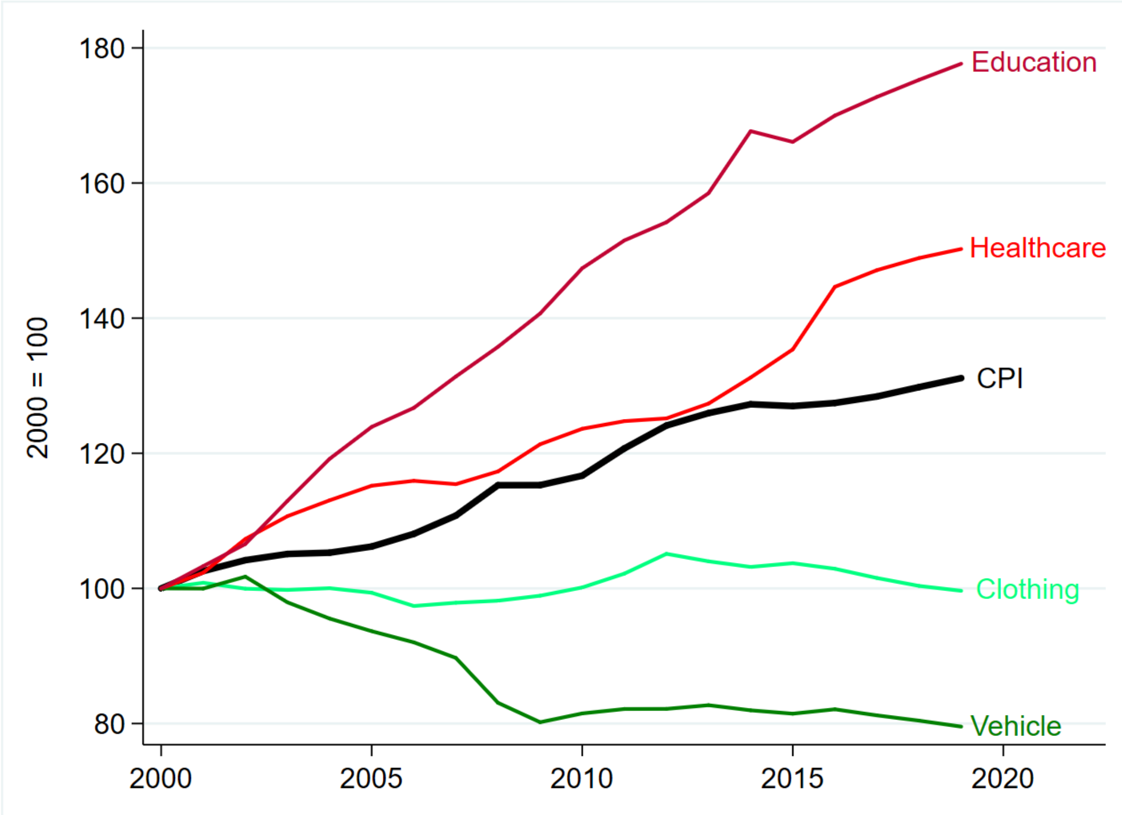

Figure 3.1 *Price indices in different industries in Finland 2000-2019.*

Figure 3.1 shows the evolution of the Finnish consumer price index and four other price indices, of which the red line graphs are usually seen as part of the NPS and the green ones undoubtedly part of the PS. The real-costs increases in both education and healthcare, would thus be the difference to CPI. Additionally, it is clear that the real price changes in progressive alternatives, vehicle and clothing industries, have been negative over the time period (Statistics Finland, 2019). Figure 3.1 reflects the first proposition well. However, the unbalanced price growth in an economy is only partially attributed by differences in productivity. Sectors with low degree of competition could as easily show similar patterns, so price indicators might be insufficient. Therefore, it is important to review the other propositions of the theory and use logical reasoning to verify if the cost disease would be truly a problem.

Healthcare should show signs of Baumol's (1967) propositions for other reasons as well. First, healthcare is part service industry, which generally shows less technical improvements. Secondly, since many countries promote universal health coverage and thus subsidise the sector, the demand is likely to increase rather than vanish. This makes healthcare suitable for analysis of the third proposition, which requires third-party subsidisation or a heavily inelastic price elasticity of demand. Finally, the distribution in the labour force in Finland has changed and the healthcare sector has shown growth in recent decades (Statistics Finland, 2019), which could be explained with the third proposition.

3.2. Empirical review of the cost disease in healthcare

In the last 12 years, there has been an intense discussion about Baumol's cost disease and its genuine effect on the globally growing health expenditures. In empirical studies, the dominant method is to create a variable, which will measure one of the properties of the cost disease and see how it will affect the total health expenditures in the specified region. Baumol (1996; 2012) argues constantly that prices are one of the best indicators to show the cost disease. However, many other authors find the medical price indices to be upwardly biased, due to improved quality of health production and thus troublesome measurements of real-price increases. (Triplett, 1999; Hartwig, 2008; Colombier, 2012 & 2017).

Hartwig's (2008) study was the first systematic approach of testing Baumol's cost disease in healthcare. In his study, he tests how the unbalanced growth affects healthcare expenditures in 19 OECD countries within the time period from 1960 to 2004. His approach follows closely Baumol's (1967) theoretical structure, by constructing a variable called the *Baumol variable*, which measures the essence of the cost disease discussed in chapter 3.1. The Baumol variable is calculated by taking the difference of the growth of overall wages and growth of total labour productivity, in each country. A positive value of said variable indicates a greater wage growth than productivity growth. The method is tested econometrically with numerous regression models, where the measurement of health expenditures is the dependent variable and the Baumol variable acting as an endogenous independent variable. The models are controlled by several known determinants of health expenditures as well, such as growth of national per capita income and employment. The results of Hartwig's (2008) study suggest that Baumol's cost disease is a significant explanatory factor of rising expenditures of healthcare.

The Finnish Ministry of Finance (2013) estimate the Baumol variable's effect on health expenditures from a Finnish time series data sample between 1975 and 2011. The econometrical approach is the same as Hartwig (2008), and the results of the estimations show similar coefficients. However, the estimations of the Baumol variable are problematic according to Ministry of Finance (2013). The multicollinearity in the model is severe and the autocorrelations in the variables might distort the effect. The estimation results are quite large which gives support to the Baumol's cost disease hypothesis in the health sector, but the magnitude of the variables' effects are somewhat unreliable.

Colombier (2012) uses a similar model as Hartwig (2008) to test the increasing health expenditures in OECD regions between 1965 and 2007, with a small but important correction. A closer examination of Hartwig's (2008) study reveals that it measures an assumption where all labour is allocated to the healthcare sector, which represents the non-progressive one. Thus, Colombier (2012) estimates econometrically the same model, corrected by the healthcare sector's share of total labour. As expected, the result's estimates show a much lower effect of the Baumol variable than Hartwig (2008), but a significant effect nonetheless. Colombier (2012) diminishes the impact of Baumol's cost disease as a severe threat and determinant of the growing costs of health.

Bates and Santerre (2013) use the same method as Colombier (2012) to test the effect in the United States, by using a state-level panel between 1980 and 2009. The main differing factor from the abovementioned studies is that Baumol's cost disease is tested only in one individual country. The results of the estimations are quite similar to Colombier (2012); a statistically significant but economically meaningless effect. Additionally, Bates and Santerre (2013) extend the method by doing a two-stages least squares (2SLS) estimation on how the Baumol variables affect health expenditures, in order to eliminate the possible endogeneity bias that may occur in a traditional state-fixed effects method. They use the housing-price inflation as an instrument for the Baumol variable. Wage growth is one part of the Baumol variable and since wages generally correlate with the prices of housing, it should perform as a reasonable instrument in the regression. The results of the 2SLS-model yield a threefold increase in the effect of the Baumol variable, compared to the first estimations of Bates and Santerre (2013). The authors' study raises the relevance of the cost disease in health production.

Colombier's (2017) study is an addition to his earlier paper, where he follows the 2SLS method which Bates and Santerre (2013) constructed. Colombier (2017) uses data of twenty OECD countries from 1970 to 2010 and uses the growth of the manufacturing sector as an instrument for the Baumol variable. The manufacturing industry certainly affects the growth of productivity in the entire economy and is exogenous to health expenditures. Colombier's (2017) model produces higher estimations of the Baumol variable than his previous study (Colombier, 2012) provided, and the coefficient is similar to Bates and Santerre (2013).

Atanda, Menclova and Reed (2018) build a different method to characterize the cost disease, by measuring how the healthcare share of the labour force and relative prices of health goods react to productivity growth (method discussed more closely in chapter 3.3). They studied 28 OECD countries in 1995-2016, and different sectors of the U.S. economy between the period of 1947 and 2016. Their findings are neither robust nor significant.

In a later paper, Atanda and Reed (2019) even write that their results show “no evidence to support the existence of Baumol’s Cost Disease” in healthcare and generally eradicate the Baumol variable’s effect on health expenditures. Their conclusions are harshly different from others’, which is certainly interesting given that they ultimately study the same data as the prior studies in the field (Colombier, 2012 & 2017; Bates & Santerre, 2013).

3.3. The new model

The propositions in chapter 3.1 are seldom tested in health expenditure studies. The main objective tends to be to test the difference in the growth of wages and the growth of total productivity. In Atanda, Menclova and Reed (2018), the original theory was rebuilt by examining the propositions that Baumol (1967) suggested. The underlying propositions are tested by excluding the Baumol variable (see chapter 3.2), since their model is partially built by using only the total productivity of the economy. The theory (Atanda, Menclova & Reed, 2018) shows mathematically that Baumol’s propositions should be visible only with the total productivity variable at time t, excluding the variable for difference in productivity. The authors’ formalization uses the disparity in productivity between NPS and PS as well, but they argue that the variable would be incomparable in cross-country data collections. Therefore, they show how the disparity variable would change if the total economy-wide productivity changes (see chapter 3.3.2). The formalized and testable versions of the cost disease propositions are displayed in chapters 3.3.1 and 3.3.2.

