Macroeconomic Consequences of Educational Composition: Sectoral Allocation, Inequality, and Automation

Dissertation von Johanna Marion Rude

München 2024

Macroeconomic Consequences of Educational Composition: Sectoral Allocation, Inequality, and Automation

Inauguraldissertation

zur Erlangung des akademischen Grades

Doctor oeconomiae publicae (Dr. oec. publ)

an der Volkswirtschaftlichen Fakultät

der Ludwig-Maximilians-Universität München

vorgelegt von

Johanna Marion Rude

2024

Referent: Professor Dr. Uwe Sunde

Korreferent: Professor Dr. Christopher Busch

Promotionsabschlussberatung: 31. Januar 2024

Contents

List of Tables

List of Figures

Acknowledgments

First and foremost, I want to thank my supervisor Uwe Sunde. He inspired me to pursue my Dissertation and focus on the topics that most interested me. His always insightful and constructive comments provided guidance and encouragement when I most needed it. This endeavor would not have been possible without his support. I am also grateful to my second supervisor Christopher Busch, from many thoughtful conversations and discussions with whom I benefited immensely. Additionally, I want to thank Carsten Eckel, who not only completed my dissertation committee but contributed to my decision to pursue a career in economics at our first encounter over 12 years ago.

I am greatly indebted to the *Studienstiftung des Deutschen Volkes*, which provided financial, moral, and ideational support throughout the course of my studies. I am also grateful for a Scholarship from the Egon Sohmen Graduate Center.

Furthermore, I want to thank the Department of Economics, especially Sabine Njoku and Caroline Benkert for their support in all matters administrative.

I want to express my appreciation for the academic community at LMU and my colleagues at the Chair for Population Economics in particular. Their constant support has been a source of inspiration and motivation and made the experience of pursuing a Dissertation a rewarding one. Special thanks go to my very reliable co-author Lukas Weber, who was equally realistic and optimistic about our joint work. It was a joy working with you.

Words cannot express my gratitude to my family. To my parents, who always believed in me. And of course to my husband and children, bringing unconditional love, joy, and purpose into my life, putting into perspective all professional problems and issues.

Preface

In the realm of macroeconomics, the pivotal role of education in influencing economic outcomes has long been recognized. However, viewing education as a homogeneous factor fails to capture the influence that different education levels and skills can have on macroeconomic aggregates. Understanding the intricate relationship of the household sector's educational attainment with macroeconomic aggregates is of great importance. The education level of households shapes, among other things, their consumption patterns, savings behavior, labor supply, and the overall income distribution. Taking into account heterogeneous education levels when analyzing macroeconomic aggregates is of paramount importance as it offers a more comprehensive and nuanced understanding of the complex relationship between the household sector and macroeconomic outcomes. It provides valuable insights into important mechanisms through which the household sector influences broader economic indicators and the overall performance of an economy.

One critical aspect of the household sector's influence lies in its consumption expenditure, which constitutes a significant component of the gross domestic product (GDP). This thesis explores the determinants and drivers of aggregate demand, emphasizing the role of consumer spending in shaping the overall demand for goods and services, while taking into account the educational composition. The spending decisions made by households have direct implications for production levels, wages, prices, aggregate demand, and economic growth. By investigating the factors influencing consumption patterns, such as preferences and income levels, this research aims to provide insights into the mechanisms through which changes in aggregate demand can influence economic performance and employment dynamics.

Furthermore, the labor supply by households has far-reaching effects on employment, wages, income distribution, capital accumulation, and productivity. Understanding the factors that influence the size and the skill level of the labor force, such as education levels and demographic change, is vital for comprehending overall macroeconomic performance. Analyzing the household sector's impact on the labor market provides insights into the formation of key economic indicators and the potential policy implications for promoting economic growth.

This thesis consists of three chapters, which take into account the impact of factors such as household education, preferences, income distribution, technological advancements, and demographic changes on macroeconomic outcomes. While each chapter is self-contained and can be read independently, they all contribute to the literature by highlighting the importance of the household sector in influencing macroeconomic aggregates. First of all, this thesis highlights the role of consumer preferences in allocating production factors between industries, which depends crucially on different education levels in the labor force. It also demonstrates how consumer preferences impact skill-dependent income inequality by shaping aggregate demand. And finally, this research shows that accounting for the level of education and the effect it has on the labor force is crucial in determining the scope for automating labor. In all three chapters, the role of the household sector and its heterogeneous education level is analyzed both empirically and through the lens of theoretical frameworks, trying to illuminate the mechanisms through which the household sector shapes macroeconomic aggregates.

The analysis in Chapter 1, which is joint work with Lukas Weber, starts with the observation, that expenditures on healthcare and employment in the healthcare sector have been steadily increasing across OECD countries for many years. This shift of expenditure and employment towards a consistently found to be less productive sector has often been associated with the idea of Baumol's (1967) cost disease. It describes a situation in which production factors are allocated towards a less productive sector, despite this reallocation being economically disadvantageous. The first chapter, titled "Is Baumol's Cost Disease Really a Disease?", investigates if diagnosing the healthcare sector with suffering from a cost disease is an apt description of the observed reallocation.

The novel feature of the paper is to introduce a microeconomic foundation to the theoretical analysis of the healthcare sector. We show analytically in a model that the demand side is very important in determining equilibrium quantities and prices. Even if there is unequal technological progress in the two sectors, the unchanged demand of households dictates that the output level of the two sectors remains constant. This leads to the prima facie unintuitive result of factor allocation towards the less productive sector, in this case, healthcare. We show that this is the case under innocuous assumptions if goods are complements. We supplement the new theoretical results by testing implications from the model empirically. Specifically, we use household-level data to estimate the elasticity of substitution between healthcare consumption and all other consumption. We find robust evidence for the complementarity of healthcare consumption and all other consumption. This is the only condition necessary to rationalize an allocation of production factors towards the less productive sector. Our model predicts that the reallocation of production factors is driven by a reallocation of unskilled labor. We find evidence of exactly that in both German and US data. Both the relative share of unskilled labor and the skill premium paid increased more in the healthcare sector than in the rest of the economy. We conclude that the reallocation of resources towards the healthcare sector is driven by the demand side of the economy and is not inefficient from an economic point of view. Therefore, describing the healthcare sector as suffering from a cost disease is unfitting, as such a description does not appreciate the role of the demand side.

Chapter 2 analyzes the relationship between income inequality and aggregate demand. Economists have observed an increase in the skill premium in particular and income inequality in general at least over the last five decades. Economic analysis often makes the simplifying assumption that the demand side of the economy can be described by the actions of one representative agent. If that is the case, income inequality and changes therein are inconsequential to macroeconomic aggregates. In the second chapter "Income Inequality and Aggregate Demand" this is investigated. Starting with an empirical analysis, it is shown that increases in income inequality are associated with decreased aggregate consumption in both US state-level data and German data. From the analysis, an interesting pattern of the association between income inequality and consumption expenditure differing systematically across consumption categories arises. Both findings are incompatible with the hypothesis of one representative agent.

In a theoretical analysis, the effect of an exogenous skill-biased technological change on equilibrium prices and expenditure shares is derived for the case of homothetic CES preferences and the case of non-homothetic CES preferences. In both cases, equilibrium prices and expenditure shares are affected via a supply-side channel. In the case of non-homothetic CES preferences, equilibrium prices and expenditure shares are also affected via a demand-side channel, due to changes in income inequality. The direction and size of that demand-side channel depend on preference parameters and remain an empirical question. The comparison of model predictions under homothetic and non-homothetic preferences results in estimation equations that allow testing for non-homotheticity in consumption data. Empirical results indicate that preferences are indeed non-homothetic. Furthermore, the non-homothetic CES preferences are well suited to explain the distinct pattern observed between consumption categories and income inequality. These findings are derived using both US and German data. In addition, using US data, a quantification of the novel demand-side channel is done. The results suggest, that the demand-side channel ameliorates exogenous changes in income inequality and is non-trivial in size.

A prominent topic in economic analysis that is connected to the concept of skill-biased technological change and income inequality is automation. In Chapter 3, titled "Demographic Change, Automation and the Role of Education", demographic change and how it affects labor supply and automation is analyzed. The new feature in the analysis is accounting for the role of education in the relationship between demographic change and automation. It is illustrated in a theoretical model featuring skilled and unskilled labor and automation capital and traditional capital. In line with the economic literature and empirical findings, it is assumed in the model that automation capital is a close substitute for unskilled labor and a complement to skilled labor. If labor supply by households decreases, for example, due to demographic change, the model states that the optimal level of automation capital increases. However, this relationship depends crucially on the level of education in the workforce.

Motivated by this novel prediction derived from the model, a new data set allowing for testing of the prediction is constructed. Patent data are combined with an automation classification to arrive at a novel measure of automation. In a series of analyses, evidence for the theoretical prediction is found. While there is a negative relationship between automation capital and population growth, the results corroborate the theoretical prediction that it is crucial to account for the role of education in that relationship. Doing so yields highly significant results which suggest that population growth is negatively correlated with automation, but that this is only true if the workforce consists of predominantly unskilled workers.

Chapter 1

Is Baumol's Cost Disease Really a Disease? Healthcare Expenditure and Factor Reallocation[1]

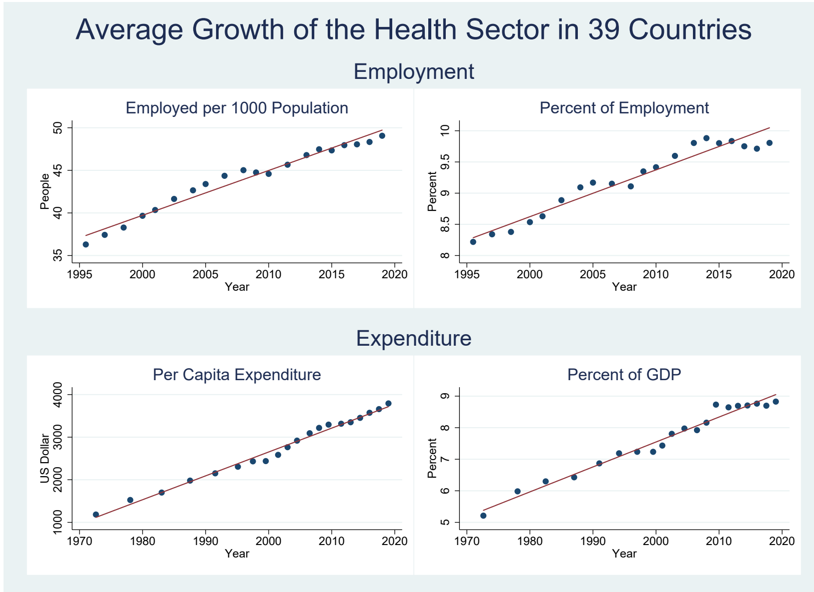
1.1 Introduction

For many decades, healthcare expenditures as a share of GDP have been continuously on the rise in OECD countries. At the same time, employment in the health sector relative to the rest of the economy has also increased, see Figure 1.1.[2] Moreover, there is wide agreement that productivity growth in the health sector relative to the rest of the economy is lower (see Sheiner and Malinovskaya (2016) and Okunade and Osmani (2018)).

One concept that has been used in the past to study these patterns is Baumol's cost disease (Baumol (1967)) - if productivity growth in one sector is higher than in the other and wages in both sectors are positively related, then this entails that the production costs and prices in the less productive sector will grow relative to the more productive sector (see

[1]This chapter is joint work with Lukas Weber

[2]In our empirical analysis, we will focus on the case of Germany. All four trends considered in Figure 1.1 are the same for Germany. The corresponding Figure A.1 can be found in the Appendix.



This figure provides a graphical illustration of the trend in employment and expenditure in the healthcare sector. Data from the OECD on 39 countries are combined, which are listed in Appendix A.2.1.

Figure 1.1: *Employment and Expenditure in the Healthcare Sector*

also Nordhaus (2008)). Multiple empirical studies have presented evidence that Baumol's cost disease is indeed partly responsible for the increase in healthcare expenditures as a share of GDP (see, for example, Hartwig (2008), Bates and Santerre (2013), Hartwig and Sturm (2014), and Colombier (2017)). Inspired by these findings, a large literature on how best to contain the expenditure disease in the healthcare sector emerged (for a review, see Stadhouders *et al.* (2019)).

However, an open question that remains in this context is whether the rise in health expenditures as a share of GDP and the reallocation of labor to the health sector combined with lower productivity growth in the health sector relative to the rest of the economy is necessarily inefficient or a "disease" and directly warrants government intervention. In this paper, we study this question in more detail and attempt to provide a potential answer to it.

To that end, we build on Acemoglu and Guerrieri (2008) and construct a microfounded two-sector closed economy general equilibrium model, and show under which conditions this model can rationalize the stylized facts presented before. In contrast to Baumol (1967), we explicitly model the demand side and thus the demand for the different goods. We assume preferences are homothetic and therefore rule out any effect operating through the income elasticity of demand.[3] An increase in the level of productivity in the non-health sector leads to an *income* and a *substitution* effect. The reallocation of the flexible production factor, as well as whether healthcare expenditures as a share of GDP increase in response to productivity growth in the non-health sector, depends on which effect dominates. We show that if health and non-health goods are complements the income effect dominates the substitution effect, leading to a reallocation of production factors from the non-health sector to the health sector and an increase in the share of healthcare expenditures as a share of GDP. If they are substitutes, the substitution effect dominates the income effect, and the opposite occurs. In case the elasticity of substitution is one, the two effects exactly offset each other, and the allocation of production factors remains unchanged. Therefore, the central parameter in our framework is the elasticity of substitution between the two goods, which governs whether health and non-health goods are complements or substitutes.

This entails, that our model, in contrast to the one proposed in Baumol (1967), has additional testable implications that can be examined using available data, i.e., the value of the elasticity of substitution between health and non-health consumption.[4]

Our theory does not depend on any forms of frictions or rigidities to rationalize the empirical findings and thus suggests that the patterns observed in the data are *potentially* optimal from the perspective of a utility-maximizing representative household, i.e., it is optimal to spend a larger fraction of nominal income on the good that is produced in the relatively less productive sector and allocate more production factors to the relatively less

[3]If preferences are non-homothetic and health consumption constitutes a luxury good, an increase in the share of expenditures devoted to health consumption could be explained by higher income levels. However, studies such as Martín *et al.* (2011) and Ke *et al.* (2011) have found income elasticities with respect to health consumption of less than one, i.e., they found evidence that health consumption is not a luxury good.

[4]Baumol (1967) predicts that wages increase in excess of productivity growth in the stagnant sector, and this is how the theory is often tested empirically (see, for example, Hartwig (2008)). Our theory can make the same prediction if we assume that there is a flexible production factor. However, in our model, this ultimately depends on the value of the elasticity of substitution between health and non-health consumption, giving us an additional testable implication.

productive sector.[5] Therefore, our theory warrants caution when regarding the rise in health expenditures as a share of GDP combined with lower productivity growth in the healthcare sector relative to the rest of the economy as problematic or inefficient.

Whether the pattern in the data is indeed optimal from the perspective of the representative household depends, as mentioned before, on the value of the elasticity of substitution between health and non-health goods. More specifically, we require that the elasticity of substitution between health and non-health goods is below one, i.e., health and non-health goods are complements, in order for our model to rationalize the stylized facts and for the pattern observed in the data to be line with the behavior of a utility-maximizing representative household.

We, therefore, proceed to estimate the elasticity of substitution using German household-level data. Our estimates suggest that the elasticity of substitution is below one, which supports our theory. Moreover, the model makes contrasting predictions regarding the skill premium in the health and non-health sectors, depending on the value of the elasticity of substitution. More specifically, if the elasticity of substitution is below one, an increase in the level of productivity in the non-health sector relative to the health sector leads to a higher skill premium in the health sector relative to the non-health sector. This provides us with an additional possibility to assess the validity of our model. Using German wage data, we show that the data supports the prediction our model makes if the elasticity of substitution between health and non-health goods is below one. Subsequently, we extend our analysis to the US, where we find similar patterns.

This paper is related to a large literature on health economics. The existing literature is largely concerned with identifying the determinants of healthcare spending (see for example Erdil and Yetkiner (2009), De Meijer *et al.* (2013), Baltagi *et al.* (2017) and You and Okunade (2017)) or, relatedly, productivity growth in the healthcare sector (see for example Dunn *et al.* (2022), Cutler *et al.* (2022) and Chernew and Newhouse (2011) for a review). In this strand of

[5] Our model does not feature any imperfections or externalities, and thus the competitive equilibrium is Pareto efficient by the 1st Welfare Theorem.

the literature, of which Getzen and Okunade (2017) provide a concise review, determinants of healthcare spending are analyzed on the macro level. This paper in contrast suggests a microeconomic explanation for increased healthcare spending.

This paper also relates to the large literature on structural change and non-balanced growth (see Herrendorf *et al.* (2014) for an overview). This literature seeks to understand structural change through mechanisms that either pertain to the supply side or the demand side. Theories concentrating on the supply side focus on factors such as differences in rates of technological progress and capital intensities (see, for example, Baumol (1967), Ngai and Pissarides (2007), Acemoglu and Guerrieri (2008), and Duarte and Restuccia (2010)). In contrast, theories focusing on the demand side emphasize the role of non-homothetic preferences, i.e., the income elasticity of demand differs across income groups (see, for example, Kongsamut *et al.* (2001), Boppart (2014), Alonso-Carrera and Raurich (2015), and Comin *et al.* (2021)). In this paper, we contribute to the literature by attempting to combine the two views. To that end, we assume households consume two different goods but otherwise have standard homothetic preferences. If productivity growth in the two sectors differs, this can lead to a reallocation of production factors from one sector to the other. The direction of reallocation is solely determined by the preferences of the households, namely, by the elasticity of substitution. Thus, we highlight the importance of another elasticity, i.e., the elasticity of substitution, relative to the income elasticity of demand, in contributing to explaining structural change.

The rest of the paper is structured as follows. Section 1.2 introduces the model. In Section 1.2.4 we derive the theoretical results that serve as testable predictions. Section 1.3 empirically tests the predictions made by the model and Section 1.4 concludes.

1.2 Theory

1.2.1 Production

We consider a closed economy with no capital. Each good is produced using high- and low-skilled labor with a constant returns to scale production technology. Sector 1 produces good 1 and Sector 2 produces good 2.[6]

The production function for good j with $j \in \{1,2\}$ is given as

$$Y_{j,t} = L_{j,t}^{\alpha_j}(A_{j,t}H_{j,t})^{1-\alpha_j}, \tag{1.1}$$

with $\alpha_j \in (0,1)$.

There are three groups of households: engineers and doctors, who together constitute high-skilled labor and low-skilled workers. Engineers work in Sector 1, i.e., the non-health sector, and doctors work in Sector 2, i.e., the health sector. We assume that high-skilled labor cannot switch sectors.[7] Becoming a high-skilled worker requires acquiring occupation-specific skills through, for example, university studies, which takes time. In our model, we consider the short- and medium-run and therefore assume that workers can't acquire additional occupational skills.[8] Assuming all high-skilled labor is employed implies

$$\begin{aligned} N_t^e &= H_{1,t}, \\ N_t^d &= H_{2,t}. \end{aligned}$$

We assume that the low-skilled labor supply is fixed and denoted by N_t^l. Moreover, we

---

[6]Throughout the chapter, we use the term good, but this is only done for simplicity and does not imply that we only consider physical products.

[7]While assuming that high-skilled labor cannot switch sectors is certainly an overly restrictive and simplifying assumption, there is evidence for labor mobility to decrease with the education level, which, presumably, is a proxy for skill level. Mincer and Jovanovic (1981) note that the probability to switch jobs is negatively predicted by an individual's education level. In addition, Kambourov and Manovskii (2009) find evidence for occupation-specific human capital. Neffke *et al.* (2017) find that it is mainly workers with low wages in low-skill occupations that change their employment across the industry classification system.

[8]Our main results, except for Proposition 3, do not depend on the assumption that high-skilled labor is immobile; see Section A.1.3 in the Appendix.

assume that, unlike the other production factors, low-skilled labor is fully mobile, i.e., can switch between sectors at no cost.

An equilibrium in the market for low-skilled labor requires

$$N_t^l = L_{1,t} + L_{2,t},$$

where $L_{1,t}$ and $L_{2,t}$, respectively, denote the number of low-skilled workers employed in either sector.

To keep the production side as simple as possible, we assume firms operate under perfect competition and thus take all prices as given and make zero profits in equilibrium. The profit maximization problem of each sector is given as

$$\max_{L_{1,t},H_{1,t}} \pi_{1,t} = p_{1,t} L_{1,t}^{\alpha_1} (A_{1,t} H_{1,t})^{1-\alpha_1} - W_{1,t}^l L_{1,t} - W_{1,t}^h H_{1,t}, \tag{1.2}$$

$$\max_{L_{2,t},H_{2,t}} \pi_{2,t} = p_{2,t} L_{2,t}^{\alpha_2} (A_{2,t} H_{2,t})^{1-\alpha_2} - W_{2,t}^l L_{2,t} - W_{2,t}^h H_{2,t}. \tag{1.3}$$

Good 1 is used as the numeraire, and thus $p_{1,t} \equiv 1$.

Define the nominal wage of high-skilled labor, i.e., in terms of the numeraire, of each group as[9]

$$w_t^e = \frac{W_{1,t}^h}{p_{1,t}} = w_{1,t}^h,$$
$$w_t^d = \frac{W_{2,t}^h}{p_{1,t}} = w_{2,t}^h,$$

and the nominal wage of low-skilled labor in each sector as

$$w_{1,t}^l = \frac{W_{1,t}^l}{p_{1,t}},$$
$$w_{2,t}^l = \frac{W_{2,t}^l}{p_{1,t}}.$$

Using $p_t = \frac{p_{2,t}}{p_{1,t}}$, nominal wages can be written as

[9] To find the real wage rate, we need to calculate a price index, which depends on prices and the structure, as well as parameters, of the utility function; see Section 1.2.2.

$$\begin{aligned}
w^h_{1,t} =&(1-\alpha_1)L^{\alpha_1}_{1,t}A^{1-\alpha_1}_{1,t}H^{-\alpha_1}_{1,t},\\
w^h_{2,t} =&p_t(1-\alpha_2)L^{\alpha_2}_{2,t}A^{1-\alpha_2}_{2,t}H^{-\alpha_2}_{2,t},\\
w^l_{1,t} =&\alpha_1 L^{\alpha_1-1}_{1,t}(A_{1,t}H_{1,t})^{1-\alpha_1},\\
w^l_{2,t} =&p_t\alpha_2 L^{\alpha_2-1}_{2,t}(A_{2,t}H_{2,t})^{1-\alpha_2}.
\end{aligned}$$

Aggregate nominal income of each group is given as[10]

$$\begin{aligned}
I^e_t N^e_t =&Y_{1,t} - w^l_{1,t}L_{1,t},\\
I^d_t N^d_t =&p_t Y_{2,t} - w^l_{2,t}L_{2,t},\\
I^l_t N^l_t =&w^l_{1,t}L_{1,t} + w^l_{2,t}L_{2,t}.
\end{aligned}$$

Aggregating over the three groups yields aggregate production

$$I^e_t N^e_t + I^d_t N^d_t + I^l_t N^l_t = Y_{1,t} + p_t Y_{2,t}.$$

1.2.2 Households

Preferences are homothetic, and a household of group i with $i \in \{e, d, l\}$ consumes a final good c^i_t that is produced by combing two goods, i.e., good 1 and good 2, using a CES aggregator. This gives rise to the following maximization problem in nominal terms

$$\begin{aligned}
&\max_{c^i_{1,t}, c^i_{2,t}} c^i_t(c^i_{1,t}, c^i_{2,t}) = \left(\gamma^{\frac{1}{\theta}}(c^i_{1,t})^{\frac{\theta-1}{\theta}} + (1-\gamma)^{\frac{1}{\theta}}(c^i_{2,t})^{\frac{\theta-1}{\theta}}\right)^{\frac{\theta}{\theta-1}} \quad i \in \{e, d, l\}\\
&\text{s.t. } c^i_{1,t} + p_t c^i_{2,t} = I^i_t,
\end{aligned} \tag{1.4}$$

with $\theta \in (0, \infty)$. θ denotes the elasticity of substitution between the two goods. For $\theta \in (0, 1)$, the two goods are complements, and for $\theta \in (1, \infty)$, the two goods are substitutes.

The optimal demand for either good is given as

[10] $p_t Y_{2,t} = w^l_{2,t}L_{2,t} + w^h_{2,t}H_{2,t}$ due to perfect competition and constant returns to scale.

$$
\begin{aligned}
c^i_{1,t} =& \gamma \frac{I^i_t}{\gamma + (1-\gamma) p_t^{1-\theta}}, && (1.5)\\
c^i_{2,t} =& (1-\gamma) p_t^{-\theta} \frac{I^i_t}{\gamma + (1-\gamma) p_t^{1-\theta}}, && (1.6)
\end{aligned}
$$

where $p_t = \frac{p_{2,t}}{p_{1,t}}$ is the relative or nominal price of good $c_{2,t}$.

The price index, i.e., the price of one unit of c^i_t, is given as

$$
\mathcal{P}_t = \left(\gamma + (1-\gamma) p_t^{1-\theta}\right)^{\frac{1}{1-\theta}}.
$$

We assume that preferences are the same across groups, i.e., all households are symmetric.[11]

1.2.3 Equilibrium

Market clearing requires that, for each good, demand be equal to supply

$$
\begin{aligned}
Y_{1,t} =& \sum_i \gamma \frac{I^i_t}{\gamma + (1-\gamma) p_t^{1-\theta}} N^i_t,\\
Y_{2,t} =& \sum_i (1-\gamma) p_t^{-\theta} \frac{I^i_t}{\gamma + (1-\gamma) p_t^{1-\theta}} N^i_t.
\end{aligned}
$$

We can combine the equilibrium conditions of the two goods markets

$$
\begin{aligned}
\frac{Y_{2,t}}{Y_{1,t}} =& \frac{(1-\gamma) p_t^{-\theta} \sum_i I^i_t N^i_t}{\gamma \sum_i I^i_t N^i_t}\\
p_t^{\theta} \frac{Y_{2,t}}{Y_{1,t}} =& \frac{1-\gamma}{\gamma}.
\end{aligned}
\qquad (1.7)
$$

As low-skilled labor is fully mobile, we require an additional equation that determines the equilibrium division of low-skilled labor between the two sectors, i.e., we need to determine the equilibrium values of $L_{1,t}$ and $L_{2,t}$. Full mobility implies that the nominal

[11]See Section A.1.4 in the Appendix for a short discussion of how heterogeneous preferences could affect the model.

wage rate in both sectors needs to be equal

$$w^l_{1,t} = w^l_{2,t} = w^l_t.$$

In equilibrium, firms maximize their profits, households maximize their utility, all markets clear, and the wage rate of low-skilled labor has to be equal across both sectors. We can characterize the equilibrium as a system of two non-linear equations

$$\begin{aligned} F \equiv & p_t^{\theta} \frac{(L_t - L_{1,t})^{\alpha_2} (A_{2,t} H_{2,t})^{1-\alpha_2}}{L_{1,t}^{\alpha_1} (A_{1,t} H_{1,t})^{1-\alpha_1}} - \frac{1-\gamma}{\gamma} = 0 \\ F \equiv & p_t^{\theta} \frac{Y_{2,t}}{Y_{1,t}} - \frac{1}{\gamma} + 1 = 0, \end{aligned} \tag{1.8}$$

$$\begin{aligned} G \equiv & w^l_{1,t} - w^l_{2,t} = 0 \\ G \equiv & \alpha_1 L_{1,t}^{\alpha_1 - 1} (A_{1,t} H_{1,t})^{1-\alpha_1} - p_t \alpha_2 (L_t - L_{1,t})^{\alpha_2 - 1} (A_{2,t} H_{2,t})^{1-\alpha_2} = 0 \\ G \equiv & \alpha_1 \frac{Y_{1,t}}{L_{1,t}} - p_t \alpha_2 \frac{Y_{2,t}}{L_{2,t}} = 0, \end{aligned} \tag{1.9}$$

with p_t and $L_{1,t}$ as the endogenous variables, where p_t is the relative price of good 2 and $L_{1,t}$ the number of low-skilled workers employed in Sector 1. Equation 1.8 determines the relative price p_t such that the demand and supply for both goods are equalized. Equation 1.9 is only present if low-skilled labor is mobile.[12] It ensures that the wage in either sector is equal for low-skilled workers.

1.2.4 Results

Lemma 1. *An increase in $A_{1,t}$ leads ceteris paribus to an increase in the relative price of good 2, i.e., p_t.*

Proof. See Section A.1.1 in the Appendix. □

A higher level of productivity in Sector 1 relative to Sector 2 entails that good 1 becomes

[12] Would we assume that low-skilled labor cannot switch between sectors, the equilibrium could be characterized by only one equation, i.e., Equation 1.8.

relatively more abundant and good 2 relatively more scarce.[13] Thus, the relative price of good 2 will increase. This is in line with the empirical evidence presented in Nordhaus (2008).

Proposition 1. *An increase in $A_{j,t}$ has the following effect on $L_{j,t}$*

$$\frac{\partial L_{j,t}}{\partial A_{j,t}} \begin{cases} < 0 & \text{if } \theta < 1, \\ > 0 & \text{if } \theta > 1, \\ = 0 & \text{if } \theta = 1. \end{cases}$$

Proof. See Section A.1.1 in the Appendix. □

Proposition 1 states that an increase in the level of productivity in Sector 1, i.e., the non-health sector, can either lead to an inflow or outflow of low-skilled labor from this sector, depending on whether the two consumption goods are complements, i.e., $\theta \in (0,1)$, or substitutes, i.e., $\theta \in (1, \infty)$. Moreover, this also implies that if $\theta \in (0,1)$, a fall in the productivity level of the health sector due to, for example, exogenous distortions or inefficiencies, would lead to a reallocation of low-skilled labor to the health sector.

The economic intuition behind this result is that an increase in $A_{1,t}$ increases $w^l_{1,t}$ directly through a *scale* effect. In addition, it leads to a rise in p_t, which increases $w^l_{2,t}$, i.e., good 1 becomes more abundant, and thus the inverse of its relative price increases, through an indirect *price* effect. In equilibrium, the no-arbitrage condition must be satisfied, i.e., $w^l_{1,t} = w^l_{2,t}$, thus as low-skilled labor is fully flexible, it will switch between sectors if the *scale* effect is larger or smaller than the *price* effect. For $\theta = 1$ the two effects exactly offset each other; for $\theta < 1$, i.e., the two goods being complements, the *price* effect dominates the *scale* effect, which leads to an outflow of low-skilled labor from Sector 1, which increases $w^l_{1,t}$ and reduces $w^l_{2,t}$. For $\theta > 1$, i.e., the goods being substitutes, the *scale* effect dominates the *price* effect, which leads to an outflow of low-skilled labor from Sector 2 and a corresponding inflow into Sector 1, which reduces $w^l_{1,t}$ and increases $w^l_{2,t}$.

[13]Productivity does not only encompass the level of technology but also other factors that determine how efficiently the input factors can be combined in the production process. A fall in $A_{2,t}$, i.e., the healthcare sector becoming less efficient, would yield the same qualitative results.

We can also interpret our results in terms of an income and a substitution effect. An increase in $A_{1,t}$ makes good 2 more expensive relative to good 1, and thus consumers will consume more of the relatively cheaper good; this is the *substitution* effect. Moreover, a higher level of $A_{1,t}$ also makes the economy altogether richer.[14] This leads to an *income* effect, i.e., households will demand more of both goods. Which effect dominates depends on the elasticity of substitution between the two goods. For $\theta \in (0, 1)$, the income effect dominates the substitution effect, and to satisfy the additional demand for good 2, low-skilled labor is transferred from Sector 1 to Sector 2. If $\theta \in (1, \infty)$, the substitution effect dominates the income effect, leading to a reallocation of low-skilled labor to Sector 1 to meet the additional demand for good 1. For $\theta = 1$, i.e., log utility, the two effects cancel each other out.

Unlike Baumol (1967), we provide a micro-founded theory and explicitly model how the flexible production factor is allocated between the two sectors. A reallocation of production factors from the sector that experiences an increase in productivity relative to the other sector might at first seem counterintuitive, as it reduces overall physical output, i.e., $Y_{1,t} + Y_{2,t}$. However, the utility of households in this economy is *not* necessarily maximized by maximizing the physical production of the two goods; that would only be the case if the two goods are perfect substitutes, i.e., $\theta \to \infty$. Rather, households want to consume an optimal relative bundle of the two goods, which depends on preferences and relative prices as well as the elasticity of substitution.[15]

Moreover, recall that our model does not feature any form of imperfections or externalities, and thus the competitive equilibrium derived here is Pareto efficient by the 1st Welfare Theorem. Therefore, if the economy devotes more income and resources to the less productive sector, i.e., the health sector, this does not necessarily mean that the economy suffers from a form of inefficiency or "disease" that warrants government intervention. Rather, it could be the case that preferences, i.e., the elasticity of substitution, are such that

[14] As preferences are homothetic and the same for all groups, the distribution of the additional income is not relevant.

[15] Combing the first order conditions of the representative household yields $\frac{c_{1,t}}{c_{2,t}} = \frac{\gamma}{1-\gamma}\left(\frac{p_{2,t}}{p_{1,t}}\right)^{\theta}$.

the income effect dominates the substitution effect.

Proposition 2. *If $\alpha_1 = \alpha_2 = \alpha$, an increase in $A_{1,t}$ always has the following effect on the share of good 1 in nominal GDP, i.e., $\xi_t = \frac{Y_{1,t}}{Y_{1,t}+p_t Y_{2,t}}$,*

$$\frac{\partial \xi_t}{\partial A_{1,t}} \begin{cases} < 0 & \text{if } \theta < 1, \\ > 0 & \text{if } \theta > 1, \\ = 0 & \text{if } \theta = 1. \end{cases}$$

Proof. See Section A.1.2 in the Appendix. □

There are three channels through which an increase in $A_{1,t}$ can influence ξ_t in this model. First, directly by increasing the output produced in Sector 1. Second, by triggering a reallocation of low-skilled labor from one sector to another. Third, by influencing the relative price of good 2 and thus affecting the nominal value of output produced in Sector 2. The first and third channels have opposite effects on ξ_t. The sign of the effect of the second channel on ξ_t depends on the value of θ. The result of the above proposition remains unchanged if we assume that, unlike in Baumol (1967), all production factors are immobile.

Let $\phi_{j,t}$ denote the skill premium in sector j. Thus, the skill premium in Sector 1 is given as

$$\phi_{1,t} = \frac{w^h_{1,t}}{w^l_{1,t}} = \frac{1-\alpha_1}{\alpha_1}\frac{L_{1,t}}{H_{1,t}}, \tag{1.10}$$

and in Sector 2 as

$$\phi_{2,t} = \frac{w^h_{2,t}}{w^l_{2,t}} = \frac{1-\alpha_2}{\alpha_2}\frac{L_{2,t}}{H_{2,t}} = \frac{1-\alpha_2}{\alpha_2}\frac{L_t - L_{1,t}}{H_{2,t}}. \tag{1.11}$$

Lemma 2. *An increase in $L_{j,t}$ leads ceteris paribus to a higher skill premium in sector j and a lower skill premium in sector k for $j \neq k$.*

Proof. Follows from $\frac{\partial \phi_{1,t}}{\partial L_{1,t}} > 0$ and $\frac{\partial \phi_{2,t}}{\partial L_{1,t}} < 0$. □

The elasticity of substitution between high- and low-skilled labor is one, and the production function has positive but decreasing returns to scale with respect to high- and low-skilled labor, respectively, i.e., $Y_{L_{j,t}} > 0$ and $Y_{L_{j,t}L_{j,t}} < 0$. In addition, the cross derivatives are positive, i.e., $Y_{H_{j,t}L_{j,t}} > 0$. Thus, an inflow of low-skilled labor will decrease the wage rate of low-skilled labor and increase the wage rate of high-skilled labor, as they are complemented by the additional low-skilled workers. Hence, if low-skilled workers switch from the non-health sector to the health sector, this increases the wage rate of high-skilled workers in the health sector and decreases the wage rate of low-skilled workers in the health sector, and vice versa for the wage rate in the non-health sector.

Proposition 3. *An increase in $A_{j,t}$ leads ceteris paribus to a lower skill premium in sector j and a higher skill premium in sector k if $\theta \in (0, 1)$ and to a higher skill premium in sector j and a lower skill premium in sector k if $\theta \in (1, \infty)$ for $j \neq k$.*

Proof. Follows from Proposition 1 and Lemma 2. □

Therefore, a rise in the skill premium in the health sector relative to the rest of the economy can be explained by an increase in the level of technology in the non-health sector if $\theta \in (0, 1)$. The intuition for this result is that for $\theta \in (0, 1)$ an increase in $A_{1,t}$ leads to an outflow of low-skilled labor from the non-health sector and a corresponding inflow of low-skilled labor into the health sector. Thus, the ratio of low-skilled workers to high-skilled workers increases in the health sector. As this ratio governs the skill premium in our model, the change therein leads to a rise in the skill premium in the health sector relative to the rest of the economy.

As discussed in the introduction, productivity growth in the non-health sector seems to be stronger than in the health sector. Moreover, we observe a rise in health expenditures as a share of GDP and an increase in the share of workers employed in the healthcare sector. Similar to Baumol (1967), our model can potentially replicate these empirical findings. The sufficient condition for our model to do so is that the elasticity of substitution between health and non-health consumption, i.e., θ, is below one. However, in contrast to the former, our model also provides us with additional testable implications that can be tested using available data. The first is whether the elasticity of substitution between health and non-health consumption is indeed below one. The second is whether the skill premium in the health sector has increased relative to the rest of the economy.

1.3 Testable Model Implications

The model described in Section 1.2 can be falsified by testing its implications empirically along two lines. First, the model predicts that a reallocation of resources towards the less productive sector, as documented in the introduction, depends on the parameter value of θ. Specifically, the resource reallocation is expected to occur if the two consumption goods considered, in this case, healthcare and all other consumption, are complements. This is equivalent to $\theta < 1$, which is a necessary condition for the mechanism proposed in the model to explain the empirical facts highlighted in the introduction. Using data to test if indeed $\theta < 1$, the model can be falsified. And second, the model predicts that given $\theta < 1$ and higher mobility of unskilled than skilled labor, both the share of unskilled labor and the skill premium in the health sector increase. In the model, this is due to a shift of unskilled labor from the non-health to the health sector. To assess the model's validity and assumptions, both aspects are addressed in this section.

1.3.1 Preference Estimation

In the introduction, we documented a shift of resources toward the health sector. This reallocation took place despite lower productivity growth in the health sector than in the rest of the economy. The model in Section 1.2 provides a micro foundation for the mechanisms that can rationalize this finding. It predicts that a shift of resources towards the less productive sector occurs only if the goods produced in the less productive sector are complements to the goods produced in the other sector. The crucial parameter and its restriction to see such a reallocation is $\theta < 1$.

The FOCs from the household maximization can be used to derive the optimal consumption ratio of c_1 and c_2.

$$\frac{c_2}{c_1} = \frac{1-\gamma}{\gamma}\left(\frac{p_1}{p_2}\right)^{\theta}$$
$$\ln\left(\frac{c_2}{c_1}\right) = \ln\left(\frac{1-\gamma}{\gamma}\right) + \theta \ln\left(\frac{p_1}{p_2}\right)$$

This log-linearized ratio can be used to motivate an estimation equation. Of course, other factors besides relative prices and the substitution parameter θ may influence the optimal ratio. We assume that these are captured by the error term ε. The estimation equation is given by

$$\ln\left(\frac{c_{2,t}}{c_{1,t}}\right) = \ln\left(\frac{1-\gamma}{\gamma}\right) + \theta \ln\left(\frac{p_{1,t}}{p_{2,t}}\right) + \varepsilon_t, \tag{1.12}$$

where ε_t is the error term and $\ln\left(\frac{1-\gamma}{\gamma}\right)$ is the constant.