3.3.1. The relative productivity method

The first assumption, similar to Baumol (1967), is that the only input factor to the production is labour. Let the production of healthcare (H) and the remaining economy (NH) be written as

$$Y_H = L_H \tag{3.15}$$

$$Y_{NH} = \phi L_{NH} \ , \tag{3.16}$$

where ϕ is defined as the relative productivity between these sectors, measured in terms of output per labour as

$$\phi = \frac{Y_{NH}/L_{NH}}{Y_H/L_H} = \frac{Y_{NH}L_H}{Y_H L_{NH}} \ . \tag{3.17}$$

The productivity of healthcare is thus generalized to one, and $\phi > 1$ suggests that the productivity in NH is higher than in healthcare. The variable ϕ measures thus the productivity gap and is one of the key elements in this analysis.

Furthermore, by assuming an equilibrium where demand is equal to the supply in the rest of the economy, we define the percentage of NH demand of the total output as a variable k, so that

$$kY = Y_{NH} \tag{3.18}$$

and the equation (3.16) can be substituted to (3.18) as

$$kY = \phi L_{NH} \ . \tag{3.19}$$

Naturally, the whole labour supply L is the sum of both sectors' labour and the production Y is a function of labour productivity, so that

$$L = L_H + L_{NH} \tag{3.20}$$

$$Y = L_H + \phi L_{NH} \ . \tag{3.21}$$

By writing the whole NH labour supply given equation (3.19) as

$$L_{NH} = \frac{kY}{\phi}\ , \tag{3.22}$$

we get L_{NH} and L_H as functions of each other:

$$L_{NH} = \left(\frac{k}{\phi}\right)(L_H + \phi L_{NH})$$

$$L_{NH} = \frac{k}{\phi(1-k)} L_H \tag{3.23}$$

$$L_H = \frac{\phi(1-k)}{k} L_{NH}\ . \tag{3.24}$$

By substituting equation (3.20) in equations (3.23) and (3.24), and dividing both sides with the total labour, we get both sectors' share of the labour force with

$$\frac{L_{NH}}{L} = \frac{k}{k + \phi(1-k)} \tag{3.25}$$

$$\frac{L_H}{L} = \frac{\phi(1-k)}{k + \phi(1-k)}\ . \tag{3.26}$$

We have now one measurable hypothesis of the cost disease, discussed in chapter 3.1. The third proposition in Baumol's (1967) suggests that if the ratio of production between NPS and PS is held constant, the employment in the economy must move towards the NPS. The new model (Atanda, Menclova & Reed, 2018) shows this proposition as well. The first derivative of equation (3.26) with respect to ϕ can be shown as

$$\frac{\partial\left(\frac{L_H}{L}\right)}{\partial\phi} = \frac{k(1-k)}{(k+\phi(1-k))^2}$$

and since both k and ϕ are by definition greater than zero, we get

$$\frac{\partial\left(\frac{L_H}{L}\right)}{\partial\phi} > 0 \ . \tag{3.27}$$

Therefore, if the supply of health is kept constant, a change in relative productivity would result in a growing portion of the population working in healthcare. Additionally, the properties of the equation anticipates that a negative change in $\frac{L_H}{L}$ with respect to k, or

$$\frac{\partial\left(\frac{L_H}{L}\right)}{\partial k} = \frac{-\phi}{\left(k+\phi(1-k)\right)^2} < 0 \ , \tag{3.28}$$

when $\phi > 0$. Equation (3.28) measures the market effect if the demand for NH-goods grows. The change is quite intuitive; if the production in either sector rises, more labour is required in that sector when the production gap is held constant.

In conclusion, we express the H proportion of the labour force as an equation of demanded ratio of NH goods in the economy (k), and a measurement of the disparity in productivity (ϕ), with equation (3.26). This measures elements of *P2* and *P3* as discussed in chapter 3.1. As a result, the work done by Atanda, Menclova and Reed (2018) gives us equations (3.27) and (3.28) as two testable consequences of Baumol’s cost disease.

3.3.2. The total productivity method

Atanda, Menclova and Reed (2018) construct their econometrical methods without the disparity variable (ϕ). The argument is that the necessary comparable data to generate the variable is unavailable for a cross-country panel and would consequently be pointless. Their analysis, however, is generated from all the variables in chapter 3.3 and uses the same theoretic assumptions, but the main variable to measure the Baumol's cost disease is the cross-sector productivity in each country. The mathematical proof presented in the study is built the following way:

Given the equation (3.21), we can obtain the weighted average productivity in the economy with

$$Prod = \frac{Y}{L} = \frac{\phi L_{NH} + L_H}{L} = \phi\left(\frac{L_{NH}}{L}\right) + \frac{L_H}{L} \qquad (3.29)$$

and the labour shares in both sectors are in itself functions of ϕ, which means that

$$Prod = f(\phi). \qquad (3.30)$$

If we only can use the economy-wide productivity ($Prod$) to demonstrate the effect in equation (3.27), we will need to prove that

$$\frac{\partial\left(\frac{L_H}{L}\right)}{\partial Prod} = \frac{\partial\left(\frac{L_H}{L}\right)}{\partial \phi} \cdot \frac{\partial(\phi)}{\partial Prod} > 0 \qquad (3.31)$$

in order to estimate that there is a positive change in the health share of the labour force if the productivity increases. Since $\frac{\partial\left(\frac{L_H}{L}\right)}{\partial \phi}$ is already shown before, sufficient proof of equation (3.31) can by substitution of equations (3.25) and (3.26) be derived the following way:

$$Prod = \frac{Y}{L} = \phi\left(\frac{L_{NH}}{L}\right) + \frac{L_H}{L} = \phi\frac{k}{k + \phi(1-k)} + \frac{\phi(1-k)}{k + \phi(1-k)}$$
$$= \frac{\phi}{k + \phi(1-k)} \tag{3.32}$$

and

$$\frac{\partial Prod}{\partial \phi} = \frac{k}{\left(k + \phi(1-k)\right)^2} \ . \tag{3.33}$$

The variable k is by definition positive, which means that $\frac{\partial Prod}{\partial \phi} > 0$ for every $\phi \in \mathbb{R}_+$. The $Prod$ function is unique and strictly increasing. Therefore, there exists an inverse function $f^{-1}(\phi)$ and it is derived as

$$f^{-1}(\phi) = \phi = \frac{Prod \cdot k}{1 - Prod \cdot (1-k)} \ . \tag{3.34}$$

The derivate of ϕ with respect to $Prod$ is

$$\frac{\partial \phi}{\partial Prod} = \frac{k}{\left(1 - Prod \cdot (1-k)\right)^2} \tag{3.35}$$

and given the properties of the used variables, $\frac{\partial \phi}{\partial Prod}$ is greater than zero as well. Therefore, the inequality in equation (3.31) holds, an increase in the total productivity will result in a higher share of labour in healthcare as well, similar to the effect in equation (3.27).

With this proof (Atanda, Menclova & Reed, 2018), we can measure the same effect in healthcare labour by a change in the total productivity ($Prod$), as we would with the relative productivity between health and non-health sector. This is extremely useful when estimating effects from a cross-country data panel, since the output or labour in the health sector is rarely available. Baumol's third proposition will be tested in this study by estimating if both $\frac{\partial\left(\frac{L_H}{L}\right)}{\partial \phi}$ and $\frac{\partial\left(\frac{L_H}{L}\right)}{\partial Prod}$ is greater than zero.

4. Method

This chapter presents the methodological approach to evaluate if the Finnish healthcare suffers from Baumol's cost disease. Chapter 4.1 describes and displays the gathered data, and the following part shows how it is applied in an econometric analysis.

4.1. Data

The data used in this study is a strongly balanced panel of Finnish sub-regions between 2000 and 2016 (Statistics Finland, 2019; Finnish institute for health and welfare, 2020). Finland is divided into 19 regions, which each contains between one to seven sub-regions. The total number adds up to 70 sub-regions, 67 of which are on the mainland Finland. Figure 4.1 shows the geographical division of Finnish sub-regions.

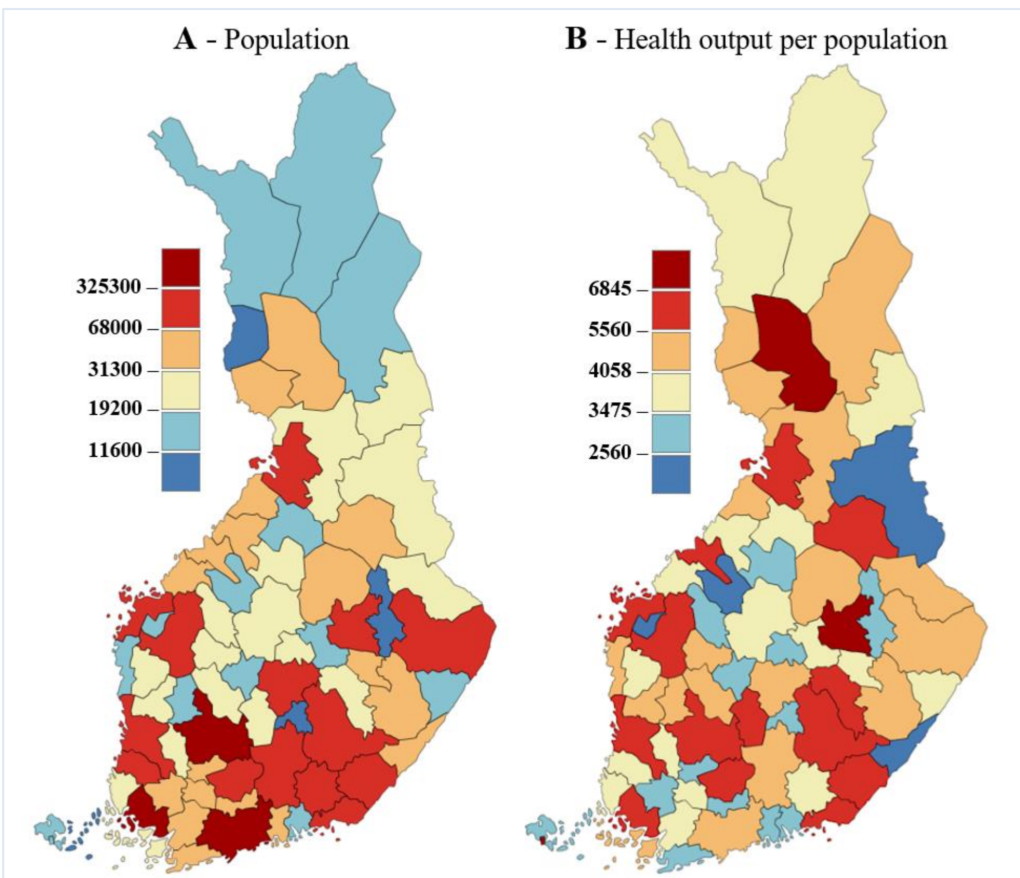

Figure 4.1 *Population (A) and health output in EUR (B) in Finnish sub-regions in 2016.*

Note: the colour distribution is divided into the 5th, 25th, 50th, 75th and 95th percentiles.