Data

Equation (1.12) demonstrates how the elasticity of substitution between healthcare spending and all other consumption spending can be estimated. Using microdata, it can be tested if $\theta < 1$, implying that healthcare consumption is complementary to all other consumption. Specifically, to estimate θ in microdata, variation in both prices and quantities at the household level is necessary. These requirements are met by the German EVS data provided by the Statistisches Bundesamt. It is a triennial household-level survey, providing detailed information on household expenditures, as well as socioeconomic information, for roughly 40,000 representative households in each wave. In addition to reporting very granular expenditure data, the EVS also provides the user with transparently combined aggregate measures for different spending categories, one of them being healthcare. For the estimation, the EVS waves of 2003 and 2018 are used.

The Statistisches Bundesamt collects the EVS data with the primary purpose of constructing inflation measures from it. The official price data, also obtainable from the Statistisches Bundesamt, are derived from the EVS data. We, therefore, use and combine two data sets from the same data source. This guarantees a correspondence of available price sub-indices and consumption categories in the EVS. Since price data is indispensable for the estimation proposed, this constitutes a considerable advantage of using EVS data. For the estimation, it is essential to obtain price variation at the household level. The household-level price data is constructed by weighting the official prices of the sub-categories of consumption with the household-specific shares of expenditure devoted to each sub-category of consumption. Importantly, the data only covers expenditures made by the household. For healthcare expenditures, this means that only those expenditures that are not covered by health insurance are recorded in the EVS. This poses a problem for identification, which is discussed in the next section.

Identification

A common problem with measuring household-level healthcare expenditures is that healthcare spending is often at least partially covered by private or public health insurance. Therefore, healthcare spending by households is likely to be underestimated. In Germany, public health insurance is mandatory and it arguably covers most if not even all necessary treatments. If households report private healthcare spending, it is for services above and beyond the quite generous basic coverage. Analytically, mandatory healthcare insurance can be modeled as the opposite of a subsistence constraint. This is in analogy to the class of Stone-Geary utility functions (going back to Geary (1950) and Stone (1954)). In that case, household preferences are given by

$$\max_{c_{1,t},c_{2,t}} c_t(c_{1,t},c_{2,t}) = \left(\gamma^{\frac{1}{\theta}}(c_{1,t})^{\frac{\theta-1}{\theta}} + (1-\gamma)^{\frac{1}{\theta}}(\kappa \cdot c_{2,t} + (1-\kappa)\overline{x}_2)^{\frac{\theta-1}{\theta}}\right)^{\frac{\theta}{\theta-1}}$$

$$\text{s.t. } p_{1,t}c_{1,t} + p_{2,t}c_{2,t} = I_t \text{ and } c_{2,t} \geq 0$$

where $(1-\kappa)\overline{x}_2$ refers to healthcare spending covered by insurance, which is paid for through taxation, and I_t denotes the net-of-tax income. $\kappa \cdot c_2$ refers to healthcare spending on top of items covered by health insurance. Total healthcare consumption by the household is given by $\kappa \cdot c_{2,t} + (1-\kappa)\overline{x}_2$. In the presence of $\overline{x}_2$, the optimal value of c_2 can be zero, requiring an additional non-negativity constraint in the household maximization problem. For estimation, only households reporting positive private expenditures on healthcare are used, such that the non-negativity is met by all observations included in the estimation. c_1 refers to all other consumption. Analogous to the above, an estimation equation for θ can be derived from the FOCs:

$$\kappa \cdot \frac{\kappa \cdot c_2 + (1-\kappa) \cdot \overline{x}_2}{c_1} = \frac{1-\gamma}{\gamma} \left(\frac{p_1}{p_2} \right)^{\theta} \tag{FOCs}$$

$$\ln \left(\frac{\kappa \cdot c_2 + (1-\kappa) \cdot \overline{x}_2}{c_1} \right) = \ln \left(\frac{1-\gamma}{\gamma} \right) - \ln(\kappa) + \theta \ln \left(\frac{p_1}{p_2} \right) \tag{a}$$

$$\ln \left(\frac{c_2}{c_1} \right) = \ln \left(\frac{1-\gamma}{\gamma} \right) - \ln(\kappa) + \theta_b \ln \left(\frac{p_1}{p_2} \right) \tag{b}$$

Ideally, we would like to estimate Equation (a), which theoretically is guaranteed to result in an unbiased estimate of θ. Since we do not observe $\overline{x}_2$, the only equation we can estimate is Equation (b). This results in an unbiased estimate of θ if the healthcare costs covered by public health insurance are as price sensitive as private healthcare spending. Mathematically, the coefficient of interest is defined as follows:

$$\theta = \frac{\kappa \cdot Cov\left(\frac{c_2}{c_1}, \frac{p_1}{p_2}\right) + (1-\kappa) \cdot Cov\left(\frac{\overline{x}_2}{c_1}, \frac{p_1}{p_2}\right)}{Var\left(\frac{p_1}{p_2}\right)}.$$

The coefficient that can be estimated given the available data is θ_b, which is defined as

$$\theta_b = \frac{Cov\left(\frac{c_2}{c_1}, \frac{p_1}{p_2}\right)}{Var\left(\frac{p_1}{p_2}\right)}.$$

The bias of the estimated coefficient, θ_b, relative to the true coefficient of interest, θ, can be derived mathematically. The estimated coefficient is upward biased whenever

$$\frac{Cov\left(\frac{c_2}{c_1}, \frac{p_1}{p_2}\right)}{Var\left(\frac{p_1}{p_2}\right)} > \frac{\kappa \cdot Cov\left(\frac{c_2}{c_1}, \frac{p_1}{p_2}\right) + (1-\kappa) \cdot Cov\left(\frac{\overline{x}_2}{c_1}, \frac{p_1}{p_2}\right)}{Var\left(\frac{p_1}{p_2}\right)}$$

$$\Longleftrightarrow$$

$$Cov\left(\frac{c_2}{c_1}, \frac{p_1}{p_2}\right) > Cov\left(\frac{\overline{x}_2}{c_1}, \frac{p_1}{p_2}\right).$$

A higher covariance between private healthcare spending and relative prices than between insurance-covered healthcare spending and relative prices is a sufficient condition for the estimated value of $\hat{\theta}$ to be upward biased. The inequality of covariances is likely to hold for two reasons. One, it holds if people are less price-conscious when seeking insurance-covered treatments than when seeking medical treatments which have to be paid for privately. Given that people don't even learn about the costs they cause when seeking treatment covered by health insurance, it seems safe to assume that that is the case.[16] Two, the coverage of medical treatments by public health insurance is likely to be less price sensitive than people when deciding to get elective procedures for which they have to pay the costs themselves. There are binding regulations determining which medical treatments have to be covered by public health insurance. The procedure to change these regulations is lengthy and generally not initiated by price changes.[17] Therefore, the covariance of public health insurance coverage and treatment costs is likely to be lower than that of elective healthcare expenditures and treatment costs. This results in $Cov(c_2, \frac{p_1}{p_2}) > Cov(\overline{x}_2, \frac{p_1}{p_2})$.

The bias increases in the difference in price variability of $\frac{\overline{x}_2}{c_1}$ and $\frac{c_2}{c_1}$, expressed above by the respective covariances. In addition, note that the bias increases in $(1-\kappa)$, assuming that $Cov\left(\overline{x}_2, \frac{p_1}{p_2}\right) > 0$. Effectively, the bias results from an estimation that disregards an

[16]The official website of the German public health insurance details, among other things, the contributions to and benefits of the public health insurance in Germany (only available in German, for a short link see `https://t.ly/wJ78z`).

[17]The German government mandates the Federal Joint Committee (Gemeinsamer Bundesausschuss) to determine the benefits and tariffs of the statutory health insurance funds. Details on its mandate and operation can be found on its official website (only available in German, for a short link see `https://t.ly/xvwPw`)

unobserved part of healthcare consumption that has a lower price sensitivity than the observed part of healthcare consumption. If the observed share of overall healthcare consumption increases, the estimation bias decreases. The upward bias can be directly derived as

$$\theta_b = \hat{\theta} = \left(\theta - (1-\kappa) \frac{Cov\left(\overline{x}_2, \frac{p_1}{p_2}\right)}{Var\left(\frac{p_1}{p_2}\right)} \right) \frac{1}{\kappa}.$$

For $\lim_{\kappa \to 1} \hat{\theta} = \theta$, whereas $\lim_{\kappa \to 0} \hat{\theta} = \infty$.

When consuming out-of-pocket healthcare, the basic healthcare needs of consumers in Germany have already been met by public health insurance. Estimating the empirical model given by (1.12) (which is equivalent to Equation (b)), this is not accounted for. Thus $\hat{\theta}$ is biased upward in the presence of relatively price-inelastic, mandatory, and sufficiently generous healthcare insurance. The bias invariably works against finding complementarity of healthcare spending and all other consumer spending.[18]

Eliminating the bias and obtaining unbiased estimates would require data on both the health insurance premium directly subtracted from income, as well as a monetary estimate of the health care sought out but paid for by the insurance on the individual level. Unfortunately, this is not possible due to data availability. Based on Equation (1.12), we proceed to estimate θ using the German EVS data. Keeping in mind the upward bias mandatory health insurance exudes on the estimated coefficient, this estimation can still provide us with insightful results.

Estimating θ

For the estimation of Equation (1.12), the aggregated value for health spending relative to the rest of consumption spending is analyzed. If using aggregated values, the price for health spending is constant across all observations, as variation in the composition of health

[18]Note that the bias described here is different from a classical measurement error in the dependent variable. This would require the measurement error $\overline{x}_2$ to be independent of c_2. The generosity of the public health insurance coverage however is very likely to be correlated with private healthcare spending, such that the problem at hand cannot adequately be described with classical measurement error.

spending across individuals cannot be used. This implies that only cross-year analysis is feasible. Results from the structural equation estimation are reported in Table 1.1.

The main regression result is reported in column one, using the whole pooled EVS sample. In addition, columns two, three, and four report the estimated elasticity of substitution for subsamples divided along the income distribution. In theory, we would expect both preference parameters γ (estimated indirectly by the constant) and θ to be constant across all subsamples. The theory is derived with the clearly simplifying assumption of a representative agent, such that obtaining non-varying estimates of the two preference parameters in survey data is unrealistic. If, however, the estimated values of the parameters are reasonably stable across subsamples, it suggests some robustness of the results. In particular, it is of special interest to see if θ is estimated to be above or below the value of one.

Table 1.1: *Estimating θ by Income Group*

	All	Bottom 50%	Next 40%	Top 10%
$\hat{\theta}$	0.017	0.149	0.161	1.331
	[-0.19,0.22]	[-0.12,0.42]	[-0.17,0.49]	[0.58,2.08]
Constant	-3.973	-4.087	-3.949	-3.661
	[-3.99,-3.96]	[-4.11,-4.07]	[-3.97,-3.93]	[-3.71,-3.61]
Observations	77,501	37,089	32,219	8,193

Note: Dependent variable is the log ratio of health to all other expenditures as reported in the 2003 and 2018 waves of the EVS. Significance stars are suppressed because they are not informative in this context. The numbers in brackets report the 95%-confidence interval. The constant represents the estimate of $\ln\left(\frac{1-\gamma}{\gamma}\right)$.

From the reported confidence intervals it is quite clear that $\hat{\theta}$ is estimated to be smaller than one, except in the subset of the Top 10% of highest income households. As detailed in the previous section, the estimated coefficients reported in Table 1.1 are biased upward because of the broad coverage public health insurance provides in Germany.

The finding of increasing estimated values of $\hat{\theta}$ along the income distribution can be rationalized with $\bar{x}_2$, the basic coverage of healthcare costs provided by health insurance, being less relevant as income increases. For low levels of income, the amount of healthcare covered by insurance, $(1-\kappa)\overline{x}_2$, may be larger than optimal from the household's point of view, resulting in the non-negativity constraint of c_2 to be binding, such that $c_2^* \leq 0$. However, as income increases, households may want to consume more healthcare than covered by health insurance, such that the non-negativity constraint of c_2 is no longer binding. Assuming a fixed $\overline{x}_2$ across all households, the share of healthcare costs covered by insurance decreases as income increases. This leads to an increase in the upward bias of the estimated coefficient $\hat{\theta}$, as argued above.

The mathematical explanation can be supplemented by intuitive reasoning, why the upward bias is higher, the higher the household income is. The majority of healthcare expenditures by low-income households, if not zero, is likely to be primarily due to co-payments on drugs, dentures, and other basic medical needs, which households have to make irrespective of their price. Households with higher incomes in contrast may decide to get elective medical treatments such as teeth beautification, skin care, or plastic surgery. This intuition is supported by the expenditure elasticities of healthcare ($\epsilon_{health} = 1.17$) and all other consumption ($\epsilon_{other} = 0.97$).[19] The fact that the expenditure elasticity of healthcare is larger than one whereas the expenditure elasticity of all other consumption is smaller than one signifies that healthcare is a luxury good. As a standard CES-utility function in principle cannot accommodate expenditure elasticities that are different from one, two remarks are in order: One, the expenditure elasticities are once more estimated without taking the fixed amount of healthcare provided by insurance into account. This results in an upward bias in the estimated expenditure elasticity of health consumption. Therefore, the difference between expenditure elasticities of total health consumption and all other consumption is likely to be smaller in reality. It highlights again the problems for empirical analysis caused by an only partial observation of healthcare consumption. Two, the finding of non-unit expenditure elasticities implies once more, that in reality preferences cannot be perfectly described by the representative agent with a CES-utility function. Nevertheless,

[19]The expenditure elasticities are estimated by regressing the log of healthcare expenditures on the log of aggregate expenditures. Both 95%-confidence intervals of the estimated expenditure elasticities exclude the value one with $ci_{health} = [1.15, 1.18]$ and $ci_{other} = [0.968, 0.971]$.

the model described in Section 1.2 yields valuable insights in highlighting the role played by the demand side in general and the elasticity of substitution in particular when analyzing factor reallocation across sectors.

As these estimates are based on individual consumption expenditures, noise in the data may attenuate the estimated coefficients. However, for the estimation, the consumption aggregates of the EVS were used and merged with official price data on the same consumption aggregates published by the Statistisches Bundesamt. This leaves no room for own interpretation or discretion about the handling of the data. Therefore, any measurement error exerting attenuation bias would lie with the Statistisches Bundesamt. While classical measurement error cannot be ruled out completely, it is likely to be much smaller than the structural upward bias discussed in the previous section. If anything, we expect the estimates to be upward biased.

Alternative θ Estimation

The estimation equation is derived from the FOCs and therefore under the implicit assumption of constant income. Furthermore, there is no theoretical reason for including income as a control variable when estimating θ, as preferences are assumed to be homothetic. The variation of results across columns reported in Table 1.1 however indicates that the relationship between the consumption ratio and the price ratio changes with income. To investigate and control the role of income in the estimation results, we repeat the estimation, this time including income as an explanatory variable. If preferences are indeed homothetic, we would expect the corresponding coefficient $\hat{\beta}$ to equal to zero.

$$\ln\left(\frac{c_{2,t}}{c_{1,t}}\right) = \ln\left(\frac{1-\gamma}{\gamma}\right) + \theta \ln\left(\frac{p_{1,t}}{p_{2,t}}\right) + \beta \ln(income) + \varepsilon_t \tag{1.13}$$

The results of estimating Equation (1.13), reported in Table 1.2, confirm that income plays a role in the relationship between the consumption ratio and the price ratio. The estimates for θ go up if income is included as a control variable. However, they remain smaller than one except in the subsample of the Top 10% of the income distribution, where

Table 1.2: *Estimating θ with Income Effect*

	All	Bottom 50%	Next 40%	Top 10%
Theta	0.221	0.138	0.287	1.437
	[0.02,0.42]	[-0.13,0.41]	[-0.04,0.62]	[0.68,2.19]
Beta	0.222	0.087	0.407	0.273
	[0.21,0.24]	[0.05,0.12]	[0.33,0.49]	[0.14,0.41]
Constant	-6.018	-4.841	-7.829	-6.442
	[-6.17,-5.87]	[-5.12,-4.56]	[-8.61,-7.05]	[-7.83,-5.05]
Observations	77,473	37,061	32,219	8,193

Note: Dependent variable is the log ratio of health to all other expenditures as reported in the 2003 and 2018 waves of the EVS. Significance stars are suppressed because they are not informative in this context. The numbers in brackets report the 95%-confidence interval. The constant represents the estimate of $\ln\left(\frac{1-\gamma}{\gamma}\right)$.

it is estimated to be larger than one, as in the baseline regression. Income however is positively correlated with the share of consumption made up by healthcare. This once again indicates that healthcare is a luxury good. In our model, we can rationalize a reallocation of resources towards the healthcare sector if $\theta < 1$ under homothetic preferences. The finding reported in Table 1.2 shows that some of the reallocations towards the healthcare sector may be driven by healthcare being a luxury good. Assuming non-homothetic preferences would thus facilitate modeling a reallocation. Our model however can explain the empirical facts with a minimum of free parameters. While non-homothetic preferences may be part of the story, our model can explain the empirical facts using homothetic preferences, which continue to be the benchmark case in economic models.

The structural estimation in (1.12) can also be separated out and reformulated, imposing equality and opposite signs for the two price coefficients. The new estimation equation is given by

$$\ln(c_{2,t}) = \ln\left(\frac{1-\gamma}{\gamma}\right) + \theta\ln(p_{1,t}) - \theta\ln(p_{2,t}) + \eta\ln(c_{1,t}) + \varepsilon_t. \tag{1.14}$$

This does not address the problem of a structural bias in the θ estimate but allows for a more flexible and intuitive estimation. The relationship can be estimated by putting a constraint on the coefficients of $\ln(p_{2,t})$ and $\ln(p_{1,t})$ to be of the same magnitude but have different signs. Results are reported in Table 1.3.

Table 1.3: *Alternative Estimation of θ by Income Group*

	All	Bottom 50%	Next 40%	Top 10%
Other Price	0.092	0.633	0.739	1.782
	[-0.11,0.30]	[0.36,0.90]	[0.41,1.07]	[1.04,2.53]
Health Price	-0.092	-0.633	-0.739	-1.782
	[-0.30,0.11]	[-0.90,-0.36]	[-1.07,-0.41]	[-2.53,-1.04]
Other Consumption	0.895	0.701	0.387	0.294
	[0.88,0.91]	[0.67,0.73]	[0.34,0.43]	[0.21,0.38]
Constant	-3.035	-1.548	1.648	3.052
	[-3.20,-2.87]	[-1.80,-1.29]	[1.25,2.05]	[2.21,3.89]
Observations	77,501	37,089	32,219	8,193

Note: Dependent variable is log healthcare expenditures as reported in the 2003 and 2018 waves of the EVS. The two price coefficients are constrained to be equal but of opposite signs. Significance stars are suppressed because they are not informative in this context. The numbers in brackets report the 95%-confidence interval. The constant represents the estimate of $\ln\left(\frac{1-\gamma}{\gamma}\right)$.

In this setup, $\hat{\theta}$ is the estimated coefficient of Other Price, reported in the first row of Table 1.3. As expected from the previous regressions, it is estimated to be smaller than 1, again indicating that health consumption and other consumption are complements. This is true for the pooled sample as well as the Bottom 50% of the income distribution. As already seen in Table 1.1, the estimated $\hat{\theta}$ increases over the income distribution, which is what would be expected, given the structure of the upward bias discussed earlier. While the estimated $\hat{\theta}$ remains larger than one for the Top 10% of the income distribution, it is still estimated to be smaller than one for the Bottom 90% and the pooled sample, which is reassuring.

1.3.2 Factor Reallocation

Under two assumptions, the model predicts a reallocation of unskilled labor to the less productive sector. Assumption one is that the two goods produced are complements, which is the case if $\theta < 1$, supportive evidence of which is presented in the previous section. Assumption two is that unskilled labor is more mobile than skilled labor. This assumption is based on findings in the literature investigating labor mobility. That labor mobility is negatively predicted by education is an empirical finding already shown for the US by Mincer and Jovanovic (1981). This finding is confirmed by Kambourov and Manovskii (2009), who use US data from 1968-1993 to argue that human capital is occupation specific. Using German social security records from 1999-2008, Neffke *et al.* (2017) report that workers in high-income segments switch industries less often than those in low-income segments. Furthermore, if high-income workers do switch industries, they tend to switch to industries that are closely related to their origin industry. In summary, there is ample evidence based on data from the US and Germany, that higher education results in less labor mobility in the sense of sectoral switches.

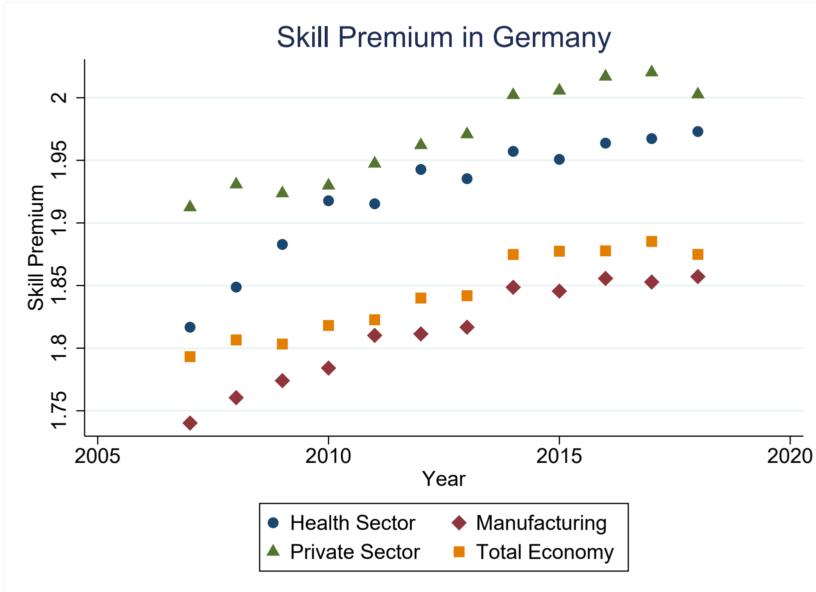
The Case of Germany

In this section, we test if there indeed was a reallocation of unskilled labor to the healthcare sector, focusing on the case of Germany. Data for this analysis is taken from the German Statistical Office.[20] Optimally, we would like to investigate data spanning the period 2003-2018, such that it is the same as the period over which the preference parameter θ is estimated. However, data is only available as far back as 2007, reducing statistical power.

In the statistic, it is distinguished between five skill levels. For the purpose of this analysis, the two top skill levels are aggregated into a high-skilled group, with the remaining three skill levels aggregated into a low-skilled group. The high-skilled group comprises workers in management positions and specialized positions who have graduated from college and/or have many years of experience and expert knowledge. This definition of high-skilled labor is in line with occupation-specific human capital accumulation found to make employment switches across sectors less likely.

There is barely a change in the share of skilled and unskilled labor in Germany from 2007 to 2018, displayed in Figure A.2. The share of high-skilled labor in the healthcare sector decreased from 40.1% in 2007 to 37.0% in 2018. This is the first indicative evidence of an increase in low-skilled labor in the healthcare sector, as predicted by the model. In

[20] https://www.statistischebibliothek.de/mir/receive/DESerie_mods_00000301



This figure provides a graphical illustration of the skill premia paid in the healthcare sector, the manufacturing sector, the private sector, and the overall economy, respectively. Data is taken from the German Statistical Office for the years 2007-2018.

Figure 1.2: *Skill Premium in Germany in Different Sectors and the Overall Economy*

the overall economy, the share of high-skilled labor increased slightly from 35.8% in 2007 to 36.0% in 2018. Given these small changes, direct analysis of employment shares by skill level is unlikely to yield meaningful results. Instead of measuring the reallocation of different kinds of labor into or out of the healthcare sector, we measure labor mobility indirectly via a skill premium. If the skill premium in one sector increases, it indicates that unskilled labor increases by more than skilled labor, relative to the respective demand for the different kinds of labor in that sector. One advantage of using this measure is that aggregate data is sufficient to investigate the relative mobility of labor rather than requiring individual-level data. A second advantage is that it measures supply relative to the demand for the two kinds of labor, which makes the measure robust to potential structural changes and trends, such as an overall increased supply of skilled labor. The model predicts that the skill premium increased by more in the healthcare sector than in the rest of the economy, which is captured by the parameter ϕ in Section 1.2.4.

Figure 1.2 displays the skill premium paid to high-skilled employees in Germany in

Table 1.4: *Estimating the Time Trends of Skill Premia*

	Total	Private	Manu	Health
Year	0.00903***	0.0105***	0.0109***	0.0127***
	(10.96)	(10.24)	(12.23)	(7.61)
Constant	-16.32***	-19.15***	-20.12***	-23.66***
	(-9.84)	(-9.28)	(-11.22)	(-7.03)
R^2	0.92	0.91	0.94	0.85
Observations	12	12	12	12

Note: Dependent variable is the skill premium in the overall economy, the private sector, the manufacturing sector, and the health sector, respectively. Results are obtained using German employment Data from 2007-2018. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics in parentheses.

different sectors. There is a general increase in the skill premium from 2007 to 2018 in Germany, illustrated by the squares in Figure 1.2. While all sectors considered experience an increase in the skill premium, the increase is fastest in the healthcare sector, illustrated by the dots in Figure 1.2. These results are indicative of a reallocation of unskilled labor to the healthcare sector, as predicted by the model.[21]

By separately regressing the skill premium for total employment and employment in different sectors on a time variable, it can be tested if there is a statistical difference between the skill premium increase in the different sectors. The regression results are reported in Table 1.4. As foreshadowed by the graphical illustration, the time trend for the skill premium is the steepest in the healthcare sector. A Wald-test of similarity indicates that the null hypothesis of similar trends can be rejected at a $p - value = 0.06$ for total employment. The difference between the time trend in the healthcare sector and the private sector, and the healthcare sector and the manufacturing sector is not statistically significant, with respective p-values of $p - value_{Private} = 0.30$ and $p - value_{Manu} = 0.37$.

The lack of statistical significance between the healthcare sector and the other two sectors may be owing to the short period with available data. It may also be due to the highly

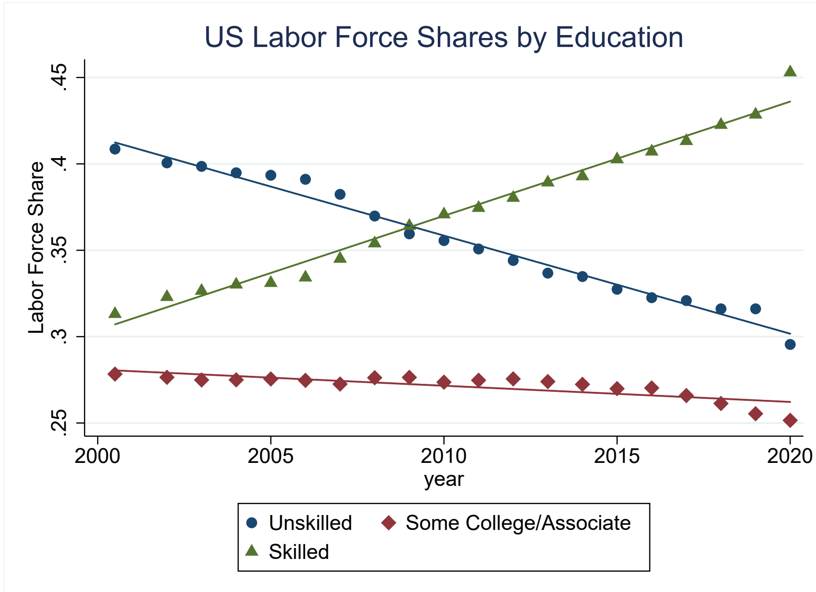
[21] All the results presented in this section are calculated based on employment numbers for Germany. There are some particularities with employment in the healthcare sector in Germany, none of which pose a threat to our identification strategy. For details, see Appendix A.2.2.

regulated labor market in Germany. While providing protection for workers, it reduces the flexibility with which any sector can react to changes in labor demand. The reallocation of unskilled labor, which in principle could easily switch into the healthcare sector to help meet increased demand for healthcare, is thus inhibited by the strong German labor protection laws. This is likely to reduce the expected skill premium increase in the healthcare sector in Germany and works against finding a statistically significant difference between the healthcare sector and other sectors. Both aspects work against finding a statistically significant difference in the time trends of skill premia. That we find (partially) statistically significant results despite these caveats emphasizes the relevance of the model's implications.

Extending the Analysis to the USA

In this section, we investigate if there was a reallocation of unskilled labor towards the health sector in the US. The purpose of this section is twofold. One, by replicating findings regarding the skill premium found for Germany using US data, the relevance and plausibility of the model is once more demonstrated. Both the healthcare system and the labor market regulation in the US are very different from the ones in Germany. Showing the specific pattern in the skill premium to hold in two distinct countries makes external validity and general applicability of the model likely. Two, the US data covers a longer period and there is a larger variation in skill shares of employment over time, rendering analyses of changes in the share of unskilled labor meaningful. First, we check if a reallocation of unskilled labor towards the health sector took place. The testable implication is that the share of unskilled labor rose faster in the health sector than in the rest of the economy. Second and as discussed before, we analyze if the skill premium increased by more in the health sector than in the rest of the economy, resulting from a reallocation of unskilled labor towards the health sector.

Each year, the US Bureau of Labor Statistics releases wage data for different education levels in the whole US economy. According to the data, the share of unskilled labor (measured as the share of workers with a high school degree or less) decreased from 39.9%



This figure provides a graphical illustration of the trend in employment shares of three different skill groups. It is based on data from the US Bureau of Labor Statistics. Workers with a high school education or less and no professional training are classified to be unskilled and workers with at least a Bachelor's degree are classified to be skilled.

Figure 1.3: *Skilled and Unskilled Workers*

to 31.6% from 2003 to 2018, a decrease of 21%. At the same time, the share of skilled labor increased from 32.7% to 42.3%. Workers with some college experience or an associate's degree are not included in either group, as it is unclear which category they belong to. The share of that in-between education group is rather large, on average making up 27% of the labor force. However, this share stays quite constant over time, changing from 27% to 26% between 2003 and 2018. The change over time for each skill group is depicted in Figure 1.3. Contrary to the case of Germany, there is a considerable trend in the shares of differently skilled labor in the total labor force. Overall, unskilled labor decreased, accompanied by a simultaneous increase in skilled labor across all sectors of the US economy.

The statistics cited in the previous paragraph clearly show an increasing time trend in the share of skilled labor across all sectors. To analyze how the share of skilled labor and the skill premium paid changed over the same time within the health sector, a different data set from the US Bureau of Labor Statistics has to be employed, which reports employment and wage statistics for different occupational groups.[22] To distinguish between skilled and unskilled labor, the occupational group "Healthcare practitioners and technical occupations" is compared to the "Healthcare support occupations". The employment share of the unskilled group decreased from 34.2% to 32.3% in the health sector. Thus the share of unskilled employment decreased from 2003 to 2018 by 6% in the health sector, which is much lower than the 21% recorded for the overall economy. The shares of unskilled labor in the overall economy and the health sector are displayed in the first row of Table 1.5. The respective growth rates are reported in the second row. The fact that the share of unskilled labor decreased by less in the health sector than in the rest of the economy is in line with the prediction made by the model given $\theta < 1$.

Next, it is informative to compare the change in the skill premium between the overall economy and the health sector. As noted before, the share of skilled labor increased in the overall economy in the period 2003-2018. At the same time, the skill premium of college graduates relative to workers with a high school diploma or less increased from 1.87 to 1.91 or by 2.1% across all sectors.[23] When adding those workers with some college or an associate's degree to the unskilled labor forces, thus comparing college graduates to all other workers, the considered changes are of similar magnitude.[24] In the health sector, the skill premium increased from 2.11 to 2.23 or by 5.7% in the same period.[25] So while the skill premium increased in both the overall economy and the health sector, the increase was stronger in the health sector. The skill premium as well as its growth rate in the overall economy and the health sector are displayed in the third and fourth row of Table 1.5, respectively. The fact that the skill premium increased by more in the health sector than in the overall economy is in line with the model prediction for $\theta < 1$.

Instead of looking at the share of unskilled labor and the skill premium in each sector

[22]The data is taken from `https://www.bls.gov/oes/current/oes_nat.htm`.

[23]The skill premium is measured at the median of the respective education group's wage distribution.

[24]The skill premium of college graduates relative to all other workers changed from 1.70 to 1.76 or by 3.5%. The share of unskilled labor, including all but college graduates, changed from 67.3% to 57.7% or by 14%.

[25]Again, the skill premium is measured for median wages.

Table 1.5: *US Labor Force Changes 2003-2018*

	Overall Economy		Health Sector	
	2003	2018	2003	2018
Unskilled Labor Force	39.9%	31.6%	34.2%	32.3%
Δ		-21%		-6%
Skill Premium	1.87	1.91	2.11	2.23
Δ		+2.1%		+5.7%

Note: Calculations based on data from the US Bureau of Labor Statistics.

Table 1.6: *Ratios of Key Indicators*

	2003	2018
Unskilled Labor Force Ratio	0.857	1.022
Skill Premium Ratio	1.128	1.168

Note: Calculations based on Table 1.5, which summarizes data from the US Bureau of Labor Statistics. The ratios implicitly account for time trends and compositional changes in the labor force.

separately, the ratio of these indicators can be constructed for each year. While this may be a less intuitive measure, it has the advantage of being unaffected by overall time trends. Specifically, the ratio is constructed as $Ratio_{measure,t} = \frac{\text{measure}_{health,t}}{\text{measure}_{overall,t}}$, with measure referring either to the share of the unskilled labor force or the skill premium. Table 1.6 displays the ratio of unskilled labor and the skill premium in the health sector relative to the overall economy for 2003 and 2018. For example, the skill premium in the health sector was 1.128 times larger than the skill premium in the overall economy in 2003. In 2018, it was 1.168 times larger in the health sector than in the overall economy. This indicates that the skill premium increased faster in the health sector than in the overall economy. The same is true for the ratio of unskilled labor force shares. By comparing the ratios across time, time trends and overall compositional changes in the labor force are implicitly accounted for.

Overall, there is supportive evidence for the testable model implication of a factor reallocation and ensuing changes in factor remuneration, summarized in Tables 1.5 and 1.6. Taking the time trends for both the unskilled labor share and skill premium into account, the development in the health sector of both measures is in line with the model predictions if $\theta < 1$.

One drawback of the US data used here is that the health sector of course is included in the data on the overall economy. The limited data availability prohibits the direct comparison between the overall economy excluding the health sector and the health sector. It implies that the actual difference between the two groups is larger than identified in the imperfect data, which works against finding any differences between the two groups compared here.

To facilitate the comparison of the results concerning factor reallocation in German and US data, Table 1.5 and Table 1.6 are replicated using the German data from Section 1.3.2. The tables can be found in Appendix Section A.2.3 as Table A.1 and Table A.2. Compared to the case of Germany, the increase in the share of unskilled labor in the health sector relative to the overall economy is more pronounced in the US data. In contrast, the increase in the skill premium paid in the health sector relative to the overall economy is a bit smaller in the US than in the German data. The overall pattern of a more-than-average increase in both the unskilled labor share and the skill premium in the health sector is present in both German and US data. This is remarkable, given the very different labor markets, especially with regard to labor protection laws, and healthcare systems in the two countries.

1.4 Conclusion

Spending on healthcare as a share of GDP has steadily increased for at least 50 years across 39 countries with available data. Employment in the health sector has mirrored the increase in spending, documenting a reallocation of labor towards the health sector. These two phenomena have been extensively studied by economists, and different explanations for the "excess growth" have been proposed and analyzed (see Getzen (2016) for a review of the literature). In the quest for explanations, the focus has been on macroeconomic variables like income per capita.

The health sector and its increasing share in GDP are often associated with Baumol's cost disease, a phrase based on Baumol (1967). We construct a micro-founded theory that can rationalize the empirical findings and provides us with additional testable implications that can be evaluated using available data.

We show that if the level of productivity increases in one sector relative to the other, this gives rise to a substitution effect and an income effect. The substitution effect entails that more resources flow into the more productive sector, whereas the income effect encompasses the opposite. Which of the two effects dominates depends on the elasticity of substitution between health and non-health consumption. If the elasticity of substitution between health and non-health goods is less than one, i.e., $\theta \in (0, 1)$, for which we provide empirical evidence, higher productivity growth in one sector relative to the other leads to an outflow of the flexible production factor from the more productive sector. Moreover, this can potentially increase the share of the less productive sector in terms of nominal GDP. Therefore, unequal productivity growth increases the relative price of the good produced in the relatively less productive sector and leads to a reallocation of production factors from the relatively more productive sector to the relatively less productive sector. This is in line with Baumol's cost disease. However, in this case, the term "cost disease" might be misplaced because the outcome, i.e., a reallocation of production factors from the more productive sector to the less productive sector, is optimal from the perspective of a representative utility-maximizing household. Of course, this does not necessarily mean that spending an ever-larger fraction of income and production factors on healthcare is always optimal from a welfare perspective. Nonetheless, our model highlights that the intuition that reallocating production factors from the relatively more productive sector to the relatively less productive sector is inefficient or constitutes a "disease", as it will lower overall physical output, is not a priori correct and thus does not directly warrant government intervention.

Chapter 2

Income Inequality and Aggregate Demand

2.1 Introduction

It is a well-documented fact that income inequality has increased over the last decades, as discussed for example by Piketty and Goldhammer (2014) and Saez and Zucman (2020). The 2007-2008 financial crisis and its broader economic ramifications made income inequality a topic of public interest. This was manifested for example in the Occupy Wall Street movement in 2011, which forced politicians to confront the issue. In a speech given at the White House, Krueger (2012), then Chairman of the President's Council of Economic Advisers, argued that a redistribution of income could boost aggregate demand. He invokes the idea of a dwindling Middle Class harming aggregate demand and sees a possible "latent pressure" on aggregate demand, caused by income inequality. The notion that income inequality affects aggregate demand is not new. There is a large literature analyzing how income inequality relates to economic growth, mostly finding a negative relationship (see, for example, Persson and Tabellini (1994) Murphy *et al.* (1989), and Berg *et al.* (2012)). The mechanism put forward by economic theory is that variations in consumption patterns across income groups can influence the overall level of demand in the economy and through it economic growth.

This paper suggests an additional channel through which income inequality affects aggregate demand, namely by influencing its composition. In a first step, the relationship between income inequality and aggregate demand is investigated empirically, using US state-level expenditure data from 1997-2018. The results indicate, that income inequality and aggregated personal consumption expenditures are negatively correlated. The US data covers not only personal consumption expenditure aggregates but also reports consumption expenditures at a more disaggregated level. Analyzing the subcategories of consumption expenditures shows that the negative effect between income inequality and aggregate demand is solely driven by demand for services, which in the aggregate even overcompensates a positive effect of income inequality on goods consumption. Both the negative relationship between income inequality and aggregate demand and the distinctive pattern emerging from analyzing demand subcategories is also present in German EVS data from 2003 and 2018. While the former finding confirms the intuition voiced for example by Krueger (2012), the finding of the robust, distinctive pattern in the relationship between income inequality and consumption subcategories is a novel empirical finding.

The empirical results reported in this paper suggest that income inequality affects both the level and the composition of aggregate demand. To gain a better understanding of the mechanism through which inequality affects demand for different consumption subcategories, a theoretical model is formulated. Following the example of Comin *et al.* (2021), a model featuring non-homothetic preferences over two types of goods is developed. The model abstracts from the effect income inequality has on the level of aggregate demand, but it is well suited to illustrate how income inequality changes the composition of aggregate demand. Specifically, it illustrates how the non-homotheticity of preferences opens up a demand-side channel through which income inequality affects aggregate demand, which can have an amplifying or ameliorating effect on income inequality. In addition, and depending on the income elasticity of the consumption categories services, durable, and non-durable goods, the model can explain why an increase in income inequality increases demand for goods but decreases demand for services. Specifically, for the model to explain the empirical finding, the income elasticity of services has to lie between those of durable and non-durable goods. In that case, households with decreasing income consume more non-durable goods and fewer services, whereas households with increasing income consume more durable goods and fewer services, resulting in the pattern observed in the data.

The model suggests that income inequality affects demand composition because income elasticities vary across consumption categories. In the next step, income elasticities of the consumption categories are estimated using US data and the approach proposed by Aguiar and Bils (2015) as well as German EVS data and the approach proposed by Comin *et al.* (2021). This is analogous to estimating the marginal propensity to spend on different consumption categories, which is constant across income groups. In that regard, the estimation approach is distinct from the one used in previous literature, which has focused on estimating marginal propensities to consume at different levels of income. Irrespective of the data used and the empirical strategy employed, the income elasticities are indeed estimated to increase from non-durable goods to services to durable goods. Thus, the estimated income elasticities are such that the model can explain the decreased demand for services and the increase in demand for goods, both durable and non-durable.

In the final step, the novel demand-side channel emerging from the model is quantified in an attempt to demonstrate its importance. To that end, it is analyzed if changes in aggregate consumption composition driven by income inequality reinforce or dampen wage inequality, which is measured by the skill premium. The results indicate that income inequality-driven changes in consumption composition ameliorate the original increase in income inequality, by increasing demand for goods produced in industries paying relatively low skill premia. The back-of-the-envelope calculation indicates that income inequality decreased by about 0.2 percentage points due to the dampening influence of changes in consumption composition for every percentage point increase in income inequality.

There is a large literature investigating how income inequality affects the level of aggregate demand. In their analysis of marginal propensities to consume, Fisher *et al.* (2020)

find that marginal propensities to consume differ systematically across wealth and income quintiles. They conclude that it is crucial to account for income and wealth distributions to calculate the effect of, for example, fiscal stimulus, and increases in income per capita in general, on aggregate expenditure. In a series of papers, Mian *et al.* (2020), Mian *et al.* (2021a), and Mian *et al.* (2021b) use non-homothetic preference to build models explaining how income inequality affects aggregate economic outcomes such as household borrowing, interest rates, and wealth inequality. The models and accompanying empirical findings highlight how high-income households, with greater purchasing power, allocate a larger share of their income to investments and savings, thereby dampening aggregate demand. In a similar vein, Corneo (2018) develops a simple microeconomic model to analyze the effect of increasing income inequality on aggregate demand. Other papers analyzing the inequality consumption nexus theoretically are Auclert and Rognlie (2017) and Bilbiie *et al.* (2022). Only a few papers are trying to estimate the effect of income inequality on aggregate demand at the macro level. Stockhammer and Wildauer (2016) fail to find a significant relationship in a sample of OECD countries. Crespo Cuaresma *et al.* (2018) regress average propensity to consume on income inequality and find, if anything, a positive relationship which they interpret as evidence against income inequality negatively affecting aggregate demand. In contrast, Brown (2004) does find a significantly negative relationship between consumption expenditures and income inequality. The estimates are derived using only US data and time series analysis.

There is ample and constantly increasing evidence for non-homotheticities in consumer preferences. For example, Straub (2019) finds an income elasticity of 0.7 over his preferred averaging period of nine years. The estimated income elasticity is well below one, the value to be expected in the case of homothetic preferences. The income elasticity estimate in this paper is slightly lower at 0.5, but given the shorter time horizon, it is even higher than what Straub (2019) finds for short time periods. Aguiar and Bils (2015) analyze to which extent consumption inequality mirrors income inequality and conclude that the relationship is quite strong. Their estimation approach relies on relative expenditures on necessities and luxuries, implying that they base their analysis on non-homothetic preferences. Comin *et al.* (2021) introduce a non-homothetic CES utility function and demonstrate how its parameters can be estimated. In a direct comparison of the same non-homothetic CES utility to standard homothetic CES utility, this paper demonstrates how to test for non-homotheticity empirically. The results indicate that consumer preferences are indeed non-homothetic.

This paper uses non-homothetic preferences to illustrate how income inequality affects aggregate demand. Intuitively, high-income households allocate a larger share of their income to luxury goods whereas lower-income households, facing limited resources, often prioritize necessities and essential goods. Changes in income inequality thus result in shifts in the composition of goods and services demanded, thereby impacting specific industries or sectors. There is a pertaining literature analyzing how changes in aggregate demand can impact other macroeconomic aggregates. These shifts can have broader economic ramifications, including the demand for differently skilled labor inputs. In the structural change literature, these aspects play an important role (see for example Boppart (2014), Cravino and Sotelo (2019), Comin *et al.* (2020) Comin *et al.* (2021), and Buera *et al.* (2022)). Furthermore, these studies often find that structural change is associated with changes in income inequality and in particular wage polarization (see Autor *et al.* (2005a), Autor *et al.* (2005b), Autor *et al.* (2006), and Bárány and Siegel (2018), all using US data). Goos and Manning (2007) show the same pattern for the UK. Spitz-Oener (2006) and Dustmann *et al.* (2009) show that this also holds for Germany, a country previously singled out to have the least wage polarization. The pertaining literature and shortcomings thereof are thoroughly discussed in Acemoglu and Autor (2010).

The literature review suggests that income inequality, aggregate demand, and consumption patterns are intricately linked factors that shape economic dynamics and outcomes. This paper demonstrates that a comprehensive understanding of the relationship between these elements requires accounting for the role of non-homothetic preferences – the idea that individuals' consumption patterns change with variations in their income levels. In the case of homothetic preferences, changes in income inequality do not affect aggregate demand.

Non-homothetic preferences introduce a crucial dimension to the study of income inequality and its impact on aggregate demand, as they affect not only the magnitude but also the composition of consumption across income groups. This in turn affects income inequality through the demand-side channel. The paper is structured as follows. In Section 2.2, the relationship between income inequality and aggregate demand as well as subcategories of consumption is estimated empirically. Section 2.3 introduces the model featuring both homothetic and non-homothetic preferences. Subsequently, estimated income elasticities of different consumption subcategories are presented in Section 2.4. Finally, the demand-side channel is quantified in Section 2.5. Section 2.6 concludes.

2.2 Estimating Aggregate Demand

The empirical analysis in this section examines the correlation between income inequality and aggregate demand, as well as various subcategories of consumption. Throughout the paper, empirical analysis is conducted using data from the US and Germany. Both are briefly described in the following. Subsequently, the empirical identification strategy is discussed and estimation results, which are derived from regressions using the two distinct data sets, are presented.

2.2.1 Data

To analyze the relationship between income inequality and aggregate demand, US state-level data from 1997-2018 provided by the Bureau of Economic Analysis (BEA) is combined with data from the World Inequality Database (WID).[1] The BEA provides information on personal consumption expenditures, both in absolute and in per capita terms. Total personal consumption expenditures are further broken down into 15 subcategories as classified in the

[1] The former data can be found on the website of the BEA (for a short link see `https://t.ly/BOzPa`), the latter on the website of the WID: `https://wid.world/country/usa/`.

National Income and Product Accounts (NIPA).[2] Additionally, data on income and income per capita can be obtained from the BEA. The focus throughout the analysis will be on per capita terms. As an inequality measure, the share of income going to the Top 10% of the income distribution is used. The data available at the WID is prepared and continually updated by Mark Frank, see Frank (2009). Importantly, it is calculated at the state level for each year, such that the income inequality measures varies across states and time.

The German data comes from the EVS, which is a triennial, repeated cross-sectional household-level survey conducted by the German Statistisches Bundesamt. It reports detailed consumption expenditures as well as socio-demographic information for roughly 40,000 households in each wave. For the analysis in this study, the 2003 and 2018 waves are used. The survey reports the Bundesland of residence for each household, as well as the quarter in which the data was collected. With that information, it is possible to construct a Bundesland-level panel by aggregating the household-level information at the Bundesland-Quarter level. This results in information on each Bundesland for a total of 8 quarters in 2003 and 2018, making panel estimation at the Bundesland-level possible.

2.2.2 Identification

The goal is to identify the effect income inequality has on personal consumption expenditure and its subcategories. With data available at the state level, state-fixed effects can be included in the regression. This is a first step in the direction of identification, as state-fixed effects act as a catch-all for omitted variables that are constant over time at the state level. Likewise, time-fixed effects are included to account for time variation which is constant across states, such as macroeconomic shocks.

The baseline estimation regresses consumption per capita measures on a variable measuring inequality, income per capita, state- and time-fixed effects, and an error term. Specifically,

[2] These categories are Services (further broken down into Food, Housing, Health, Insurance, Recreation, Transports, Other), Durable Goods (further broken down into Furnishing, Recreation, Vehicles and Other), and Non-durable Goods (broken down into Clothing, Food, Gasoline and Other).

the estimation equation is given by

$$log(PCE_{s,t}) = \alpha + \beta_1 \cdot Top10\%_{s,t} + \beta_2 \cdot log(Income_{s,t}) + FE_s + FE_t + \varepsilon_{s,t}, \tag{2.1}$$

where subscript s refers to state and subscript t refers to time, measured in years.

As dependent variable, aggregated personal consumption expenditure at the state level is used. In addition, subcategories of consumption, such as services and durable and non-durable goods, again aggregated at the state level, are used as dependent variables. The main explanatory variable is $Top10\%$, which measures the share of total income going to the Top 10% of the income distribution. Its effect on consumption expenditures is measured by the coefficient β_1. The variable $Top10\%$ is defined on the range $[0; 100]$. If the share of income going to the Top 10% increases by one percentage point, consumption is estimated to increase by $\hat{\beta_1}\%$.

For the estimation to yield any results, both the outcome variable and the inequality measure have to vary across states and time, such that the variation is picked up by neither fixed effect. For inequality to vary, the income distribution has to change. Thus variation in income inequality across states and time requires variation in the income distribution across states and time. By including income per capita as a control variable, the effect changes in the income distribution have on consumption via income inequality can be distinguished from all other potential effects changes in the income distribution have on consumption, but which are unrelated to the inequality measure. If for example income per capita in one state increases faster than in another state but the share of income going to the Top 10% of the income distribution remains unchanged in both states, this effect is picked up by the coefficient of income per capita, rather then incorrectly attributing any potential effect this has on consumption expenditure to changes in income inequality. Due to the log-log relationship of income per capita and the dependent variable, the coefficient β_2 can be interpreted as an income elasticity. If income per capita increases by 1%, consumption expenditure is estimated to increase by $\hat{\beta_2}\%$.

2.2.3 Baseline Results

Table 2.1 reports regression results obtained from estimation in US state-level data. Ceteris paribus, a one percentage point increase in the share of income going to the Top 10% is associated with a decline of personal consumption expenditures by 0.115%, as shown in column (1). In the US, the share of income going to the Top 10% increased steadily from 32.7% in 1970 to 50.5% in 2018. According to the estimation results, this increase was quantitatively accompanied by an estimated decline in personal consumption expenditures of 2.05%.

Decomposing consumption into services, durable and non-durable goods shows that the negative relationship between income inequality and consumption expenditure is entirely driven by services (see column (1), column (2) and column (3) of Table 2.1, respectively). Both durable and non-durable goods consumption is positively correlated with income inequality. Due to services making up a larger share of overall consumption, the negative effect of inequality on service consumption dominates the positive effect of inequality on goods consumption.[3]

Turning next to the effect of income on different expenditure categories, one obvious result is that the estimated coefficient $\hat{\beta_2}$, which resembles an income elasticity, is well below 1 for aggregate total personal consumption expenditure. This finding is based on aggregate data, such that it is not clear, that there is a direct correspondence between the coefficient $\hat{\beta_2}$ and individual income elasticities. Nevertheless, an estimated income elasticity well below one is in line with the central finding by Straub (2019), implying non-homotheticity of preferences. Interestingly, the income elasticity varies considerably across the broad consumption categories. Services have the lowest income elasticity at 0.374, followed by non-durable goods at 0.513 and durable goods at 0.963. This implies that the effect of an overall increase in income per capita will differ across consumption categories.

[3] On average, service consumption accounts for 65%, durable goods consumption for 12%, and non-durable goods consumption for 23% of total consumption. Over time, the share of service consumption increases, whereas the share of both types of goods consumption decreases. For a visualization, see Figure B.1 in the Appendix.

Table 2.1: *Aggregate Personal Consumption Expenditures and Inequality*

	log(PCE)	log(Services)	log(Durable)	log(Nondurable)
Top 10%	-0.115***	-0.279***	0.213***	0.337***
	(-3.54)	(-7.44)	(3.38)	(6.88)
log(Income pc)	0.488***	0.374***	0.963***	0.513***
	(33.51)	(22.30)	(34.25)	(23.43)
Time FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
R^2	0.99	0.99	0.94	0.98
Observations	1,144	1,144	1,144	1,144

Note: The dependent variables and income per capita are used as reported by the BEA at the US-state level, using data from 1997-2018. The variable Top10% reports the share of income going to the top 10% of the income distribution, as reported by Mark Frank. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics in parentheses.

2.2.4 Results using German Data

In this section, the robustness of the results reported in Table 2.1 is tested. This is done by repeating the baseline regression in a panel dataset constructed from the 2003 and 2018 waves of the German EVS. To that end, a Bundesland-quarter level panel is constructed from the 2003 and 2018 German EVS waves. The results from estimating Equation (2.1) in the EVS panel are reported in Table 2.2. Compared to the US state-level data, all estimated effects of an increase in income inequality on consumption are much larger in magnitude. The overall negative effect on personal consumption expenditures can be replicated in significance but is five times larger. In the German data, it is not just driven by the negative effect on services, but income inequality is also negatively correlated with the consumption of non-durable goods. The coefficient in the case of durable goods consumption is in this case also positive but insignificant.

The estimated income elasticities however are similar to those found in US data. This is true both for the size of the elasticities, as well as their ordering with respect to size. Again, services are estimated to have the lowest income elasticity, followed by non-durable goods and finally durable goods. Since this will be important in Section 2.4.2, note that the estimated income elasticities for services and non-durable goods are not significantly

Table 2.2: *Personal Consumption Expenditures and Inequality, EVS Data*

	log(PCE)	log(Services)	log(Durable)	log(Nondurable)
Top 10%	-0.578**	-0.758***	0.853	-0.764***
	(-2.52)	(-3.47)	(0.76)	(-3.37)
log(Income pc)	0.519***	0.511***	0.997**	0.586***
	(6.48)	(6.70)	(2.56)	(7.41)
Quarter FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
R^2	0.98	0.99	0.14	0.91
Observations	127	127	127	127

Note: All variables are based on the German EVS waves from 2003 and 2018. The raw data were used to construct a Bundesland-quarter level panel, which is used for estimation. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics in parentheses.

different from each other. A Hausman-style test results in a $p - value = 0.38$.

Overall, the negative relationship between income inequality and personal consumption expenditures found in US state-level data can be replicated in the German EVS data. In both cases, the effect seems to be driven by service consumption. Furthermore, the income elasticities of the consumption categories vary, indicating that the underlying preferences generating such demand patterns are non-homothetic.

2.3 Theory

The model considered in the following is static. Since there is no time dimension to the model, it abstracts from savings by households. Therefore it is assumed, that households spend all of their income, which consists solely of labor income. There is a mass $N = 1$ of infinitely lived households that are endowed with one unit of labor, which they supply inelastically. There are two types of households, which differ in their skill endowment $s \in \{l, h\}$. Let γ denote the share of the population with skill level $s = h$ and $(1 - \gamma)$ denote the share of the population with skill level $s = l$. Aggregating across all individuals, this yields $H = \int h^i di = h \cdot \gamma \cdot N = h \cdot \gamma$ and $L = \int l^i di = l \cdot (1 - \gamma) \cdot N = l \cdot (1 - \gamma)$. The share γ is assumed to be exogenously given and constant over time throughout the ensuing analysis. Likewise, the skill levels l and s are exogenously given and constant over time. Given different marginal products for the two labor inputs L and H, households potentially receive different levels of labor remuneration and thus income. There is no capital in the model.

2.3.1 Production

Each consumption good is produced by a different industry, all using a linear production technology. To simplify the exposition, the case of two competitive industries is considered.[4] The two industries $i \in \{1, 2\}$ produce the two different consumption goods C_1 and C_2. Profits in both industries are zero due to perfect competition.

Both industries employ labor, but Industry 1 uses only high-skilled labor H whereas Industry 2 uses only low-skilled labor L. Additionally, the two industries use technology A_i, which is assumed to differ across industries. The two production functions can be specified as

$$Y_1 = A_H \cdot H$$
$$Y_2 = A_L \cdot L.$$

The two kinds of labor receive their respective marginal product as remuneration. Let good 2 be the numeraire, and thus $p_2 \equiv 1$ and $p = \frac{p_1}{p_2}$ denote the relative price of good 1. Using the FOCs for the two kinds of labor input, the nominal wage rates paid in the two industries, in terms of the numeraire, can be expressed as the following ratio, which is equivalent to the skill premium:

$$\frac{w_H}{w_L} = \frac{A_H}{A_L} p. \tag{2.2}$$

The two wage rates and the skill premium depend on the two production technologies A_H and A_L and the equilibrium relative price p, which, among other things, depends on the

[4]The analysis can be extended to the case of I industries, complicating the analysis but not changing the nature of the results derived in the following.

relative supply of the two kinds of labor L and H. For simplicity, the supply of both kinds of labor is assumed to be constant. Therefore, the skill premium changes if either the relative production technology or relative prices change.[5]

Aggregated nominal income of each household group is given as

$$E_h = w_H \cdot H = Y_1 p$$
$$E_l = w_L \cdot L = Y_2.$$

Note that the income of both groups depends on the respective population share, as $Y_1(H) = Y_1(h \cdot \gamma)$ and $Y_2(L) = Y_2(l \cdot (1-\gamma))$. Without loss of generality, assume that $\frac{E_h}{\gamma} > \frac{E_l}{(1-\gamma)}$ throughout. It is equivalent to stating that income per capita is higher in the group of high skilled households than in the group of low skilled households. Aggregating income across the two household groups yields aggregate production

$$E_h + E_l = Y_1 p + Y_2.$$

2.3.2 Homothetic Preferences

For the benchmark case, all households, independent of their skill level s, are assumed to have the same CES-utility function. Thus, preferences are homothetic and independent of income levels. Specifically, let

$$\mathcal{U} = \left[\zeta_1^{\frac{1}{\sigma}} c_1^{\frac{\sigma-1}{\sigma}} + \zeta_2^{\frac{1}{\sigma}} c_2^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$$

describe the utility function of all households. The weight attached to each consumption good is given by ζ_i and σ denotes the elasticity of substitution between goods c_1 and c_2. The consumer's optimization problem can be set up as a maximization over a consumption

[5]Here, production is modeled to take place without capital. Alternatively, both industries could use capital, the total supply and allocation of which across industries is assumed to be constant in the short term. In that case, capitalists would be introduced to the model as a third type of household. Capitalists then supply capital to both industries and receive the return on capital, which they consume outside of the model economy. The production function of Industry 1 in that case is given by $Y_1 = A_H K_H^{\alpha} H^{1-\alpha}$ and that for Industry 2 by $Y_2 = A_L K_L^{\alpha} L^{1-\alpha}$. With capital supply and allocation fixed in the short run and capital returns irrelevant for aggregate consumption, the model dynamics are unchanged by the introduction of capital.

bundle (c_1^s, c_2^s), subject to the budget constraint $E_s = pc_1^s + c_2^s$, where E_s denotes total expenditure, given by $E_h = w_H \cdot H \cdot p$ and $E_l = w_L \cdot L$. All households face the same set of prices, which they take as given.

$$\max_{c_1^s, c_2^s} \mathcal{L} = \left[\zeta_1^{\frac{1}{\sigma}} (c_1^s)^{\frac{\sigma-1}{\sigma}} + \zeta_2^{\frac{1}{\sigma}} (c_2^s)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{1-\sigma}} - \lambda_s \left[E_s - pc_1^s - c_2^s\right] \quad s \in \{h, l\}$$

The first order conditions with respect to c_1^s and c_2^s are given by

$$\frac{\partial \mathcal{L}}{\partial c_1^s} = (\zeta_1)^{\frac{1}{\sigma}} \frac{\sigma - 1}{\sigma} (c_1^s)^{-\frac{1}{\sigma}} + \lambda_s p \overset{!}{=} 0$$
$$\frac{\partial \mathcal{L}}{\partial c_2^s} = (\zeta_2)^{\frac{1}{\sigma}} \frac{\sigma - 1}{\sigma} (c_2^s)^{-\frac{1}{\sigma}} + \lambda_s \overset{!}{=} 0.$$

Together with the budget constraint, the optimal ratio of consumption expenditure can be derived from the FOCs as

$$\frac{c_1^s}{c_2^s} = \frac{\zeta_1}{\zeta_2} p^{-\sigma}. \tag{2.3}$$

Note that the optimal ratio of consumption expenditure is independent of the type of household. The corresponding price index of one unit of utility is the same for both types of household and given by $\mathcal{P} = \left(\zeta_1 p^{1-\sigma} + \zeta_2\right)^{\frac{1}{1-\sigma}}$. The optimal demand for either good is given as

$$c_1^s = \zeta_1 p^{-\sigma} \cdot \frac{E_s}{\zeta_1 p^{1-\sigma} + \zeta_2}$$
$$c_2^s = \zeta_2 \cdot \frac{E_s}{\zeta_1 p^{1-\sigma} + \zeta_2}.$$

Testable Implications

The optimal ratio of consumption expressed in (2.3) can be used to derive a structural estimation equation, such that the ratio of preference parameters (ζ_1/ζ_2) and the elasticity of substitution σ can be estimated in consumption data. To do so, both sides of Equation (2.3) are multiplied by the relative price p to arrive at a ratio of expenditure shares:

$$\frac{\omega_1^s}{\omega_2^s} = \frac{\zeta_1}{\zeta_2} p^{1-\sigma}.$$

Taking the log, this yields an equation which can be estimated.

$$\log\left(\frac{\omega_1^s}{\omega_2^s}\right) = \log\left(\frac{\zeta_1}{\zeta_2}\right) + (1-\sigma)\log(p) \tag{2.4}$$

Equilibrium

In equilibrium, production factors are paid their marginal product, firms make zero profits, households maximize their utility, and the relative price p is such that the market for both consumption goods clears. Aggregate demand for both goods can be derived by summing demand across skill groups.

$$C_1 = \frac{\zeta_1 p^{-\sigma}}{(\zeta_1 p^{1-\sigma} + \zeta_2)}(E_h + E_l)$$
$$C_2 = \frac{\zeta_2}{(\zeta_1 p^{1-\sigma} + \zeta_2)}(E_h + E_l).$$

Market clearing requires that, for each good, demand be equal to supply

$$Y_1 = C_1 = \frac{\zeta_1 p^{-\sigma}}{(\zeta_1 p^{1-\sigma} + \zeta_2)}(E_h + E_l)$$
$$Y_2 = C_2 = \frac{\zeta_2}{(\zeta_1 p^{1-\sigma} + \zeta_2)}(E_h + E_l).$$

Note, that since preferences are homothetic, the optimal ratio of consumption is equal for both types of households. Furthermore, homothetic preferences imply that the optimal ratio of consumption at the aggregate level is independent of the aggregate level of income and the income distribution. It implies that

$$\frac{C_1}{C_2} = \frac{\zeta_1}{\zeta_2}p^{-\sigma},$$

where C_i denotes the aggregate level of consumption of good i. Therefore, the equilibrium condition can be stated as

$$\frac{Y_1}{Y_2} = \frac{\zeta_1}{\zeta_2}p^{-\sigma}. \tag{2.5}$$

Since output in both industries is a function of γ, with $Y_1(h\cdot\gamma)$ and $Y_2(l\cdot(1-\gamma))$, the equilibrium condition stated in (2.5) is an implicit function of γ. For comparative static

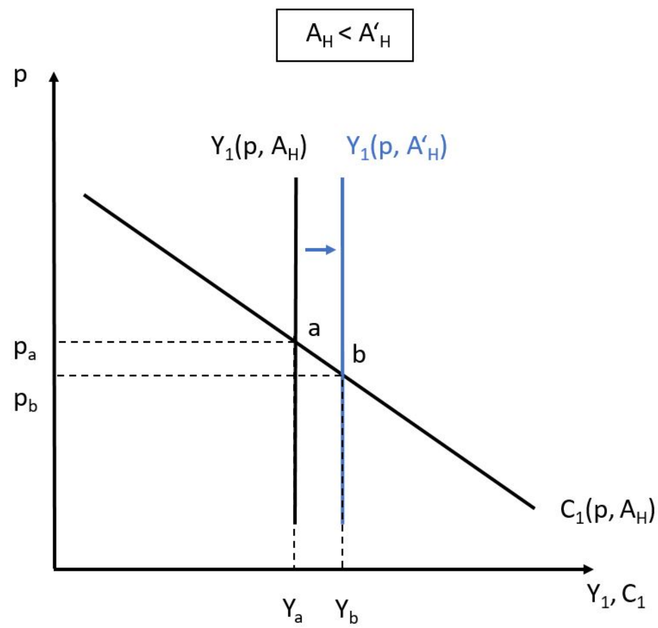


Figure 2.1: *An Exogenous Positive Increase in A_H*

analyses, the equilibrium can also be stated as a structural equation:

$$F \equiv \frac{A_H \cdot H}{A_L \cdot L} - \frac{\zeta_1}{\zeta_2} p^{-\sigma} = 0. \tag{2.6}$$

Comparative Static

Now consider an exogenous increase in A_H and how it affects different aspects of the equilibrium. Firstly, it affects the output of Industry 1, Y_1, which follows directly from the production function. Graphically, this is captured in Figure 2.1 by the shift of the supply curve of good 1 to the right. The output quantity increases from Y_a to Y_b.

Secondly, the increase in A_H affects equilibrium prices. Graphically, this is captured by the lower relative equilibrium price of good 1, p_b, in Figure 2.1. Analytically, this can be calculated using the implicit function theorem and Equation (2.6). For the derivation, see Appendix B.2.2. Specifically,

$$\frac{dp}{dA_H} = -\frac{\partial F/\partial A_H}{\partial F/\partial p} < 0.$$

Besides affecting the equilibrium quantity and price of good 1, the increase in A_H also affects the income distribution.[6] The effect can best be illustrated with the Gini coefficient. The Gini coefficient is a measure of inequality defined over the domain $G \in [0,1]$, with a value of 1 describing the most unequal distribution of income and a value of 0 describing a completely equal distribution of income. In the case of just two different groups, it is calculated as the (absolute) difference between the income share and the population share of the population group with higher per capita income. Thus, under the premise of $\frac{E_h}{\gamma} > \frac{E_l}{1-\gamma}$, the Gini coefficient can be calculated as

$$G = \frac{A_H H \cdot p}{A_H H \cdot p + A_L L} - \gamma.$$

An increase in A_H unequivocally decreases income inequality for a given γ, as a negative derivative demonstrates:

$$\frac{\partial G}{\partial A_H} = -\frac{H \cdot p \cdot A_L \cdot L(1-\sigma)}{(A_H H \cdot p + A_L L)^2} < 0.$$

The negative sign results from a negative effect of A_H on the income of high-skilled households.[7] This is driven by the price effect, which dominates the scale effect of A_H if and only if $\sigma < 1$ is assumed, which implies that c_1 and c_2 are complements. Evidence for complementarity and thus the assumption that $\sigma < 1$ is presented in Section 2.4.2. If instead $\frac{E_h}{\gamma} < \frac{E_l}{1-\gamma}$, such that the group of lows skilled workers has a higher income per capita, the increase in A_H increases income inequality.

The ratio of aggregate demand is independent of the income distribution because preferences are homothetic. Therefore, this decrease in inequality is irrelevant to aggregate demand. An exogenous increase in A_H thus only affects the equilibrium by changing the supply side, causing an adjustment in equilibrium prices.

[6]One example for how a shift in technology can affect income inequality is automation of labor, discussed for example by Acemoglu and Restrepo (2022b).

[7]For a derivation of the effect of A_H on income, see Appendix B.2.2.

2.3.3 Non-homothetic Preferences

Deviating from the benchmark case discussed in the previous section, in this section preferences are assumed to be of the non-homothetic CES type. This class of preferences goes back to work by Hanoch (1975) and Sato (1977), who noted that the standard CES function is a very restrictive way to describe preferences. By assuming that preferences are directly explicitly additive, the income effects of all goods are implicitly constrained to be equal to one. Introducing the notion of direct implicit additivity, Hanoch (1975) describes a class of preferences that still exhibit constant elasticity of substitution while allowing for non-constant income effects, resulting in non-homotheticity. Preferences are defined to be directly implicitly additive if the direct utility function is implicitly additive. This class of preferences, also referred to as Implicit CES (Matsuyama (2022)) is growing in popularity in economic research and accordingly has been used to study a variety of economic issues.[8]

Standard assumptions are put on the utility function $\mathcal{U}(\mathbf{c}, \mathcal{I})$, namely that it is continuously and monotonically increasing and concave in income denoted by $\mathcal{I}$, such that $\partial\mathcal{U}(\mathbf{c}, \mathcal{I})/\partial\mathcal{I} > 0$, $\partial^2\mathcal{U}(\mathbf{c}, \mathcal{I})/\partial\mathcal{I}^2 < 0$, and continuously and monotonically increasing in all consumption goods c_i, such that $\partial\mathcal{U}(\mathbf{c}, \mathcal{I})/\partial c_i > 0 \; \forall c_i \in \mathbf{c}$. Due to income effects differing across consumption goods, the utility of household type s can only be implicitly defined as

$$\sum_{i \in I} \left(U_s^{\varepsilon_i} \zeta_i\right)^{\frac{1}{\sigma}} \left(c_i^s\right)^{\frac{\sigma-1}{\sigma}} = 1, \quad i \in \{1, 2\}, \quad s \in (h, l). \tag{2.7}$$

Within the respective skill groups, households are assumed to be homogeneous. Equation (2.7) describes an indirect utility function that is already optimized. To see that, note that each summand in Equation (2.7) corresponds to the optimal expenditure share of the respective consumption good, denoted by ω_i^s. Each summand is therefore equivalent to the optimal expenditure share $\omega_i^s = \left(U_s^{\varepsilon_i} \zeta_i\right)^{\frac{1}{\sigma}} \left(c_i^s\right)^{\frac{\sigma-1}{\sigma}} \; \forall \; i \in I$. Anticipating the discussion on page 60 ff, note that a normalization of the preference parameters $\zeta_i, i \in \{1, 2\}$ and income elasticity parameters $\varepsilon_i, i \in \{1, 2\}$ renders utility as described by (2.7) cardinal. The

[8]For example, Bohr *et al.* (2021) study directed technical change, Comin *et al.* (2021) look at structural transformation and Fujiwara and Matsuyama (2022) analyze the effect of a technology-gap on premature deindustrialization.

same applies to the cost-of-living index denoted by $\mathcal{P}_s$ for each household type $s \in (h, l)$. With U_s and $\mathcal{P}_s$ cardinal, it follows that the utility level of household type s is given by total expenditures E_s divided by the price index $\mathcal{P}_s$, such that $U_s = \frac{E_s}{\mathcal{P}_s}$. Note the different notation used for the household type specific, and thus carrying a subscript, maximum attainable utility level U_s, and the general utility function $\mathcal{U}(\mathbf{c}, \mathcal{I})$. Only the former will be relevant for the ensuing analysis.

The difference between these preferences and standard homothetic preferences is the weight with which the different consumption goods enter utility. As in the homothetic benchmark case discussed above, the utility weight consists of $\zeta_i^{\frac{1}{\sigma}}$, which varies across consumption goods as indicated by its subscript i. In addition, the weight also consists of $U_s^{\varepsilon_i \cdot \frac{1}{\sigma}}$, which depends on the consumption good specific elasticity parameter ε_i and the utility level U_s, where U_s refers to the maximum utility level obtainable for given income level I and relative prices, as expressed by $p \equiv \frac{p_1}{p_2}$. If income elasticities vary across consumption goods, which is expressed by the subscript i, these preferences are non-homothetic. To see why, note that if the utility level increases, for example due to increased income, the relative weight of the different consumption goods changes if $\varepsilon_1 \neq \varepsilon_2$. As in the standard CES case, σ governs the elasticity of substitution between goods and is assumed to be constant. In the following, the analysis will center around a two-goods scenario. It can be extended to the I goods case, for which the same results can be derived.

The consumer's optimization problem can be set up as a maximization of the implicit utility as defined in (2.7) over a consumption bundle (c_1^s, c_2^s) subject to a standard budget constraint, where $E_s = pc_1^s + c_2^s$ denotes total expenditure and is given by $E_h = w_H \cdot H \cdot p$ and $E_l = w_L \cdot L$, respectively. As in the homothetic case, all households face the same set of prices, which they take as given.

$$\max_{c_1^s, c_2^s} \mathcal{L} = \left[(U_s^{\varepsilon_1} \zeta_1)^{\frac{1}{\sigma}} (c_1^s)^{\frac{\sigma-1}{\sigma}} + (U_s^{\varepsilon_2} \zeta_2)^{\frac{1}{\sigma}} (c_2^s)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{1-\sigma}} - \lambda_s \left[E_s - pc_1^s - c_2^s \right] \quad s \in \{h, l\}$$

The first order conditions with respect to c_1^s and c_2^s are given by

$$\frac{\partial \mathcal{L}}{\partial c_1^s} = (U_s^{\varepsilon_1}\zeta_1)^{\frac{1}{\sigma}}\frac{\sigma-1}{\sigma}(c_1^s)^{-\frac{1}{\sigma}} + \lambda_s p \overset{!}{=} 0$$
$$\frac{\partial \mathcal{L}}{\partial c_2^s} = (U_s^{\varepsilon_2}\zeta_2)^{\frac{1}{\sigma}}\frac{\sigma-1}{\sigma}(c_2^s)^{-\frac{1}{\sigma}} + \lambda_s \overset{!}{=} 0$$

Plugging $\omega_1^s = \left(U_s^{\varepsilon_1}\zeta_1\right)^{\frac{1}{\sigma}}\left(c_1^s\right)^{\frac{\sigma-1}{\sigma}}$ into $\frac{\partial \mathcal{L}}{\partial c_1^s}$ and rearranging yields

$$c_1^s p = \omega_1^s \frac{1}{\lambda_s}\frac{1-\sigma}{\sigma}$$

Note that in the two goods case $\omega_1^s + \omega_2^s = 1$ and $pc_1^s + c_2^s = E_s$. From this it follows that $E_s = \frac{1}{\lambda_s}\frac{1-\sigma}{\sigma}$. Plugging in and rearranging results in an expression for the Hicksian demand. This can be done analogously for good 2, such that the respective Hicksian demands are given by

$$\begin{aligned} c_1^s &= \zeta_1\left(\frac{E_s}{p}\right)^{\sigma} U_s^{\varepsilon_1} \\ c_2^s &= \zeta_2 E_s^{\sigma} U_s^{\varepsilon_2}. \end{aligned} \tag{2.8}$$

The price index for one unit of utility is given by $\mathcal{P}_s = \left(U_s^{\varepsilon_1}\zeta_1 p_1^{1-\sigma} + U_s^{\varepsilon_2}\zeta_2\right)^{\frac{1}{1-\sigma}}$. Note that as the ratio of optimal consumption depends on the household type $s \in \{h, l\}$, the price index is different for each household type.

The Marshallian demand for either consumption good can be derived by combining $\frac{\partial \mathcal{L}}{\partial c_1^s}$ and $\frac{\partial \mathcal{L}}{\partial c_2^s}$ and plugging into the budget constraint of the household. The second equality is derived using the definition of $\mathcal{P}_s$ and holds due to normalization of parameters, such that U_s and $\mathcal{P}_s$ are cardinal and $E_s = U_s \cdot \mathcal{P}_s$ holds.

$$\begin{aligned} c_1^s &= \frac{E_s}{\zeta_1 p^{1-\sigma}U_s^{\varepsilon_1} + \zeta_2 U_s^{\varepsilon_2}} \cdot \zeta_1 U_s^{\varepsilon_1} p^{-\sigma} = \zeta_1 U_s^{1+\varepsilon_1} E_s^{\sigma} p^{-\sigma} \\ c_2^s &= \frac{E_s}{\zeta_2 p^{1-\sigma}U_s^{\varepsilon_1} + \zeta_2 U_s^{\varepsilon_2}} \cdot \zeta_2 U_s^{\varepsilon_2} = \zeta_2 U_s^{1+\varepsilon_2} E_s^{\sigma} \end{aligned} \tag{2.9}$$

The optimal ratio of expenditure shares and consumption goods is given by

$$\frac{\omega_1^s}{\omega_2^s} = \frac{\zeta_1}{\zeta_2} p^{1-\sigma} U_s^{\varepsilon_1 - \varepsilon_2} \tag{2.10}$$

$$\frac{c_1^s}{c_2^s} = \frac{\zeta_1}{\zeta_2} p^{-\sigma} U_s^{\varepsilon_1 - \varepsilon_2}. \tag{2.11}$$

The ratios illustrate the two important features of this form of preferences: One, the constant parameter σ governs the elasticity of substitution between goods. Intuitively, if prices change at the same rate, this cancels out and relative demand is not affected. Only a change in relative prices affects relative demand, and the size and direction of that effect depend on σ, the elasticity of substitution. And two, the expenditure for good c_1 relative to good c_2 increases (decreases) as income and with it the overall utility level U_s increases, if and only if $\varepsilon_1 > \varepsilon_2$ ($\varepsilon_1 < \varepsilon_2$). Thus, goods can be ranked according to their income elasticity parameter ε_i from "most like a necessity" to "most like a luxury". The additional assumption of $\varepsilon_i > 0 \quad \forall i \in I$ guarantees that the absolute consumption level of all goods increases as the overall utility level U_s increases. Analytically, the change in the optimal ratio of goods consumption at the household level as utility level U_s changes is given by the derivative of (2.11) with respect to U_s:

$$\frac{\partial \left(c_1/c_2\right)}{\partial U_s} = \frac{\zeta_1}{\zeta_2} p^{-\sigma} (\varepsilon_1 - \varepsilon_2) \cdot U_s^{\varepsilon_1 - \varepsilon_2 - 1} \tag{2.12}$$

$$\frac{\partial \left(c_1/c_2\right)}{\partial U_s} = \begin{cases} > 0 \text{ if } \varepsilon_1 > \varepsilon_2 \\ < 0 \text{ if } \varepsilon_1 < \varepsilon_2 \end{cases}$$

Testable Implications

The preference structure given in (2.7) and the subsequent derivations can be used to derive a structural equation which facilitates testing for non-homotheticity of preferences in consumption data. In addition, the structural equation shows that without loss of generality, all preference parameters ζ_i and expenditure elasticity ε_i can be normalized by dividing by the preference parameter and expenditure elasticity of one good $i \in I = \{1, 2\}$.

The Hicksian demand expressed in (2.8) can be reformulated to arrive at an expression for U_s. For ease of exposition, this is done for consumption good c_2. Since good 2 is the numeraire good, $\omega_2^s = c_2^s / E_s$ holds. Reformulating yields

$$\varepsilon_2 \log(U_s) = log\left(\frac{\omega_2^s}{\zeta_2}\right) + (1-\sigma)\log(E_s). \tag{2.13}$$

Taking the log of the optimal expenditure share ratio in (2.10) and plugging in the expression for U_s derived in (2.13) results in an expression for the optimal expenditure ratio that is independent of the utility level U_s.

$$\log\left(\frac{\omega_1^s}{\omega_2^s}\right) = (1-\sigma)\log(p) + \left(\frac{\varepsilon_1-\varepsilon_2}{\varepsilon_2}\right)(1-\sigma)\log(E_s) + \left(\frac{\varepsilon_1-\varepsilon_2}{\varepsilon_2}\right)\log(\omega_2^s) - \\ - \left(\frac{\varepsilon_1-\varepsilon_2}{\varepsilon_2}\right)\log(\zeta_2) + \log\left(\frac{\zeta_1}{\zeta_2}\right) \tag{2.14}$$

The reformulation demonstrates that the optimal expenditure ratio depends on the relative size of the ε_is and ζ_is but not on their absolute value. Thus, any normalization of those parameters is an isoelastic transformation of the utility function and leads to obersvationally equivalent utility maximization outcomes. Therefore, let $\varepsilon_2 \equiv 1$ and $\zeta_2 \equiv 1$. This cardinalizes the utility function and the price index faced by households, such that $E_s = U_s \cdot P_s$ holds.