The Finnish sub-regions, shown in Figure 4.1, contain between two to seventeen municipalities, depending on population and land area. The dark red-coloured areas in 4.1A; Helsinki, Tampere and Turku sub-regions, represent 40 % of the nation's population. Finnish inhabitants are somewhat clustered around the most populated municipalities and centred more towards the southern part of the country.

Regarding healthcare, the Finnish municipalities are responsible for primary care in their own regions and the provision and responsibility of specialised care services are divided into 20 Hospital districts nationally (Local and Regional Government Finland, 2020; Keskimäki et al, 2019). Considering the regional size, sub-regions are conveniently located between municipalities and hospital regions. This will hopefully capture elements of both primary care and specialised care in the regressions in this thesis. In Finland, sufficient data at the municipal level is unavailable and gathering data at the hospital level would drastically decrease the number of observations and potentially place too much emphasis on specialised care provision. Either way, the Finnish health-system is highly decentralised, which allows sizeable variations in interregional health statistics.

The main variables in this study, which are used in the models in chapter 3.3, are the ratio of labour working in healthcare, the total production of non-healthcare goods and the relative labour productivity between non-healthcare and healthcare production. One strength in this study is that it uses data provided by Statistics Finland's (2019) of total production, divided separately by industry sector in the Finnish economy. In this study, the healthcare sector is represented by the sum of total production in human health and social work activities (Statistics Finland, 2019). Figure 4.1 B illustrates the variation of healthcare output per population.

As shown and discussed above, there are clear differences in health related variables between Finnish sub-regions. Interestingly, the variation is geographically disseminated and there are no clear differences between hospital districts. Therefore, it is useful to measure the effects of the cost disease with a cross-region fixed effects model. Fixed effects models are often used in health economics studies because they can measure more causal effects for variable coefficients by excluding the time and cross-regional effects. In cross-country panels, the relative productivity variable is yet to be observable, since the data of production in each sector is unavailable and often non-comparable between countries (Atanda, Menclova & Reed, 2018). The cross-regional Finnish data, however, is gathered and measured by the same statistics institution.

Table 4.1 Descriptive statistics of Finnish sub-regions 2016.

Variable	All economic regions			Regions divided by income[1]			
				High		Low	
	Mean	min / max	Std.dev	Mean	Std.dev	Mean	Std.dev
LHL , %	15.9	10.4 / 23.0	2.7	15.9	2.7	16.1	2.6
ϕ	2.44	1.41 / 8.34	0.85	2.47	0.88	2.28	0.55
k , %	92.3	86.7 / 98.2	2.3	92.4	2.4	91.9	2.0
% of private health production	31.9	3.4 / 77.1	7.3	32.3	6.3	29.7	1.1
Costs of specialised healthcare per capita	1198	1006 / 1744	182	1174	162	1325	225
Cancer index [2]	100	64 / 112	9.0	101	7.5	92	12.2
Circulatory disease index [2]	100	45 / 150	12.5	94	15.7	108	21.5
Population [3], x 1000	78.6	2.12 / 1519	22.5	132	256.9	25.1	14.3
GDP per capita	39 350	18 998 / 74 160	9517	41 705	8498	26 950	2063
Age > 64 , %	20.9	14.7 / 35.7	4.5	19.8	3.8	26.8	3.8
Unemployment, %	13.3	2.7 / 19.3	2.5	13.3	2.3	13.3	3.2
Observations	70	-	-	35	-	35	-

Note: Values are weighted by population in each region. Monetary values measured in EUR. [1] The division is calculated from median values GDP per capita. [2] Higher values indicate higher level of morbidity, index 100 = total population in Finland, 2016. [3] Average values unweighted by population.

Table 4.1 presents data on variables likely to affect the Finnish healthcare sector, for the whole population and separately for high- and low-income regions. Most variables are weighted by population, since the variation is larger in small sample groups, and the regions with low population will unlikely represent the average population in Finland. The first three variables are applied in the theoretical method of measuring Baumol's cost disease (see chapter 3.3). The variable LHL is calculated by dividing the number of employed individuals in the healthcare sector by the total labour force in each region. In absolute terms, approximately 400 000 people are working in the health sector in Finland and the variation between regions is actually quite small. The variable ϕ measures the disparity in productivity between non-health and health sectors. As we can see in Table 4.1, the labour productivity in the non-health sector is considerably higher. The average productivity in the rest of the economy is more than two times higher than in healthcare, which gives support to the main hypothesis in this study. The variable k represents the demand for non-health goods. k is calculated by subtracting the healthcare output from total output and dividing the difference by the total output in the economy. Thus, the share of demand for health goods is represented by $(1-k)$. The demand for healthcare goods adds up to nearly 10 % of the total demand and the regional variation is rather large.

The next four variables in Table 4.1 represent the institutional variables in healthcare. Private health production counts for almost a third of the total health production in Finland, thereby making it a crucial factor in the analysis. In the last twenty years, the private health sector has grown by ten percentage points. The costs of specialised care varies substantially in different regions and the mean value is quite high, compared to the total health outputs per capita in Figure 4.1 B. Variables *Circulatory disease index* and *Cancer index* measure the morbidity level in sub-regions of the most common causes of death in Finland (Keskimäki et al, 2019). Both indices differ considerably between high- and low-income regions, cancer being more common in wealthier regions and circulatory diseases affecting more low-income regions.

The rest of the rows in Table 4.1 show the demographic and economic variables. The demographic variables are often regarded essential influencing factors in healthcare system analyses (Hitiris, 1997) and especially income considered as the most important determinant of health expenditures (Gerdtham & Jönsson, 2000; Baltagi & Moscone, 2010; Benavides, 2018). The age structure, population and income are therefore controlled for in this study, as well. The population in Finland has grown only slightly in the last two decades, at a yearly growth rate of 0.4 percent. Concurrently, the age structure has changed drastically. The proportion of elderly population has grown with five percentage points between 2000 and 2016, and accounts now for approximately one fifth of the Finnish population. In low-income regions, the same statistic is more than a 25 %. In some regions the ratio of over 64-year-old inhabitants is nearly 40 %. The variation in demographic variables between sub-regions is thus quite large.

4.2. Primary empirical model

This study will use a two-way fixed effects regression model, to test Baumol's (1967) cost disease hypothesis, i.e. the equations (3.27) and (3.28) discussed in chapter 3.3. According to the theoretical framework (Atanda, Menclova & Reed, 2018), there are two affecting variables in the healthcare share of the labour force equation (see eq. (3.26)), thus giving two endogenous variables in the regression. For the Finnish healthcare system, the measured model is specified the following way:

$$LHL_{it} = \beta_1 \phi_{it} + \beta_2 k_{it} + \upsilon \cdot Z_{it} + u_{it} \quad , \tag{4.1}$$

where the error term is sum of non-observable region-specific effects and secular year-specific effects and individual time trends, defined as

$$u_{it} = \alpha_i + \eta_t + \mu_i \cdot t + \varepsilon_{it} \quad .$$

LHL is the proportion of the labour force working in the healthcare sector, ϕ is the relative productivity between sectors and k the ratio of production in the non-healthcare sector.

All endogenous variables are explained in greater detail in chapter 3.3 and Table A1. Equation (4.1) includes also a dot product of a vector of exogenous explanatory variables, Z, and a vector of coefficients, υ. The vector Z contains several variables (see Table 4.1 & Table A1), which are used in numerous prior health economics studies, that reviews the determinants of healthcare expenditures (Gerdtham & Jönsson, 2000; Hartwig, 2008; Colombier, 2012; Bates & Santerre, 2013). The econometric model uses only unique observations in region i at time t ($i = 1, \dots, 67; t = 1, \dots, 17$) and excludes the time and regional invariant effects of healthcare. Thus, any economic trends affecting the whole country, such as inflation and price indices, cannot be controlled for and will not be included in the model. Additionally, by allowing the model to be controlled for individual time trends will absorb the heterogeneity that might arise with time between regions. (Allison, 2009; Wooldridge, 2007 & 2012; Dynarski, Jacob & Kreisman, 2018).

4.3. Other variable measurements

Due to insufficient data, Atanda, Menclova and Reed (2018) measures Baumol's cost disease with the same model as in equation (4.1), but with different endogenous variables. As we saw in chapter 3.3.2, proposition 3 in Baumol's (1967) theory can be measured with the economy-wide productivity. Atanda, Menclova and Reed (2018) utilize this property and substitute the relative productivity variable (equation (3.17)) with the total productivity in different countries, measured as the ratio of GDP to the number of hours worked. In this study, the same detail will be applied i.e. substituting ϕ_{it} in equation (4.1) with the variable $Prod_{it}$, which is measured as

$$Prod = \frac{GDP}{total\ hours\ worked} \ . \tag{4.2}$$

Additionally, dividing the production by the total number of working hours might raise the accuracy of calculating the productivity. By measuring the required time of producing one unit of goods instead of number of employees in the sector, eliminates the possible omitted effect that could emerge from differences in length of standard working days between sectors. The standard equilibrium for working hours might even be different in each sub-region depending on the economic incentives and demographics (Alesina, Glaeser & Sacerdote, 2005).

5. Results

This chapter presents the results of the econometric models presented in chapter 4.2. The results in this chapter follow closely the rebuilt theory by Atanda, Menclova and Reed (2018), which is presented in chapter 3.3.