Equation (2.14) is used in Section 2.4.2 to estimate the income elasticity parameters of different consumption goods. The estimation equation is derived by using (for example) good 2 as a base good, such that $\varepsilon_2 \equiv 1$. To clarify notation, the price normalization is abandoned for this example. In that case, (2.14) simplifies to

$$\log\left(\frac{\omega_1^s}{\omega_b^s}\right) = (1-\sigma)\log\left(\frac{p_1}{p_b}\right) + (\varepsilon_1 - 1)(1-\sigma)\log\left(\frac{E}{p_b}\right) + (\varepsilon_1 - 1)\log(\omega_b^s) - \\ -(\varepsilon_1 - 1)\log(\zeta_b) + \log\left(\frac{\zeta_1}{\zeta_b}\right).$$

Except for the terms $\frac{\zeta_1}{\zeta_b}$ and ζ_b, which will be subsumed in an estimated constant, all terms consisting of parameters can in principle be estimated. Defining one consumption category to be the base category, the equation consists only of observable variables and can be estimated.

Estimating the empirical counterpart of (2.14) is informative in two respects. If the estimated coefficients of $\log\left(\frac{E}{p_b}\right)$ and $\log(\omega_b^s)$ are statistically significantly different from zero, it indicates that consumer preferences are indeed non-homothetic. This can be seen when comparing Equation (2.4), derived from the model featuring homothetic preferences, to Equation (2.14). In addition, the empirical estimates of the preference parameters ε_i and σ will be helpful to determine if the model is consistent with the empirical findings reported in Section 2.2.

Equilibrium

The model is closed by requiring market clearing. This imposes equality of aggregate demand and aggregate supply in each industry and consumption category. The aggregation process of the demand side is more complex if, as is the case here, preferences are non-homothetic. Taking into account that households within a given skill group are homogeneous, aggregate demand for good i can be derived by summing Marshallian demand, as given by (2.9), across skill groups.

$$C_1 = c_1^h + c_1^l = \zeta_1 p^{-\sigma} \left(E_h^\sigma U_h^{1+\varepsilon_1} + E_l^\sigma U_l^{1+\varepsilon_1} \right)$$
$$C_2 = c_2^h + c_2^l = \zeta_2 \left(E_h^\sigma U_h^{1+\varepsilon_2} + E_l^\sigma U_l^{1+\varepsilon_2} \right)$$

From this, the ratio of aggregate demand for the two goods can be derived as

$$\frac{C_1}{C_2} = \frac{\zeta_1}{\zeta_2} p^{-\sigma} \frac{E_h^\sigma U_h^{1+\varepsilon_1} + E_l^\sigma U_l^{1+\varepsilon_1}}{E_h^\sigma U_h^{1+\varepsilon_2} + E_l^\sigma U_l^{1+\varepsilon_2}}.$$

In the case of non-homothetic preferences, the ratio of aggregate demand thus depends on the income and expenditure levels of both types of households $E_s,\ s \in (h,l)$, as well as their utility levels U_s and the expenditure elasticities $\varepsilon_i,\ i \in (1,2)$, besides the pure taste parameters ζ_i, and the price ratio. The price ratio is the slack parameter that adjusts to equalize aggregate demand and aggregate supply in the equilibrium.

In general, the ratio of aggregate demand for good 1 and good 2 resulting from non-homothetic preferences differs from the ratio of aggregate demand if preferences are homothetic. Indeed, the ratios coincide if and only if $U_h^{\varepsilon_1} = U_h^{\varepsilon_2} \cap U_l^{\varepsilon_1} = U_l^{\varepsilon_2}$ or $U_h = U_l$. This is equivalent to requiring either $\varepsilon_i = 0 \ \forall \ i \in (1,2)$, or $\varepsilon_1 = \varepsilon_2$, or $U_h = U_l$. The latter is equivalent to both household types receiving exactly the same income per capita and is therefore a special case which is unlikely to be given in reality. $\varepsilon_i = 0 \ \forall \ i \in (1,2)$ implies that the utility level enters the preferences with a power of zero, rendering preferences homothetic. $\varepsilon_1 = \varepsilon_2$ implies that the utility weight of all goods is independent of the utility level, which once again renders preferences homothetic.[9] Therefore, $\varepsilon_1 \neq \varepsilon_2$ in combination with $U_l \neq U_h$ is a sufficient condition for the ratio of aggregate demand given non-homothetic preferences to differ from the ratio of aggregate demand given homothetic preferences.

Market clearing requires that, for each good, demand be equal to supply

$$Y_1 = \zeta_1 p^{-\sigma} \left(E_h^\sigma U_h^{1+\varepsilon_1} + E_l^\sigma U_l^{1+\varepsilon_1} \right)$$
$$Y_2 = \zeta_2 \left(E_h^\sigma U_h^{1+\varepsilon_2} + E_l^\sigma U_l^{1+\varepsilon_2} \right).$$

These conditions for the two goods markets can be combined and expressed as a ratio, such that the equilibrium condition can a be stated as

$$\frac{Y_1}{Y_2} = \frac{\zeta_1}{\zeta_2} p^{-\sigma} \frac{E_h^\sigma U_h^{1+\varepsilon_1} + E_l^\sigma U_l^{1+\varepsilon_1}}{E_h^\sigma U_h^{1+\varepsilon_2} + E_l^\sigma U_l^{1+\varepsilon_2}}.$$

To facilitate comparative static analyses, the equilibrium can also be stated as a structural equation:

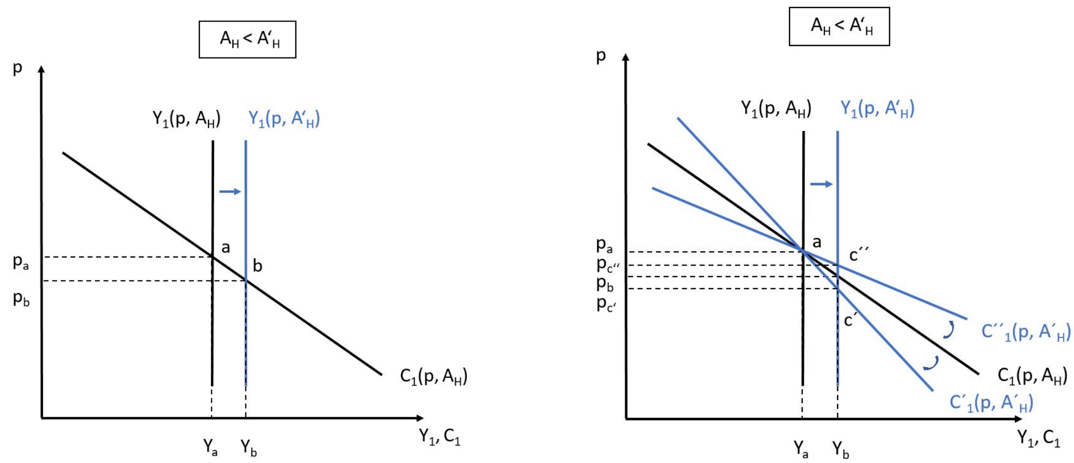
$$F \equiv \frac{A_H H}{A_L L} - \frac{\zeta_1}{\zeta_2} p^{-\sigma} \frac{E_h^\sigma U_h^{1+\varepsilon_1} + E_l^\sigma U_l^{1+\varepsilon_1}}{E_h^\sigma U_h^{1+\varepsilon_2} + E_l^\sigma U_l^{1+\varepsilon_2}} = 0. \tag{2.15}$$

When comparing the structural equations for the case of homothetic preferences (Equation (2.6)) and non-homothetic preferences (Equation (2.15)), it is obvious that comparative static analyses are more intricate in the non-homothetic case compared to the homothetic benchmark case. The reason is the income-dependent preference structure.

[9] In the more general I-good case, relative aggregate demand is unaffected by price changes if either $\varepsilon_i = 0 \ \forall \ i$ or $\varepsilon_i = \varepsilon_j \ \forall \ i \neq j$. The same logic applies, such that in both cases preferences are homothetic.

Comparative Static

Consider again an exogenous increase in production technology of Industry 1, A_H, which results in a higher output of Industry 1, Y_1. This is captured graphically in Subfigure 2.2a, which is equivalent to Figure 2.1. Again, the increased supply of Y_1 causes its relative price p to decrease, which is denoted by p_b in Subfigure 2.2a.



(a) *Supply Side Effect* **(b)** *Demand Side Effect*

Figure 2.2: *Decomposing the Supply Side Effect and the Demand Side Effect*

As in the case of homothetic preferences, the increase in A_H affects the income distribution. Specifically, income inequality is reduced by the increase in A_H. As preferences are now assumed to be non-homothetic, the change in the income distribution affects aggregate demand. This change in aggregate demand in turn affects equilibrium prices. This channel will in the following be referred to as the demand-side channel. It is illustrated graphically in Subfigure 2.2b.

In which direction aggregate demand is shifted by the non-homotheticity of preferences is a priori unclear. The sign of the effect depends on the derivative of the non-homothetic part of the structural equation pinning down equilibrium prices with respect to A_H. It also depends on the income and utility levels of the two types of households, as they both depend on A_H. It is thus determined by

$$\partial \left(\frac{E_h^\sigma U_h^{1+\varepsilon_1} + E_l^\sigma U_l^{1+\varepsilon_1}}{E_h^\sigma U_h^{1+\varepsilon_2} + E_l^\sigma U_l^{1+\varepsilon_2}} \right) / \partial A_H \lesseqgtr 0. \tag{2.16}$$

For the demand side channel to be inactive, it is required that the term is equal to zero. Trivially, this is the case if $\varepsilon_1 = \varepsilon_2$, as in that case aggregate demand is the same under homothetic and non-homothetic preferences. If however $\varepsilon_1 \neq \varepsilon_2$, the term is generally not zero, such that non-homotheticity of preferences is a sufficient condition for the demand-side channel to be active.[10]

Specifically, if the derivative has a positive sign, meaning that demand for good 1 increases relative to demand for good 2 if A_H increases, then $\partial F / \partial A_H$ is lower than if preferences are homothetic. In other words, the equilibrium price $p_{c''}$ given non-homothetic preferences is higher than the equilibrium price given homothetic preferences. As an example, consider an exogenous increase in A_H. The resulting increased output in Industry 1, Y_1, reduces the relative price of good 1. The ensuing change in relative prices leads to adjustments of relative demand. Under the innocuous assumption of $\frac{E_h}{\gamma} > \frac{E_l}{1-\gamma}$ and $\gamma = const$, the change in A_H and relative prices reduces the inequality between h-types and l-types by impacting E_h negatively. This leads to an additional change in aggregate demand due to the non-homotheticity of preferences, the size and direction of which is given by (2.16). If the derivative given in (2.16) is positive, the change in income inequality increases relative demand for good 1 more than proportionally. This is equivalent to an anticlockwise rotation of the demand curve, causing an additional upward price adjustment of the relative price p, such that the resulting equilibrium price $p \equiv p_1 / p_2$ is higher than the equilibrium price when preferences are homothetic. The negative effect of an increase in A_H on p is in that case ameliorated by the demand-side channel. This corresponds to the line C_1'' in Subfigure 2.2b. The opposite is true if the derivative in (2.16) has a negative sign, which corresponds to the line C_1' in Subfigure 2.2b.

To summarize, the essence of the model can be described as follows. Given a fixed

[10] For more detail and mathematical derivations, see Appendix B.2.2.

supply of input factors and non-homothetic preferences as described by Equation (2.7), an exogenous increase in A_H affects the equilibrium price p via two channels. On the one hand, an increased supply of Y_1 results in a decreased equilibrium price p. This is equivalent to the model dynamics if preferences are homothetic and illustrated in Subfigure 2.2a. On the other hand, the increase in A_H affects the income distribution and with it aggregate demand, which is particular to the model featuring non-homothetic preferences. The demand-side channel also affects the equilibrium price p, which is illustrated in Subfigure 2.2b by a rotation of the demand curve. The non-homothetic model thus demonstrates how changes in the income distribution affect aggregate demand. It also illustrates how aggregate demand affects equilibrium prices and with it income inequality, suggesting a feedback loop. Ultimately, the direction and size of the demand-side channel and how it affects income inequality is an empirical question.

2.3.4 Discussion

The purpose of the model described above is to point out the existence of a demand-side channel, which is active if preferences are non-homothetic. This is illustrated by comparing the model dynamics if preferences are non-homothetic to the model dynamics in the benchmark case with homothetic preferences. An active demand-side channel rotates the demand curve but does not affect the supply curve. Therefore, equilibrium prices and expenditure shares are different in the homothetic and non-homothetic models. The direction and magnitude of the demand curve rotation depend on the severity of the non-homotheticity, which in the model can be proxied by $(\varepsilon_1 - \varepsilon_2)$, the difference in income elasticities.

The crucial element for the demand-side channel to be active is a difference in income elasticities between different consumption goods, captured by $(\varepsilon_1 - \varepsilon_2) \neq 0$. However, the effect of the non-homotheticity also depends on the severity of income inequality. With uniformly distributed income, the effect of the non-homotheticity is minimized. An increase in income inequality increases the effect of the non-homotheticity on relative expenditure shares. Thus, the demand-side channel, which is equivalent to a rotation of the demand curve, is an increasing function of income inequality. From this insight, a further testable implication of the model can be derived, namely that the effect of income inequality on consumption categories is, in general, non-linear.

In the analysis above, it is considered how the model dynamics change in reaction to an increase in A_H. This is equivalent to a reduction in income inequality, given that the high-skilled households have higher per capita earnings than the low-skilled households before the increase in A_H. The demand-side channel goes in the opposite direction if the increase in A_H results in higher income inequality. How A_H affects income inequality ultimately depends on the population share of high-skilled workers γ, which for the analysis is held constant. For high levels of γ, the per capita income of high-skilled households can be lower than the per capita income of low-skilled households. In that case, an increase in A_H increases income inequality in the model.

As pointed out before, the juxtaposition of a model with homothetic preferences and a model with non-homothetic preferences allows to test for non-homotheticity of preferences directly. The preference parameters of both the homothetic model and the non-homothetic model can be estimated in suitable data. As the estimation equations derived from the models are quite similar (see Equation (2.4) and Equation (2.14)), statistical significance (or lack thereof) of the coefficients indicating non-homotheticity in the underlying preferences is informative as to which model is better equipped to describe the issue of interest in the real world.

2.4 Estimating Non-homotheticity

Before turning to the estimation of how changes in aggregate demand affect income inequality, the key part of the model, non-homotheticity of preferences, is tested empirically. In Comin *et al.* (2021), an empirical strategy for estimating non-homothetic CES from observable variables is developed. For the proposed estimation approach, data with variation across observations in both consumption quantities and prices is needed. Unfortunately, no data on prices at the state and disaggregation level corresponding to the BEA consumption data is publicly available, impeding such an analysis at the US state level.

Nevertheless, it is possible to estimate expenditure elasticities for the consumption subcategories reported in the BEA data by following the approach of Aguiar and Bils (2015). It can be used to see if expenditure elasticities vary across consumption categories, implying non-homothetic preferences, and results in crude estimates of the size of the different income elasticities.

In addition to estimating expenditure elasticities at the state level, this section also reports non-homotheticity parameters estimated in German EVS data. The aim of that exercise is threefold. One, if there is evidence for non-homotheticity in both US and German data, it emphasizes the necessity for using non-homothetic preferences to model economic relationships if consumer behavior plays a role. Two, the EVS data combined with disaggregated price data allows for the estimation of consumption good specific expenditure elasticities as proposed by Comin *et al.* (2021). And lastly, equipped with estimated expenditure elasticities of the consumption subcategories, the ability of the model proposed in Section 2.3 to explain the different correlations of subcategories of consumption with income inequality as described in Section 2.2 can be determined.

2.4.1 Expenditure Elasticities

Following the approach proposed by Aguiar and Bils (2015), expenditure elasticities for different consumption subcategories are estimated using the same state-level data as in Section 2.2.[11] By nature of expenditure elasticities, the sum of expenditure elasticities of consumption subcategories weighted by the respective expenditure shares of those subcategories is equal to the expenditure elasticity of the aggregate, which by definition is equal to one. The estimation is a log-linear approximation to Engel curves. As noted

[11]The proposed approach is changed only slightly to account for the fact that the data used here is not on the household but rather on the state level. Specifically, instead of using a good-time-fixed effect as originally proposed, a time-fixed effect is used instead. Additionally, state-fixed effects are used instead of a vector of demographic controls at the household level proposed in the original paper.

for example by Banks *et al.* (1997) and Battistin and Nadai (2015), it is required to include a quadratic term of expenditures in the estimation of Engel curves to arrive at unbiased estimates for expenditure shares. The goal of the estimation here is to infer one elasticity parameter for each consumption category. For the analysis at hand, it is not important how the expenditure elasticity may change along the income distribution. To simplify the interpretation of results it is therefore abstained from using a quadratic expenditure term in the estimation, such that the estimation equation is given by

$$\log(x_{sit}) - \log(\overline{x}_{it}) = \alpha_i + \beta_i \cdot \log(pce_{st}) + FE_{is} + FE_{it} + \varepsilon_{sit}, \qquad (2.17)$$

where x_{sit} is the consumption of good i in state s at time t and $\overline{x}_{it}$ is the average consumption of good i at time t across all states. Additionally, state- and time-fixed effects are included in the regression to account for state- or time-specific effects. As the regression is estimated independently for each consumption good category, the state- and time-fixed effects are allowed to vary across consumption categories, as indicated by the subscript i. The coefficient β_i represents the estimated expenditure elasticity for each consumption category.[12]

Table 2.3: *Estimated Expenditure Elasticities*

	log(Services)	log(Durable)	log(Nondurable)
log(pce)	0.902***	1.571***	0.925***
	(73.51)	(44.56)	(33.00)
Time FE	Yes	Yes	Yes
State FE	Yes	Yes	Yes
R^2	1.00	0.94	0.93
Observations	1,144	1,144	1,144

Note: SAEXP Data at the US state level from 1997-2018 is used for estimation. The dependent variable is given by as $\log(x_{sit}) - \log(\overline{x}_{it})$, where x_{sit} is the consumption good i in state s at time t and $\overline{x}_{it}$ is the average consumption of good i at time t across all states. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics in parentheses.

[12]In reality, the consumption of different categories is related. To take that into account, the three equations of services, durable and non-durable goods are additionally estimated in a Seemingly Unrelated Regression Estimation. Results are reported in Table B.4 in the Appendix. The regression results in exactly the same point estimates for the respective expenditure elasticities. The pertaining estimated confidence intervals are slightly wider, but statistical significance of all coefficients remains unchanged.

Results are reported in Table 2.3. The first observation is that the three consumption categories included in the table are estimated to have different expenditure elasticities. The finding of varying expenditure elasticities supports the modeling choice of non-homothetic preferences. Second, the estimated expenditure elasticities are much higher than the income elasticities reported in Table 2.1. Intuitively, income elasticities are affected by savings, which decrease the income elasticity of all consumption categories. Expenditure elasticities in contrast are unaffected by savings. If, as is the case here, savings are not the main focus of analysis, expenditure and income elasticities are equally informative, as they provide a ranking of consumption categories along the necessity-luxury spectrum. Indeed, the ranking of consumption categories is similar in Table 2.1 and Table 2.3, both suggesting $\varepsilon_{services} < \varepsilon_{nondurable} < \varepsilon_{durable}$. While in Table 2.1 estimates of income elasticities are reported, which are all lower than one, Table 2.3 reports expenditure elasticities, which naturally are higher than income elasticities due to excluding the issue of savings and how it affects consumption decisions. Therefore, a comparison of absolute size of income- and expenditure elasticities is not informative in this context.

In contrast to the income elasticities, the inequality between the expenditure elasticity of services and nondurable goods is quite weak in Table 2.3. Indeed, a Hausman-style test indicates that the two expenditure elasticities are not statistically different ($p - value = 0.65$). Upon further inspection, the slight difference in expenditure elasticity of services and non-durable goods suggested by the results reported in Table 2.3 is almost entirely driven by housing consumption. The estimated expenditure elasticities for all 15 consumption subcategories with available data are reported in Tables B.5, B.6, and B.7 in the Appendix.

Table 2.4 reports the results from running the same regression but using the household-level EVS data for estimation. The findings are quite similar to those in the US state-level data. However, in this case, the estimated expenditure elasticities of services and non-durable goods are significantly different, the estimated expenditure elasticity of non-durable goods being lower than that of services. A Hausman test of statistical difference reports a $p - value = 0.00$. The difference in results using aggregated and household-level data could

Table 2.4: *Estimated Expenditure Elasticities, German EVS data*

	log(Services)	log(Durable)	log(Nondurable)
log(pce)	0.876***	1.743***	0.637***
	(523.76)	(250.69)	(238.66)
Time FE	Yes	Yes	Yes
R^2	0.82	0.50	0.63
Observations	84,970	83,986	84,969

Note: German EVS data from 2003 and 2018 at the household level are used for estimation. The dependent variable is given by as $\log(x_{jit}) - \log(\overline{x}_{it})$, where x_{jit} is the consumption good i by household j at time t and $\overline{x}_{it}$ is the average consumption of good i at time t across all households. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics in parentheses.

be due to a systematic bias because of aggregation. To address the issue, the estimation is repeated using the panel constructed from the German EVS data. This leads to very similar results as when using the non-aggregated EVS data, which are reported in Table 2.5. The expenditure elasticities vary even stronger across the consumption categories than when using the raw EVS data. The difference between the service elasticity and non-durable goods elasticity is now larger in magnitude, but no longer statistically significant. That, however, is very likely due to the much-reduced sample size. The $p - value = 0.17$ is quite low, considering the small sample size.

2.4.2 Non-homotheticity in German Data

In the previous section, using US state-level data and German Bundesland-level as well as household-level data, it has already been shown that expenditure elasticities vary across the consumption categories services, durable and non-durable goods. The reported findings are indicative of non-homothetic preferences. In this section, relative income elasticities and the elasticity of substitution, which is constant across consumption categories, are estimated following the approach proposed by Comin *et al.* (2021). The estimation requires prices and consumption quantities to vary at the same level of observation, for example at the state level. Unfortunately, there is no data reporting consumer prices at the level of disaggregation needed to employ the proposed strategy at the US state level. Instead, the

Table 2.5: *Estimated Expenditure Elasticities, German EVS panel*

	log(Services)	log(Durable)	log(Nondurable)
log(pce)	0.859***	2.811***	0.377***
	(12.08)	(9.41)	(3.00)
Quarter FE	Yes	Yes	Yes
State FE	Yes	Yes	Yes
R^2	0.97	0.76	0.78
Observations	128	128	128

Note: All variables are based on the German EVS waves from 2003 and 2018. The dependent variable is given by as $\log(x_{sit}) - \log(\overline{x}_{it})$, where x_{sit} is the consumption good i in Bundesland s at time t and $\overline{x}_{it}$ is the average consumption of good i at time t across all Bundslander. The raw data were used to construct a Bundesland-quarter level panel, which is used for estimation. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics in parentheses.

parameters of interest are estimated in German EVS data from the 2003 and 2018 waves. The household expenditures reported in the German EVS data can be aggregated into the same consumption categories as the US state-level data.

Using the EVS data for estimation has two advantages. One, it overcomes the missing price data problem inherent in the US state-level data, such that it is feasible to estimate the structural estimation proposed by Comin *et al.* (2021). And two, it complements the findings reported in previous sections. For all relationships analyzed so far, similar results to the ones found in the US state-level data can be reported for the German EVS data as well. That the findings of interest can be found in both US state-level and German household-level data is reassuring. Being able to show that the patterns are present in both data sets speaks to their overall relevance and robustness.

Data

To facilitate the estimation, price data at the household level is necessary. This is achieved by merging the EVS data with official price data reported by the Statistisches Bundesamt. The official price data is derived from the EVS, which results in a perfect correspondence of price data and consumption data categories. The price data is available at a very fine disaggregation level. By taking into account how much of each consumption good a household consumes, the price at the disaggregation level of the 15 consumption categories discussed earlier varies at the household level.[13]

The estimation includes control variables at the household level. Specifically, the household size, age of the head of household, and the number of earners in the household are used. The household size dummy is constructed as follows: it takes on the value of 1 if the household size is smaller than three, the value of 2 if the household size is between 3 and 4, and the value of 3 if there are more than 5 household members. The dummy reporting the number of earners takes on the values of zero, one, and two, where two includes all households which have at least two earners.

Estimation Strategy

For estimation, the strategy developed in Comin *et al.* (2021) is used. They demonstrate how the relative income elasticity and constant elasticity of substitution across consumption categories can be estimated in a structural equation. It is equivalent to Equation (2.14) derived in Section 2.3.3.

The substitution parameter σ as well as the income elasticity parameters ε_i can be estimated using the following equation:

$$\log\left(\frac{\omega_{i,n}}{\omega_{b,n}}\right) = (1-\sigma)\log\left(\frac{p_{i,n}}{p_{b,n}}\right) + (1-\sigma)(\varepsilon_i - 1)\log\left(\frac{E_n}{p_{b,n}}\right) + (\varepsilon_i - 1)\log(\omega_{b,n}) + \\ + \beta_i' X_n + v_{i,n}$$

$\omega_{i,n}$ is consumption category i's share of total consumption by household n and $\omega_{b,n}$ is the share of total consumption spend on the base consumption category b by household n. Likewise, $p_{i,n}$ denotes the price of consumption category i faced by household n and $p_{b,n}$ the price of the base consumption category b faced by household n. E_n denotes total expenditure on consumption by household n and X_n is a vector of household-specific characteristics,

[13] Figure B.2 in the Appendix illustrates the construction of household-level price data.

which are a dummy measuring the household size, the age of the head of household, and a dummy denoting the number of earners in the household. In addition, a year dummy is included to account for the fact that the data comes from two waves of the EVS. It is included to control for year-fixed effects.

The structural equation is estimated using a Generalized Method of Moments (GMM) estimator. It allows imposing constraints on the estimated coefficients, which makes the estimation feasible. Following Aguiar and Bils (2015) and Comin *et al.* (2021), total household expenditure is instrumented for by total household income and the quintile of the income distribution in which the household's income lies. This is done to minimize the effect measurement error has on overall household expenditure, which is calculated by aggregating all reported expenditures. Household income is determined in a separate survey question and is likely to be measured with less error. As household income is correlated with household expenditures, it provides a valid instrument for total expenditure, without the inherent measurement error.

Results

Estimation results are reported in Table 2.6. The estimation can be used to infer *relative* expenditure elasticities, relative referring to the expenditure elasticity of the base category b. In each column of Table 2.6, the estimation results from using the category which denominates the column as a base category are reported. If, for example, expenditures on non-durable goods are used as a base category, its expenditure elasticity is normalized to one, which results in the table entry $\varepsilon_{non-dur} - 1 = 0$. Relative to that, the expenditure elasticity of services is higher at $\varepsilon_{services} - 1 = 0.23$ and that of durable goods even higher at $\varepsilon_{durable} - 1 = 1.04$. The overall substitution parameter is estimated to be $\sigma = 0.29$, indicating that all goods are complements.

The estimation is carried out using each of the three broad consumption categories in turn as the base category. The three estimates of the substitution parameter are reasonably similar in size and, importantly, all indicate that the consumption categories are complements.

Table 2.6: *Estimating Income Elasticities in German EVS data*

	Non-durable	Services	Durable
σ	0.29	0.45	0.23
	[0.271, 0.316]	[0.431, 0.474]	[0.195, 0.265]
$\varepsilon_{non-dur} - 1$	0	-0.21	-0.81
		[-0.214, -0.200]	[-0.836, -0.782]
$\varepsilon_{services} - 1$	0.23	0	-0.46
	[0.220, 0.238]		[-0.483 -0.442]
$\varepsilon_{durable} - 1$	1.04	0.51	0
	[1.019, 1.070]	[0.495, 0.528]	

Note: Estimation in German EVS data from the 2003 and 2018 waves. Results are derived using a GMM estimator. 95%-confidence intervals are reported in brackets.

Furthermore, the ordering of the expenditure elasticities is consistent across the use of different base categories. Since only relative expenditure elasticities are estimated, variations in the size of the estimated expenditure elasticities are irrelevant.

Taken together, the results from estimating expenditure elasticities using two different estimation approaches in two different data sets indicate that expenditure elasticities vary across consumption categories. Furthermore, the expenditure elasticities can quite consistently be ranked to increase from non-durable goods, over services to durable goods consumption, such that $\varepsilon_{non-dur} < \varepsilon_{services} < \varepsilon_{durable}$. This suggests that the model proposed in Section 2.3 is consistent with the pattern found in Section 2.2. If income inequality increases, low-income households increase their relative consumption of non-durable goods and reduce their relative consumption of durable goods and services. High-income households instead increase their relative consumption of durable goods and reduce their relative consumption of non-durable goods and services if there is an increase in inequality. At the aggregate level, this results in increased consumption of both durable and non-durable goods and decreased consumption of services. Note that this analysis, as well as the model, abstracts from the effects increases in income inequality have on the level of aggregate consumption due to differences in the propensity to save. It is possible to include that channel in the model by treating saving as another consumption good with a high expenditure elasticity. This extension of the model is described in Appendix B.2.1.

2.4.3 Estimating Non-linearity of Income Inequality Effects

As discussed in Section 2.3.4, the model predicts that the magnitude of the demand-side channel depends on the severity of income inequality. From this, a testable implication arises, namely that income inequality has a non-linear effect on the expenditure shares of different consumption goods.

The most straightforward way to test the model implication is by running regressions similar to those specified in Equation (2.1). The only difference is, that in addition to a linear term of the inequality measure, the estimation equation includes an additional quadratic term of the inequality measure.

$$\begin{aligned} log(Exp.-share_{s,t}) = \alpha + \beta_1 \cdot Top10\%_{s,t} + \beta_3 \cdot (Top10\%_{s,t})^2 + \beta_2 \cdot log(Income_{s,t}) + \\ + FE_s + FE_t + \varepsilon_{s,t} \end{aligned} \tag{2.18}$$

The dependent variable in that case is the log of expenditure share for different consumption categories. The expenditure share is calculated as the expenditure of category i in state s and year t divided by total personal consumption expenditures in state s and year t. The model predicts that the estimated coefficient $\hat{\beta}_3$ is significantly different from zero.

The BEA data reports expenditures on 15 different subcategories of consumption. Regression results are reported in Tables B.8, B.9, and B.10 in the Appendix. In 14 out of 15 regressions, the estimated coefficient $\hat{\beta}_3$ is statistically different from zero, indicating that the relationship between expenditure shares and income inequality is indeed non-linear for all consumption categories, except housing. This finding is in line with the prediction made by the model discussed in Section 2.3.

2.5 Quantifying the Demand-side Channel

The model described in Section 2.3 illustrates how an exogenous increase in income inequality can change aggregate consumption. If the production of consumption goods differs with respect to the skill premium paid in the producing industries, these changes in aggregate consumption affect the economy-wide skill premium. The demand channel can therefore amplify or attenuate the exogenous shock to income inequality. Increases in income inequality have been well documented. In Section 2.2, the effects of rising income inequality on aggregate demand have been explored. This section in turn will analyze how the demand shift caused by increased income inequality affects the skill premium paid and thus wage inequality. The effect size and direction is a priori unclear. On the one hand, increased income inequality may shift demand towards sectors with a relatively low skill premium, thereby attenuating income inequality. On the other hand, increased inequality can reinforce income inequality if it shifts demand towards goods produced predominantly in sectors paying a high skill premium.

To estimate the size and direction of a demand-side channel on income inequality, information on the skill premium at different levels of aggregation is needed. First, wage inequality at the industry level, and second, wage inequality at the consumption category level has to be known or estimated. Additionally, the change in demand at the consumption good level due to an increase in income inequality has to be known. The last part has already been estimated in Section 2.2. The goal of this section is to calculate the wage inequality at the industry and consumption category levels. To do so, the skill premium at the industry level has to be aggregated first at the consumption category level and then at the economy-wide level. In the aggregation process, it is crucial to use appropriate weighting schemes.

2.5.1 Data

This section describes how different data sets are merged to arrive at a mapping of skill use at the industry level, where it is routinely recorded, to the consumption category level. It follows the approach proposed and described by Buera *et al.* (2022).

A reliable measure of wage inequality is the skill premium paid in different industries. Information on skill use at the industry level is made available by EU KLEMS for different countries and years.[14] In the following, data on US industries in 2008 is used for analysis. At the isic3 industry level, the employment- and wage share of three different skill levels in total industry employment is reported. The educational attainment of workers is classified into "University graduates" "Intermediate" and "No formal qualifications". The skill premium at the industry levels is constructed by dividing the wage share of university graduates by their employment share.

To map the skill use at the industry level to the consumption category level, consumption goods categories have to be matched to industry levels. The BEA provides a mapping of Personal Consumption Expenditures categories along NIPA lines to NAICS codes at the industry level.[15] The most recent such mapping is available for 2012. For example, all the industry sectors contributing to the final consumption category "Vehicles" and their respective input values are listed.

To merge the labor input data provided by EU KLEMS with industry output data provided by the BEA, isic3 codes have to be mapped to NAICS codes. While no official mapping between isic3 and NAICS codes exists, there is a clear correspondence in almost all cases. By matching the industry codes used in the EU KLEMS data to NAICS codes used in the BEA dataset, the use of high- medium- and low-skilled labor, the respective wage shares, and the resulting skill premium can, in principle, be calculated at the consumption good level.

[14] The data can be downloaded from `https://dataverse.nl/dataset.xhtml?persistentId=doi:10.34894/MGSB4H`

[15] Source: `https://www.bea.gov/industry/industry-underlying-estimates`

2.5.2 Weighting

In general, the production of the consumption categories considered requires input from different industries. Ideally, the skill premium at the consumption category level would be calculated as a weighted sum of the skill premium paid in the input industries, the weight consisting of the labor intensity of an input industry and its input share at the consumption good level. Unfortunately, EU KLEMS does not report labor intensity or overall employment numbers at the industry level, preventing the implementation of this first-best solution.

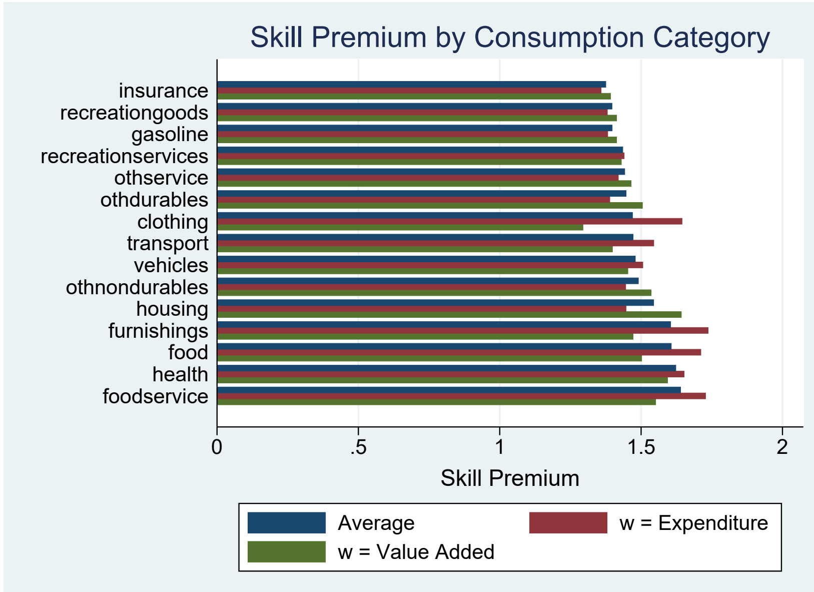
Still, there are two feasible approaches to aggregate the skill premium paid in input sectors to the consumption good level. One, the skill premium at the industry level can be weighted by that industry's share in overall input used for the production of a consumption category. In that case, the labor intensity at the industry level is disregarded in the calculation. Two, the skill premium at the industry level can be weighted by that industry's share of the total value added produced by all input industries. This approach yields a good approximation of the first best solution under the assumption that value-added and labor intensity are positively correlated. However, it does not take into account the share of input coming from the single industries at the consumption good level. As a last, additional, approach to calculating the skill premium at the consumption good level, the average of the two previously described skill premium measures can be taken at the consumption category level. In the absence of an economically meaningful guideline as to how to best calculate the average of the other two weights, both receive, somewhat arbitrarily, equal weight. In mathematical terms, the three calculations can be formalized as follows:

$$\text{skill premium}_{Input,i} = \sum_{j}^{J} \frac{input_{j,i}}{input_i} \cdot \text{skill premium}_j,$$

where $input_i$ is the sum of all inputs used to produce consumption category i and $input_{j,i}$ is the input of industry j used by consumption category i.

$$\text{skill premium}_{VA,i} = \sum_{j}^{J} \frac{VA_j}{VA_i} \cdot \text{skill premium}_j,$$

where VA_i is the sum of value added by all industries which are used to produce consump-



This figure provides a graphical illustration of skill premium associated with each consumption category considered. The red bars refer to the skill premium calculation using expenditure weights for the accumulation across industries. The green bars refer to skill premium calculation using value added weights for the accumulation across industries. The blue bars show the average of the two skill premium calculations at the consumption category level.

Figure 2.3: *Calculated Skill Premium in the US for Different Consumption Categories*

tion category i. Finally, taking the average results in

$$\text{skill premium}_{mean,i} = \frac{1}{2}\left(\text{skill premium}_{VA,i} + \text{skill premium}_{Input,i}\right).$$

Depending on the weighting scheme, the skill premium associated with the different consumption categories varies slightly. The mean across all consumption categories ranges from 1.47 to 1.52. The consumption category associated with the lowest skill premium according to the average measure is insurance, with a skill premium of 1.38, and the highest average skill premium is 1.64, paid for providing food services. The sensitivity to the weighting scheme used for aggregation is surprisingly low.[16] The three different measures of the skill premium at the consumption good category level are depicted in Figure 2.3.

[16]Using weights derived from value-added results in a coefficient of variation of 0.09, which is higher than that of the calculated skill premium when using input share weights, in which case it is 0.06.

2.5.3 Quantification

We are now equipped with a skill premium measure at the consumption category level. From Section 2.2, the marginal effect an increase in income inequality has on the demand for 15 different consumption categories is known. Combining those two statistics, it can be calculated how an exogenous increase in income inequality affects the economy-wide skill premium by aggregating the changes in demand for each consumption category.

The marginal effect increased consumption of a specific consumption category c_i has on the overall skill premium paid in the economy can now be calculated as

$$\frac{\partial \text{ Skill Premium}}{\partial c_i} \approx \text{skill premium}_i - \frac{\sum_{j \neq i}^{J} \text{skill premium}_j \cdot w_j}{\sum_{j \neq i}^{J} w_j} \tag{2.19}$$

where skill premium$_i$ refers to the skill premium paid in the production of consumption category i. Intuitively, the effect an increased consumption of good i has on the economy-wide skill premium depends on the difference between the skill premium in that category and the (weighted) skill premium in all other categories. The skill premia associated with the production of all other consumption categories should be weighted in the aggregation process, which is expressed by including the weights w_j in Equation 2.19. Analogously to before, two economically meaningful weights are worth considering. One, a consumption category's share of total consumption, and two, a consumption category's share of total value added. The first weight takes into account how important each category is for aggregate demand, reflecting the overall resources going into producing the goods in each consumption category. The second weight is likely to be a more specific measure of labor input. It is unclear which, if any, of the two weighting schemes is preferable. In any case, the weighting schemes are positively correlated with a correlation coefficient of $\rho = 0.24$. A third weighting scheme can be constructed by using the average of the two possible weighting schemes. This average weight has a correlation of $\rho = 0.81$ with the consumption share weight and a correlation of $\rho = 0.77$ with the value-added weight.[17]

[17] The three different weights are constructed as follows: $w_{Exp} = \frac{c_i}{\sum_i^J c_i}$, where c_i refers to the consumption of

The marginal effect an increase in income inequality has on demand for consumption category i has been estimated in Section 2.2. The overall effect an increase in income inequality has on the skill premium can thus be approximated by the following calculation:

$$\frac{\partial \text{ Skill Premium}}{\partial \text{ Inequaltiy}} \approx \sum_{i}^{I} \frac{\partial \text{ Skill Premium}}{\partial c_i} \cdot \frac{\partial c_i}{\partial \text{ Inequality}}$$

As discussed previously, two possible weighting schemes can be applied when summing across input industries and also consumption categories. Taking the average of these two weighting schemes generates a third weighting scheme. The estimates for the effect a change in aggregate demand has on the economy-wide skill premium using the three different weighting schemes are very similar and reported in Table 2.7. The estimated effect size of the demand side channel is in the range of $[-0.22; -0.20]$. The effect of an exogenous increase in income inequality by 1 percentage point, via the demand channel, thus is estimated to reduce the overall skill premium by 0.2 percentage points. Changes in aggregate demand hence attenuate changes in income inequality.