5.1. The relative productivity method

Table 5.1 summarises the determinants of health sector labour share with three different OLS estimation methods: a region-fixed effect estimation (FE), a two-way fixed effect (2WFE) estimation, which includes year dummy variables, and finally a fixed effect estimation with region-specific time trends (FErTT). All three models show the raw effect of the variables, which theoretically should affect the dependant variable (Atanda, Menclova and Reed, 2018), followed by a set of control variables affecting both the demand and supply of health production. All estimations are controlled for demographic and economic variables as well. The logarithmic dependent variables are chosen in non-percentage variables by tentatively maximising the coefficient of determination (R^2) in the model and applying it in numerically large observation values. GDP/cap growth rate is measured with the first difference of the natural logarithm of per capita income ($d.\ln(GDP\ per\ capita)$), which is a standard way of controlling the effect of income elasticity in recent health economics studies. It mitigates the problem of non-stationary time series and spurious effects. The coefficients show the significance level with cluster-robust standard errors, which are found to be more reliable when the regression sample is clustered by regions, much like the Finnish data set does. (Barros, 1998; Bowerman, OConnell & Koehler, 2005; Abadie, Athey, Imbens, & Wooldridge, 2017)

In the first column in Table 5.1, both main variables are significantly different from zero. The effect of the relative productivity (ϕ) is positive and share of demand for non-health goods (k) is negative, which is in line with the predictions in equations (3.27) and (3.28). By adding the control variables in column 2, the coefficient of the main variables remains virtually identical to the first model. All institutional variables have a small insignificant effect. Growth of GDP per capita is negative but insignificant, which is unusual given the strong relationship between health expenditure and income, but the effect is similar to what Atanda, Menclova and Reed (2018) estimated. The age structure seems to have a quite large and significant effect on the labour in healthcare. The higher the proportion of young and elder population, the larger share of health personnel is required in the region. The positive effect is intuitive, since high birth rate regions employ more gynaecologists and paediatricians in maternity and child health clinics and the older population generally require care more frequently. The estimated effect of unemployment in the second column in Table 5.1 is statistically no different from zero. In other words, the results suggest that the health sector is unaffected by the changes in the total labour force. High public indebtedness should intuitively affect the spending and saving in each municipality and thus the planning of primary care provision. However, there is no evidence that indebtedness would affect health production in column 2 either.

The models in the third and fourth column in Table 5.1 include a set of year dummies to exclude the time-specific effects in Finnish healthcare. The effects of the first two variables are both in line with the tested theory and roughly the same as in the regular fixed effect model. The effect of the share of under 15-year-old and over 64-year-old population is slightly larger and remains significant. Otherwise, the control variables show similar results as the first two columns. The F-test of the time dummies are jointly significantly equal to zero, indicating that the 2WFE is more appropriate than the FE model (Baum, 2006). However, R^2 is only slightly greater in the 2WFE alternative, which means that time-specific changes offer only little explanation of the variation in LHL.

The third model combines the fixed effects with individual time trend effects. Therefore, the model allows variables to grow at different rates in each region and control for unobserved time trends that might affect health sector share of the labour force within every region (Bates & Santerre, 2015). The FErTT models in Table 5.1 show a slightly larger positive effect for the relative productivity variable, but ultimately the significant coefficient remains close to the estimations in columns 1–4.

Table 5.1 Fixed effects estimations with relative productivity. Dependant variable: Healthcare sector share of the labour force (LHL).

	Fixed effects		Two-way fixed effects		Fixed effects with region specific time trends	
VARIABLES	(FE)		(2WFE)		(FErTT)	
	(1)	(2)	(3)	(4)	(5)	(6)
ln (ϕ)	0.0760**	0.0708**	0.0690**	0.0720**	0.0810**	0.0793**
	(0.0051)	(0.0059)	(0.0057)	(0.0054)	(0.0048)	(0.0049)
k	– 1.386**	– 1.332**	– 1.167**	– 1.340**	– 1.254**	– 1.429**
	(0.0877)	(0.0872)	(0.1190)	(0.0873)	(0.1380)	(0.0845)
Private health production		– 0.008		– 0.012		– 0.018**
		(0.008)		(0.008)		(0.004)
ln (costs of specialised care/capita)		– 0.002		– 0.005		0.002
		(0.004)		(0.005)		(0.002)
ln (sickness allowances)		– 0.006		– 0.003		0.004
		(0.006)		(0.007)		(0.005)
ln (cancer index)		– 0.001		– 0.001		– 0.002
		(0.004)		(0.004)		(0.004)
ln (circulatory disease index)		0.001		0.004		0.001
		(0.003)		(0.003)		(0.002)
GDP/cap growth rate		– 0.002		– 0.001		– 0.001
		(0.002)		(0.002)		(0.001)
Age > 64		0.141**		0.242**		0.101
		(0.035)		(0.058)		(0.062)
Age < 15		0.218*		0.366**		0.015
		(0.106)		(0.113)		(0.093)
Unemployment		0.009		0.010		0.024
		(0.033)		(0.058)		(0.017)
Relative public indebtedness		0.002		– 0.001		– 0.003
		(0.005)		(0.006)		(0.005)
Constant	1.368**	1.300**	1.165**	1.241**	1.241**	1.353**
	(0.080)	(0.099)	(0.109)	(0.108)	(0.126)	(0.081)
Observations	1 190	1 005	1 190	1 005	1 190	1 005
Adj. R^2	0.698	0.790	0.740	0.798	0.966	0.978
Number of regions	70	67	70	67	70	67
Time fixed effects	No	No	Yes	Yes	Yes	Yes
Indiv. region time trends	No	No	No	No	Yes	Yes
F-test (time effects)	-	-	p =.000	p =.001	p =.000	p =.000
F-test (linear time trends)	-	-	-	-	p =.000	p =.000

Note: Cluster-robust standard errors in parentheses. F-test: tests if the specified variables are jointly significant in the model, H0: no significant effects. ** p <0.01, * p <0.05.

The non-health sector demand is negative and significant. Given the earlier models, the private health production show interestingly a significant negative effect on health sector labour force. The effect is close to those in column 2 and 4, but the standard error in the estimation is smaller. This would indicate that the stronger private health sector would reduce the effect of Baumol's cost disease. The additional health variables continue to be insignificant, which means that the population's health status might have no effect on the number of health care personnel. The age structure has a smaller impact on LHL in the FErTT models and the rest of the economic variables remain the same. Models in column 5 and 6 are tested for an F-test for linear time trend effects as well and the trends are jointly significant.

The results show quite robust estimations in all columns (Table 5.1, see Appendix C). The effects of the control variables are mostly similar to findings in other health economics studies. The Finnish Ministry of Finance (2013) found a significantly positive relationship with health expenditure and private production, which would indicate that private production has an effect on contributing factor of health production. Private health production has negative effects in estimations in Table 5.1, of which one was clearly significant. Growth of GDP has no effect on proportion of healthcare labour in Table 5.1, which is the case in Atanda, Menclova and Reed (2018) as well. Generally, growth of income tends to affect health production positively (Barros, 1998; Gerdtham & Jönsson, 2000; Bates & Santerre, 2013), but the dependent variable is usually different. The age structure seems to have only a slight effect on health production in several studies (Ministery of Finance, 2013; Atanda, Menclova and Reed, 2018; Colombier (2017). However, the estimations in Table 5.1 show a strong positive effect on the labour in healthcare as a control. The age variables have obviously a high correlation with each other which might affect the actual coefficient in Finland (see Table A3). The regional unemployment seems to be an unimportant factor for the labour force in the health sector, which is in line with the estimates of Bates and Santerre (2013; 2015).

5.2. The total productivity method

The model, which Atanda, Menclova and Reed (2018) use to estimate the cost disease in their paper is applied in this paper as well. As discussed in chapter 4.3, the model in equation (4.1) can be tested, and the expected values of the coefficients are the same. The estimations and each variable's effects are presented in Table 5.2, exactly in the same way as in table 5.1 except for the first variable.

In Table 5.2, the two first columns present the fixed effect model without controlling for year dummies. The effect of the total productivity is significantly positive and the demand for non-health goods (k) is negative. The positive effect of ln $(Prod)$ is less than the effect of ln (ϕ) in Table 5.1, which is actually expected in the theory (see calculations in Appendix B2). The control variables are widely the same. More private health production decreases the number of health care personnel in the area and the effect is significant, unlike the corresponding coefficient in Table 5.1 which is much smaller. Otherwise, all the institutional variables are insignificant and similar to the first estimations. The growth of GDP per capita is negative and the age structure variables are significantly positive. Both unemployment and region's public indebtedness is statistically no different from zero. The R^2 value in columns (1) and (2) are approximately half of which they are in Table 5.1, which could be anticipated with a more generalised version of the productivity measurement of healthcare production.

Columns (3) and (4) show the estimations with time fixed effects (2WFE). The raw model with the two endogenous variables fades the effect of total productivity. The effect is close to zero and insignificant. However, by adding the group of control variables, the effect of the total productivity increases to a similar level as in column (2). The variable k has a negative significant effect and show more robust effects than $Prod$. Private health production is close to its estimations in columns (1) and (2) and significant. Specialised healthcare costs, sickness allowances on the morbidity indexes are insignificant in column (3) and (4). Economic and demographic variables are roughly the same as in the first two columns.

Table 5.2 Fixed effects estimations with total productivity. Dependant variable: Healthcare sector share of the labour force (LHL).

	Fixed effects		Two-way fixed effects		Fixed effect with region specific time trends	
VARIABLES	(FE)		(2WFE)		(FErTT)	
	(1)	(2)	(3)	(4)	(5)	(6)
ln $(Prod)$	0.0227**	0.0106**	– 0.0001	0.0114**	0.0004	0.0124**
	(0.0020)	(0.0036)	(0.0035)	(0.0038)	(0.0039)	(0.0044)
k	– 0.339**	– 0.414**	– 0.206**	– 0.405**	– 0.200**	– 0.430**
	(0.0239)	(0.0322)	(0.0265)	(0.0332)	(0.0276)	(0.0347)
Private health production		– 0.043**		– 0.046**		– 0.057**
		(0.006)		(0.007)		(0.007)
ln (costs of specialised care/capita)		– 0.004		– 0.007		– 0.000
		(0.003)		(0.005)		(0.004)
ln (sickness allowances)		– 0.010		– 0.002		0.002
		(0.006)		(0.007)		(0.006)
ln (cancer index)		– 0.001		0.000		– 0.006
		(0.004)		(0.004)		(0.004)
ln (circulatory disease index)		0.004		0.006		0.009**
		(0.003)		(0.003)		(0.003)
GDP/cap growth rate		– 0.007		– 0.004		– 0.006
		(0.004)		(0.004)		(0.003)
Age > 64		0.282**		0.336**		0.207**
		(0.033)		(0.053)		(0.069)
Age < 15		0.347**		0.432**		0.379**
		(0.073)		(0.087)		(0.126)
Unemployment		– 0.008		0.053		0.004
		(0.023)		(0.035)		(0.025)
Relative public indebtedness		0.007		0.007		– 0.006
		(0.004)		(0.005)		(0.005)
Constant	0.380**	0.444**	0.329**	0.369**	0.331**	0.385**
	(0.026)	(0.051)	(0.025)	(0.059)	(0.028)	(0.055)
Observations	1 190	1 005	1 190	1 005	1 190	1 005
Adj. R^2	0.259	0.444	0.362	0.445	0.897	0.926
Number Regions	70	67	70	67	70	67
Time fixed effects	No	No	Yes	Yes	Yes	Yes
Indiv. region time trends	No	No	No	No	Yes	Yes
F-test (time effects)	-	-	p =.000	p =.003	p =.007	p =.000
F-test (time trends)	-	-	-	-	p =.006	p =.000

Note: Cluster-robust standard errors in parentheses. F-test: tests if the specified variables are jointly significant in the model, H0: no significant effects. ** p <0.01, * p <0.05.