Table 2.7: *Results from Quantifying the Demand Side Channel*

	Weighting Scheme		
	Expenditure	Value Added	Average
$\frac{\partial \text{ Skill Premium}}{\partial \text{ Inequaltiy}}$	-0.200	-0.215	-0.199

Note: Data from EU KLEMS and the BEA are combined to estimated marginal effects of income inequality on consumption categories and, subsequently, the aggregated skill premium. The aggregation is carried out using three different weighting schemes; the result of each is reported in the thus named column.

The results of this back-of-the-envelope quantification suggest that the additional effect the demand side channel has on aggregate demand reduces the the effect of exogenous shifts in income inequality. Both the size and the direction of the effect seem reasonable. It would be surprising to find that the additional effect caused by the demand side channel is

category i; $w_{VA} = \frac{VA_i}{\sum_i^I VA_i}$, where VA_i is the value added produced by all industries contributing to producing good i; $w_{mean} = \frac{1}{2}\left(w_{Exp} + w_{VA}\right)$.

larger than the exogenous shock triggering the changes in aggregate demand. Therefore, an effect size smaller than one is in line with intuition. Regarding the direction of the effect, no theoretical prior exists. Depending on the expenditure elasticities and skill premia paid in the producing industries, an attenuation or amplification of the original shift is possible.

Based on the quantification results, industries in which a lower skill premium is paid apparently benefit from the shift in consumption correlated with increased income inequality. The changed consumption composition thus attenuates income inequality. From 1970 to 2018, income inequality increased by 17.8 percentage points in the US. This implies that the economy wide skill premium increased by 3.56 percentage points *less* due to the demand-side channel.

2.5.4 Discussion

The quantification done in the previous section can only be regarded as a first-order approximation of the potentially non-linear effect of income inequality on expenditures. Indeed, including a quadratic term of the income inequality measure in the regressions discussed in Section 2.2 suggests that the effect of income inequality on consumption expenditures is non-linear for almost all consumption categories, as indicated by highly significant coefficients of the quadratic term. This is visualized in Figure B.3 in the Appendix.

An increase in the inequality measure by 17.8 points cannot be considered a marginal increase. Therefore, using estimated marginal effects to calculate the effect inequality has on the skill premium through changes in consumption expenditures can only result in a crude approximation of the true effect. The aim of the analysis carried out above is to sense the order of magnitude that is plausible. So while the quantification exercise is unlikely to reveal the exact size of the demand side channel, it is nevertheless informative. Besides providing a first approximation of both the size and the direction of the true effect, it demonstrates the existence of the proposed demand-side channel and with that highlights a so far under-researched aspect of inequality.

2.6 Conclusion

In conclusion, this paper has delved into the intricate relationship between aggregate demand, non-homothetic preferences, and income inequality. By conducting different empirical analyses and using a model to on the one hand explain novel empirical findings and on the other hand derive further testable implications from it, the multifaceted dynamics and interdependencies among these factors were uncovered.

The analysis reveals that income inequality is related to both the magnitude and composition of aggregate demand. In the first step, the relationship between income inequality and aggregate demand is estimated empirically. There is evidence for the expected negative relationship. This addresses the first aspect of how income inequality affects aggregate demand if preferences are non-homothetic, namely that it influences the overall level of aggregate demand. As a byproduct of that analysis, the interesting pattern of the reduction in aggregate demand being exclusively due to decreased service consumption is detected. This addresses the second aspect of how income inequality can affect aggregate demand, namely by changing the composition of aggregate demand.

Focusing on that second aspect, a theoretical model is proposed featuring non-homothetic preferences and linking income inequality and aggregate demand. It can explain the finding of an unequal response to changes in income inequality across consumption categories. This is conditional on income elasticities increasing in a certain order, for which there is indeed evidence in US and German data. The model also illustrates how changes in aggregate demand can have profound implications for specific industries and sectors and, relatedly, income inequality. A change in the composition of demand can potentially amplify or ameliorate a first shock to inequality by affecting the average skill premium paid in the economy. This demand side channel emphasizes that income inequality and non-homothetic preferences should be considered not only in the context of their impact on aggregate demand but also in their role as potential drivers of structural change.

By combining different data sets, a back-of-the-envelope calculation to quantify the demand side channel is done. It suggests that a 1 percentage point increase in income inequality reduces wage inequality by 0.2 percentage points via the demand side channel. While the quantification is insightful and gives a first impression of the effect size and direction, its exact value is of secondary importance. The main purpose of the quantification exercise is to illustrate the existence of the demand side channel and to emphasize the conceptual contribution made in this paper by pointing out its existence in the first place.

Chapter 3

Demographic Change, Automation, and the Role of Education

3.1 Introduction

Demographic change and in particular population aging have put labor markets under pressure in the past decades. Japan is a cautionary example for the rest of the Western world, which is on the same trajectory as Japan was four decades ago.[1] According to data from the World Bank, the share of countries experiencing negative working-age population growth was 7% in 1990 but has increased to 34% in 2015.[2] Balakrishnan *et al.* (2015) calculate for the US economy that aging is responsible for 50% of the decline in the labor force participation rate from 2007-2013, a trend likely to continue. The concern with an aging population is that it reduces the labor force, thus impeding growth (for an overview of the literature on aging and economic growth see Bloom *et al.* (2010)). Kotschy and Sunde (2018) explore how the interplay of population aging and human capital accumulation affect economic growth, concluding that there is potential to offset the negative effects of population aging by increasing education levels. Another potential remedy to counteract the effects of population

[1] The Economist (05.12.2019): "Japan's economic troubles offer a glimpse of a sobering future"

[2] For an illustration of the trend using World Bank data, see Figure C.1 in the Appendix.

aging may be close at hand: Japan, the prime example of the adverse effects of a declining population on labor markets, is reported to have successfully invested in automation technology, thereby mitigating the negative impact of the population shrinking on economic growth.[3] This suggests increased automation as a potential solution for problems caused by labor shortages, especially in capital-rich countries facing a decreasing labor force (see also Acemoglu and Restrepo (2017)).

Automation of course is one of the most prevalent topics in the 21st century. However, there are clear limits to how well human labor can be substituted for by machines. It is generally agreed upon in previous reporting and research, that mainly low-skilled jobs are threatened by automation (e.g. Brynjolfsson and McAfee (2011) Frey and Osborne (2017), Nedelkoska and Quintini (2018), De Vries *et al.* (2020), Acemoglu and Loebbing (2022)). This literature emphasizes the role education and skills play in the discussion of automation potential. It also highlights a shortcoming in the literature analyzing the relationship between demographic change and automation. This shortcoming is a failure to account for the role of education.

This paper extends the existing theory relating population changes and automation by including education in a general equilibrium model. Subsequently, the comparative statics derived from the model are tested empirically. For the empirical part of the paper, a new, freely available measure for automation is constructed. It can facilitate further research contributing to the literature studying the effect of demographics on automation specifically, and automation more generally. There is of course a large literature on structural change and how it depends on human capital and thus education (see for example Teixeira and Queirós (2016), Cruz (2019), and Porzio *et al.* (2022)). However, papers in that literature do not analyze the interplay of demographic change, technology, and human capital. One exception is the working paper by Peralta and Gil (2022), who propose a theoretical model in which individuals choose education and fertility in the presence of automation. The paper analyzes how demographic change affects automation, human capital, and the skill

[3] The Economist (27.02.2013): "Doing more with less?"

premium, but not how education impacts the effect demographic change has on automation.

In independent papers, Abeliansky and Prettner (2021) and Acemoglu and Restrepo (2022a) study the effect of demographic change on automation. This paper differs from those papers in two aspects. First, this paper shows both theoretically and empirically that education has an important impact on the relationship between demographic change and automation. The theoretical prediction regarding the relationship between demographic change and automation changes drastically when extending the model proposed by Abeliansky and Prettner (2021) to include education.[4] While Acemoglu and Restrepo (2022a) acknowledge that education may influence the relationship between demographic change and automation, they do not explore the impact theoretically or empirically.

In previous papers, the relationship between automation and population growth has been found to be negative. The intuition behind that finding is that automation capital can substitute for unskilled labor as a production input. Intuitively, the labor of an unskilled worker at a production line can easily be replaced by an industrial robot. In the case of negative population growth, a shortage in the labor market can be compensated for by increased use of automation capital. In contrast, it is much harder to replace skilled labor with machines. Quite to the contrary, machines such as personal computers are likely to increase the output of skilled labor. In summary, automation capital generally acts as a substitute for unskilled labor and as a complement to skilled labor.

This intuitive understanding of the changing nature of automation capital, depending on the skill level of the labor force, is formalized in an analytically tractable model. The model is used to show the mechanism through which population growth affects automation capital and how that mechanism depends on the education level of the population.

In the case of negative population growth causing a shortage in unskilled labor, automation capital can help to ameliorate the negative effect of unskilled labor shortage on output. If instead negative population growth causes a shortage in skilled labor, the need

[4] In an extension of their model, Abeliansky and Prettner (2021) differentiate between two skill groups as input factors, but do not analyze the impact changes in education levels have on the relationship between automation and demographic change.

for automation capital decreases. The mathematical characterization of the mechanism highlights how the relationship between population growth and automation depends on the education level of the population. These results are derived under the assumption of an exogenously given level of education in the population and a fixed stock of capital for a given period, which can be allocated between automation uses and traditional uses within a period.

The second contribution of this paper is the combination of patent data with a classification of patents into automation and non-automation categories, thereby constructing a novel cross-country panel of an automation measure. This data set is subsequently used to test the new theoretical predictions. Acemoglu and Restrepo (2022a) also use patent data to measure innovation in automation technology but focus on a very narrow definition of automation patents.[5] Additionally, by relying solely on USPTO data, they only use a subsample of patents filed worldwide, diminishing the patent measure's reliability and representativity. As a more general caveat, their reported evidence comes from estimating long-time differences. As a consequence, they do not use fixed effects and rely on quite small sample sizes of 60 and 31 for their regressions. The first aspect raises questions about omitted variable bias and the second leads to statistically insignificant results.

In comparison, the novel cross-country panel used in this paper has two advantages. One, a much larger database of patents is used. Two, the definition of what constitutes an automation patent is broader and thus well suited to explore the relationship between demographic change and automation in general, instead of being limited to exploring the narrow relationship between demographic change and industrial robot utilization. This aspect seems especially relevant given the rapid advances and utilization of software in production processes. The validity and suitability of appropriately classified patent data as an automation measure are demonstrated by testing similar hypotheses on the relationship between demographic change and automation and arriving at the same results as when using data on industrial robot shipments provided by the International Federation of

[5] They only use patents classified as 901 under the USPTO as a measure for automation innovation.

Robotics.

This paper is related to the growing literature on automation and its economic effects (see e.g. Dechezleprêtre *et al.* (2019), Prettner and Bloom (2020), Krenz *et al.* (2021), and Mann and Püttmann (2023)). While the effect of automation on wages and employment of different skill levels has been studied before (e.g. Acemoglu and Restrepo (2018), Graetz and Michaels (2018)), to the best of my knowledge the effect of education on automation has not.

The existing literature relies heavily on data gathered by the International Federation of Robotics (IFR).[6] It reports the yearly delivery of "multipurpose manipulating industrial robots" for several countries, starting in 1993. This data set has two main drawbacks. First, the data only starts in 1993, and second, it can only be obtained for a high fee, possibly deterring some researchers from engaging with the automation topic. Another feature of the IFR data set is its uniqueness, which guarantees consistency in the automation measure used in the economic literature. On the one hand, this means that different studies and the findings therein can easily be compared, on the other hand, it inhibits testing for out-of-sample consistency of any findings. This paper contributes to the automation literature by making use of a novel and more comprehensive automation measurement.

The paper is structured as follows. In Section 3.2, the model proposed by Abeliansky and Prettner (2021) is extended to include education. The theoretical analysis suggests that a higher education level reduces the possibility to automate labor as a response to a decreasing labor force. Section 3.3 derives an estimation equation and details the construction of the new data set. Section 3.4 consists of three parts. In Section 3.4.1, results emphasizing the relevance of including education in any reduced form estimation analyzing the relationship between demographic change and automation are presented. The robustness of the results is tested along several lines in Section 3.4.2. Section 3.4.3 shows a replication of the analysis done by Abeliansky and Prettner (2021) using patent data to measure automation. It once again emphasizes the importance to account for education, and, by reproducing the

[6]For an exception see Dechezleprêtre *et al.* (2019) and Mann and Püttmann (2023), who use patent data.

original results, shows that patent data provides an apt measure of automation. Section 3.5 concludes.

3.2 Theory

This section outlines a neo-classical, general equilibrium model that illustrates a channel through which education affects the relationship between demographic change and automation. To do so, the standard neo-classical model is extended in two ways. One, there are two types of labor used to produce output, skilled labor and unskilled labor.[7] Importantly, the share of skilled labor in the labor force is assumed to be exogenously given and not determined by an endogenous choice of households. And two, there are two types of capital used for production, traditional capital and automation capital. Capital is assumed to be fully mobile between traditional uses and automation uses within a period. The model consists of two parts, an intertemporal utility maximization by a representative household, which pins down the capital stock available for production in each period, and an intratemporal output maximization by a representative firm, which takes the available capital stock as given. The intertemporal utility maximization problem is not specific to the novel channel proposed here, such that its discussion is relegated to the Appendix C.1.1. Its main contribution is to show that the equilibrium capital stock per capita available for production in each period is independent of population growth and, absent technological growth, constant over time. The rest of this section will focus on discussing the intratemporal maximization of output, given demographic change and the possibility to use capital for automation purposes, highlighting the role of education.

[7]This is the crucial aspect in which the model proposed here differs from the one put forward by Abeliansky and Prettner (2021). The model by Abeliansky and Prettner (2021) is a limit case of the model proposed here and discussed in more detail in Section 3.2.3.

3.2.1 Basic Assumptions

Time is discrete and indexed by $t = 0, 1, 2, ...$ In each period, labor services, capital services, and final output are traded. There is a continuum of infinitely lived households with mass N_t, who are endowed with one unit of labor each. Population grows at rate n_t between time t and time $t+1$. Households differ in their skill level $S \in \{L, H\}$, where $L_t = (1-e) \cdot N_t$ and $H_t = e \cdot N_t$, with $e \in [0; 1]$, refer to the unskilled labor force and the skilled labor force, respectively. Importantly, the share of educated population e is modeled to be exogenously given and constant across time, which is why it has no subscript t. Besides being endowed with labor, households also own all capital. Other than in their skill level, households are assumed to be identical. Households maximize their lifetime utility by choosing consumption and investment optimally, taking prices as given. The intertemporal utility maximization results in a constant equilibrium level of capital per capita $\tilde{k}_t$, which is owned by the households and available for production in each period.

3.2.2 Production

Firms operate under perfect competition, take prices as given and make zero profits in equilibrium. In the following, the actions of one representative firm are considered. Output is produced by combining traditional capital K, automation capital P, and skilled and unskilled labor $H = e \cdot N$ and $L = (1-e) \cdot N$, where e refers to the share of the skilled labor force and is constrained by $e \in [0; 1]$. Capital is assumed to be mobile between traditional uses and automation uses. Due to capital mobility, the overall capital stock $\tilde{K}_t$ is divided between automation uses P_t and traditional uses K_t such that output is maximized and $\tilde{K}_t = K_t + P_t$. The production function is assumed to be Cobb-Douglas, ensuring analytical tractability. Specifically, consider a constant returns to scale, nested Cobb-Douglas production function of the form:

$$F(K_t, P_t, N_t) = K_t^{\alpha} \left(\left((1-e) N_t + P_t \right)^{\beta} (eN_t)^{1-\beta} \right)^{1-\alpha}. \tag{3.1}$$

In the way automation capital P is introduced to the production function, it is a perfect substitute for unskilled labor $(1-e)N$ but acts as a complement to skilled labor eN. This modeling choice is justified for example by the findings presented in Griliches (1969), and, more recently, Krusell *et al.* (2000), Acemoglu and Restrepo (2020), and Prettner and Strulik (2020).

The firm's maximization problem is given by

$$\begin{aligned}\max_{K_t,P_t} \pi_t = \rho_t(Y_t - r_t^{trad}K_t - r_t^{auto}P_t - w_{H,t}(eN_t) - w_{L,t}(1-e)N_t)\\ s.t.\ Y_t = F(K_t, P_t, N_t, e)\\ \tilde{K}_t = K_t + P_t,\end{aligned} \tag{3.2}$$

where ρ_t refers to the market price of output Y_t, which is normalized to one in the following, such that $\rho \equiv 1$. $w_{H,t}$ refers to the wage rate of skilled labor, and $w_{L,t}$ refers to the wage rate of unskilled labor in period t. The firm takes ρ_t, $w_{H,t}$, $w_{L,t}$, $\tilde{K}_t$ and N_t as given and faces a static optimization problem. Therefore, time subscripts are dropped for the following analysis whenever possible.

The equilibrium wage rates w_H and w_L are given by the marginal product of the respective labor input, using the definition of $H = e \cdot N$ and $L = (1-e) \cdot N$.

$$w_H = \frac{\partial Y}{\partial H} = (1-\alpha)(1-\beta)\frac{Y}{H}$$
$$w_L = \frac{\partial Y}{\partial L} = (1-\alpha)\beta\frac{Y}{L+P}$$

$$\begin{aligned}\frac{w_H}{w_L} = \frac{1-\beta}{\beta} \cdot \frac{L+P}{H}\\ \frac{\partial(w_H/w_L)}{\partial P} = \frac{1-\beta}{\beta} \cdot \frac{L}{H} > 0\end{aligned} \tag{3.3}$$

The ratio of $\frac{w_H}{w_L}$ measures the skill premium paid to skilled labor. If automation capital P increases, the skill premium increases as well. So increased utilization of automation capital affects the skill premium in a similar way as skill-biased technological change does.[8]

[8]For an overview of the pertinent literature, see Violante (2008).

To determine how changes in the size of the labor force affect the optimal distribution of $\tilde{K}_t$ between K_t and P_t from the firm's point of view, the assumption of full mobility of capital is used. The return on automation capital P is given by its marginal product:

$$r^{auto} = \frac{\partial Y}{\partial P} = (1-\alpha)\beta \frac{Y}{(1-e)N+P}. \tag{3.4}$$

Likewise, the return on traditional capital K is given by its marginal product:

$$r^{trad} = \frac{\partial Y}{\partial K} = \alpha \frac{Y}{K}. \tag{3.5}$$

As capital is fully mobile between traditional and automation uses, the optimal allocation of $\tilde{K}$, which maximizes output, can be obtained by setting the marginal products of K_t and P_t equal and rearranging. let K^* denote the optimal amount of traditional capital, which can be derived by using the equality of marginal products and plugging in $P = \tilde{K} - K$.

$$K^* = \left((1-e)N + \tilde{K}\right) \frac{\alpha}{\alpha + (1-\alpha)\beta}$$

Analogously, this can be done for P^*, which denotes the optimal amount of automation capital, in which case $K = \tilde{K} - P$ is plugged into the equalized marginal products.

$$P^* = \tilde{K} \frac{(1-\alpha)\beta}{\alpha + (1-\alpha)\beta} - (1-e)N \frac{\alpha}{\alpha + (1-\alpha)\beta}$$

The maximum output obtainable given $\tilde{K}$ and N can be derived by plugging K^* and P^* into the production function given by (3.1):

$$Y^* = \left(\frac{\alpha}{\alpha + (1-\alpha)\beta}\right)^{\alpha} \left(\frac{(1-\alpha)\beta}{\alpha + (1-\alpha)\beta}\right)^{\beta(1-\alpha)} \left((1-e)N + \tilde{K}\right)^{\alpha + (1-\alpha)\beta} (eN)^{(1-\alpha)(1-\beta)}.$$

And finally, the equilibrium interest rate r^* can be derived by plugging either K^* or P^* into the respective marginal product and rearranging.

$$r^* = r^{trad} = r^{auto}$$
$$r^* = (\alpha + (1-\alpha)\beta)\,\frac{Y^*}{(1-e)N + \tilde{K}}$$

To see how an increase in population size N affects the two kinds of capital, consider the respective derivatives with respect to N:

$$\frac{\partial K^*}{\partial N} = (1-e)\frac{\alpha}{\alpha + (1-\alpha)\beta} \geq 0$$

$$\frac{\partial P^*}{\partial N} = -(1-e)\frac{\alpha}{\alpha + (1-\alpha)\beta} \leq 0.$$

The derivative of K^* with respect to N shows, that as the labor force N grows, more of total capital $\tilde{K}$ is used in traditional ways and not for automation purposes. The derivative of P^* with respect to N shows, that as the labor force N grows, less of total capital $\tilde{K}$ is used for automation purposes. This demonstrates clearly that as N increases, more capital is allocated towards traditional uses K and away from automation uses P. Furthermore, for $e = 0$ the size of the effect N has on the capital allocation is at its maximum, with the effect size decreasing as e increases.[9] For $e = 1$, the intra-period allocation of $\tilde{K}$ is independent of the population size N. This can also be demonstrated by looking at the ratio of K^* and P^* and its derivative with respect to N directly. Again, it is obvious that N affects the ratio most if $e = 0$ and the labor force is unskilled, whereas N does not affect the ratio if $e = 1$.

$$\frac{K^*}{P^*} = \frac{\alpha((1-e)N + \tilde{K})}{\tilde{K}(1-\alpha)\beta - \alpha(1-e)N}$$

[9]Mathematically, this can be demonstrated by looking at the limit cases of the partial derivatives of K and P with respect to N.

$$\left.\frac{\partial K}{\partial N}\right|_{e=0} > \left.\frac{\partial K}{\partial N}\right|_{e=1}$$
$$\left.\frac{\partial P}{\partial N}\right|_{e=0} < \left.\frac{\partial P}{\partial N}\right|_{e=1}$$

$$\frac{\partial(K^*/P^*)}{\partial N} = \frac{(1-e)\alpha\tilde{K}(\alpha+(1-\alpha)\beta)}{\left(\tilde{K}(1-\alpha)\beta-\alpha(1-e)N\right)^2} \geq 0$$

A change in the population size N affects the marginal products of the two types of capital, K and P, and the marginal product of the overall capital stock $\tilde{K}$. Specifically, N affects the marginal product of K and P in such a way, that it entails a reallocation of capital from automation uses towards traditional uses, as the positive sign of the derivative above demonstrates. The effect of N on the marginal product of the overall capital stock $\tilde{K}$ is universally positive. The respective derivatives are shown in Appendix C.1.2.

Note, that the marginal effect of N on $\frac{K^*}{P^*}$ is derived under the implicit assumption of $\frac{\partial \tilde{K}}{\partial N} = 0$. The optimal allocation of $\tilde{K}$ between traditional uses K and automation uses P considered here corresponds to a short-run output maximization problem, as specified in (3.2). Output is maximized in each period, taking population size N and the overall capital stock $\tilde{K}$ as given. From the intertemporal utility maximization of households discussed in Appendix C.1.1, a constant optimal per capita capital stock $\tilde{k}^*$ can be derived, which is independent of population growth. In general, this is not equivalent to a capital stock $\tilde{K}_t$ which is independent of population growth. However, if the variation in n_t (and hence also N_t) is unpredictable, which seems like a reasonable assumption, the capital stock in each period is independent of the unforeseen variation in the contemporaneous population growth and population size. Therefore, it is quite likely that $\tilde{K}_t$ does not react to unpredictable changes in n_t and treating $\tilde{K}_t$ as independent of unsystematic variation in n_t and hence also N_t is, after all, appropriate. In addition, empirical studies have found the capital stock to be quite slow in responding to shocks (see for example Ashraf *et al.* (2008), discussing the effect of demographic changes on capital accumulation). In light of that evidence, disregarding the marginal effect N_t has on $\tilde{K}_t$ is a mild assumption, given that in this model only the short run is considered. In any case, the results derived above carry through when taking into account that $\frac{\partial \tilde{K}_t}{\partial N_t} \neq 0$ under the assumption that $\tilde{K}_t > N_t$. From this, longer-run implications from the model can be derived. For a derivation and discussion of the results, see Appendix C.1.3.

Education is the new feature of the model and its effect on the return to automation capital is of special interest. The cross derivative of r^{auto} with respect to e and N is universally positive. This implies that education has the potential to mute any negative effects an increase in the population size has on the return on automation capital. The equivalent derivative for the return on traditional capital is always negative. Thus the difference in the effect population size has on the return on traditional- and automation capital diminishes as the share of the educated workforce e increases. The respective derivatives are shown in Appendix C.1.2. The effect of e on the return on the overall capital stock $\tilde{K}$ is universally positive:[10]

$$\begin{aligned}\frac{\partial r^*}{\partial e} &= (\alpha + (1-\alpha)\beta)\,\frac{1}{\left((1-e)N+\tilde{K}\right)^2}\left(\left((1-e)N+\tilde{K}\right)\cdot\frac{\partial Y^*}{\partial e} + Y^*\cdot N\right) = \\ &= (\alpha + (1-\alpha)\beta)\,\frac{1}{\left((1-e)N+\tilde{K}\right)^2}\left(\frac{1}{e}Y^*(N+\tilde{K})\right) > 0.\end{aligned}$$

3.2.3 Limit Cases of the Production Function

The elasticity of substitution between labor and automation capital depends on the skilled share of the labor force. This difference in complementarity results in a different sign of $\partial r^{auto}/\partial N$, as discussed in Appendix C.1.2, depending on whether the population is skilled or unskilled. In this setup, education plays a crucial role in determining the effect of population growth, and hence demographic change, on automation. To better understand how education influences the effect population size has on the incentive to automate, consider the two limit cases of $e = 0$, a fully unskilled labor force, and $e = 1$, a fully skilled labor force.

[10] The derivation makes use of $\frac{\partial Y}{\partial e} = Y \cdot \left(\frac{(1-\alpha)(1-\beta)(N+\tilde{K}) - eN}{e((1-e)N+\tilde{K})}\right)$.

Fully Unskilled Labor Force

With $e = 0$, the production function reduces to the one proposed by Abeliansky and Prettner (2021).

$$Y = K^{\alpha} (N + P)^{1-\alpha} \tag{3.6}$$

The marginal product of automation capital P and the marginal product of traditional capital K are given by

$$r^{auto} = \frac{\partial Y}{\partial P} = (1 - \alpha)\frac{Y}{N + P}$$

$$r^{trad} = \frac{\partial Y}{\partial K} = \alpha\frac{Y}{K}$$

Analogously to before, we can equate the marginal products due to full capital mobility and set $P = \tilde{K} - K$ and $K = \tilde{K} - P$ to derive the optimal levels of traditional capital K^*, and automation capital P^*. Plugging K^* and P^* into the production function specified in (3.6) gives the maximum output level obtainable for a given N and $\tilde{K}$.

$$K^* = \alpha(N + \tilde{K})$$

$$P^* = (1 - \alpha)\tilde{K} - \alpha N$$

$$Y^* = \alpha^{\alpha}(1 - \alpha)^{1-\alpha}(N + \tilde{K})$$

Finally, the equilibrium interest rate r^* can be derived by plugging either K^* or P^* into the respective marginal product and rearranging.

$$r^* = r^{trad} = r^{auto}$$

$$r^* = \frac{Y^*}{N + \tilde{K}}$$

To analyze how a change in N affects the optimal allocation of $\tilde{K}$ between K and P, consider the respective derivative of the optimal level with respect to N:

$$\frac{\partial K^*}{\partial N} = \alpha > 0$$
$$\frac{\partial P^*}{\partial N} = -\alpha < 0.$$

The same can be done for the ratio of optimal K^* to optimal P^*:[11]

$$\frac{K^*}{P^*} = \frac{\alpha(N + \tilde{K})}{(1-\alpha)\tilde{K} - \alpha N}$$
$$\frac{\partial (K^*/P^*)}{\partial N} = \frac{\tilde{K}}{\left((1-\alpha)\tilde{K} - \alpha N\right)^2} > 0.$$

Fully Skilled Labor Force

In the second case of $e = 1$, the whole population is educated and the elasticity of substitution between automation capital and skilled labor is equal to one.

$$Y = K^{\alpha}\left(P^{\beta} N^{1-\beta}\right)^{1-\alpha} \tag{3.7}$$

The marginal product of automation capital P and the marginal product of traditional capital K are given by

$$r^{auto} = \frac{\partial Y}{\partial P} = (1-\alpha)\beta\frac{Y}{P}$$

$$r^{trad} = \frac{\partial Y}{\partial K} = \alpha\frac{Y}{K}$$

Analogously to before, the optimal levels of K^* and P^* pinning down the optimal allocation of $\tilde{K}$ can be derived by using the full mobility of capital assumption and equating the

[11]As in Section 3.2.2, the derivative shown here disregards the effect of N on $\tilde{K}$. The reasoning for the approach is the same as above. Also, the result of $\partial(K^*/P^*)/\partial N > 0$ carries through when taking into account that $\partial\tilde{K}/\partial N \neq 0$ if $\tilde{K} > N$, as demonstrated in Appendix C.1.3.

marginal products. Plugging in $K = \tilde{K} - P$ and $P = \tilde{K} - K$ yields:

$$K^* = \frac{\alpha}{\alpha + (1-\alpha)\beta}\tilde{K}$$
$$P^* = \frac{(1-\alpha)\beta}{\alpha + (1-\alpha)\beta}\tilde{K}.$$

The maximum obtainable level of output for given N and $\tilde{K}$ is derived by plugging K^* and P^* into (3.7)

$$Y^* = \left(\frac{\alpha}{\alpha + (1-\alpha)\beta}\right)^{\alpha} \left(\frac{(1-\alpha)\beta}{\alpha + (1-\alpha)\beta}\right)^{\beta(1-\alpha)} \tilde{K}^{\alpha+(1-\alpha)\beta} N^{(1-\alpha)(1-\beta)}.$$

The equilibrium interest rate r^* is derived by plugging K^* or P^* into the respective marginal product.

$$r^* = r^{trad} = r^{auto}$$
$$r^* = (\alpha + (1-\alpha)\beta)\frac{Y^*}{\tilde{K}}$$

If the population is fully educated, skilled labor is a complementary input to both traditional capital and automation capital. An increase in the labor force due to population growth thus increases the return on automation- and traditional capital equally. The optimal ratio of traditional- and automation capital is in this case determined solely by exogenous parameters and thus independent of population size N.

$$\frac{\partial K^*}{\partial N} = 0$$
$$\frac{\partial P^*}{\partial N} = 0$$
$$\frac{\partial (K^*/P^*)}{\partial N} = 0$$

3.2.4 Population Growth

So far, it has been shown that the effect of population size on the return on automation capital depends on the education level of the labor force. The effect of population size on traditional capital and its return is always positive. The effect of population size on the return on automation capital however is positive if the labor force is skilled and negative if the labor force is unskilled. This is driven by automation capital acting as a complement to skilled labor and as a substitute for unskilled labor input. In the case of an unskilled labor force, capital is thus shifted from automation to traditional uses if the population level increases. In the case of a skilled labor force, no capital is shifted, since the marginal effects are equal in the optimum.

Next, consider how population growth affects automation capital per capita. Focusing on automation capital per capita has two advantages over considering automation capital levels. One, it is the more natural measure for cross-country analysis. And two, it is more closely linked to population growth, which is the main variable of interest when analyzing demographic change. Turning next to the effect of population growth, it can be shown that it potentially exerts two forces on automation capital per capita.

Define $y_t = \frac{Y_t}{N_t}$, $k_t = \frac{K_t}{N_t}$ and $p_t = \frac{P_t}{N_t}$. If the labor force is partially educated, output per capita is given by

$$y_t = k_t^{\alpha} \left((1 - e + p_t)^{\beta} e^{1-\beta} \right)^{1-\alpha}.$$

Again, the marginal products of the two kinds of capital have to be equal in the optimum. Equalizing the marginal products, plugging in $p = \tilde{k} - k$ or $k = \tilde{k} - p$, and rearranging, the optimal ratio of the two kinds of capital per capita can be derived as

$$\frac{k^*}{p^*} = \frac{\alpha(1 - e + \tilde{k})}{\tilde{k}(1-\alpha)\beta - \alpha(1-e)}, \tag{3.8}$$

where $\tilde{k} = \frac{\tilde{K}}{N}$ denotes the overall capital stock per capita available for production. In this intensive form formulation of the model, deviations in the population growth rate n from its balanced growth path value result in variation in $\tilde{k}$, which is constant on the balanced growth path. A negative deviation of n from its balanced growth path value leads to a positive deviation of $\tilde{k}$ from its balanced growth path value. This results in a decrease in the optimal ratio of $\frac{k^*}{p^*}$, as the negative sign of the following derivative demonstrates:

$$\frac{\partial(k^*/p^*)}{\partial \tilde{k}} = -\frac{\alpha(1-\alpha)(1-e)}{\left(\tilde{k}(1-\alpha)\beta - \alpha(1-e)\right)^2} \leq 0.$$

A negative deviation of n from its balanced growth path value thus leads to an increased share of overall available per capita capital stock $\tilde{k}$ to be allocated towards automation uses. It is obvious that an increase in the education level e ameliorates this effect. For a fully educated labor force and $e = 1$, a variation in n and $\tilde{k}$ does not affect the optimal ratio of $\frac{k}{p}$.

Intuitively, as population growth decreases, the labor force decreases which, if it is unskilled ($e = 0$), is a perfect substitute for automation capital, thereby increasing the marginal product of automation capital. As capital is allocated endogenously to traditional and automation uses, this results in a capital allocation towards automation uses.[12] If $e \in (0;1)$, population growth not only increases the substitute input for automation capital unskilled labor but also the complement input skilled labor. A higher e can therefore attenuate the positive effect lower population growth has on p. If $e = 1$, the ratio of $\frac{k}{p}$ is unaffected by an increase in n, as both kinds of capital are complements to skilled labor. For a derivation of the results and a separate discussion of the two limit cases of the production function in per capita terms, see Appendix C.1.4. Neither of the limit cases is relevant anywhere in the world. Therefore, it is important to take education levels into account when analyzing the relationship between population growth and automation capital.

3.3 Empirical Relevance

The empirical question revolves around the relationship between the incentive to automate and demographic change with a focus on how education influences that relationship. This section first describes the data used for empirical estimation in detail, with an emphasis on the newly created automation measure. In the second step, an estimation equation based on the theoretical analysis in the previous section is derived.

[12]The great advantage automation provides, is that growth can be generated simply by accumulating capital. The size of this additional growth opportunity is determined by how important automation capital is relative to traditional capital. By increasing automation capital, capital per capita can be deepened in a growth-enhancing manner. The size of p effectively measures how much use the economy makes of growth by accumulation. An increase in p thus increases the economic growth potential.

3.3.1 General Data Description

This section describes the data used to test the theoretical relationship described above empirically. If not indicated otherwise, five-year averages of all data are taken. Doing so reduces noise in the data and partially addresses timeliness concerns regarding patent filings. Given the time span of available data, this results in 9 periods of observation which can be used for estimation.

All data discussed in this paragraph is taken from the World Bank.[13] Information on both total population and population by age group is utilized. The latter is used to construct the working-age population, defined to be aged 20-64. For the regressions, the log of population growth is used as an explanatory variable. One inherent property of growth rates is, that they naturally and frequently take on negative values. To avoid the loss of many observations, the growth rates are transformed linearly by adding the absolute value of the smallest growth rate observed in the data to all observations before taking the log. This is equivalent to a linear rescaling of the variable and does not affect its correlation with any other variable, such that regression results are unaffected by the linear transformation. As a measure of savings, gross fixed capital formation as a share of GDP is used. For robustness checks, some additional variables are considered. GDP per capita is measured in 2015 US Dollars, the openness of the economy is calculated as the external balance on goods and services measured in percent of GDP, and the importance of the service sector is measured as the value added by the service sector in percent of GDP.

Education plays an important role in the theoretical predictions. Specifically, the effect of population growth on automation depends on the share of skilled labor in the economy. The education measure used comes from Barro and Lee.[14] It reports the population share with at least completed secondary education in 5-year intervals.

The core data used is patent data published by the OECD for 59 countries, starting in 1977 and ending in 2020. For a list of all countries with available patent data see Appendix

[13] The data is freely available at `https://databank.worldbank.org/`.

[14] It can be downloaded at `http://barrolee.com/`.

C.2.4. It is combined with the classification of patent categories into automation and non-automation categories developed by Dechezleprêtre *et al.* (2019) to arrive at a count of automation patents for each country-year observation with available data.

3.3.2 Patent Data

The vast majority of the literature analyzing the economic effects of automation uses data on industrial robots supplied by the IFR. In this paper, freely available patent data is used to measure automation instead. As this is fairly new, it is discussed in detail in the following.

Classifying Patent Data

Dechezleprêtre *et al.* (2019) use data from patents filed with the EPO to develop a classification of patent categories into automation and non-automation. Two different classifications are proposed. In each patent category, the share of patents described using automation keywords is calculated. The patent categories are then ordered by their share of automation patents. Two cutoff thresholds are considered to classify a patent category as an automation patent category. The stricter one defines all patent categories at or above the 95th percentile of the distribution of the automation patent shares as automation categories. The less strict one defines all patent categories at or above the 90th percentile of the automation patent shares as automation categories. This results in 5% or 10% of all patent categories being defined as automation patent categories. In a final step, all patents belonging to a thus-defined automation patent category are summed at the country-year level to arrive at a raw number of automation patents for each country-year observation. This results in the patent measures $auto95$ and $auto90$.

In addition to those two measures introduced by Dechezleprêtre *et al.* (2019), a third automation measure using patents is proposed here. By counting all patents belonging to a patent category with the highest share of automation patents, considerable noise may be introduced to the automation measures $auto90$ and $auto95$. The newly proposed measure addresses that concern. In the first step, the number of patents within each patent category is multiplied by the share of automation patents in that patent category. In the second step, the resulting number of automation patents belonging to different patent categories is then summed at the country-year level, resulting in one number of automation patents for each country-year observation. This measure is henceforth called $auto1$ and it is the preferred measure for automation patents.

Empirical Considerations

So far, empirical analyses have mainly used data on robots to measure a country's automation level. The IRF data provides information on the yearly installation of multipurpose industrial robots at the country level. Theoretically, using robot data has the advantage of directly measuring how much automation technology is employed. However, it is unclear, how long robots can operate, and at which point in time they are outdated or defunct. Thus, to estimate the stock of robots used in a country, assumptions about the service life of robots have to be made, which is complicated by a likely variation of the service life across time, due to differences in the pace of innovation, and variation in the service life across application areas of the robots. The alternative to using an inevitably noisy estimate of the robot stock is to focus on newly acquired robots. Such a measure will however vary strongly with business cycles, making averaging over several periods necessary and reducing the number of available data points.

This paper proposes an alternative measure of automation, namely automation patents. Conceptually, automation patents measure innovation in the realm of automation and provide an imperfect measure of a country's automation level, just like robots. Berkes *et al.* (2022) evoke the idea of patents as a means "to ensure that investments in new ideas can be recovered with future profits". With that concept in mind, a patent's economic value is equivalent to the present value of the innovation it is protecting. While the market value of patents is in general not known, the number of patents is a helpful, if not perfect, measure of the present value of the ideas protected by them. Automation patents, therefore, provide a measure of investment into research directed toward automation, the level of which is directly linked to the expected present value of such research. One determinant of the present value of automation patents is the demand for automation. In summary, automation patent data provides an alternative and potentially even better measure of the present value of automation in a country than the flow of industrial robots does.