The third pair of estimations in Table 5.2 show the variables' effect on healthcare share of the labour force with the FErTT model. Compared to the 2WFE models, the total productivity variable behaves in a similar way by having no impact to LHL in the raw estimation and a significant coefficient with the control variables. The demand for non-health goods has a quite strong negative effect, which is expected since the level of demand should affect the produced quantity of health, thus, affect the labour within the sector as well. The percent of private health output has a significant negative impact on the health labour force in the sixth column and the age structure remain an essential part in the estimations. Given the results in Table 5.2, a younger population requires more health care personnel than an older population on average. The rest of the controls are similar in columns (4) and (6), except for the level of circulatory disease morbidity rate which appears to have a small significant impact.

The total productivity model in Table 5.2 does not show as robust effects as the relative productivity alternative in Table 5.1. The effect of the most important variable, $Prod$, varies highly in each estimation and has both theoretically (appendix B2) and empirically a smaller effect than ϕ on LHL. The phenomenon where the estimated effect changes to significant by adding control variables is also evidence of multicollinearity. A change in most other variables are thus affecting the change in $Prod$, which in itself has a weak explanatory power. However, the results in Table 5.2 show slightly more favourable estimations of supporting the existence of Baumol's cost disease than Atanda, Menclova and Reed (2018), but the results are less convincing than the model used in Table 5.1.

5.3. Discussion

Table 5.1 seems to capture the effect of Baumol's cost disease. Both of the expected effects in equation (3.27) and (3.28) are visible regardless of the model. An average change in ϕ would thus result in an estimated 600 – 850 jobs moving either into or from the health sector in Finland, ceteris paribus. The yearly change in relative productivity has varied each year between 2000 and 2016. Figure 5.1 illustrates the yearly change in ϕ and the variation between regions.

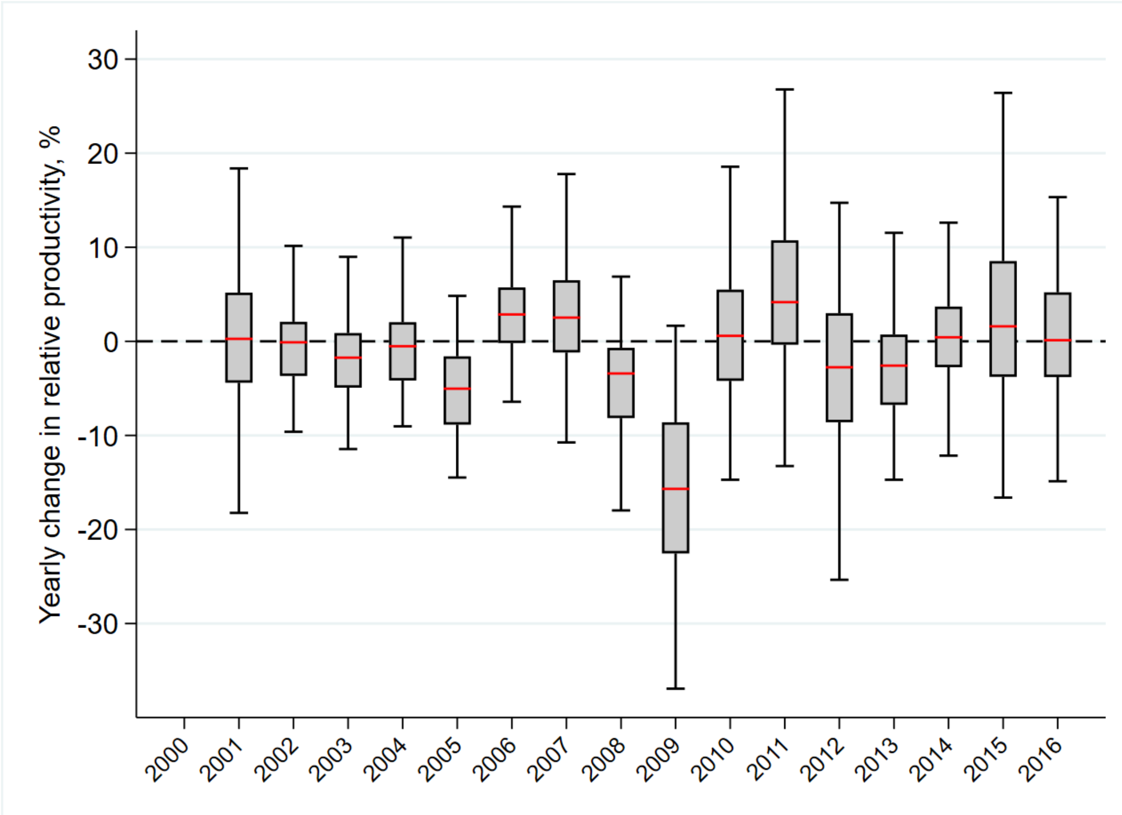

Figure 5.1 *Box graph of the yearly change in relative productivity in Finland*. **Note:** the boxes show variability in Finnish sub-regions.

The total number of professionals employed by the health sector is nearly 400 000, thus, the yearly impact of the relative productivity on the labour market is modest at best. Given the data sample used in this study, there is no clear direction in relative productivity variable in the future. The main factor of ϕ is the denominator (see eq. 3.17) since the economic downturns tend to affect more the production of the non-health sector. In the last two decades, the Finnish GDP growth has been mostly negative at two time intervals; 2008-2010 and 2012-2016. The medium-term growth rate in Finland is 1.5 % (Economic Policy Council, 2018). Due to the world-wide financial crisis in 2008 and the European debt crisis which followed soon after, the data sample of the Finnish economy (Statistics

Finland, 2019) might fail to capture the real level or actual direction of the productivity growth between sectors in the future. Nevertheless, if labour productivity grows with a different rate in non-health sectors than in healthcare in Finland a change in the labour structure is expected, which is the general idea of Baumol's third proposition (Baumol, 1967) of the cost disease. Additionally, assuming the total production will grow with its medium-term rate, the impact of the disparity in productivity would be economically noticeable in a decade.

Productivity indicators are often criticised for their inaccuracy. The relative productivity variable is an imprecise solution for measuring the real impact of the productivity gap. Using the same indicator of productivity in different sectors tends to misrepresent the actual effectiveness in each sector of the economy (Bernard & Jones, 1996; Van de Walle, 2008). Cowen (1996) argues in his paper that the main problem with studies of Baumol's cost disease is that growth in quality is often overlooked in the non-progressively growing sector. Quality improvements certainly increase the value of the produced goods, and sometimes the changes over time occur in an unmeasurable way. In healthcare, the measurement of productivity is often corrected for infant-mortality and life-expectancy (Afonso, Schuknecht & Tanzi, 2005) and there are many potential ways to improve the indicator of the actual level of performance. However, correcting the health sector productivity would further misrepresent the relative productivity variable, especially if the labour productivity in the rest of the economy would be left uncorrected. Even though the productivity values in equations (4.1) and (4.2) are somewhat flawed, the change in the labour market is identified by a change in the relative labour productivity between the studied sectors. Therefore, the method in this thesis measures the actual problem which arises with disparity in labour productivity (Baumol 1967), whether the productivity variable measures actual performance within sectors or not.

The results in Table 5.2 are derived from a more generalized model and the coefficients between the models show greater variations. A noteworthy difference between the theories used in the estimations in Table 5.1 and 5.2 is that the disparity in health sector productivity and rest of the economy is merely an assumption in the total productivity model. In an intuitive sense, productivity in the whole economy should not affect the health sector labour any more than other sectors if the productivity cannot be observed separately, even though it theoretically will (chapter 3.3.2). The difference in productivity is excluded from the measurements, which threatens the validity of the model. However, there seems to be some evidence to the cost disease argument with the total productivity nonetheless, even though the effect is smaller in magnitude than the relative productivity variable.

In Finland, health expenditures per capita have grown by 40 % between 2000 and 2016 and the real prices have grown by almost 20 % (OECD 2019; Statistics Finland, 2019). The study by Ministry of Finance (2013) suggests that health expenditures in Finland are affected by the difference in productivity, which is in line with the first proposition (*P1*) of Baumol's cost disease. The demand for health is kept constant with the national health insurance and regulated out-of-pocket payments (Keskimäki et al, 2019). The subsidisation restricts the demand for health to decline, which would be predicted by the second proposition (*P2*). Given the results in this chapter, the third proposition (*P3*) is visible as well. Thus, the empirical evidence in this study suggests that the healthcare sector in Finland would suffer Baumol's cost disease. In other words, the productivity in health sector affects healthcare costs, prices and labour, controlled for the average productivity in all other sectors of the economy. Despite the clear signs of the cost disease, the effect in the whole economy is quite small. The disparity in productivity explains a marginal part of the change in the labour market in Finland, whereas the demand for health has a sizeable explanatory power, much like the mainstream conclusions in the field of study. Nevertheless, the results suggest that Baumol's cost disease cannot be excluded as an explanation of how health sector has evolved in the past and is going to evolve in the future. The overall results are in line with Colombier (2012, 2017) and Bates and Santerre (2013, 2015). The effect of Baumol's cost disease is not quite as high as in Hartwig (2008) and Ministry of Finance (2013), and somewhat greater than in Atanda, Menclova and Reed (2018).

6. Conclusions

The aim of this thesis is to evaluate if the Finnish healthcare sector is affected by Baumol's cost disease. Baumol's cost disease proposes that a relatively non-productive sector will exhibit four consecutive market phenomena over time: 1) the cost per input unit will increase. 2) The demand will decline without third-party intervention. 3) The labour force will move from the productive sector toward the non-productive sector. 4) The costs of third-party intervention will increase and affect the overall economic growth. In countries with universal health coverage the third party is naturally the government. The repeatedly discussed problem, however, is to use reliable models and measurements to study the changes in market structures. This thesis uses a model from a recent study by Atanda, Menclova and Reed (2018), which measures the third proposition of the cost disease, controlled by the second proposition. The most important factor is the effects of the relative productivity variable, which is measured by dividing the labour productivity of the health sector by the productivity of the rest of the economy.

The results of the sub-regional-level fixed effect estimations suggest that a positive change in the relative productivity will increase the health sector share of the labour force in Finland. Additionally, a similar estimation derived from the relative productivity model suggest the economy-wide productivity will affect health production positively as well. The estimates are mostly statistically significant but trivial from the perspective of the entire labour market. Within Finland, an average change of the disparity in productivity variable affects under a thousand employees yearly.

This study contributes in two ways to the economic field of research: Firstly, Baumol's cost disease can be used as an explanation for growing health expenditures. The impact is small but visible. Secondly, despite the dispute of the theory, Baumol's (1967) framework seems to explain the evolution of different sectors and is relevant to this day.

7. Summary in Swedish – Svensk Sammanfattning

Finns det tecken av Baumols kostnadssjuka på den finska hälsovårdssektorn?

Hälsovårdskostnader har ökat drastiskt under de senaste 40 åren i de flesta industriländer. I Finland har totala hälsoutgifterna ökat från fyra till tio procent av BNP och utgifterna förväntas öka 2 % årligen under nästa årtionde. Inom 2000-talet har man gjort flera bidrag till hälsoekonomin om vad som orsakar kostnaderna i hälsovård. Den viktigaste determinanten som framställts till ökande kostnader är BNP per capita, vilket tyder på att ju högre inkomst desto mer hälsovård efterfrågas. Det återstående problemet är att inkomsttillväxten inte fullständigt förklarar fenomenet och att andra väsentliga variabler inte verkar påverka hälsoutgifter lika kraftigt. Varierande metoder har testats för att mäta arbetsproduktivitetens inverkan på hälsosektorn under senaste åren, med motstridande resultat. En av dessa metoder är att undersöka ifall hälsovårdssektorn drabbas av Baumols kostnadssjuka, vilket i teorin direkt skulle förklara hur produktivitetsskillnader ökar hälsoutgiftera i ett land.

I denna avhandling undersöks om det finns tecken av Baumols kostnadssjuka i finska hälsovårdssystemet, med den metod som konstruerades av Atanda, Menclova och Reed (2018). Kostnadssjukan sker då två sektorer i ekonomin har en ojämn produktivitetstillväxt, men tillväxten i lönerna är lika. Enligt Baumols (1967) formalisering visar sektorn med långsam tillväxt fyra olika förändringar på marknaden: 1) Kostnaderna ökar relativt till resten av ekonomin, 2) efterfrågan sjunker eftersom priserna måste stiga i förhållande till kostnaderna i produktionen, 3) ifall efterfrågan hålls konstant med att en tredje part subventionerar produktionen kommer en ökande andel av arbetskraften förflytta sig till den sektorn där tillväxten är långsam, 4) tillväxttakten i hela ekonomin kommer att sjunka, eftersom subventionernas mängd växer kontinuerligt. Dessa fyra effekter kallas ofta för Baumols propositioner. Första propositionen anses ofta vara viktigast, men relativt högre kostnader till resten av ekonomin kan förekomma på flera olika sätt. Därmed är det viktigt att undersöka även de andra propositionerna, så att man med högre sannolikhet kan identifiera en bransch som en Baumolsektor. I Atanda, Menclova och Reeds (2018) ekonometriskt tillämpbara modell är andelen som jobbar inom hälsovård en funktion av en variabel som beskriver efterfrågan av icke-hälsoprodukter och en variabel som mäter skillnaden i produktivitet mellan hälsovård och resten av ekonomin. Efterfrågan av icke-hälsoprodukter förväntas ha en negativ effekt och produktivitetsskillnaden en positiv effekt på andelen hälsorelaterade jobb. Intuitivt kunde hälsovården klassas som en sektor som drabbas av kostnadssjukan, eftersom tjänsteproduktion har allmänt långsammare produktivitetstillväxt, kostnaderna har okat markant och Finland erbjuder allmän hälsovård, alltså en tredje part finansierar produktionen.

Metoden som används i avhandlingen är en ekonometrisk regressionsanalys, där huvudsakliga målet är att analysera hur skillnaden i produktivitetstillväxt påverkar andelen av arbetskraft inom hälsovård. Produktivitetsskillnaden är i detta fall uträknad med att dela den totala arbetsproduktiviteten i övriga ekonomin med arbetsproduktiviteten i hälsovård. Ifall en ökning i produktivitetsskillnaden orsakar en ökning i hälsoarbetskraften, skulle detta tyda på att hälsosektorn drabbas av Baumols kostnadssjuka. Därutöver testas problemet även en annan regression som mäter ifall en ökning i totala arbetsproduktiviteten i hela regionen ökar på arbetskraften inom hälsovård. Datamaterialet är från finska ekonomiska regioner mellan åren 2000 och 2016. I studien mäts effekten med en OLS modell med fixa effekter på regionnivå. Modellen kontrolleras även för tidsfixa effekter samt regionsspecifika tidstrender, vilket kan minska på eventuell heterogenitet som kan förekomma i datamaterialet. Metoden lämpar sig för finsk data, eftersom det finns stora skillnader i hälsovariabler mellan regioner inom landet. I Finland ansvar kommunerna för primärvården och sjukhusdistrikten ansvarar för specialvård. Det saknas en tydlig beslutsfattningsstruktur mellan dessa nivåer och regionerna tillämpar tillhandahållandet av sjukvård på olika sätt. Globalt sett rangordnas det finska hälsosystemet högt, men nationella skillnaderna är stora även om alla kommuner är verksamma inom samma fundamentala hälsosystem.

Enligt regressionsresultaten ökar produktivitetsskillnaden på andelen arbetskraft som jobbar med hälsorelaterade tjänster då efterfrågan av hälsoprodukter hålls konstant, vilket tyder på att två av Baumols (1967) ursprungliga propositioner är synliga. Även totala produktiviteten ökar på hälsoarbetskraften, vilket i teorin förutspås byggd av Atanda, Menclova och Reed (2018). Resultaten i studien av Finansministeriet (Ministry of Finance, 2013) av finska hälsovårdssystemet tyder också på att skillnaden i produktivitetstillväxten orsakar högre kostnader inom hälsovård. Finansministeriets resultat och utfallen i denna studie indikerar kraftigt att hälsosektorn i Finland påverkas av Baumols kostnadssjuka. Relativa kostnader för hälsovård stiger och andelen av arbetsuppgifterna förflyttas mot hälsobranschen, på grund av att produktiviteten har en kraftigare tillväxt i resterande ekonomin. Däremot är effekten inte så stor. En medelmåttlig årlig förändring i produktivitetsskillnadsvariabeln skulle fenomenet påverka 600-850 individer varje år. Hela finska hälsovårdssystemet sysselsätter omkring 400 000 människor (Statistics Finland, 2019). Det finns inte heller en tydlig trend hur produktiviteten i hälsosektorn och i resten av ekonomin kommer att förändras i framtiden. Hälsosektorn påverkas allmänt mindre av konjunktursvängningar i ekonomin, vilket förvränger variabelvärden en del. Allmänt kan dock Baumols kostnadssjuka inte uteslutas som en förklaring till både ökade kostnader för hälsovård och förändring på hälsoarbetskraften.

References

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. (2017). When should you adjust standard errors for clustering? (No. w24003). National Bureau of Economic Research.

Afonso, A., Schuknecht, L., & Tanzi, V. (2005). Public sector efficiency: an international comparison. Public choice, 123(3-4), 321-347.

Alesina, A., Glaeser, E., & Sacerdote, B. (2005). Work and leisure in the United States and Europe: why so different?. *NBER macroeconomics annual*, *20*, 1-64.

Allison, P. D. (2009). *Fixed effects regression models*. Sage Publications, Thousand Oaks, CA.

Atanda, A., Menclova, A. K., & Reed, W. R. (2018). Is health care infected by Baumol's cost disease? Test of a new model. *Health economics*, *27*(5), 832-849.

Atanda, A., & Reed, W. R. (2019). Not Evidence for Baumol's Cost Disease. University of Canterbury, Department of Economics and Finance.

Atella, V., Belotti, F., Bojke, C., Castelli, A., Grašič, K., Kopinska, J., Montari, P. & treet, A. (2019). How health policy shapes healthcare sector productivity? Evidence from Italy and UK. *Health Policy*, *123*(1), 27-36.

Baltagi, B. H., & Moscone, F. (2010). Health care expenditure and income in the OECD reconsidered: Evidence from panel data. *Economic modelling*, *27*(4), 804-811.

Barros, P. P. (1998). The black box of health care expenditure growth determinants. *Health economics*, *7*(6), 533-544.

Bates, L. J., & Santerre, R. E. (2013). Does the US health care sector suffer from Baumol's cost disease? Evidence from the 50 states. *Journal of Health Economics*, *32*(2), 386-391.

Bates, L. J., & Santerre, R. E. (2015). Does Baumol's Cost Disease Account for Nonfederal Public-Sector Cost Growth in the United States? A New Test of an Old Idea. *Social Science Quarterly*, *96*(1), 251-260.

Baum, C. F. (2006). *An introduction to modern econometrics usingStata*. Stata press.

Baumol, W. J. (1967). Macroeconomics of unbalanced growth: the anatomy of urban crisis. *The American economic review*, *57*(3), 415-426.

Baumol, W. J. (1996). Children of performing arts, the economic dilemma: The climbing costs of health care and education. *Journal of Cultural Economics*, 183-206.

Baumol, W. J. (2012). *The cost disease: why computers get cheaper and health care doesnt*. New Haven, CT: Yale University Press.

Benavides, A. (2018). Per Capita GDP, Health Expenditures, and the Income Elasticity of Demand for Health Care in Developing Nations. *Journal of Applied Business and Economics*, *20*(2).

Bernard, A. B., & Jones, C. I. (1996). Comparing apples to oranges: productivity convergence and measurement across industries and countries. *The American Economic Review*, 1216-1238.