One concern regarding the suitability of patent data as a measure of automation is that they measure ideas, which, contrary to robots, are mobile across countries. In extreme cases, countries may adopt and use automation technology prolifically without registering any automation patents themselves. For such countries, the use of automation capital in the production process is underestimated when relying on automation patents as a measure of automation capital. That, however, is unlikely to occur for two reasons. One, there is a large literature finding that the investment required to adopt foreign technology is similar to the investment required to generate new technology (see, for example, Cohen and Levinthal (1989), Griffith *et al.* (2004) and Aghion *et al.* (2009), p. 151 ff). Therefore, it is unlikely that a country is adopting automation technology without it also generating some automation patents at the same time. This may also be related to a second aspect found in empirical studies, namely that there is a considerable time lag between a technology's invention in one country and its adoption in other countries (see Comin and Hobijn (2010) and Comin and Mestieri (2018), who find a minimum lag of adoption of 5-8 and 7-12 years, respectively). Together, these aspects make it unlikely that the mobility of ideas causes a systematic measurement error or bias in the automation measure constructed using patent data.

Patent data has the great advantage of being widely available, very granular, and detailed. Because so many details are given, most disadvantages that are inherent to patent data can be addressed and possibly dispelled completely.[15] Patent data only provides an indirect measure of automation technology, which certainly is its main disadvantage. Many patents are filed, but only a few are applicable in the industry and thus of real economic value.

[15]The OECD provides an in-depth discussion of the patent data provided by it, see the OECD Patent Statistic Manual (2009) at https://www.oecd-ilibrary.org/science-and-technology/oecd-patent-statistics-manual_9789264056442-en.

This concern can be addressed by using only patents filed under the Patent Co-operation Treaty (PCT), using only patents filed at the EPO, the Japan Patent Office (JPO), and at the USPTO at the same time (referred to as the Triadic family by OECD) or using only patents protected in at least two international patent offices worldwide, one of which within the Five IP offices (IP5), namely the EPO, JPO, USPTO, the Korean Intellectual Property Office (KIPO) and the People's Republic of China National Intellectual Property Administration (CNIPA). Filing a patent is time intensive and costly. Those patents that meet one of the three filing requirements are all but certain to be a subset of the most important and thus economically valuable patents in terms of expected present value. Focusing on this subset of patents also ensures that the patents considered are not affected by different propensities to patent across countries or industries, as only international patents are used in the first place. Another potential drawback of using patents is that changes in patent laws may affect the propensity to patent. As all patents considered here have to be filed under international laws, changes in national patent laws are likely irrelevant. Additionally, as the empirical analysis will be across time, including time-fixed effects will take care of potential problems caused by changes in international patenting laws.

In principle, PCT patents, Triadic family patents, and IP5 patents are equally suited for analysis. However, the count of IP5 patents is suited best for the analysis at hand. While today a patent filed under the PCT is automatically protected in all PCT countries, this is only the case since 2004. Before that, there were fewer member states of the PCT, and the fees for PCT patents increased in the number of countries where the patent was filed. Thus PCT patent data are very well suited for analysis starting in 2004 but less reliable before that and therefore not well suited for the research question at hand. Regarding Triadic family patents, the main drawback is its acknowledged lack of timeliness. Since the goal of this paper is to relate changes in patents and population growth over time, timeliness is relevant. This makes IP5 patents the preferred measure of patenting activity for this paper and the one used if not stated otherwise.

Patents are of course a broad measure of technological progress. Making use of the classification put forward by Dechezleprêtre *et al.* (2019), only those patents that are related to automation technologies are used in the analysis. Thus the lack of specificity can be addressed by using the vast additional data provided with each patent. This makes the invention of automation technology and therefore a competitive edge of economies directly measurable.

And lastly, from a researcher's perspective, it will always be interesting to use different measures for the same underlying object of interest. It not only justifies revisiting old ideas but makes it possible to check their robustness and therefore relevance. In this case, the new data comes with an additional advantage in the time it spans. The patent data is available as far back as 1977 and thus starts much earlier than the robot data provided by the IFR. Using patent data, research over longer time periods is possible. In summary, the classified patent data provide an interesting alternative data set for empirical analysis, which has been underutilized so far.

3.3.3 Estimation Equation

The model developed in Section 3.2 demonstrates how the relationship between automation capital density and population growth depends on the level of skills present in the labor force. It illustrates how the interaction between population growth and education affects the allocation of resources toward automation uses. Therefore, education should be included in any specification trying to estimate the relationship between automation and population growth.

The goal is to identify the effect working-age population growth has on automation density, accounting for the effect education has on this relationship. The data set used for estimation has a panel structure, such that it is possible to include country-fixed and time-fixed effects in the regression. Country-fixed effects prevent omitted variables that are constant over time at the country level to bias the coefficients of interest. Additionally, time-fixed effects, which pick up variation over time affecting all countries equally, such as macroeconomic shocks, are included in the regression. This addresses concerns that results are driven by systemic economic shocks. While the inclusion of neither fixed effect guarantees the estimates to be unbiased, it is an important step toward the identification of the true parameter values.

To test the theoretical predictions, a measure for automation in per capita terms is needed as the dependent variable. For that, the variable $auto1$, the construction of which is explained in the previous section, is divided by the working-age population to construct a per capita measure of automation. The main explanatory variables are population growth and education. The economic concern with demographic change is that it affects the size of the working-age population. For that reason, the growth rate of the working-age population is calculated and used in the regressions. The share of the working-age population with at least completed secondary education is used as a measure for education. Reflecting the new theoretical results, it is important to include an interaction term of working-age population growth and education in the regression. As a control variable, gross fixed capital formation measured in percent of GDP is included in the regression to proxy for the savings rate.

To address reverse causality concerns, all regressions use a lag of one period (which is equivalent to five years) for all explanatory variables. For interpretation purposes, the log of the dependent variable and the log of the working-age population growth rate is used in the regression. Based on these considerations, the following baseline estimation equation is derived

$$\log(p_{c,t}) = \eta_0 + \eta_1 \cdot \log(n_{c,t-1}) + \eta_2 \cdot e_{c,t-1} + \eta_3 \cdot (\log(n_{c,t-1}) \times e_{c,t-1}) + \eta_4 \cdot s_{c,t-1} + FE_c + FE_t + \varepsilon_{c,t}, \quad (3.9)$$

where $p_{c,t}$ measures automation patents per capita, $n_{c,t-1}$ refers to working-age population growth, $e_{c,t-1}$ is the share of working-age population with at least completed secondary education, and $s_{c,t-1}$ is the savings rate. The subscript c indicates that variables are measured at the country level and subscript t refers to the period of observation. $\varepsilon_{c,t}$ is the error term.

3.4 Empirical Results

This section analyzes the empirical relevance of the theoretical results derived in Section 3.2. First, results from estimating the baseline regression are reported and discussed in Section 3.4.1. Subsequently, these results are shown to be robust to using different measures for the outcome and explanatory variables and including additional control variables in Section 3.4.2. And lastly, Section 3.4.3 demonstrates the adequacy of the new automation measure proposed by replicating previous findings with this new data.

3.4.1 Main Results

The theory presented in Section 3.2 makes clear predictions about the relationship between population growth and automation capital per capita, and the crucial way in which this relationship is influenced by the overall education level. Using Equation (3.9) as an estimation equation, a fixed effects regression is run to test the model predictions. The theory makes clear predictions about the signs of the estimated coefficients. If the assumed production function is a good representation of the real world, $\hat{\eta}_1 < 0$ and $\hat{\eta}_2 > 0$ are expected. The coefficient of the interaction term is predicted to be positive $\hat{\eta}_3 > 0$ if $e \in (0;1)$. Finally, $\hat{\eta}_4 > 0$ is predicted.

Results are reported in Table 3.1. Throughout, the investment variable is included as a control variable. Its coefficient is consistently estimated to be positive and it is highly significant, as expected. To emphasize the contribution made by including an interaction term between education and population growth, the explanatory variables of interest are added one by one. In column (1) the only explanatory variable is working-age population growth. Its coefficient has the expected negative sign but remains statistically insignificant. The specification corresponds to the limit case of the production function if the whole workforce is uneducated. The coefficient of working-age population growth stays statistically insignificant when education is included as an explanatory variable in column (2), the coefficient of which is also insignificant. This specification has no clear correspondence to the theoretical hypotheses. However, a cautious interpretation of the insignificant coefficient

Table 3.1: *Working-age Population Growth and Automation Density*

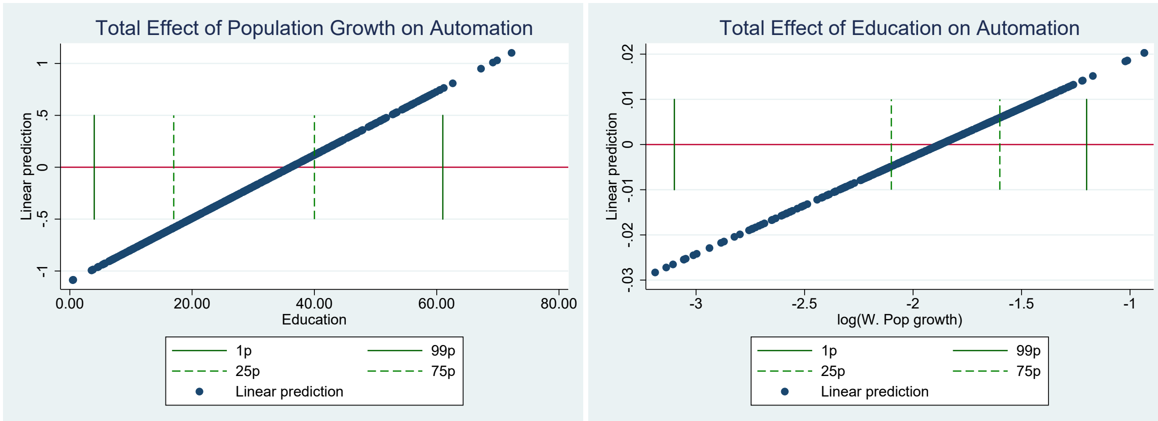
	(1)	(2)	(3)
log(W. Pop growth)	-0.1646	-0.1636	-1.0837***
	(-1.14)	(-1.13)	(-3.97)
Education		0.0004	0.0588***
		(0.07)	(3.65)
log(W. Pop growth) $\times$ Education			0.0322***
			(3.94)
Investment Share	0.0362***	0.0361***	0.0345***
	(3.47)	(3.43)	(3.36)
Time FE	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
R^2	0.60	0.60	0.62
Observations	328	325	325

Note: Dependent variable is the log of the automation measure *auto*1, constructed from patent data reported by the OECD and divided by working-age population to arrive at a per-capita measure. All explanatory variables are lagged by one period. log(W. Pop growth) is the log of working-age population growth. Education measures the share of the working-age population with at least completed secondary education as reported by Barro and Lee. Investment Share refers to gross fixed capital formation as a share of GDP. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics are reported in parentheses.

of education is that the density of automation patents does not vary with education when analyzed in isolation from population growth. In column (3), an additional interaction term between working-age population growth and education is included. This is the specification that is closest to the one derived from the theory. The theoretical hypothesis states that an increase in working-age population growth has a negative effect on automation capital per capita but that the effect is smaller the higher the education level of the workforce. This last aspect is picked up by the interaction term, which the significance of the coefficient indicates to be highly relevant.

Focusing on the estimation results reported in column (3), the results in Table 3.1 indicate that including the interaction effect is crucial. Given that the whole population is unskilled (implying $e = 0$ in the model), the coefficient of the interaction term can be neglected and working-age population growth is estimated to have a negative effect on the incentive to automate. The effect is quantitatively substantial. A 1% decrease in working-age population growth is associated with an increase of 1.1% in the automation measure. If in turn working-age population growth is zero, the effect of a 1% increase in the skilled population share is associated with an increase of the automation measure by 0.6%. The coefficient of the interaction term is positive as expected and highly significant. It attenuates the negative effect of working-age population growth on the automation measure, such that its effect is smaller in countries with a higher educated population share.

To interpret the effect of working-age population growth on automation, it may be helpful to look at how the effect changes depending on the level of education. Over the whole period 1977-2019, the education share takes on the values 16.7%, 28.2%, and 39.3% at the 25th, 50th, and 75th percentile of the distribution across countries. Over time, the average value of the education measure increased from 18% in 1977-1979 to 35% in 2016-2019. This is visualized in Figure C.3a in the Appendix. The total effect of working-age population growth at different levels of education can be calculated as $\frac{\partial automate}{\partial popgrowth} = \eta_1 + \eta_3 \cdot educ$. The result is plotted in Figure 3.1a (for the sake of clarity predicted values are plotted). The effect of working-age population growth on automation density is estimated to be negative if the



(a) *Total Effect of Population Growth* **(b)** *Total Effect of Education*

Figure 3.1: *Visualization of Regression Results in Table 3.1*

Note: The Figure illustrates the regression results reported in Table 3.1. Panel (a) shows the predicted total effect log population growth has on automation density for different values of the education variable. Panel (b) shows the predicted total effect an increase in education has on automation density for different values of working-age population growth. In both cases, solid lines mark the 1st and 99th percentile of the distribution of the variable plotted at the x-axis and dashed lines mark the 25th and 75th percentile of the distribution.

share of the educated population is low. As the share of the educated population increases, which is equivalent to a movement along the x-axis of Figure 3.1a, the size of the negative effect decreases. If the share of the educated population is at 35%, the effect of working-age population growth on automation density flips sign and becomes positive. The ambiguity of the effect working-age population growth has on automation density is relevant in the real world, as the sign of the estimated coefficient flips between the 25th and 75th percentile of the distribution of the education variable. This means that there are many countries where the effect is negative, but also many where the effect is positive. It also helps to explain why the estimated coefficient of working-age population growth in column (1) of Table 3.1 is insignificant. If education and especially the interaction between education and working-age population growth are omitted from the regression, the resulting coefficient estimates the average effect working-age population growth has on automation density, which across countries is not statistically different from zero.

Figure 3.1b visualizes the total effect of education on automation density for different values of working-age population growth. Again, the sign of the relationship changes from negative to positive between the 25th and 75th percentile of the working-age population growth distribution. The effect of education on automation density thus depends on the level of working-age population growth. Averaging it across countries experiencing different levels of working-age population growth results in an estimated average effect that is insignificant, as shown in column (2) of Table 3.1. In addition to looking at the total effect, marginal effects of population growth at certain levels of education, and vice versa, can be plotted as well. The resulting figures are relegated to the Appendix, see Figure C.2.

In summary, the empirical findings are in line with the theoretical hypotheses derived from the model. The predicted negative effect of working-age population growth on automation can be demonstrated in the data. This is especially true when the role of education is taken into account. Specifically, the model predicts a negative relationship between automation and population growth, which is mitigated by education. The empirical analysis suggests that this effect is even reversed to the positive if the education level of the workforce is sufficiently high.

3.4.2 Robustness

In this section, the robustness of the results presented above is tested. They are robust to using different measures of the education variable. Specifically, the share of the population with some tertiary education and the share of the population with completed tertiary education is considered instead of the share of the population with completed secondary education. Both are strict subsets of the originally used education measure. The results are reported in Table C.5 and Table C.6 in the Appendix. The magnitude of the coefficients of interest changes slightly but the pattern and statistical significance stay the same. The results are also robust to using total population growth instead of working-age population growth (reported in Table C.7). The coefficients are all significant at the 1% level and even higher in magnitude than in the original specification.

Alternative Patent Measures

For reasons laid out in Section 3.3.2, of the three available measures for patent data, those patents reported under the IP5 were used for analysis so far. A good robustness check is thus to analyze the theoretical relationship using those patents counted towards the Triadic Family and the PCT as well, to see if similar results can be obtained.

Reassuringly, the analysis results in very similar estimates using both alternative patent measures, which are reported in Table C.8 in the Appendix. The coefficients are estimated at the same significance level and even slightly higher in magnitude for both Triadic patent data and PCT patent data. In both cases, the coefficient of the investment share is insignificant but the point estimate remains positive.

As explained in Section 3.3.2, three different measures for automation arise from the Dechezleprêtre *et al.* (2019) classification, *auto*1, *auto*90, and *auto*95. So far, *auto*1 has been used to derive a density measure of automation. As a robustness check, the estimations reported in Section 3.4.1 are repeated using automation density measures constructed from *auto*90 and *auto*95 as the dependent variable. The results are reported in Table C.9. The main results can be replicated with the significance and magnitude of the coefficients very similar to those in the baseline regression.

Additional Control Variables

The estimations discussed so far have only included variables suggested by the theory to be of importance. Despite using time- and country-fixed effects in all specifications, there might be concerns regarding omitted variable bias. This section reports results from including several additional control variables in the regression to test the robustness of the results discussed so far.

Three additional control variables are included. One is the service share of the economy. It measures changes in the focus on manufacturing or services of individual economies which are not picked up by time-fixed effects. Two, the log of GDP is included to control for booms or recessions in individual economies not picked up by time-fixed effects. And three, the external balance as a measure of openness is included. Openness likely affects the pressure to keep up with technological advances and to stay competitive in general. Trade liberalization was ongoing in the period considered, which potentially makes openness a confounding factor if it is not included in the regression.

Results are reported in Table 3.2. The coefficients of all control variables are positive and highly significant when incorporated into the regressions. The coefficients of the main explanatory variables appear robust to the inclusion of further control variables, as displayed in column (3). The significance of the coefficients remains at the 1% level and the magnitude of the coefficients of working-age population growth, education, and the interaction effect even increase slightly.

The Relationship of Education and Population Growth

In the Unified Growth Theory, the interplay of population growth and education plays an important role. According to the literature, sustained economic growth was only made possible once fertility rates declined. With the onset of industrialization, human capital became more important and valuable. In a Unified Growth framework, the fertility choice is often described to feature a quantity-quality trade-off, referring to the number and education of children. A key assumption and often confirmed finding in this literature is a negative correlation between education levels and population growth rates.[16]

A strong negative correlation between education and population growth, as proposed in the Unified Growth Theory, potentially leads to imprecise estimates of the coefficients reported in Table 3.1. Indeed, the correlation between education and working-age population growth is quite strong at $\rho = -0.59$ across the whole sample. If some of the countries used for the estimation of the main specification are still undergoing the demographic transition, this can confound the estimates.

The data set used for estimation comprises information on 59 countries, including all

[16]The Unified Growth Theory was founded by Oded Galor. It has produced a large body of literature and is an actively researched topic in economics. For a comprehensive treatment of the theory and empirical findings see Galor (2011).

Table 3.2: *Adding Control Variables*

	(1)	(2)	(3)
log(W. Pop growth)	-0.2135	-0.2285	-1.9831***
	(-1.10)	(-1.18)	(-5.23)
Education		0.0068	0.0891***
		(1.14)	(5.39)
log(W. Pop growth)$\times$ Education			0.0451***
			(5.29)
Investment Share	0.0470***	0.0466***	0.0466***
	(3.85)	(3.82)	(4.05)
log(GDP p.c.)	1.4353***	1.4580***	1.4733***
	(5.70)	(5.77)	(6.19)
Openness	0.0514***	0.0519***	0.0466***
	(4.53)	(4.58)	(4.34)
Service Share	0.0116	0.0127	0.0194**
	(1.23)	(1.34)	(2.15)
Time FE	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
R^2	0.74	0.74	0.77
Observations	282	282	282

Note: Dependent variable is the log of the automation measure *auto*1, constructed from patent data reported by the OECD and divided by working-age population to arrive at a per-capita measure. All explanatory variables are lagged by one period. log(W. Pop growth) is the log of working-age population growth. Education measures the share of the working-age population with at least completed secondary education as reported by Barro and Lee. Investment share refers to gross fixed capital formation as a share of GDP. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics are reported in parentheses.

member countries of the G20. While some of the developing countries present in the data may still be undergoing the demographic transition at the start of the observation period in 1977, this is unlikely to be the case for the G20 member countries. Overall, the correlation between education and working-age population growth is lower in G20 member countries than in the rest of the sample countries, with respective values of $\rho = -0.47$ and $\rho = -0.62$.[17] When repeating the regressions reported in Table 3.1 in the sub-sample of G20

[17]There is a large variation in the correlation between education and population growth within the G20 member countries. For example, in Argentina, Australia, Russia, the UK, and the US, the two variables are positively correlated, and in all other countries negatively correlated. There is however no discernible or concerning pattern in the variation of the correlation.

member countries, the results (reported in Table C.10 in the Appendix) are very similar to the ones obtained when using the whole sample for estimation. Based on these results, it is unlikely that an ongoing demographic transition drives the overall results.

Time Series Analysis

In the cross-country analysis described and reported in Section 3.4.1, country-fixed effects were included in all regressions to control for unobserved heterogeneity across countries. These fixed effects take care of time-constant heterogeneity, such as cultural values or institutions, which might be correlated with the dependent and explanatory variables. Spanning 43 years, the time period used for estimation is quite long. Over such a long time span, even country characteristics considered quite stable across time, such as the education system, may change, potentially weakening the effectiveness of fixed effects in controlling for cross-country heterogeneity.

To address such concerns, this section reports results from an empirical analysis focusing on the US. The data on population size by age group and the education variable taken from Barro and Lee are only available at 5-year intervals. However, when using only 8 periods for estimation it is unlikely that reasonably reliable estimates can be obtained. Therefore, the data used so far is augmented by data on education taken from the PSID. The PSID gathers information of 5,000 representative households in the US, among other things on the highest level of education attained. Starting in 1997, the education variable is only surveyed biyearly. From the individual-level data, education measures corresponding to those provided by Barro and Lee are constructed. Reassuringly, the respective correlations are quite high at 0.76, 0.92, and 0.91 for completed secondary education, some tertiary education, and completed tertiary education, respectively. Since the correlation of the two measures of some tertiary education is the highest, this is the education variable used for analysis in the following. Data on population size by age group, from which a measure of working-age population is constructed, is only available at the 5-year interval. To avoid losing many time periods for estimation due to that data limitation, the total population size and its growth

Table 3.3: *Time Series Analysis using US Data*

	Automation Density	
	(1)	(2)
log(Pop growth)	-8.445	-3.909
	(-1.55)	(-0.59)
Education	0.994**	0.209
	(2.46)	(0.28)
log(Pop growth) × Education	0.315**	0.069
	(2.36)	(0.28)
Investment Share	-0.008	-0.012
	(-0.23)	(-0.25)
R^2	0.80	
Observations	24	13

Note: Dependent variable is the log of automation patents per capita. All explanatory variables are lagged by one period. log(Pop growth) is the log of population growth. Education measures the share of the population with some tertiary education. Investment share refers to gross fixed capital formation in % of GDP. Column (1) reports results from a regression in levels. Column (2) reports results from a regression in first differences. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics are reported in parentheses.

rate is used for the following analysis instead.

The estimation equation is the same as the one used for the panel data. Results from running an OLS regression using the US data are reported in column (1) of Table 3.3. When restricting the sample to the US data, a similar pattern of relationships is found. The coefficient of population growth is negative, though only significant at the 14% level. The coefficients of the education variable and the interaction term have the same sign as before but have a lower significance level as well. Given the small sample size, low significance levels of coefficients are not surprising.

One concern in time series analysis is a potential serial correlation of the error term. A graphical analysis of the residuals however finds no significant autocorrelation or partial autocorrelation of the error terms (see Figure C.4 in the Appendix). Several tests are available to assess if residuals from linear regression are serially correlated. Two of the most common test are Engle's Lagrange multiplier test and the Durbin-Watson test. In both cases, the null hypothesis is that there is no serial correlation in the errors. A rejection of the null hypothesis, therefore, indicates that the error terms are indeed serially correlated. When applying Engle's Lagrange multiplier test and a Durbin-Watson test to the error terms of the regression, the respective test statistics are given by $\chi^2 = 0.37$, which corresponds to a $p - value = 0.54$, and $d = 0.19$. In both cases, the null hypothesis is not rejected, signifying that there is no statistically significant evidence for serial correlation of the error terms. So neither the graphical analysis nor the analytic tests of the error terms indicate that they are serially correlated.

Another concern in time series analysis is that variables may be non-stationary, in which case a regression can lead to spurious results. The standard test for non-stationarity is the Augmented Dickey-Fuller test. Applying it to the dependent variable and all explanatory variables, the null hypothesis of non-stationarity cannot be rejected. Likewise, the Engle-Granger test, designed to test for cointegration of variables, indicates that the dependent variable and the explanatory variables are indeed cointegrated. For the first differences of all variables, the Augmented Dickey-Fuller test rejects the null hypothesis of non-stationarity. This suggests that the variables are non-stationary in levels but stationary in first difference form. Due to the only bi-yearly availability of the education measure starting in 1997, the sample available for estimation if first differences of the data are used is much reduced to only 13 observations. Results from such an estimation are reported in column (2) of Table 3.3. The signs of the estimated coefficients stay the same, the magnitude however is much reduced and none of the coefficients is significantly different from zero. Given the small sample size, this is not surprising.

The time series analysis of US data finds a similar pattern as the panel analysis in Section 3.4.1. The advantage of using only US data is that the results cannot be biased by time-varying unobserved heterogeneity across countries. The disadvantage is that the sample size is much smaller, reducing statistical power, and that time series analysis is accompanied by its own confounding factors, such as serial correlation, non-stationarity, and cointegration. Given the fact that the analysis in this section is so different from the baseline regression, it is striking that the results from it are similar to the previously reported results. In view of the differences in the empirical approach, the tentative results reported in this section reinforce the confidence that the analysis indeed reveals a systematic relationship between demographic change, automation, and education.

3.4.3 Replication

One contribution of this paper is to show that patent data combined with a classification system of patent categories into automation and non-automation classes constitute a new, appropriate, and high-quality resource to measure automation. To assess and test the usefulness of patent data, a replication study of Abeliansky and Prettner (2021) is done. This study is replicated because it addresses a similar empirical question. By trying to replicate the study with different data, both the patent data and the empirical results derived from it are tested.

The estimation equation used in Abeliansky and Prettner (2021) is:

$$\log(p_{i,t}) = a + \beta_1 \log(n_{i,t-1}) + \beta_2 \cdot \log(s_{i,t-1}) + \beta_3 \cdot \log(p_{i,t-1}) + \varepsilon_{i,t}$$

where p_t and p_{t-1} refer to an automation measure in t and $t-1$, the latter of which is only included in some specifications. n refers to population growth, and s refers to investment. Based on their theory (which omits education), the following signs are expected for the coefficients: $\hat{\beta_1} < 0$, $\hat{\beta_2} > 0$, $\hat{\beta_3} > 0$.

There are a few issues with this approach. First, their model predicts a negative effect of population growth on automation capital density. Irrespective of that, they use the growth rate of robot density (their automation measure) as a dependent variable, rather than the level of robot density. The estimation specification is thus not suited to test the theoretical hypothesis derived from their model (which corresponds to the limit case of $e = 0$ also discussed in Section 3.2.3). That they use the growth rate of robot density as a dependent variable causes a second issue, namely that the dependent variable, as well as the main explanatory variable, naturally and frequently takes on negative values. According to Abeliansky and Prettner (2021), they apply a box-cox transformation to the outcome variable and the main explanatory variable to deal with zero and negative values. However, a box-cox transformation does not alleviate the problem of zero and negative values but only ensures a zero-skew distribution of a variable, once negative and zero values are dropped.[18] Therefore, by applying a box-cox transformation, many observations will be dropped. Despite these issues, the patent data is transformed in the same way, first calculating growth rates and then applying a box-cox transformation to it, to replicate their analysis as closely as possible. Additionally and in line with the approach taken by Abeliansky and Prettner (2021), three-year rather than five-year averages of all data are taken for the replication exercise, to ensure maximal comparability of the findings.

When using patent data covering the same period as in the original study, the regression results, in particular the significance of the estimated coefficients, cannot be replicated (see Table C.1 in the Appendix). However, when running the same regressions using the whole period the patent data is available, the finding of a significantly negative relationship between population growth and automation growth can be replicated (see Table C.2 in the Appendix). This is robust to using different measures of automation patents.

Next, the baseline regression considered here is extended by successively including education and an interaction term between education and population growth, as done in Section 3.4.1. The results reported in Table C.4 in the Appendix show that including education, and especially the interaction term with population growth, is important. The coefficients of both education and the interaction term are statistically significant and positive. Additionally, the point estimate for the population growth variable has a higher significance and a higher magnitude if education and an interaction term are included. This indicates that any specification omitting education and the interaction term most likely fails to estimate the true relationship between population growth and automation.

[18] A box-cox transformation creates a new variable z in the following manner: $z = (x^{\lambda} - 1)/\lambda$. λ is chosen such that the skewness of z is zero. However, for the transformation to work, x has to be strictly positive (see Stata Manual https://www.stata.com/manuals13/rlnskew0.pdf).

According to the model put forward by Abeliansky and Prettner (2021), automation today is a function of automation in the last period. Therefore, the fixed effects regressions may be misspecified. To account for that, Abeliansky and Prettner (2021) test a dynamic specification as well, using corrected fixed effects.[19] The same is done using the patent data, results of which are reported in Table C.3 in the Appendix. As in the original paper, the magnitude of the autocorrelation coefficient is small. In the replication it is also statistically insignificant, indicating that neglecting to account for it is unproblematic, a conclusion also drawn by Abeliansky and Prettner (2021). However, in the replicated dynamic specification the estimated coefficient of population growth is smaller and no longer statistically significant, contrary to the original paper in which the coefficient of population growth stays significantly negative when including an autocorrelation term.

It should be noted here, that the model proposed by Abeliansky and Prettner (2021) can be simplified considerably when assuming a steady state. If in steady state, their derived expression for p_{t+1} can be simplified by setting $p_{t+1} = p_t = p^*$:

$$p_{t+1} = s(1-\alpha)\left(\frac{\alpha}{1-\alpha}\right)^{\alpha}\frac{1+p_t}{1+n}$$
$$p^* = s(1-\alpha)\left(\frac{\alpha}{1-\alpha}\right)^{\alpha}\frac{1+p^*}{1+n}.$$

Subsequently, the steady state condition can be solved for p^*:

$$p^* = \frac{s(1-\alpha)\left(\frac{\alpha}{1-\alpha}\right)^{\alpha}}{(1+n) - s(1-\alpha)\left(\frac{\alpha}{1-\alpha}\right)^{\alpha}}.$$

The relationship between population growth and automation density can now be derived by taking the derivative of p^* with respect to $(1+n)$:

$$\frac{\partial p^*}{\partial(1+n)} = -\frac{s(1-\alpha)\left(\frac{\alpha}{1-\alpha}\right)^{\alpha}}{\left((1+n) - s(1-\alpha)\left(\frac{\alpha}{1-\alpha}\right)^{\alpha}\right)^2} < 0.$$

By assuming that the model economy is in a steady state or on a Balanced Growth Path, an expression for p^* can thus be derived. With this, it can unambiguously be shown that

[19]For the estimation, the stata command *xtbcfe* developed by De Vos *et al.* (2015) is used.

the effect of population growth on automation density is always negative, making dynamic specifications obsolete.

In summary, the patent data can replicate the findings of Abeliansky and Prettner (2021) well. Since the empirical model is misspecified, the results in and of themselves should not be considered reliable. Beyond and more important than replicating empirical results, this section has shown that freely available patent data from the OECD are a well-suited measure of automation across countries and time.

3.5 Conclusion

Demographic change, especially the shrinking of the working-age population, poses a threat to economic growth in many developed economies. With the retirement of the baby boomer generation imminent, politicians struggle to counteract the drainage of the labor market by means of immigration, improving family and work compatibility, or encouraging the elderly to stay in the labor market longer. Another possibility to fill the void baby boomers are leaving in the labor market is by automating labor.

This paper studies the link between population growth, especially working-age population growth, and automation. Importantly, the proposed model distinguishes between skilled and unskilled labor. Assuming that automation capital is a closer substitute to unskilled labor than skilled labor, an economy's automation potential is predicted to depend on the population's education level.

The theoretical prediction is then tested empirically and verified using patent data from the OECD. A decrease in working-age population growth by 1% is associated with a 1.1% increase in automation, given the population is unskilled. The effect of population growth however depends crucially on the education level, such that the relationship between population growth and automation is even positive if a large enough share of the population is skilled.

The empirical results are derived using a new measure of automation based on patent data. While data on automation patents measures something slightly different than robot data, it still contains valuable information about automation at the country year level, which can be used for empirical analysis. In the replication exercise, I show that findings using robot data can be replicated with patent data. This automation measure based on patent data has several advantages over robot data. It covers a longer time horizon, as far back as 1977 instead of 1993. It is publicly available, meaning free of charge, contrary to the quite costly robot data. And lastly, it constitutes a broader measure of automation. Compared to robot data, patent data is much more likely to capture technological advances based on AI. In the face of rapid developments in automation in general and AI in particular, having a measure other than industrial robots for automation seems more important than ever.

The theoretical results were derived under two assumptions. One, that the education level is exogenously given and two, that capital is fully mobile between automation and traditional uses in the intratemporal maximization. In the next step, it will be interesting to explore how relaxing those assumptions affects the theoretical results. Intuitively, imposing friction in the mobility of capital between uses does not alter the results, as long as mobility of capital between uses is in general possible. The friction will lead to a sluggish response to changes in exogenous model parameters, which affects the transition between equilibria but not the direction or size of the effect population growth has on capital allocation. Endogenizing the education level of the labor force in contrast is likely to affect the model results in more complex ways. As derived in Section 3.2.2, the skill premium paid to skilled labor increases as more capital is used for automation. Therefore, an increase in automation capital, caused by changes in population growth, also affects a household's incentive to seek higher education. The interplay of demographic change, automation, and education will thus be more nuanced, once the share of educated labor in the workforce is the result of an endogenous choice by households. Extending the model by relaxing the two assumptions as discussed here is an interesting avenue for future work.

References

Abeliansky, A. L. and Prettner, K. (2021). Population Growth and Automation Density: Theory and Cross-Country Evidence. *Institut für Demographie - VID*, pp. 1–40.

Acemoglu, D. and Autor, D. (2010). *Skills, Tasks and Technologies: Implications for Employment and Earnings*. Tech. Rep. w16082, National Bureau of Economic Research, Cambridge, MA.

— and Guerrieri, V. (2008). Capital Deepening and Nonbalanced Economic Growth. *Journal of Political Economy*, **116** (3), 467–498.

— and Loebbing, J. (2022). Automation and Polarization. *Journal of Political Economy, Revise and Resubmit*.

— and Restrepo, P. (2017). Secular Stagnation? The Effect of Aging on Economic Growth in the Age of Automation. *American Economic Review*, **107** (5), 174–179.

— and — (2018). Low-Skill and High-Skill Automation. *Journal of Human Capital*, **12** (2), 204–232.

— and — (2020). Unpacking Skill Bias: Automation and New Tasks. *AEA Papers and Proceedings*, **110**, 356–361.

— and — (2022a). Demographics and Automation. *The Review of Economic Studies*, **89** (1), 1–44.

— and — (2022b). Tasks, Automation, and the Rise in U.S. Wage Inequality. *Econometrica*, **90** (5), 1973–2016.

Aghion, P., Howitt, P. and Bursztyn, L. (2009). *The Economics of Growth*. Cambridge, Mass: MIT Press.

Aguiar, M. and Bils, M. (2015). Has Consumption Inequality Mirrored Income Inequality? *American Economic Review*, **105** (9), 2725–2756.

Alonso-Carrera, J. and Raurich, X. (2015). Demand-based structural change and balanced economic growth. *Journal of Macroeconomics*, **46**, 359–374.

Ashraf, Q. H., Lester, A. and Weil, D. N. (2008). When Does Improving Health Raise GDP? *NBER Macroeconomics Annual*, **23** (1), 157–204.

Auclert, A. and Rognlie, M. (2017). Aggregate Demand and the Top 1 Percent. *American Economic Review*, **107** (5), 588–592.

Autor, D., Katz, L. and Kearney, M. (2005a). *Rising Wage Inequality: The Role of Composition and Prices*. NBER Working Papers 11628, National Bureau of Economic Research, Inc.

—, — and — (2005b). *Trends in U.S. Wage Inequality: Re-assessing the Revisionists*. NBER Working Papers 11627, National Bureau of Economic Research, Inc.

Autor, D. H., Katz, L. F. and Kearney, M. S. (2006). The Polarization of the U.S. Labor Market. *American Economic Review*, **96** (2), 189–194.

Balakrishnan, R., Dao, M., Solé, J. and Zook, J. (2015). *Recent U.S. Labor Force Dynamics: Reversible or Not?* Washington, D.C.: International Monetary Fund.

Baltagi, B. H., Lagravinese, R., Moscone, F. and Tosetti, E. (2017). Health Care Expenditure and Income: A Global Perspective. *Health Economics*, **26** (7), 863–874.

Banks, J., Blundell, R. and Lewbel, A. (1997). Quadratic Engel Curves and Consumer Demand. *The Review of Economics and Statistics*, **79** (4), 527–539.

Bárány, Z. L. and Siegel, C. (2018). Job Polarization and Structural Change. *American Economic Journal: Macroeconomics*, **10** (1), 57–89.

Bates, L. J. and Santerre, R. E. (2013). Does the U.S. Health Care Sector Suffer from Baumol's Cost Disease? Evidence from the 50 States. *Journal of Health Economics*, **32** (2), 386–391.

Battistin, E. and Nadai, M. D. (2015). Identification and Estimation of Engel Curves with Endogenous and Unobserved Expenditures. *Journal of Applied Econometrics*, **30** (3), 487–508.

Baumol, W. J. (1967). Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis. *The American Economic Review*, **57** (3), 415–426.

Berg, A., Ostry, J. D. and Zettelmeyer, J. (2012). What Makes Growth Sustained? *Journal of Development Economics*, **98** (2), 149–166.

Berkes, E., Manysheva, K. and Mestieri, M. (2022). *Global Innovation Spillovers and Productivity: Evidence from 100 Years of World Patent Data*. Tech. rep., Federal Reserve Bank of Chicago.

Bilbiie, F. O., Känzig, D. R. and Surico, P. (2022). Capital and Income Inequality: An Aggregate-Demand Complementarity. *Journal of Monetary Economics*, **126**, 154–169.

Bloom, D. E., Canning, D. and Fink, G. (2010). Implications of Population Ageing for Economic Growth. *Oxford Review of Economic Policy*, **26** (4), 583–612.

Bohr, C. E., Mestieri, M. and Yavuz, E. E. (2021). Engel's Treadmill: The Perpetual Pursuit of Cornucopia.

Boppart, T. (2014). Structural Change and the Kaldor Facts in a Growth Model With Relative Price Effects and Non-Gorman Preferences: Structural Change and the Kaldor Facts. *Econometrica*, **82** (6), 2167–2196.

Brown, C. (2004). Does Income Distribution Matter for Effective Demand? Evidence from the United States. *Review of Political Economy*, **16** (3), 291–307.

Brynjolfsson, E. and McAfee, A. P. (2011). *Race Against the Machine: How the Digital Revolution Is Accelerating Innovation, Driving Productivity, and Irreversibly Transforming Employment and the Economy*. Digital Frontier Press.

Buera, F. J., Kaboski, J. P., Rogerson, R. and Vizcaino, J. I. (2022). Skill-Biased Structural Change. *The Review of Economic Studies*, **89** (2), 592–625.

Carroll, C. (1998). *Why Do the Rich Save So Much?* Tech. Rep. w6549, National Bureau of Economic Research, Cambridge, MA.

Chernew, M. E. and Newhouse, J. P. (2011). Health Care Spending Growth. In *Handbook of Health Economics*, vol. 2, Elsevier, pp. 1–43.

Cohen, W. M. and Levinthal, D. A. (1989). Innovation and Learning: The Two Faces of R & D. *The Economic Journal*, **99** (397), 569–596.

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

Comin, D., Danieli, A. and Mestieri, M. (2020). *Income-Driven Labor Market Polarization*. Tech. Rep. w27455, National Bureau of Economic Research, Cambridge, MA.

— and Hobijn, B. (2010). An Exploration of Technology Diffusion. *American Economic Review*, **100** (5), 2031–2059.

—, Lashkari, D. and Mestieri, M. (2021). Structural Change With Long-Run Income and Price Effects. *Econometrica*, **89** (1), 311–374.

— and Mestieri, M. (2018). If Technology Has Arrived Everywhere, Why Has Income Diverged? *American Economic Journal: Macroeconomics*, **10** (3), 137–178.

Corneo, G. (2018). Time-poor, Working, Super-rich. *European Economic Review*, **101**, 1–19.