Bowerman, B. L., OConnell, R. T., & Koehler, A. B. (2005). Forecasting, time series, and regression: an applied approach. Belmont, CA: Thomson Brooks/Cole.

Colombier, C. (2012). *Drivers of health care expenditure: Does Baumol's cost disease loom large?* (No. 12-5). University of Cologne, FiFo Institute for Public Economics.

Colombier, C. (2017). Drivers of Health-Care Expenditure: What Role Does Baumol's Cost Disease Play?. *Social Science Quarterly*, *98*(5), 1603-1621.

Cowen, T. (1996). Why I do not believe in the cost-disease: Comment on Baumol. *Journal of cultural Economics*, 207-214.

Di Matteo, L. (2003). The income elasticity of health care spending. *The European Journal of health economics*, *4*(1), 20-29.

Dieleman, J. L., Sadat, N., Chang, A. Y., Fullman, N., Abbafati, C., Acharya, P. & Alkerwi, A. A. (2018). Trends in future health financing and coverage: future health spending and universal health coverage in 188 countries, 2016–40. *The Lancet*, *391*(10132), 1783-1798.

Dynarski, S., Jacob, B., & Kreisman, D. (2018). How important are fixed effects and time trends in estimating returns to schooling? Evidence from a replication of Jacobson, Lalonde, and Sullivan, 2005. *Journal of Applied Econometrics*, *33*(7), 1098-1108.

Economic Policy Council (2018). *Economic Policy Council Report 2017.* Available 2020-03.04, from https://www.talouspolitiikanarviointineuvosto.fi/wordpress/wp-content/uploads/2018/01/Report2017.pdf

European Commission (2020). Eurostat database. Available 2020-03-18, from https://ec.europa.eu/eurostat/data/database

Feldmann, H. (2013). Technological unemployment in industrial countries. *Journal of Evolutionary Economics*, *23*(5), 1099-1126.

Finnish institute for health and welfare (2020). Sotkanet.fi Statistics and Indicator Bank. Available 2019-11-04, from https://sotkanet.fi/sotkanet/en/index

Gerdtham, U. G., & Jönsson, B. (2000). International comparisons of health expenditure: theory, data and econometric analysis. In *Handbook of health economics* (Vol. 1, pp. 11-53). Elsevier.

Hartwig, J. (2008). What drives health care expenditure?—Baumol's model of 'unbalanced growth' revisited. *Journal of Health Economics*, *27*(3), 603-623.

Hitiris, T. (1997). Health care expenditure and integration in the countries of the European Union. *Applied Economics*, *29*(1), 1-6.

Iversen, T., & Lurås, H. (2002). Waiting time as a competitive device: an example from general medical practice. *International Journal of Health Care Finance and Economics*, *2*(3), 189-204.

Keskimäki, I., Tynkkynen L. K., Reissell E., Koivusalo M., Syrjä V., Vuorenkoski L., Rechel B. & Karanikolos M. (2019). Finland: Health system review. Health Systems in Transition, 2019; 21(2): 1 – 166

Khraief, N., Shahbaz, M., Heshmati, A., & Azam, M. (2018). Are unemployment rates in OECD countries stationary? Evidence from univariate and panel unit root tests. *The North American Journal of Economics and Finance*.

Local and Regional Government Finland (2020). *Sosiaali- ja terveysasiat*. Available 2020-01-09, from https://www.kuntaliitto.fi/sosiaali-ja-terveysasiat

Lucas, R. E., & Rapping, L. A. (1969). Price expectations and the Phillips curve. *The American Economic Review*, *59*(3), 342-350.

Medeiros, J., & Schwierz, C. (2013). *Estimating the drivers and projecting long-term public health expenditure in the European Union: Baumol's" cost disease"* *revisited* (No. 507). Directorate General Economic and Financial Affairs (DG ECFIN), European Commission.

Ministry of Finance (2013). Terveysmenojen kasvu. Available 2020-01-19, from https://vm.fi/keskustelualoitteet

Ministry of Social Affairs and Health (2016a). Sosiaali- ja terveydenhuollon kustannusvaikuttava ja tehokas ohjaus. Available 2020-03-14, from https://stm.fi/julkaisu?pubid=URN:ISBN:978-952-00-3696-6

Ministry of Social Affairs and Health (2016b). International expert panel pre- review of health and social care reform in Finland. Available 202-03-25, from https://stm.fi/julkaisu?pubid=URN:ISBN:978-952-00-3848-9

OECD (2019). OECD.Stat database. Available 2019-12-10, from https://stats.oecd.org/

Sinervo, T, Tynkkynen, L-K, Vehko, T (2016). Mitä kuuluu perusterveydenhuolto? Valinnanvapaus ja integraatio palveluiden kehittämisen polttopisteessä. Helsinki: Finnish Institute for Health and Welfare. Available 2020-02-18, from https://www.julkari.fi/handle/10024/131276

Statistics Finland (2019). Statistics Finland's PxWeb databases. Available 2019-11-07, from http://pxnet2.stat.fi/PXWeb/pxweb/en/StatFin/

Teperi, J., Porter, M. E., Vuorenkoski, L., Baron, J. F. (2009). The Finnish Health Care System. A Value-Based Perspective. Helsinki: Sitra. Available 2020-03-25, from https://www.sitra.fi/en/publications/finnish-health-care-system-0/

Triplett, J. E. (Ed.). (1999). *Measuring the prices of medical treatments*. Brookings Institution Press.

Van de Walle, S. (2008). Comparing the performance of national public sectors: conceptual problems. *International Journal of Productivity and Performance Management.*

Wooldridge, J. (2007). Econometric analysis of cross section and panel data. London: MIT Pr.

Wooldridge, J. M. (2012). *Introductory Econometrics: A Modern Approach.* Ohio, USA: South-Western.

Appendix

Appendix A. Variable definitions

Table A1 Variable definitions

Variable	Measured	Definition	Source
LHL	$\frac{L_H}{L}$	Healthcare sector share of total labour.	Statistics Finland (2019)
ϕ	$\frac{\left(Y_{NH}/L_{NH}\right)}{\left(Y_H/L_H\right)}$	Relative productivity. Measures the disparity in productivity between health sector and rest of the economy (sum of all other sectors in Finland).	Statistics Finland (2019)
$Prod$	$\frac{GDP}{total\ hours\ worked}$	Productivity.	Statistics Finland (2019); Finnish institute for health and welfare (2020)
k	$\frac{Y_{NH}}{Y}$	The output share of non-health sector. Measures the demand for non-health goods.	Statistics Finland (2019)
Private health production	$\frac{Y_{H,private}}{Y_{H,total}}$	Private health output of the total health production. Measures the size of the private health sector.	Statistics Finland (2019)
Costs of specialised care/capita	$\frac{C_{Specialised\ care}}{population}$	Cost of the specialised care. Indicator of the size specialised care sector, within the health sector.	Finnish institute for health and welfare (2020)
Sickness allowances	$\frac{Recipients}{population} * 1000$	The number individuals that have received at least one day of sickness allowances.	Finnish institute for health and welfare (2020)
Cancer index	Index value	Number of new cases of cancer. 100 = total average in Finland 2016. Age-standardised.	Finnish institute for health and welfare (2020)
Circulatory disease index	Index value	Number of new diagnoses of first- and second-degree coronary diseases. 100 = total average in Finland 2016. Age-standardised.	Finnish institute for health and welfare (2020)

GDP/cap growth rate	$d.\ln\left(\frac{GDP}{population}\right)$	Growth of per capita income.	Finnish institute for health and welfare (2020)
Age > 64	$\frac{percent}{100}$	The share of the resident population that are more than 64 years old.	Finnish institute for health and welfare (2020)
Age < 15	$\frac{percent}{100}$	The share of the resident population that are less than 15 years old.	Finnish institute for health and welfare (2020)
Unemployment	$\frac{unemployed}{labour\ force}$	Rate of unemployment.	Finnish institute for health and welfare (2020)
Relative public indebtedness	$\frac{liabilities}{current\ income}$	Percent of the public sector's current income that would be required to offset the liabilities. Indicates the need of public saving.	Finnish institute for health and welfare (2020)

Appendix B. Further calculations

B1: Calculations of the fourth proposition (*P4*) of the cost disease

The fourth proposition shows how the economy-wide growth will eventually slow down (Chapter 3.1). The calculations by Baumol (1967) for *P4* is shown in the following way:

We can measure the index of total output as an index of both sector's weighted outputs as

$$I = B_{NPS}Y_{NPS}(t) + B_{PS}Y_{PS}(t) = B_{NPS}aL_{NPS}(t) + B_{PS}bL_{PS}(t)e^{rt} \quad , \qquad \text{(B1.1)}$$

where B_{NPS} and B_{PS} are the sectors' weights. By substituting equations (3.12) and (3.13) to (B1.1), we get

$$I = B_{NPS}a\frac{\gamma L(t)e^{rt}}{1+\gamma e^{rt}} + B_{PS}b\frac{L(t)}{1+\gamma e^{rt}}e^{rt} \quad . \qquad \text{(B1.2)}$$

If we assume that the total labour supply is a constant over time, we can rewrite equation (B1.2) as

$$I = \frac{Re^{rt}}{1+\gamma e^{rt}} \quad , R = L(t)(a\gamma B_{NPS} + bB_{PS}) \quad . \qquad \text{(B1.3)}$$

Additionally, a change in I with respect to t can be written as

$$\frac{\partial I}{\partial t} = \frac{rRe^{rt}}{(1+\gamma e^{rt})^2} \qquad \text{(B1.4)}$$

and a percentage change in I is shown as

$$\frac{\partial I/\partial t}{I} = \frac{\frac{rRe^{rt}}{(1+\gamma e^{rt})^2}}{\frac{Re^{rt}}{1+\gamma e^{rt}}} = \frac{r}{1+\gamma e^{rt}} \quad . \qquad \text{(B1.5)}$$

Finally, $\lim_{t\to\infty}\left(\frac{\partial I/\partial t}{I}\right) = 0$, meaning that the growth of the index function will slow down.

B2: Calculation of the total productivity method effect

The differences in the coefficients in Table 5.1 and Table 5.2 can theoretically be explained. The difference is the effects $\frac{\partial\left(\frac{L_H}{L}\right)}{\partial \phi}$ and $\frac{\partial\left(\frac{L_H}{L}\right)}{\partial Prod}$ in the theory (Atanda, Menclova & Reed, 2018). The application in the econometric model (eq. 4.1) in this study uses both variables. The calculations can be shown the following way:

In equation (3.31), the derivative of $\frac{L_H}{L}$ with respect to $Prod$ is shown as

$$\frac{\partial\left(\frac{L_H}{L}\right)}{\partial Prod} = \frac{\partial\left(\frac{L_H}{L}\right)}{\partial \phi} \cdot \frac{\partial(\phi)}{\partial Prod} > 0 \ . \tag{B2.1}$$

To evaluate when $\frac{\partial\left(\frac{L_H}{L}\right)}{\partial \phi} > \frac{\partial\left(\frac{L_H}{L}\right)}{\partial Prod}$, we only need to show that $\frac{\partial(\phi)}{\partial Prod} < 1$. Given equation (3.35), we see that

$$\frac{\partial \phi}{\partial Prod} = \frac{k}{\left(1 - Prod \cdot (1-k)\right)^2} \tag{B2.2}$$

and the condition when equation (B2.2) is less than one is calculated by

$$\frac{k}{\left(1 - Prod \cdot (1-k)\right)^2} < 1 \Leftrightarrow k < \left(1 - Prod \cdot (1-k)\right)^2$$

$$\Leftrightarrow \sqrt{k} < 1 - Prod \cdot (1-k) \ or -\sqrt{k} > 1 - Prod \cdot (1-k)$$

$$\Leftrightarrow Prod < \frac{1-\sqrt{k}}{1-k} \ or \ Prod > \frac{1+\sqrt{k}}{1-k} \ . \tag{B2.3}$$

These are the restricting conditions where $\frac{\partial\left(\frac{L_H}{L}\right)}{\partial \phi} > \frac{\partial\left(\frac{L_H}{L}\right)}{\partial Prod}$. By calculating the difference between these restrictions and the actual data (Statistics Finland, 2019) used in the econometric estimations (Table 5.1 & Table 5.2), we can see if the theoretically expected coefficient should be lower as well. Therefore, suppose that the restriction variable $Prod_{res}$ is calculated for every observation in the data set as

$$Prod_{res|it} = \frac{1+\sqrt{k_{it}}}{1-k_{it}} \quad ,$$

where i and t are the notations for each region and year respectively (the upper limit for the inequality (B2.3) is used, since there are no $Prod$ values below one). Now we can test the difference and ratio between the actual and the restricting value to identify if the condition in (B2.3) is true. The tests are:

$$Test_1: \quad Prod_{it} - Prod_{res|it} > 0$$

$$Test_2: \quad \frac{Prod_{it}}{Prod_{res|it}} > 1$$

The results of the tests:

Test	Obs.	Mean	Std.dev	99 % confidence interval	% observations failing the test
$Test_1$	1005	8.37	0.60	[6.81 , 9.93]	25.07 %
$Test_2$	1005	1.47	0.02	[1.42 , 1.52]	25.07 %

Both statistic tests are significantly true, which means that the expected change in $\frac{\partial\left(\frac{L_H}{L}\right)}{\partial\phi}$ is greater than in $\frac{\partial\left(\frac{L_H}{L}\right)}{\partial Prod}$.

Appendix C. Alternative measurements and robustness tests

Table A2 Hospital-district-level fixed effects estimations with relative productivity. Dependant variable: Healthcare sector share of the labour force (LHL).

VARIABLES	(FE)		(2WFE)		(FErTT)	
ln (ϕ)	0.0862**	0.0636**	0.0684**	0.0722**	0.0914**	0.0760**
	(0.0115)	(0.0175)	(0.0160)	(0.0151)	(0.0138)	(0.0192)
k	– 1.588**	– 1.090**	– 1.091**	– 1.183**	– 1.525**	– 1.379**
	(0.0939)	(0.1890)	(0.247)	(0.157)	(0.111)	(0.169)
% private health production		– 0.023 (0.016)		– 0.035* (0.017)		– 0.043* (0.019)
ln (costs of specialised care/capita)		0.014* (0.007)		– 0.007 (0.005)		0.001 (0.005)
ln (sickness allowances)		– 0.026**		– 0.013		– 0.007
		(0.006)		(0.008)		(0.006)
ln (cancer index)		– 0.000		– 0.004		0.003
		(0.006)		(0.006)		(0.004)
ln (circulatory disease index)		0.001 (0.003)		0.009* (0.003)		0.002 (0.003)
GDP/cap growth rate		– 0.001		– 0.005		– 0.002
		(0.004)		(0.005)		(0.003)
Age > 64		0.143*		0.341**		0.064
		(0.063)		(0.115)		(0.069)
Age < 15		0.435*		0.727**		0.352*
		(0.154)		(0.206)		(0.144)
% unemployment		– 0.043*		0.014		– 0.044*
		(0.017)		(0.034)		(0.020)
Relative public indebtedness		– 0.007 (0.012)		– 0.012 (0.013)		– 0.011 (0.010)
Constant	1.552**	1.043**	1.098**	1.090**	1.053**	– 0.47
	(0.078)	(0.196)	(0.218)	(0.169)	(0.411)	(1.135)
Observations	357	300	357	300	357	300
Adj. R^2	0.885	0.915	0.908	0.926	0.985	0.985
Number of regions	21	20	21	20	21	20
Time fixed effects	No	No	Yes	Yes	Yes	Yes
Indiv. region time trends	No	No	No	No	Yes	Yes
F-test (time effects)	-	-	p =.000	p =.000	p =.438	p =.000
F-test (time trends)	-	-	-	-	p =.421	p =.000

Note: Cluster-robust standard errors in parentheses. F-test: tests if the specified variables are jointly significant in the model, H0: no significant effects. ** p <0.01, * p <0.05.

Table A3 Robustness tests of the two-way fixed effects estimations with relative productivity. Dependant variable: Healthcare sector share of the labour force (LHL).

VARIABLES	(FE)		(2WFE)		(FErTT)	
ln (ϕ)	0.0715**	0.0701**	0.0720**	0.0697**	0.0705**	0.0708**
	(0.0065)	(0.0058)	(0.0062)	(0.0054)	(0.0053)	(0.0047)
k	– 1.351**	– 1.204**	– 1.241**	– 1.170**	– 1.184**	– 1.165**
	(0.0901)	(0.138)	(0.1220)	(0.118)	(0.116)	(0.119)
% private health production	– 0.007 (0.008)		– 0.008 (0.008)			
ln (costs of specialised care/capita)	– 0.007 (0.005)		– 0.006 (0.005)			
ln (sickness allowances)		– 0.018*	– 0.012			
		(0.008)	(0.007)			
ln (cancer index)		– 0.001	– 0.000			
		(0.004)	(0.004)			
ln (circulatory disease index)		0.004 (0.004)	0.003 (0.003)			
GDP/cap growth rate				– 0.005	– 0.004	– 0.004
				(0.003)	(0.003)	(0.003)
Age > 64				0.067		0.228**
				(0.054)		(0.055)
Age < 15					0.116	0.297*
					(0.090)	(0.120)
% unemployment						0.108*
						(0.050)
Relative public indebtedness						0.001 (0.006)
Constant	1.383**	1.257**	1.414**	1.158**	1.161**	1.055**
	(0.084)	(0.131)	(0.097)	(0.111)	(0.106)	(0.115)
Observations	1 139	1 035	1 005	1 120	1 120	1 120
Adj. R^2	0.780	0.741	0.781	0.734	0.735	0.752
Number of regions	67	69	67	70	70	70
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Indiv. region time trends	No	No	No	No	No	No
F-test (time effects)	p =.000	p =.000	p =.000	p =.000	p =.000	p =.000
F-test (time trends)	-	-	-	-	-	-

Note: Cluster-robust standard errors in parentheses. F-test: tests if the specified variables are jointly significant in the model, H0: no significant effects. ** p <0.01, * p <0.05.

Table A4 Fixed effects estimations with relative productivity, each sector's productivity measured as output per hours worked. Dependant variable: Healthcare sector share of the labour force (LHL).

VARIABLES	(FE)		(2WFE)		(FErTT)	
$\ln(\phi^{hours})$	0.0633**	0.0563**	0.0575**	0.0571**	0.0650**	0.0598**
	(0.0041)	(0.0051)	(0.0052)	(0.0051)	(0.0051)	(0.0053)
k	– 1.154**	– 1.121**	– 0.9940**	– 1.135**	– 1.025**	– 1.177**
	(0.0814)	(0.0928)	(0.1220)	(0.098)	(0.152)	(0.110)
% private health production		– 0.021* (0.008)		– 0.023* (0.009)		– 0.033** (0.008)
ln (costs of specialised care/capita)		– 0.007 (0.004)		– 0.004 (0.005)		0.004 (0.003)
ln (sickness allowances)		– 0.002		– 0.002		0.006
		(0.007)		(0.008)		(0.006)
ln (cancer index)		– 0.001		– 0.001		– 0.003
		(0.004)		(0.004)		(0.004)
ln (circulatory disease index)		0.003 (0.003)		0.002 (0.004)		0.005 (0.003)
GDP/cap growth rate		– 0.001		– 0.001		– 0.000
		(0.002)		(0.002)		(0.001)
Age > 64		0.201**		0.301**		0.227**
		(0.035)		(0.065)		(0.073)
Age < 15		0.190		0.293*		0.142
		(0.106)		(0.117)		(0.131)
% unemployment		0.024		0.082*		0.022
		(0.029)		(0.046)		(0.022)
Relative public indebtedness		0.003 (0.006)		0.003 (0.006)		– 0.006 (0.005)
Constant	1.172**	1.121**	1.022**	1.072**	1.050**	1.073**
	(0.075)	(0.104)	(0.112)	(0.114)	(0.140)	(0.103)
Observations	1 190	1 005	1 190	1 005	1 190	1 005
Adj. R^2	0.643	0.725	0.678	0.729	0.947	0.964
Number of regions	70	67	70	67	70	67
Time fixed effects	No	No	Yes	Yes	Yes	Yes
Indiv. region time trends	No	No	No	No	Yes	Yes
F-test (time effects)	-	-	p =.000	p =.006	p =.746	p =.000
F-test (time trends)	-	-	-	-	p =.000	p =.000

Note: Cluster-robust standard errors in parentheses. F-test: tests if the specified variables are jointly significant in the model, H0: no significant effects. ** p <0.01, * p <0.05.