Cravino, J. and Sotelo, S. (2019). Trade-Induced Structural Change and the Skill Premium. *American Economic Journal: Macroeconomics*, **11** (3), 289–326.

Crespo Cuaresma, J., Kubala, J. and Petrikova, K. (2018). Does income inequality affect aggregate consumption? Revisiting the evidence. *Empirical Economics*, **55** (2), 905–912.

Cruz, E. (2019). Kuznets meets Lucas: Structural Change and Human Capital. *Oxford Economic Papers*, **71** (4), 848–873.

Cutler, D. M., Ghosh, K., Messer, K. L., Raghunathan, T., Rosen, A. B. and Stewart, S. T. (2022). A Satellite Account for Health in the United States. *American Economic Review*, **112** (2), 494–533.

De Meijer, C., Wouterse, B., Polder, J. and Koopmanschap, M. (2013). The Effect of Population Aging on Health Expenditure Growth: A Critical Review. *European Journal of Ageing*, **10** (4), 353–361.

De Vos, I., Everaert, G. and Ruyssen, I. (2015). Bootstrap-based Bias Correction and Inference for Dynamic Panels with Fixed Effects. *The Stata Journal: Promoting communications on statistics and Stata*, **15** (4), 986–1018.

De Vries, G. J., Gentile, E., Miroudot, S. and Wacker, K. M. (2020). The Rise of Robots and the Fall of Routine Jobs. *Labour Economics*, **66**, 101885.

Dechezleprêtre, A., Hemous, D., Olsen, M. and Zanella, C. (2019). Automating Labor: Evidence From Firm-Level Patent Data. *SSRN Electronic Journal*.

Duarte, M. and Restuccia, D. (2010). The Role of the Structural Transformation in Aggregate Productivity. *The Quarterly Journal of Economics*, **125** (1), 129–173.

Dunn, A., Hall, A. and Dauda, S. (2022). Are Medical Care Prices Still Declining? A Re-Examination Based on Cost-Effectiveness Studies. *Econometrica*, **90** (2), 859–886.

Dustmann, C., Ludsteck, J. and Schönberg, U. (2009). Revisiting the German Wage Structure. *The Quarterly Journal of Economics*, **124** (2), 843–881.

Dynan, K. E., Skinner, J. and Zeldes, S. P. (2004). Do the Rich Save More? *Journal of Political Economy*, **112** (2), 397–444.

Erdil, E. and Yetkiner, I. H. (2009). The Granger-causality Between Health Care Expenditure and Output: A Panel Data Approach. *Applied Economics*, **41** (4), 511–518.

Fisher, J. D., Johnson, D. S., Smeeding, T. M. and Thompson, J. P. (2020). Estimating the Marginal Propensity to Consume Using the Distributions of Income, Consumption, and Wealth. *Journal of Macroeconomics*, **65**, 103218.

Frank, M. W. (2009). Inequality and Growth in the United States: Evidence from a new State-Level Panel of Income Inequality Measures. *Economic Inquiry*, **47** (1), 55–68.

Frey, C. B. and Osborne, M. A. (2017). The Future of Employment: How Susceptible are Jobs to Computerisation? *Technological Forecasting and Social Change*, **114**, 254–280.

Fujiwara, I. and Matsuyama, K. (2022). A Technology-Gap Model of Premature Deindustrialization. *STEG and CEPR Working Paper Series*.

Galor, O. (2011). *Unified Growth Theory*. Princeton, N.J.: Princeton University Press.

Geary, R. C. (1950). A Note on "A Constant-Utility Index of the Cost of Living". *The Review of Economic Studies*, **18** (1), 65.

Getzen, T. E. (2016). Measuring and Forecasting Global Health Expenditures. In *World Scientific Handbook of Global Health Economics and Public Policy*, pp. 177–215.

— and Okunade, A. A. (2017). Symposium Introduction: Papers on 'Modeling National Health Expenditures': Symposium Introduction: Modeling National Health Expenditures. *Health Economics*, **26** (7), 827–833.

Goos, M. and Manning, A. (2007). Lousy and Lovely Jobs: The Rising Polarization of Work in Britain. *The Review of Economics and Statistics*, **89** (1), 118–133.

Graetz, G. and Michaels, G. (2018). Robots at Work. *The Review of Economics and Statistics*, **100** (5), 753–768.

Griffith, R., Redding, S. and Reenen, J. V. (2004). Mapping the Two Faces of R&D: Productivity Growth in a Panel of OECD Industries. *Review of Economics and Statistics*, **86** (4), 883–895.

Griliches, Z. (1969). Capital-Skill Complementarity. *The Review of Economics and Statistics*, **51** (4), 465.

Hanoch, G. (1975). Production and Demand Models with Direct or Indirect Implicit Additivity. *Econometrica*, **43** (3), 395.

Hartwig, J. (2008). What Drives Health Care Expenditure?—Baumol's Model of 'Unbalanced Growth' Revisited. *Journal of Health Economics*, **27** (3), 603–623.

— and Sturm, J.-E. (2014). Robust Determinants of Health Care Expenditure Growth. *Applied Economics*, **46** (36), 4455–4474.

Herrendorf, B., Rogerson, R. and Valentinyi, Á. (2014). Growth and Structural Transformation. In *Handbook of Economic Growth*, vol. 2, Elsevier, pp. 855–941.

Kambourov, G. and Manovskii, I. (2009). Occupational Specificity of Human Capital. *International Economic Review*, **50** (1), 63–115.

Ke, X., Saksena, P. and Holly, A. (2011). The Determinants of Health Expenditure: A Country-level Panel Data Analysis. Geneva: World Health Organisazion, pp. 1–28.

Kongsamut, P., Rebelo, S. and Xie, D. (2001). Beyond Balanced Growth. *The Review of Economic Studies*, **68** (4), 869–882.

Kotschy, R. and Sunde, U. (2018). Can Education Compensate the Effect of Population Ageing on Macroeconomic Performance? *Economic Policy*, **33** (96), 587–634.

Krenz, A., Prettner, K. and Strulik, H. (2021). Robots, Reshoring, and the Lot of Low-Skilled Workers. *European Economic Review*, **136**, 103744.

Krueger, A. B. (2012). The Rise and Consequences of Inequality in the United States.

Krusell, P., Ohanian, L. E., Rios-Rull, J.-V. and Violante, G. L. (2000). Capital-skill Complementarity and Inequality: A Macroeconomic Analysis. *Econometrica*, **68** (5), 1029–1053.

Mai, C.-M. and Marder-Puch, K. (2013). Selbstständigkeit in Deutschland. *Statistisches Bundesamt, Wirtschaft und Statistik*, (Juli 2013), 482–497.

Mann, K. and Püttmann, L. (2023). Benign Effects of Automation: New Evidence From Patent Texts. *The Review of Economics and Statistics*, pp. 1–18.

Martín, J. J. M., Puerto López Del Amo González, M. and Dolores Cano García, M. (2011). Review of the Literature on the Determinants of Healthcare Expenditure. *Applied Economics*, **43** (1), 19–46.

Matsuyama, K. (2022). Non-CES Aggregators: A Guided Tour. *Annual Review of Economics, forthcoming*.

Mian, A., Straub, L. and Sufi, A. (2020). *The Saving Glut of the Rich*. Tech. Rep. w26941, National Bureau of Economic Research, Cambridge, MA.

—, — and — (2021a). Indebted Demand. *The Quarterly Journal of Economics*, **136** (4), 2243–2307.

—, — and — (2021b). What Explains the Decline in r*? Rising Income Inequality versus Demographic Shifts. *Proceedings of the 2021 Jackson Hole Symposium*.

Mincer, J. and Jovanovic, B. (1981). Labor Mobility and Wages. In *Studies in Labor Markets*, National Bureau of Economic Research, Inc, pp. 21–64.

Murphy, K. M., Shleifer, A. and Vishny, R. (1989). Income Distribution, Market Size, and Industrialization. *The Quarterly Journal of Economics*, **104** (3), 537.

Nedelkoska, L. and Quintini, G. (2018). *Automation, Skills Use and Training*. OECD Social, Employment and Migration Working Papers 202.

Neffke, F. M., Otto, A. and Weyh, A. (2017). Inter-industry Labor Flows. *Journal of Economic Behavior & Organization*, **142**, 275–292.

Ngai, L. R. and Pissarides, C. A. (2007). Structural Change in a Multisector Model of Growth. *American Economic Review*, **97** (1), 429–443.

Nordhaus, W. D. (2008). Baumol's Diseases: A Macroeconomic Perspective. *The B.E. Journal of Macroeconomics*, **8** (1).

Okunade, A. A. and Osmani, A. R. (2018). *Technology, Productivity, and Costs in Healthcare*, Oxford Research Encyclopedia of Economics and Finance, Oxford University Press, pp. 1–21.

Peralta, C. and Gil, P. (2022). *Automation, Education, and Population: Dynamic Effects in an OLG Growth and Fertility Model*. Preprint, In Review.

Persson, T. and Tabellini, G. (1994). Is Inequality Harmful for Growth? *The American Economic Review*, **84** (3), 600–621.

Piketty, T. and Goldhammer, A. (2014). *Capital in the Twenty-First Century*. Harvard University Press.

Porzio, T., Rossi, F. and Santangelo, G. (2022). The Human Side of Structural Transformation. *American Economic Review*, **112** (8), 2774–2814.

Prettner, K. and Bloom, D. E. (2020). *Automation and Its Macroeconomic Consequences*. London, United Kingdom ; San Diego, CA, United States: Academic Press, and imprint of Elsevier.

— and Strulik, H. (2020). Innovation, Automation, and Inequality: Policy Challenges in the Race Against the Machine. *Journal of Monetary Economics*, **116**, 249–265.

Saez, E. and Stantcheva, S. (2018). A Simpler Theory of Optimal Capital Taxation. *Journal of Public Economics*, **162**, 120–142.

— and Zucman, G. (2020). The Rise of Income and Wealth Inequality in America: Evidence from Distributional Macroeconomic Accounts. *Journal of Economic Perspectives*, **34** (4), 3–26.

Sato, R. (1977). Homothetic and Non-Homothetic CES Production Functions. *The American Economic Review*, **67** (4), 559–569.

Sheiner, L. M. and Malinovskaya, A. (2016). Measuing Productivity in Healthcare: An Analysis of the Literature. Hutchins Center on Fiscal and Monetary Policy at Brookings.

Spitz-Oener, A. (2006). Technical Change, Job Tasks, and Rising Educational Demands: Looking outside the Wage Structure. *Journal of Labor Economics*, **24** (2), 235–270.

Stadhouders, N., Kruse, F., Tanke, M., Koolman, X. and Jeurissen, P. (2019). Effective Healthcare Cost-containment Policies: A Systematic Review. *Health Policy*, **123** (1), 71–79.

Stockhammer, E. and Wildauer, R. (2016). Debt-driven Growth? Wealth, Distribution and Demand in OECD Countries. *Cambridge Journal of Economics*, **40** (6), 1609–1634.

Stone, R. (1954). Linear Expenditure Systems and Demand Analysis: An Application to the Pattern of British Demand. *The Economic Journal*, **64** (255), 511.

Straub, L. (2019). Consumption, Savings, and the Distribution of Permanent Income. *Econometirca, Revise and Resubmit*.

Teixeira, A. A. and Queirós, A. S. (2016). Economic Growth, Human Capital and Structural Change: A Dynamic Panel Data Analysis. *Research Policy*, **45** (8), 1636–1648.

Violante, G. L. (2008). Skill-Biased Technical Change. In Palgrave Macmillan (ed.), *The New Palgrave Dictionary of Economics*, London: Palgrave Macmillan UK, pp. 1–6.

You, X. and Okunade, A. A. (2017). Income and Technology as Drivers of Australian Healthcare Expenditures: Determinants of Healthcare Expenditures in Australia. *Health Economics*, **26** (7), 853–862.

Appendix A

Appendix to Chapter 1

A.1 Theory Appendix

A.1.1 Proof of Proposition 1 and 2

The effect of an increase in $A_{1,t}$ on p_t is given as

$$\frac{\partial p_t}{\partial A_{1,t}} = \frac{\begin{vmatrix} -\frac{\partial F}{\partial A_{1,t}} & \frac{\partial F}{\partial L_{1,t}} \\ -\frac{\partial G}{\partial A_{1,t}} & \frac{\partial G}{\partial L_{1,t}} \end{vmatrix}}{\begin{vmatrix} \frac{\partial F}{\partial p_t} & \frac{\partial F}{\partial L_{1,t}} \\ \frac{\partial G}{\partial p_t} & \frac{\partial G}{\partial L_{1,t}} \end{vmatrix}} > 0.$$

The effect on an increase in $H_{1,t}/A_{1,t}$ on $L_{1,t}$ is given as

$$\frac{\partial L_{1,t}}{\partial H_{1,t}} = \frac{\begin{vmatrix} \frac{\partial F}{\partial p_t} & -\frac{\partial F}{\partial H_{1,t}} \\ \frac{\partial G}{\partial p_t} & -\frac{\partial G}{\partial H_{1,t}} \end{vmatrix}}{\begin{vmatrix} \frac{\partial F}{\partial p_t} & \frac{\partial F}{\partial L_{1,t}} \\ \frac{\partial G}{\partial p_t} & \frac{\partial G}{\partial L_{1,t}} \end{vmatrix}} = \frac{\partial L_{1,t}}{\partial H_{1,t}} = \frac{\partial L_{1,t}}{\partial A_{1,t}} \begin{cases} > 0 & \text{if } \theta > 1, \\ < 0 & \text{if } \theta < 1, \\ = 0 & \text{if } \theta = 1. \end{cases}$$

The effect on an increase in $H_{2,t}/A_{2,t}$ on $L_{1,t}$ is given as

$$\frac{\partial L_{1,t}}{\partial H_{2,t}} = \frac{\begin{vmatrix} \frac{\partial F}{\partial p_t} & -\frac{\partial F}{\partial H_{2,t}} \\ \frac{\partial G}{\partial p_t} & -\frac{\partial G}{\partial H_{2,t}} \end{vmatrix}}{\begin{vmatrix} \frac{\partial F}{\partial p_t} & \frac{\partial F}{\partial L_{1,t}} \\ \frac{\partial G}{\partial p_t} & \frac{\partial G}{\partial L_{1,t}} \end{vmatrix}} = \frac{\partial L_{1,t}}{\partial H_{2,t}} = \frac{\partial L_{1,t}}{\partial A_{2,t}} \begin{cases} < 0 & \text{if } \theta > 1, \\ > 0 & \text{if } \theta < 1, \\ = 0 & \text{if } \theta = 1. \end{cases}$$

The effect of an increase in γ, i.e., the preference for good one, on p_t and $L_{1,t}$ is given as

$$\frac{\partial p_t}{\partial \gamma} = \frac{\begin{vmatrix} -\frac{\partial F}{\partial \gamma} & \frac{\partial F}{\partial L_{1,t}} \\ -\frac{\partial G}{\partial \gamma} & \frac{\partial G}{\partial L_{1,t}} \end{vmatrix}}{\begin{vmatrix} \frac{\partial F}{\partial p_t} & \frac{\partial F}{\partial L_{1,t}} \\ \frac{\partial G}{\partial p_t} & \frac{\partial G}{\partial L_{1,t}} \end{vmatrix}} < 0, \quad \frac{\partial L_{1,t}}{\partial \gamma} = \frac{\begin{vmatrix} \frac{\partial F}{\partial p_t} & -\frac{\partial F}{\partial \gamma} \\ \frac{\partial G}{\partial p_t} & -\frac{\partial G}{\partial \gamma} \end{vmatrix}}{\begin{vmatrix} \frac{\partial F}{\partial p_t} & \frac{\partial F}{\partial L_{1,t}} \\ \frac{\partial G}{\partial p_t} & \frac{\partial G}{\partial L_{1,t}} \end{vmatrix}} > 0.$$

If households have a higher preference for good 1, then the relative price of good 2 will fall and more low-skilled labor will flow into Sector 1.

Recall, from Section 1.2.3, that the equilibrium can be characterized as follows

$$\begin{aligned} F \equiv & p_t^{\theta} \frac{Y_{2,t}}{Y_{1,t}} - \frac{1}{\gamma} + 1 = 0, \\ G \equiv & \alpha_1 \frac{Y_{1,t}}{L_{1,t}} - p_t \alpha_2 \frac{Y_{2,t}}{L_{2,t}} = 0, \end{aligned}$$

$$\begin{aligned} &\frac{\partial F}{\partial p_t} > 0, \quad \frac{\partial F}{\partial L_{1,t}} < 0, \quad \frac{\partial F}{\partial H_{1,t}} = \frac{\partial F}{\partial A_{1,t}} < 0, \quad \frac{\partial F}{\partial H_{2,t}} = \frac{\partial F}{\partial A_{2,t}} > 0, \quad \frac{\partial F}{\partial \gamma} > 0, \\ &\frac{\partial G}{\partial p_t} < 0, \quad \frac{\partial G}{\partial L_{1,t}} < 0, \quad \frac{\partial G}{\partial H_{1,t}} = \frac{\partial G}{\partial A_{1,t}} > 0, \quad \frac{\partial G}{\partial H_{2,t}} = \frac{\partial G}{\partial A_{2,t}} < 0, \quad \frac{\partial G}{\partial \gamma} = 0. \end{aligned}$$

$$\begin{vmatrix} \frac{\partial F}{\partial p_t} & \frac{\partial F}{\partial L_{1,t}} \\ \frac{\partial G}{\partial p_t} & \frac{\partial G}{\partial L_{1,t}} \end{vmatrix} = \frac{\partial F}{\partial p_t}\frac{\partial G}{\partial L_{1,t}} - \frac{\partial F}{\partial L_{1,t}}\frac{\partial G}{\partial p_t} < 0.$$

$$\begin{vmatrix} -\frac{\partial F}{\partial A_{1,t}} & \frac{\partial F}{\partial L_{1,t}} \\ -\frac{\partial G}{\partial A_{1,t}} & \frac{\partial G}{\partial L_{1,t}} \end{vmatrix} = \left(-\frac{\partial F}{\partial A_{1,t}}\right)\frac{\partial G}{\partial L_{1,t}} + \frac{\partial F}{\partial L_{1,t}}\frac{\partial G}{\partial A_{1,t}} < 0.$$

$$\begin{aligned}
\begin{vmatrix} \frac{\partial F}{\partial p_t} & -\frac{\partial F}{\partial H_{1,t}} \\ \frac{\partial G}{\partial p_t} & -\frac{\partial G}{\partial H_{1,t}} \end{vmatrix} =& \frac{\partial F}{\partial p_t}\left(-\frac{\partial G}{\partial H_{1,t}}\right) + \frac{\partial F}{\partial H_{1,t}}\frac{\partial G}{\partial p_t} \\
=& -\theta p_t^{\theta-1}\frac{Y_{2,t}}{Y_{1,t}}\alpha_1\frac{1}{L_{1,t}}\frac{\partial Y_{1,t}}{\partial H_{1,t}} + p_t^{\theta}\frac{Y_{2,t}}{Y_{1,t}}\frac{1}{Y_{1,t}}\frac{\partial Y_{1,t}}{\partial H_{1,t}}\alpha_2\frac{Y_{2,t}}{L_{2,t}} \gtreqless 0 \\
=& p_t\alpha_2\frac{Y_{2,t}}{L_{2,t}} - \theta\alpha_1\frac{Y_{1,t}}{L_{1,t}} \gtreqless 0 \\
=& \frac{p_t\alpha_2\frac{Y_{2,t}}{L_{2,t}}}{\alpha_1\frac{Y_{1,t}}{L_{1,t}}} - \theta \gtreqless 0 \\
=& 1-\theta \gtreqless 0.
\end{aligned}$$

$$\begin{aligned}
\begin{vmatrix} \frac{\partial F}{\partial p_t} & -\frac{\partial F}{\partial H_{2,t}} \\ \frac{\partial G}{\partial p_t} & -\frac{\partial G}{\partial H_{2,t}} \end{vmatrix} =& \frac{\partial F}{\partial p_t}\left(-\frac{\partial G}{\partial H_{2,t}}\right) + \frac{\partial F}{\partial H_{2,t}}\frac{\partial G}{\partial p_t} \\
=& \theta p_t^{\theta-1}\frac{Y_{2,t}}{Y_{1,t}}p_t\alpha_2\frac{1}{L_{1,t}}\frac{\partial Y_{2,t}}{\partial H_{2,t}} - p_t^{\theta}\frac{1}{Y_{1,t}}\frac{\partial Y_{2,t}}{\partial H_{2,t}}\alpha_2\frac{Y_{2,t}}{L_{2,t}} \gtreqless 0 \\
=& \theta - 1 \gtreqless 0.
\end{aligned}$$

A.1.2 Proof of Proposition 3

Let ξ_t denotes the share of good 1 in nominal GDP

$$\xi_t = \frac{Y_{1,t}}{Y_{1,t} + p_t Y_{2,t}}.$$

There are three channels through which an increase in $A_{1,t}$ can influence ξ in this model. First, directly by increasing the output produced in Sector 1. Second, by triggering a reallocation of low-skilled labor from one sector to the other. Third, by influencing the relative price of good 2 and thus affecting the nominal value of output produced in Sector 2.

We assume first that low-skilled labor cannot switch sectors. This simplifies the analysis, as we only have one equilibrium condition in this case

$$p_t = \left(\frac{1-\gamma}{\gamma}\frac{Y_{1,t}}{Y_{2,t}}\right)^{\frac{1}{\theta}}.$$

$$\begin{aligned}
\frac{\partial p_t}{\partial A_{1,t}} =& \frac{1}{\theta}\left(\frac{1-\gamma}{\gamma}\frac{Y_{1,t}}{Y_{2,t}}\right)^{\frac{1}{\theta}} \frac{1}{Y_{1,t}}\frac{\partial Y_{1,t}}{\partial A_{1,t}} \\
\frac{\partial p_t}{\partial A_{1,t}} =& \frac{1}{\theta} p_t \frac{1}{Y_{1,t}}\frac{\partial Y_{1,t}}{\partial A_{1,t}}.
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \xi_t}{\partial A_{1,t}} =& \frac{\frac{\partial Y_{1,t}}{\partial A_{1,t}} p_t Y_{2,t} - Y_{1,t}\frac{\partial p_t}{\partial A_{1,t}} Y_{2,t}}{(Y_{1,t}+p_t Y_{2,t})^2} \gtreqless 0 \\
\frac{\partial \xi_t}{\partial A_{1,t}} =& \frac{\frac{p_t Y_{1,t} Y_{2,t}}{A_{1,t}}\left(\frac{\partial Y_{1,t}}{\partial A_{1,t}}\frac{A_{1,t}}{Y_{1,t}} - \frac{\partial p_t}{\partial A_{1,t}}\frac{A_{1,t}}{p_t}\right)}{(Y_{1,t}+p_t Y_{2,t})^2} \gtreqless 0 \\
\frac{\partial \xi_t}{\partial A_{1,t}} =& \frac{\frac{\partial Y_{1,t}}{\partial A_{1,t}} p_t Y_{2,t} - Y_{1,t}\frac{1}{\theta} p_t \frac{1}{Y_{1,t}}\frac{\partial Y_{1,t}}{\partial A_{1,t}} Y_{2,t}}{(Y_{1,t}+p_t Y_{2,t})^2} \gtreqless 0 \\
\frac{\partial \xi_t}{\partial A_{1,t}} =& \frac{\left(\frac{\partial Y_{1,t}}{\partial A_{1,t}} p_t Y_{2,t}\right)\left(1-\frac{1}{\theta}\right)}{(Y_{1,t}+p_t Y_{2,t})^2} \gtreqless 0.
\end{aligned}$$

$$\frac{\partial \xi_t}{\partial A_{1,t}} \begin{cases} < 0 & \text{if } \theta < 1, \\ > 0 & \text{if } \theta > 1, \\ = 0 & \text{if } \theta = 1. \end{cases}$$

If the two goods are substitutes an increase in $A_{1,t}$ leads to an increase of good 1 as a share of nominal GDP. Consequently, for $\theta < 1$, i.e., the two goods are complements, an increase in $A_{1,t}$ leads to an increase in the share of good 2 in nominal GDP.

In case low-skilled labor is fully mobile, the effect of $A_{1,t}$ on ξ_t is given as

$$\frac{\partial \xi_t}{\partial A_{1,t}} = \left(p_t \frac{\partial L_{1,t}}{\partial A_{1,t}} \left(Y_{2,t} \frac{\partial Y_{1,t}}{\partial L_{1,t}} - Y_{1,t} \frac{\partial Y_{2,t}}{\partial L_{1,t}} \right) + \frac{p_t Y_{1,t} Y_{2,t}}{A_{1,t}} \left(\frac{\partial Y_{1,t}}{\partial A_{1,t}} \frac{A_{1,t}}{Y_{1,t}} - \frac{\partial p_t}{\partial A_{1,t}} \frac{A_{1,t}}{p_t} \right) \right) \cdot \frac{1}{(Y_{1,t} + p_t Y_{2,t})^2}$$

$$\frac{\partial \xi_t}{\partial A_{1,t}} = \left(\underbrace{p_t \frac{\partial L_{1,t}}{\partial A_{1,t}}}_{\gtreqless 0} \underbrace{\left(Y_{2,t} \frac{\partial Y_{1,t}}{\partial L_{1,t}} - Y_{1,t} \frac{\partial Y_{2,t}}{\partial L_{1,t}} \right)}_{>0} + \frac{p_t Y_{1,t} Y_{2,t}}{A_{1,t}} \underbrace{\left((1 - \alpha_1) - \underbrace{\frac{\partial p_t}{\partial A_{1,t}} \frac{A_{1,t}}{p_t}}_{>0} \right)}_{\gtreqless 0} \right) \cdot \frac{1}{(Y_{1,t} + p_t Y_{2,t})^2} \gtreqless 0.$$

The first term captures the effect of the reallocation of low-skilled labor that follows the increase in $A_{1,t}$. Depending on the elasticity of substitution this term can be positive or negative. The second term consists of two elements with opposite signs. The first part captures the increase in output in Sector 1 due to the increase in $A_{1,t}$ and is thus positive. The second part, which is the same as in the case when low-skilled labor is immobile, captures the effect of the increase in $A_{1,t}$ on the relative price. It has a negative effect on ξ_t because an increase in $A_{1,t}$ makes good 1 relative more abundant to good 2 and this will increase the relative price of good 2, i.e., good 2 becomes more expensive and good 1 less expensive.

Assume $\alpha_1 = \alpha_2 = \alpha$, this entails that we can combine the equilibrium conditions and solve for p_t, which is given as

$$p_t = \left(\frac{1-\gamma}{\gamma} \right)^{\frac{1-\alpha}{\alpha+\theta(1-\alpha)}} \left(\frac{A_{2,t} H_{2,t}}{A_{1,t} H_{2,t}} \right)^{\frac{1-\alpha}{\alpha+\theta(1-\alpha)}}.$$

Using this, we can express $\frac{\partial p_t}{\partial A_{1,t}} \frac{A_{1,t}}{p_t}$ as

$$\frac{\partial p_t}{\partial A_{1,t}} \frac{A_{1,t}}{p_t} = \frac{1-\alpha}{\alpha + \theta(1-\alpha)}.$$

The sign of the second term of $\frac{\partial \xi_t}{\partial A_{1,t}}$ is determined by

$$\alpha + \theta(1-\alpha) - 1 \gtreqless 0,$$

which is zero for $\theta = 1$, smaller than zero for $\theta \in (0,1)$, and larger than zero for $\theta > 1$.

Proof. For $\theta = 0$, we have $\alpha - 1 < 0$, as $\alpha \in (0,1)$. For $\theta = 1$, we have $1 - 1 = 0$. As $\alpha + \theta(1-\alpha) - 1$ is strictly increasing in θ it follows that for $\theta \in (0,1)$, $\alpha + \theta(1-\alpha) - 1 < 0$ and for $\theta > 1$, $\alpha + \theta(1-\alpha) - 1 > 0$. □

Therefore, it follows that an increase in $A_{1,t}$ has the following effect on the share of Sector 1 in nominal GDP

$$\frac{\partial \xi_t}{\partial A_{1,t}} \begin{cases} < 0 & \text{if } \theta < 1, \\ > 0 & \text{if } \theta > 1, \\ = 0 & \text{if } \theta = 1. \end{cases}$$

A.1.3 Full Factor Mobility

Consider a situation in which all production factors are fully mobile, except for the level of technology. For simplicity we assume each sector only produces with one production factor, but the production function has constant returns to scale in that factor. The production function for good j with $j \in \{1,2\}$ is given as

$$Y_{j,t} = A_{j,t} L_{j,t}.$$

The equilibrium can again be characterized by a system of two equations

$$
\begin{aligned}
F \equiv & p_t^{\theta} \frac{A_{2,t}(L_t - L_{1,t})}{A_{1,t} L_{1,t}} - \frac{1-\gamma}{\gamma} = 0 \\
F \equiv & p_t^{\theta} \frac{Y_{2,t}}{Y_{1,t}} - \frac{1}{\gamma} + 1 = 0, \\
G \equiv & w_{1,t}^{l} - w_{2,t}^{l} = 0 \\
G \equiv & A_{1,t} - p_t A_{2,t} = 0 \\
G \equiv & \frac{Y_{1,t}}{L_{1,t}} - p_t \frac{Y_{2,t}}{L_{2,t}} = 0.
\end{aligned}
$$

This entails

$$
\begin{aligned}
&\frac{\partial F}{\partial p_t} > 0, \quad \frac{\partial F}{\partial L_{1,t}} < 0, \quad \frac{\partial F}{\partial A_{1,t}} < 0, \quad \frac{\partial F}{\partial A_{2,t}} > 0, \quad \frac{\partial F}{\partial \gamma} > 0, \\
&\frac{\partial G}{\partial p_t} < 0, \quad \frac{\partial G}{\partial L_{1,t}} = 0, \quad \frac{\partial G}{\partial A_{1,t}} > 0, \quad \frac{\partial G}{\partial A_{2,t}} < 0, \quad \frac{\partial G}{\partial \gamma} = 0.
\end{aligned}
$$

And therefore, the results of the comparative statics are the same as in Section A.1.1.

A.1.4 Heterogeneous Preferences

Assume households face the same maximization problem as before, except now preferences over the two goods are heterogeneous, i.e., γ^i can now potentially differ across groups. To simplify the analysis, we further assume that all labor is immobile, i.e., low-skilled workers cannot switch sectors. This allows us to express the equilibrium as one equation.

$$
\begin{aligned}
&\max_{c_{1,t}^i, c_{2,t}^i} c_t^i(c_{1,t}^i, c_{2,t}^i) = \left((\gamma^i)^{\frac{1}{\theta}} (c_{1,t}^i)^{\frac{\theta-1}{\theta}} + (1-\gamma^i)^{\frac{1}{\theta}} (c_{2,t}^i)^{\frac{\theta-1}{\theta}} \right)^{\frac{\theta}{\theta-1}} \quad i \in \{e, d, l\} \\
&\text{s.t. } c_{1,t}^i + p_t c_{2,t}^i = I_t^i
\end{aligned}
$$

with $\theta \in (0, \infty)$.

Market clearing requires

$$\begin{aligned}
\frac{Y_{2,t}}{Y_{1,t}} =& \frac{\sum_i (1-\gamma^i) p_t^{-\theta} \frac{I_t^i}{\gamma^i + (1-\gamma^i) p_t^{1-\theta}} N_t^i}{\sum_i \gamma^i \frac{I_t^i}{\gamma^i + (1-\gamma^i) p_t^{1-\theta}} N_t^i} \\
p_t^\theta \frac{Y_{2,t}}{Y_{1,t}} =& \frac{\sum_i (1-\gamma^i) \frac{I_t^i}{\gamma^i + (1-\gamma^i)} N_t^i}{\sum_i \gamma^i \frac{I_t^i}{\gamma^i + (1-\gamma^i)} N_t^i} \\
p_t^\theta \frac{Y_{2,t}}{Y_{1,t}} =& \frac{\sum_i (1-\gamma^i) I_t^i N_t^i}{\sum_i \gamma^i I_t^i N_t^i} \\
p_t^\theta \frac{Y_{2,t}}{Y_{1,t}} =& \frac{\sum_i \mathcal{I}_t^i (Y_{1,t}, Y_{2,t}, p_t)}{\sum_i \gamma^i \mathcal{I}_t^i (Y_{1,t}, Y_{2,t}, p_t)} - 1,
\end{aligned}$$

where $\mathcal{I}_t^i (Y_{1,t}, Y_{2,t}, p_t) = I_t^i N_t^i$ denotes the aggregate income of group i.

The case of homogeneous preferences can be derived by assuming γ^i is the same for all groups.[1]

$$F \equiv p_t^\theta \underbrace{\frac{Y_{2,t}}{Y_{1,t}}}_{\text{relative supply}} - \underbrace{\frac{Y_{1,t} + p_t Y_{2,t}}{\sum_i \gamma^i \mathcal{I}_t^i (Y_{1,t}, Y_{2,t}, p_t)}}_{\text{demand composition}} + 1,$$

where we use the fact that $\sum_i \mathcal{I}_t^i (Y_{1,t}, Y_{2,t}, p_t) = Y_{1,t} + p_t Y_{2,t}$. A change in $Y_{j,t}$ with $j \in \{1, 2\}$ has an effect on the relative price p_t through the relative supply as well as by altering the demand composition. The latter channel only exists in a model with heterogeneous preferences.

[1]This yields $p_t^\theta \frac{Y_{2,t}}{Y_{1,t}} = \frac{1}{\gamma} - 1$.

A.2 General Appendix

A.2.1 List of Countries

Australia, Austria, Belgium, Canada, Chile, Colombia, Costa Rica, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Latvia, Lithuania, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Russia, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Türkiye, United Kingdom, United States

A.2.2 Employment in the Healthcare Sector in Germany

This section details some of the particularities of employment in the health sector in Germany. Self-employment is quite common in the health sector in Germany. In 2012, 4.7% of all self-employed in Germany were physicians and pharmacists, making it the occupational group with the fifth most self-employed persons (see Mai and Marder-Puch (2013), p. 490, only available in German). The income of self-employed persons in general is difficult to pin down. Nevertheless, the net income of self-employed physicians' offices in Germany in 2015 is reported to have been €192,000.[2] In comparison, employed physicians earned between €57,000 and €125,000 in 2019, according to the relevant collective labor agreement.[3] Given these numbers, it seems likely that physicians earn even more than suggested in the employment data. Thus the skill premium and its increase over the year is likely underestimated in the employment data.

---

[2] This is according to the Statistische Bundesamt, Fachserie 2 Reihe 1.6.1, p.19, only available in German. For a short link see `https://t.ly/BTkRI`

[3] See: `https://www.marburger-bund.de/bundesverband/tarifvertraege`

A.2.3 German Labor Force Changes 2007-2018

Table A.1 and Table A.2 display changes in the German labor force, analogously to Tables 1.5 and 1.6 in Section 1.3.2. They display the same statistics and ratios using German data. The goal is to facilitate the comparison of results found in the German and US data. Deviating from the US data, workers with medium skill levels are counted towards unskilled workers, such that the share of the unskilled labor force both in the overall economy and the health sector is larger in Table A.1 than in Table 1.5. This however is irrelevant to the derived results, as the focus of the analysis is on relative, rather than absolute changes in labor force shares.

Table A.1: *German Labor Force Changes 2007-2018*

	Overall Economy		Health Sector	
	2007	2018	2007	2018
Unskilled Labor Force	64.7%	64.7%	57.5%	60.3%
Δ		0%		+4.9%
Skill Premium	1.75	1.83	1.79	1.96
Δ		+4.6%		+9.1%

Note: Calculations based on data from the German Statistical Office.

Table A.2: *German Ratios of Key Indicators*

	2007	2018
Unskilled Labor Force Ratio	0.889	0.932
Skill Premium Ratio	1.023	1.071

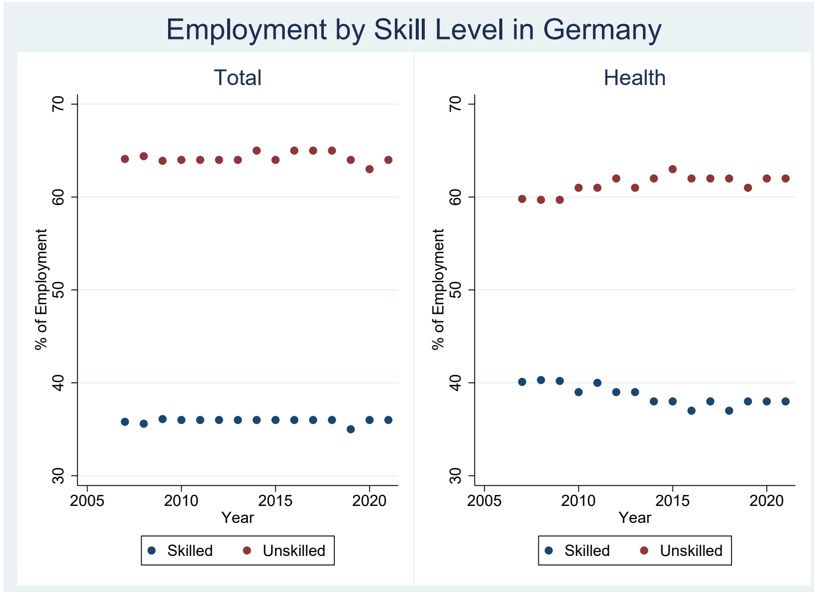
Note: Calculations based on Table A.1, which summarizes data from the German Statistical Office. The ratios implicitly account for time trends and compositional changes in the labor force.

A.2.4 Additional Graphs

This figure provides a graphical illustration of the trend in employment and expenditure in the health sector in Germany, based on data provided by the OECD

Figure A.1: *Employment and Expenditure in the Health Sector*

Figure A.1 illustrates the share of employment in the health sector as well as the share of overall expenditure going towards healthcare in Germany. Both measures have been continually on the rise in absolute as well as relative terms.



This figure provides a graphical illustration of the trend in the employment shares of skilled and unskilled workers in the overall economy and the health sector in Germany. The data used are provided by the German Statistical Office. Workers are classified as skilled if they are university graduates and unskilled otherwise.

Figure A.2: *The Share of Skilled and Unskilled Employment in Germany*

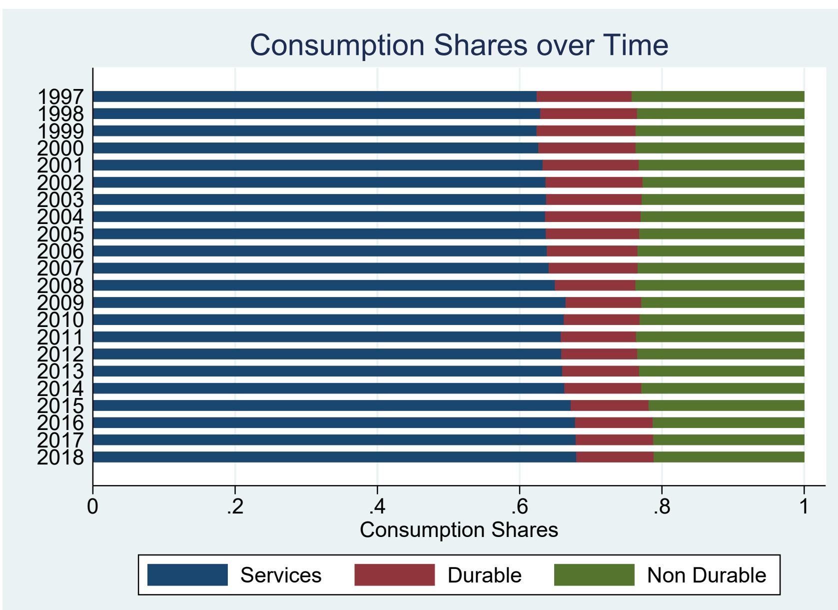
Figure A.2 illustrates the share of high- and low skilled labor for the total economy and the health sector in Germany from 2007 to 2018. While there is little to no change in the total economy, there is a slight upward trend for low skilled labor in the health sector. The left panel of Figure A.2 is in stark contrast to Figure 1.3 depicting the case of the US, which saw a marked increase in the share of high skilled labor. No figure equivalent to the right panel of Figure A.2 exists for the US, due to a lack of available data.

Appendix B

Appendix to Chapter 2

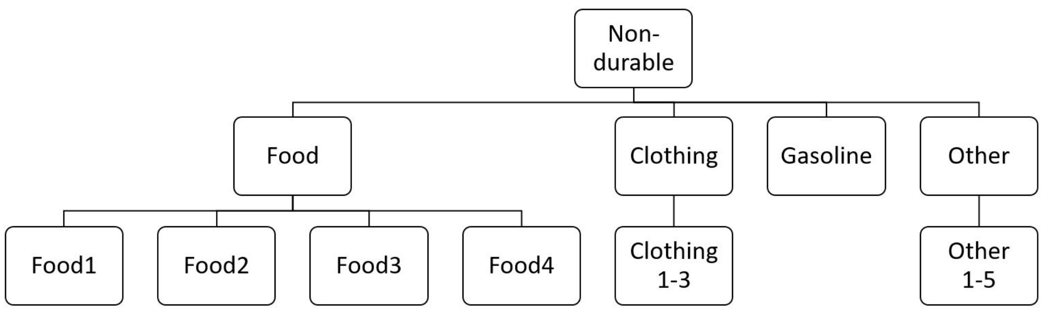
B.1 General Appendix

B.1.1 Additional Graphs



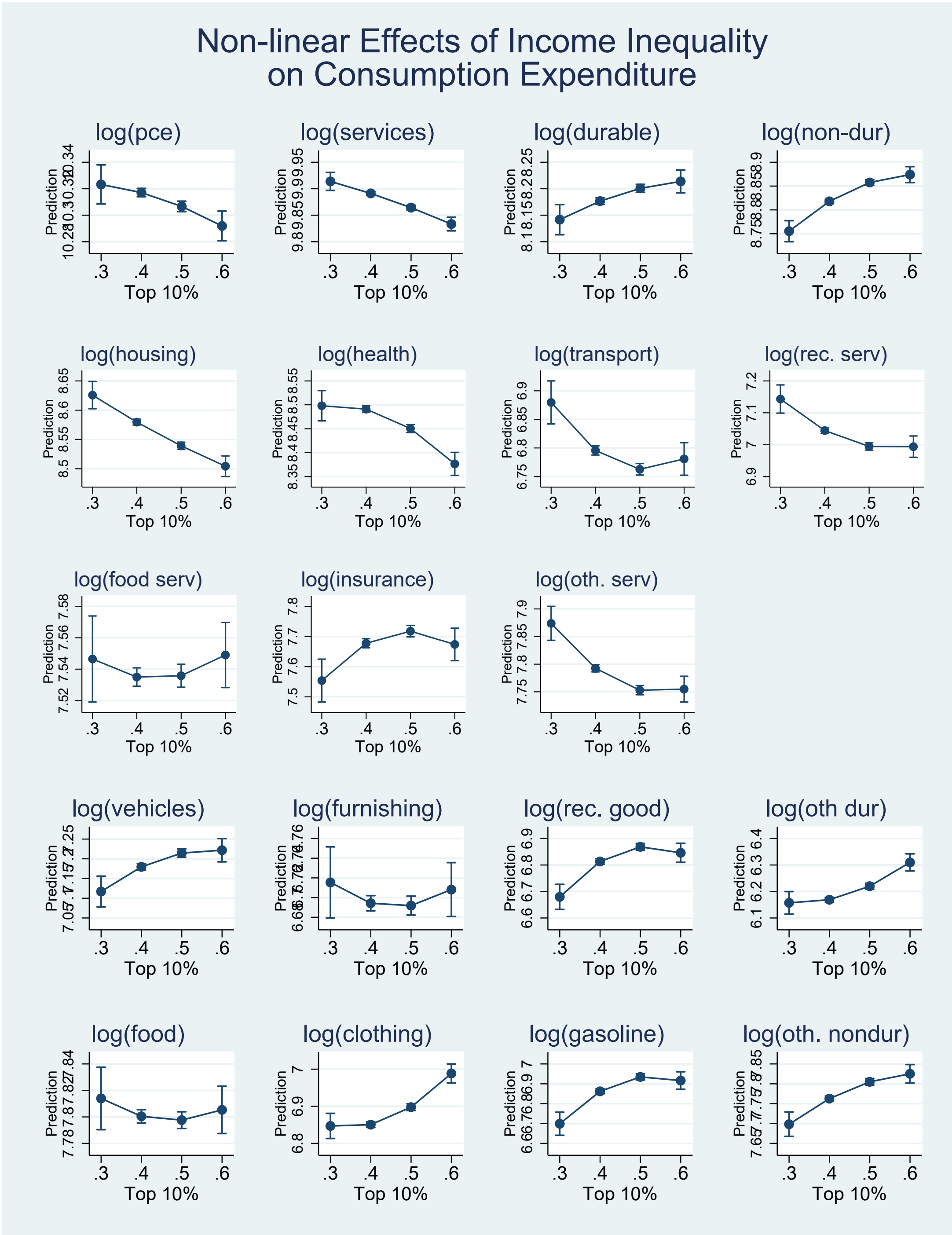
This figure provides a graphical illustration of how the consumption shares of services, durable and non-durable goods changed over time. Data is taken from the BEA.

Figure B.1: *Visualization of Consumption Shares over Time*



This figure illustrates why the price for one consumption category (in this example non-durbale goods) faced by households varies at the household level. Because the sub-categories Food1, Food2, Food3 and Food4, which have different prices, may be consumed in different quantities by households, the resulting price index for food and non-durable goods varies at the household level.

Figure B.2: *Construction of Household-level Price Data*



This figure provides a graphical illustration of the non-linear effect income inequality has on different consumption subcategories. The logged consumption categories are regressed on a linear and a quadratic term of the variable Top 10%, which measures the share of income going to the top 10% of the income distribution. Additionally, log income per capita and state and year fixed effects are included. Consumption expenditure data is taken from the BEA from 1997-2018 and income inequality data from Mark Frank.

Figure B.3: *Visualization of Non-linearity in the Inequality-Consumption Expenditure Relationship*

B.1.2 Additional Regression Results

Inequality Regressions for Sub-Categories

Table B.1: *Personal Consumption Expenditures and Inequality, Durable Goods*

	log(Vehicles)	log(Furnishing)	log(Recreation)	log(Other)
Top 10%	0.279***	0.024	0.355***	0.609***
	(3.24)	(0.29)	(3.36)	(6.46)
log(Income pc)	1.115***	1.206***	0.952***	0.303***
	(28.94)	(33.82)	(20.13)	(7.18)
Time FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
R^2	0.86	0.92	0.90	0.95
Observations	1,144	1,144	1,144	1,144

Note: The dependent variables and income per capita are used as reported by the BEA at the US-state level, using data form 1997-2018. The variable Top 10% reports the share of income going to the top 10% of the income distribution, as reported by Mark Frank. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics in parentheses.

Table B.1 reports the regression results for subcategories of durable consumption on income inequality, measured by the share of income going to the Top 10% of the income distribution, income per capita, and time- and state-fixed effects.

Table B.2 reports the regression results for subcategories of non-durable consumption on income inequality, measured by the share of income going to the Top 10% of the income distribution, income per capita, and time- and state-fixed effects.

Table B.3 reports the regression results for subcategories of service consumption on income inequality, measured by the share of income going to the Top 10% of the income distribution, income per capita, and time- and state-fixed effects.

Table B.2: *Personal Consumption Expenditures and Inequality, Non-durable Goods*

	log(Food)	log(Clothing)	log(Gasoline)	log(Other)
Top 10%	-0.002	0.583***	0.494***	0.366***
	(-0.03)	(7.74)	(3.81)	(5.40)
log(Income pc)	0.351***	0.415***	0.532***	0.402***
	(15.10)	(12.34)	(9.18)	(13.25)
Time FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
R^2	0.97	0.74	0.95	0.97
Observations	1,144	1,144	1,144	1,144

Note: The dependent variables and income per capita are used as reported by the BEA at the US-state level, using data form 1997-2018. The variable Top 10% reports the share of income going to the top 10% of the income distribution, as reported by Mark Frank. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics in parentheses.

Table B.3: *Personal Consumption Expenditures and Inequality, Service Goods*

	log(Housing)	log(Health)	log(Transport)	log(Recreation)	log(Food)	log(Insurance)	log(Other)
Top 10%	-0.391***	-0.491***	-0.199**	-0.373***	0.040	0.186	-0.291***
	(-7.60)	(-6.99)	(-2.38)	(-3.80)	(0.66)	(1.18)	(-4.24)
log(Income pc)	0.172***	0.286***	1.020***	0.593***	0.715***	0.607***	0.187***
	(7.51)	(9.11)	(27.23)	(13.54)	(26.46)	(8.59)	(6.10)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.98	0.98	0.92	0.95	0.98	0.89	0.96
Observations	1,144	1,144	1,144	1,144	1,144	1,144	1,144

Note: The dependent variables and income per capita are used as reported by the BEA at the US-state level, using data form 1997-2018. The variable Top 10% reports the share of income going to the top 10% of the income distribution, as reported by Mark Frank. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics in parentheses.

Expenditure Elasticities, different Specification and Subcategories

Table B.4: *Estimated Expenditure Elasticities using SURE*

	log(Services)	log(Durable)	log(Nondurable)
log(pce)	0.902***	1.571***	0.925***
	(76.01)	(46.08)	(34.12)
Time FE	Yes	Yes	Yes
State FE	Yes	Yes	Yes
R^2	1.00	0.94	0.93
Observations	1,144		

Note: SAEXP Data at the US state level from 1997-2018 is used for estimation. Results are obtained running a seemingly unrelated regression estimation. This is the reason for why observations are only reported in the first column. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics in parentheses.

Table B.4 reports results corresponding to the ones reported in Table 2.3 in Section 2.4.1. Here, the the single estimations for each consumption category is estimated in a Seemingly Unrelated Regression Estimation. The resulting expenditure elasticities are not affected. Only the t-statistics are slightly lower, which does not affect significance, though.

Table B.5: *Estimated Expenditure Elasticities, Durable Goods*

	Vehicles	Furnishing	Recreation	Other
log(pce)	1.739***	1.696***	1.556***	1.034***
	(32.61)	(31.35)	(23.58)	(18.53)
Constant	-16.878***	-16.652***	-15.437***	-10.289***
	(-32.57)	(-31.69)	(-24.08)	(-18.98)
Time FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
R^2	0.89	0.89	0.93	0.95
Observations	1,144	1,144	1,144	1,144

Note: SAEXP Data at the US state level from 1997-2018 is used for estimation. The dependent variable is given by as $\log(x_{sit}) - \log(\overline{x}_{it})$, where x_{sit} is the consumption good i in state s at time t and $\overline{x}_{it}$ is the average consumption of good i at time t across all states. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics in parentheses.

Table B.5, Table B.6 and Table B.7 report estimated expenditure elasticities for subcategories of consumption.

Table B.6: *Estimated Expenditure Elasticities, Non-durable Goods*

	Food	Clothing	Gasoline	Other
log(pce)	0.579***	0.695***	1.370***	0.806***
	(17.57)	(14.08)	(17.59)	(19.40)
Constant	-5.716***	-6.834***	-13.391***	-7.763***
	(-17.84)	(-14.25)	(-17.70)	(-19.23)
Time FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
R^2	0.94	0.94	0.93	0.91
Observations	1,144	1,144	1,144	1,144

Note: SAEXP Data at the US state level from 1997-2018 is used for estimation. The dependent variable is given by as $\log(x_{sit}) - \log(\overline{x}_{it})$, where x_{sit} is the consumption good i in state s at time t and $\overline{x}_{it}$ is the average consumption of good i at time t across all states. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics in parentheses.

Table B.7: *Estimated Expenditure Elasticities, Service Goods*

	Housing	Health	Transport	Recreation	Food	Insurance	Other	Services wo Ho.
log(pce)	0.524***	0.742***	1.528***	0.856***	1.016***	1.836***	0.324***	1.011***
	(16.67)	(17.25)	(28.32)	(13.26)	(25.28)	(20.24)	(7.24)	(60.73)
Constant	-5.404***	-7.338***	-15.300***	-8.544***	-10.115***	-17.940***	-3.537***	-10.064***
	(-17.68)	(-17.55)	(-29.20)	(-13.62)	(-25.90)	(-20.36)	(-8.13)	(-62.20)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.98	0.94	0.98	0.95	0.95	0.90	0.99	0.99
Observations	1,144	1,144	1,144	1,144	1,144	1,144	1,144	1,144

Note: SAEXP Data at the US state level from 1997-2018 is used for estimation. The dependent variable is given by as $\log(x_{sit}) - \log(\overline{x}_{it})$, where x_{sit} is the consumption good i in state s at time t and $\overline{x}_{it}$ is the average consumption of good i at time t across all states. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics in parentheses.

Non-linear Effects of Income Inequality on Expenditure Shares

Table B.8: *Personal Consumption Expenditures Shares and Non-linear Inequality Effects, Durable Goods*

	log(Vehicles)	log(Furnishing)	log(Recreation)	log(Other)
Top 10%	1.502***	-0.956**	3.980***	-1.341***
	(3.07)	(-2.02)	(6.42)	(-2.60)
$(\text{Top } 10\%)^2$	-1.165**	1.151**	-3.690***	2.172***
	(-2.29)	(2.34)	(-5.73)	(4.05)
log(Income pc)	0.621***	0.724***	0.443***	-0.173***
	(18.05)	(21.82)	(10.18)	(-4.77)
Time FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
R^2	0.95	0.89	0.71	0.31
Observations	1,144	1,144	1,144	1,144

Note: The dependent variables are log expenditure shares of the respective consumption categories. The dependent variables and income per capita are used as reported by the BEA at the US-state level, using data form 1997-2018. The variable Top 10% reports the share of income going to the top 10% of the income distribution, as reported by Mark Frank. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics in parentheses.

Table B.8, Table B.9, and Table B.10 report the estimated effect of income inequality and income inequality squared on the log expenditure share of different subcategories of consumption goods. Note, that different to before, the coefficient of log income per capita can no longer be interpreted as an income elasticity, because the dependent variable is an expenditure share. For nearly all subcategories, income inequality seems to have a non-linear effect on the respective expenditure shares, as indicated by statistical significance of 14 out of 15 estimated coefficients of the quadratic income inequality variable.

Table B.9: *Personal Consumption Expenditures Shares and Non-linear Inequality Effects, Non-durable Goods*

	log(Food)	log(Clothing)	log(Gasoline)	log(Other)
Top 10%	-0.596*	-1.582***	4.721***	1.327***
	(-1.75)	(-3.40)	(6.29)	(3.34)
(Top 10%)^2	0.746**	2.398***	-4.323***	-0.888**
	(2.12)	(4.96)	(-5.55)	(-2.15)
log(Income pc)	-0.133***	-0.059*	0.020	-0.091***
	(-5.55)	(-1.81)	(0.38)	(-3.27)
Time FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
R^2	0.73	0.93	0.90	0.26
Observations	1,144	1,144	1,144	1,144

Note: The dependent variables are log expenditure shares of the respective consumption categories. The dependent variables and income per capita are used as reported by the BEA at the US-state level, using data form 1997-2018. The variable Top 10% reports the share of income going to the top 10% of the income distribution, as reported by Mark Frank. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics in parentheses.

Table B.10: *Personal Consumption Expenditures Shares and Non-linear Inequality Effects, Service Goods*

	log(Housing)	log(Health)	log(Transport)	log(Rec.)	log(Food)	log(Insurance)	log(Other)
Top 10%	-0.747**	1.010**	-2.716***	-2.792***	-0.636*	4.081***	-2.361***
	(-2.45)	(2.46)	(-5.52)	(-4.47)	(-1.68)	(4.48)	(-5.15)
(Top 10%)^2	0.496	-1.457***	2.768***	2.665***	0.833**	-3.974***	2.298***
	(1.57)	(-3.42)	(5.41)	(4.11)	(2.12)	(-4.20)	(4.83)
log(Income pc)	-0.313***	-0.210***	0.548***	0.120***	0.232***	0.096	-0.288***
	(-14.59)	(-7.28)	(15.83)	(2.74)	(8.72)	(1.51)	(-8.95)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.33	0.82	0.78	0.19	0.55	0.23	0.43
Observations	1,144	1,144	1,144	1,144	1,144	1,144	1,144

Note: The dependent variables are log expenditure shares of the respective consumption categories. The dependent variables and income per capita are used as reported by the BEA at the US-state level, using data form 1997-2018. The variable Top 10% reports the share of income going to the top 10% of the income distribution, as reported by Mark Frank. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics in parentheses.

B.2 Theory Appendix

B.2.1 Savings in the Non-homothetic Model

The non-homothetic model detailed in Section 2.3 can be extended to encompass a savings decisions by households as well. The most straightforward way to do so is to include savings as one of the consumption goods over which the household maximizes utility. This may, for example, be driven by a preference for wealth, as proposed by, among others, Carroll (1998), Dynan *et al.* (2004), Saez and Stantcheva (2018), and Mian *et al.* (2021a).

In that case, the household optimizes the implicit utility as defined in Equation (2.7) over a consumption bundle (c_1, c_2, s), which now includes savings. The budget constraint is now defined over total income, rather than total expenditure, where $\mathcal{I} = p_1c_1 + p_2c_2 + s$ denotes total income.

$$\max_{c_1,c_2,s} \mathcal{L} = \left[(U^{\varepsilon_1}\zeta_1)^{\frac{1}{\sigma}} c_1^{\frac{\sigma-1}{\sigma}} + (U^{\varepsilon_2}\zeta_2)^{\frac{1}{\sigma}} c_2^{\frac{\sigma-1}{\sigma}} + (U^{\varepsilon_s}\zeta_s)^{\frac{1}{\sigma}} s^{\frac{\sigma-1}{\sigma}} \right] - \lambda \left[E - p_1c_1 - p_2c_2 - s\right]$$

Analogously to the case of consumption goods, the share of overall income used for savings is denoted by $\omega_s = (\zeta_s U^{\varepsilon_s})\, s^{\frac{\sigma-1}{\sigma}}$.

To determine if the share allocated towards savings increases as the income level and with it utility increases, consider the following derivative:

$$\frac{\partial \frac{\omega_s}{\omega_1+\omega_2}}{\partial U} = \frac{1}{(\omega_1+\omega_2)^2} \cdot \frac{1}{\sigma} \cdot \omega_s \cdot U^{-1} \left[\omega_1(\varepsilon_s - \varepsilon_1) + \omega_2(\varepsilon_s - \varepsilon_2)\right].$$

It follows, that $(\varepsilon_s > \varepsilon_1) \cap (\varepsilon_s > \varepsilon_2)$ is a sufficient condition for the share of income devoted to saving to increase relative to the share of income devoted to consumption as the income level and with it utility increases. It is equivalent to the savings rate being convex in income and the consumption rate being concave in income. Hence if income inequality increases, high-skilled households will increase their savings by more than low-skilled households will decrease their savings, resulting in an increase in aggregate savings and a decrease in aggregate consumption.

B.2.2 Comparative Static

The Homothetic Case

The equilibrium condition is defined by the structural Equation (2.6)

$$F \equiv \frac{A_H H}{A_L L} - \frac{\zeta_1}{\zeta_2} p^{-\sigma} = 0.$$

The effect of an increase in A_H on p is given as

$$\frac{\partial p}{\partial A_H} = -\frac{\partial F / \partial A_H}{\partial F / \partial p}$$

The derivatives are given by

$$\frac{\partial F}{\partial p} = \sigma \frac{\zeta_1}{\zeta_2} p^{-\sigma-1}$$

$$\frac{\partial F}{\partial A_H} = \frac{H}{A_L L}$$

Plugging in and making use of the fact that $\frac{H}{A_L L} = \frac{\zeta_1}{\zeta_2} p^{-\sigma} \frac{1}{A_H}$, it can be derived that

$$\frac{\partial p}{\partial A_H} = -\frac{p}{A_H} \frac{1}{\sigma} < 0.$$

The effect of an increase in A_H on the expenditure (which in the absence of savings is equivalent to income) of high-skilled households and low-skilled households is given by:

$$\frac{\partial E_h}{\partial A_H} = H \left(p + \frac{\partial p}{\partial A_H} \right) = H \cdot p \frac{\sigma - 1}{\sigma} < 0$$

$$\frac{\partial E_l}{\partial A_H} = 0$$

The Non-homothetic Case

The equilibrium in the case of non-homothetic preferences can be described by the structural equation (2.15), reproduced here:

$$F \equiv \frac{A_H H}{A_L L} - \frac{\zeta_1}{\zeta_2} p^{-\sigma} \frac{E_h^\sigma U_h^{1+\varepsilon_1} + E_l^\sigma U_l^{1+\varepsilon_1}}{E_h^\sigma U_h^{1+\varepsilon_2} + E_l^\sigma U_l^{1+\varepsilon_2}} = 0.$$

The effect of an increase of A_H on the equilibrium price can be calculated using the Implicit Function Theorem as

$$\frac{\partial p}{\partial A_H} = -\frac{\partial F/\partial A_H}{\partial F/\partial p}$$

Compared to the benchmark case, the comparative statics are more intricate at the demand side if preferences are non-homothetic. Specifically, the ratio of aggregate demand changes due to changes in the term

$$\frac{E_h^\sigma U_h^{1+\varepsilon_1} + E_l^\sigma U_l^{1+\varepsilon_1}}{E_h^\sigma U_h^{1+\varepsilon_2} + E_l^\sigma U_l^{1+\varepsilon_2}} = \frac{(A_H H \cdot p)^\sigma U_h^{1+\varepsilon_1} + (A_L L)^\sigma U_l^{1+\varepsilon_1}}{(A_H H \cdot p)^\sigma U_h^{1+\varepsilon_2} + (A_L L)^\sigma U_l^{1+\varepsilon_2}}.$$

$$\begin{aligned} &\partial \left(\frac{(A_H H \cdot p)^\sigma U_h^{1+\varepsilon_1} + (A_L L)^\sigma U_l^{1+\varepsilon_1}}{(A_H H \cdot p)^\sigma U_h^{1+\varepsilon_2} + (A_L L)^\sigma U_l^{1+\varepsilon_2}} \right) / \partial A_H \cdot \left((A_H H \cdot p)^\sigma U_h^{1+\varepsilon_2} + (A_L L)^\sigma U_l^{1+\varepsilon_2} \right)^2 = \\ &= \left[(A_H H \cdot p)^\sigma \left((A_H H \cdot p)^\sigma U_h^{1+\varepsilon_2} + (A_L L)^\sigma U_l^{1+\varepsilon_2} \right) U_h^{1+\varepsilon_1} \left(\frac{\sigma}{A_H} + (1+\varepsilon_1) U_h^{-1} \frac{\partial U_h}{\partial A_H} \right) \right] - \\ &- \left[(A_H H \cdot p)^\sigma \left((A_H H \cdot p)^\sigma U_h^{1+\varepsilon_1} + (A_L L)^\sigma U_l^{1+\varepsilon} \right) U_h^{1+\varepsilon_2} \left(\frac{\sigma}{A_H} + (1+\varepsilon_2) U_h^{-1} \frac{\partial U_h}{\partial A_H} \right) \right] \end{aligned} \tag{B.1}$$

For ease of notation, define

$$N \equiv (A_H H \cdot p)^\sigma U_h^{1+\varepsilon_2} + (A_L L)^\sigma U_l^{1+\varepsilon_2}$$
$$Z \equiv (A_H H \cdot p)^\sigma U_h^{1+\varepsilon_1} + (A_L L)^\sigma U_l^{1+\varepsilon_1}.$$

Then (B.1) can be rewritten as

$$\frac{\partial (Z/N)}{\partial A_H} = \frac{1}{N^2} (A_H H \cdot p)^\sigma \left[\frac{\sigma}{A_H} \left(N \cdot U_h^{1+\varepsilon_1} - Z \cdot U_h^{1+\varepsilon_2} \right) + \frac{\partial U_h}{\partial A_H} \left(N \cdot U_h^{\varepsilon_1} (1+\varepsilon_1) - Z \cdot U_h^{\varepsilon_2} (1+\varepsilon_2) \right) \right].$$

The equivalent derivative with respect to p is give by

$$\frac{\partial(Z/N)}{\partial p} = \frac{1}{N^2}(A_H H \cdot p)^\sigma \left[\frac{\sigma}{p}\left(N \cdot U_h^{1+\varepsilon_1} - Z \cdot U_h^{1+\varepsilon_2}\right) + \frac{\partial U_h}{\partial p}\left(N \cdot U_h^{\varepsilon_1}(1+\varepsilon_1) - Z \cdot U_h^{\varepsilon_2}(1+\varepsilon_2)\right)\right].$$

For ease of notation, define

$$B \equiv \left(N \cdot U_h^{1+\varepsilon_1} - Z \cdot U_h^{1+\varepsilon_2}\right)$$
$$D \equiv \left(N \cdot U_h^{\varepsilon_1}(1-\varepsilon_1) - Z \cdot U_h^{\varepsilon_2}(1+\varepsilon_2)\right)$$

Making use of N, Z, B and D, the terms of interest can be simplified to

$$\frac{\partial(Z/N)}{\partial A_H} = \frac{1}{N^2}(A_H H \cdot p)^\sigma \left[\frac{\sigma}{A_H}B + \frac{\partial U_h}{\partial A_H}D\right]$$
$$\frac{\partial(Z/N)}{\partial p} = \frac{1}{N^2}(A_H H \cdot p)^\sigma \left[\frac{\sigma}{p}B + \frac{\partial U_h}{\partial p}D\right].$$

$$\frac{\partial F}{\partial A_H} = \frac{\zeta_1}{\zeta_2}p^{-\sigma}\left(\frac{Z}{A_H} - \frac{1}{N}\right)\frac{1}{N}(A_H H \cdot p)^\sigma \left[\frac{\sigma}{A_H}B + \frac{\partial U_h}{\partial A_H}D\right]$$
$$\frac{\partial F}{\partial p} = \frac{\zeta_1}{\zeta_2}p^{-\sigma}\left(\frac{Z \cdot \sigma}{p} - \frac{1}{N}\right)\frac{1}{N}(A_H H \cdot p)^\sigma \left[\frac{\sigma}{p}B + \frac{\partial U_h}{\partial p}D\right]$$

$$\frac{\partial p}{\partial A_H} = -\frac{\partial F/\partial A_H}{\partial F/\partial p} = -\frac{\left(\frac{Z}{A_H} - \frac{1}{N}\right)}{\left(\frac{Z \cdot \sigma}{p} - \frac{1}{N}\right)} \cdot \frac{\left[\frac{\sigma}{A_H}B + \frac{\partial U_h}{\partial A_H}D\right]}{\left[\frac{\sigma}{p}B + \frac{\partial U_h}{\partial p}D\right]}$$

From this derivation it is obvious, that in general

$$\left.\frac{\partial p}{\partial A_H}\right|_{homothetic} \neq \left.\frac{\partial p}{\partial A_H}\right|_{non-homothetic}$$

It is not clear, if the effect of A_H on p is higher or lower if preferences are non-homothetic than in the homothetic benchmark case. This depends on the sign and magnitude of B and D, which in turn depend on the sign and magnitude of $\varepsilon_1 - \varepsilon_2$.

Appendix C

Appendix to Chapter 3

C.1 Theory Appendix

C.1.1 Intertemporal Maximization

Consider an infinite horizon economy in discrete time indexed by $t = 0, 1, 2, ...$ with competitively producing firms and one representative household. Population grows at rate n_t, such that $N_{t+1} = N_t(1 + n_t)$. Both the household and the firms have perfect foresight and there is no risk in the model.

Firms

Firms produce output using labor L_t and capital $\tilde{K}_t$ as inputs, with the production function given by[1]

$$Y_t = F(L_t, \tilde{K}_t),$$

which in intensive form with $y_t = \frac{Y_t}{L_t}$ and $\tilde{k}_t = \frac{\tilde{K}_t}{L_t}$ can be written as

$$y_t = f(\tilde{k}_t).$$

[1]In Section 3.2, it is important to distinguish between traditional capital K and automation capital P, which together make up the total capital stock $\tilde{K}$. It is this stock of total capital $\tilde{K}$ which is determined in the intertemporal maximization discussed here.

Firms operate under perfect competition and make zero profits in equilibrium. Therefore, production factors are paid their marginal products, such that $R_t = f'(\tilde{k}_t)$ and $w_t = f(\tilde{k}_t) - f'(\tilde{k}_t)\tilde{k}_t$. Firms maximize their profit, taking prices $\{p_t, w_t, r_t\}_{t=0}^{\infty}$ as given. The firms' maximization problem is given by

$$\max_{y_t, \tilde{k}_t, n_t} \pi_t = \sum_{t=0}^{\infty} p_t (y_t - r_t \tilde{k}_t - w_t)$$

$$s.t.\ y_t = f(\tilde{k}_t)$$

where $r_t = R_t - \delta$ is the interest rate net of depreciation and $\delta \in (0,1)$. Since the input factors labor and capital are owned by the households, firms take them as given such that the maximization is static. How the firms solve this infinite number of static maximization problems and the implications this has for the demand of sub-classes of capital is the focus of Section 3.2.

Households

The representative household owns all production factors and rents them to the firms, receiving marginal products as remuneration in return, which constitutes its income. Output can be consumed or invested. Taking into account a constant rate of population growth $n = const$, the representative household maximizes the lifetime utility of the entire dynasty by choosing consumption and savings optimally. The lifetime utility of the dynasty is given by

$$\sum_{t=0}^{\infty} (1+n)^t N_0 \left(\frac{1}{1+\rho} \right)^t u(c_t),$$

where $u(c_t)$ denotes the instantaneous per capita utility and ρ is the rate of time preference. For simplicity, assume $u(c_t) = log(c_t)$. In that case, the intertemporal elasticity of substitution is equal to one. The budget constraint faced by the household is given by the law of motion of capital

$$\tilde{K}_{t+1} = \tilde{K}_t(1-\delta) + F(\tilde{K}_t, L_t) - c_t L_t.$$

Dividing both sides by $L_{t+1} = L_t(1+n)$ gives the law of motion of the per capita capital stock

$$(1+n)\tilde{k}_{t+1} = \tilde{k}_t(1-\delta) + f(\tilde{k}_t) - c_t.$$

By choosing c_t, the household also determines $\tilde{k}_{t+1}$ through the law of motion. The intertemporal maximization problem of the household can be solved using a Lagrangian, which is set up and solved in the next section.

Solving the Household Problem

The intertemporal maximization problem of the household can be solved using a Lagrangian, which is set up as follows:

$$\max_{c_t, c_{t+1}, k_{t+1}} \mathcal{L} = \sum_{t=0}^{\infty}(1+n)^t N_0 \left(\frac{1}{1+\rho}\right)^t u(c_t) + \lambda_t(\tilde{k}_t(1-\delta) + f(\tilde{k}_t) - c_t - \tilde{k}_{t+1}(1+n))$$

The first-order conditions are

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial c_t} &= (1+n)^t N_0 \left(\frac{1}{1+\rho}\right)^t u'(c_t) - \lambda_t \stackrel{!}{=} 0,\\
\frac{\partial \mathcal{L}}{\partial c_{t+1}} &= (1+n)^{t+1} N_0 \left(\frac{1}{1+\rho}\right)^{t+1} u'(c_{t+1}) - \lambda_{t+1} \stackrel{!}{=} 0,\\
\frac{\partial \mathcal{L}}{\partial \tilde{k}_{t+1}} &= -\lambda_t(1+n) + \lambda_{t+1}((1-\delta) + f'(\tilde{k}_{t+1})) \stackrel{!}{=} 0.
\end{aligned}$$

In addition to the first order conditions, a terminal value condition is necessary, with can simply be stated as $\tilde{k}_\infty \geq 0$. It ensures that the representative household does not accumulate negative wealth and it is also referred to as a "No-Ponzi Game" condition.

Combining the first-order conditions and rearranging, the Euler Equation can be derived:

$$\begin{aligned}
\lambda_t &= \lambda_{t+1}\frac{(1-\delta) + f'(\tilde{k}_{t+1})}{1+n}\\
\left(\frac{1}{1+\rho}\right)^t u'(c_t) = \lambda_t &= \lambda_{t+1}\frac{(1-\delta) + f'(\tilde{k}_{t+1})}{1+n}\\
\frac{u'(c_t)}{u'(c_{t+1})} &= \left(\frac{1}{1+\rho}\right)((1-\delta) + f'(\tilde{k}_{t+1}))
\end{aligned}$$

Intuitively, if the time discount rate ρ increases, the marginal utility of consumption in period $t+1$ decreases relative to the marginal utility of consumption in period t, which, under the assumption of decreasing marginal utility, is equivalent to an increase in per capita consumption c_t relative to per capita consumption c_{t+1}. If however additional production possibility with one more unit of capital in $t+1$, given by $f'(\tilde{k}_{t+1})$, increases, per capita consumption in period t decreases relative to per capita consumption in period $t+1$.

Equilibrium and Steady State

Let p_t denote the price of the final output and consumption good, w_t the wage rate paid for labor services and r_t the rental rate for capital paid in period t. Taking the prices $\{p_t, w_t, r_t\}_{t=0}^{\infty}$ as given, firms maximize their profits and the household maximizes its intertemporal utility. In equilibrium, the markets for input factors capital and labor, and the final consumption good clear.

In steady state, the capital-labor ratio $\tilde{k}^*$ is constant, as well as consumption c^* and output per capita $f(\tilde{k}^*)$. The Euler Equation together with the law of motion of capital fully describes the dynamics of the neoclassical growth model.[2]

$$\frac{u'(c^*)}{u'(c^*)} = \left(\frac{1}{1+\rho}\right)((1-\delta) + f'(\tilde{k}^*)) = 1$$
$$(1+n)\tilde{k}^* = \tilde{k}^*(1-\delta) + f(\tilde{k}^*) - c^*$$

From the law of motion of capital, the steady state level of consumption can be derived as

$$c^* = f(\tilde{k}^*) - (\delta + n)\tilde{k}^*.$$

The marginal product of the steady state per capita capital stock $\tilde{k}^*$ can be derived as

$$f'(\tilde{k}^*) = \delta + \rho$$

Note, that if $f(\tilde{k}_t)$ is concave, its inverse exists, which is denoted by $f^{-1}(\tilde{k}_t)$. Making use of

[2]Note, that due to $u(c_t) = \log(c_t)$, the intertemporal elasticity of substitution is equal to one and does not scale the Euler Equation.

the invertibility of $f'(\tilde{k}_t)$ and rearranging yields the equilibrium capital stock per capita

$$\tilde{k}^* = f^{'-1}(\delta + \rho).$$

A change in the population growth rate n reduces the equilibrium consumption level c^*. The representative household reacts to a change in n since it maximizes the lifetime utility of the entire dynasty. A change in n however leaves the equilibrium per capita capital level $\tilde{k}^*$ unchanged, as n does not factor into its equilibrium level. To see why, note that the Euler Equation consists of exogenous and invariant parameters δ and ρ as well as the marginal product of the equilibrium capital stock $f(\tilde{k}^*)$. For consumption to be constant, the marginal product of capital has to be equal to some constant value given by $(\delta + \rho)$. Therefore, only one value of $\tilde{k}$ is consistent with constant consumption levels, which is given by $\tilde{k}^*$.

C.1.2 The effect of N and e on K, P, and $\tilde{K}$

To distinguish the effect N has on the marginal product of the overall capital stock $\tilde{K}$ which is optimally allocated between K and P, such that the marginal products of K and P are equalized, and the effect N has on the two types of capital K and P, respectively, K and P have to be analyzed in isolation. This allows to derive the effect N has on K and P if the allocation of $\tilde{K}$ between capital uses does not change.

To derive the effect N has on the marginal product of traditional capital K, the value of K is held constant. This isolates the effect N has on the marginal product of K, by shutting down the effect N has on the allocation of $\tilde{K}$ between K and P. The same holds for the derivation of the effect N has on automation capital P. The respective derivatives are given by

$$\frac{\partial r^{trad}}{\partial N} = \alpha(1-\alpha)\frac{Y}{K \cdot N}\left((1-e-e\beta)N + (1-\beta)P\right) > 0$$

$$\frac{\partial r^{auto}}{\partial N} = (1-\alpha)\beta\frac{Y}{((1-e)N+P)^2 N}\left((1-\alpha)(1-\beta)P - \alpha(1-e)N\right).$$

The sign of $\partial r^{trad}/\partial N$ is universally positive. The sign of $\partial r^{auto}/\partial N$ is indeterminate. On the one hand, an increase in N increases the skilled labor force, a complement to automation capital, and with it the return on automation capital. On the other hand, it increases unskilled labor, a substitute for automation capital, and with it decreases the return on automation capital. The derivative can be decomposed into the effect population size has on output and unskilled labor supply:

$$\frac{\partial r^{auto}}{\partial N} = (1-\alpha)\beta\frac{1}{((1-e)N+P)^2}\left(((1-e)N+P)\frac{\partial Y}{\partial N} - Y\frac{\partial((1-e)N+P)}{\partial N}\right).$$

Formally, the positive effect of population size on r^{auto} is captured by the first term in the large brackets, and the negative effect of population size on r^{auto} by the second term in the large brackets. Which effect dominates depends on the size of the skilled labor force.

Full mobility of capital entails, that the marginal product of both kinds of capital is equal at all times. Therefore, an increase in N is accompanied by a reallocation of $\tilde{K}$ between traditional uses and automation uses. If the ratio of K/P adjusts to the increase in N, the effect of an increase in N on the equilibrium interest rate r^* is strictly positive, reflecting the increase in the input factor labor which is available for production:

$$\frac{\partial r^*}{\partial N} = ((1-\alpha)\beta + \alpha)\left(((1-e)N + \tilde{K}) \cdot \partial Y/\partial N - Y \cdot (1-e)\right) > 0$$

$$\frac{\partial Y}{\partial N} = (1-\alpha)Y\left(\frac{\beta(1-\alpha)(1-e)\beta}{(1-\alpha)\beta+\alpha} + \frac{1-\beta}{N}\right) + \alpha Y \frac{(1-e)}{((1-\alpha)\beta+\alpha)((1-e)N+\tilde{K})} > 0$$

Holding constant the allocation of $\tilde{K}$ between traditional uses and automation uses, the cross derivatives of r^{auto} and r^{trad} with respect to N and e are given by

$$\begin{aligned}\frac{\partial^2 r^{auto}}{\partial N \partial e} = r^{auto} \cdot ((1-\alpha)(1-\beta)eN + (1-(1-\alpha)\beta)\left((1-e)N+P\right)) \cdot \\ \left(\frac{(1-e)N(1-\alpha)(1+(1-\alpha)(1-\beta))P}{((1-e)N+P)}\right) + \\ + r^{auto} \cdot N\left((1-\alpha)(1-\beta)eN + (1-(1-\alpha)\beta)(1-e)\right) > 0\end{aligned}$$

$$\frac{\partial^2 r^{trad}}{\partial N \partial e} = \alpha(1-\alpha)\frac{Y}{K \cdot N}(\beta - 1)N < 0.$$

Having shown that education influences the effect population size has on the return on automation capital and hence the incentive to automate, the direct effect of education on automation incentives is also of interest. The effect can be derived by taking the derivative of r^{auto} with respect to education, which is always positive.

$$\frac{\partial r^{auto}}{\partial e} = r^{auto} \cdot N \cdot ((1-\alpha)(1-\beta)eN + (1-(1-\alpha)\beta)((1-e)N+P)) > 0$$

Education increases r^{auto} for two reasons: First, it increases the supply of skilled labor, which is a complement to automation capital. Second, it decreases unskilled labor, which is a substitute for automation capital.

C.1.3 Taking into Account the effect of N on $\tilde{K}$

In Section 3.2.2 it has been derived that $\frac{\partial K/P}{\partial N} > 0$ under the implicit assumption of $\frac{\partial \tilde{K}}{\partial N} = 0$. Here it is shown, that the results carry through when taking into account the second-order effect of $\frac{\partial \tilde{K}}{\partial N}$.

$$\frac{K}{P} = \frac{\alpha(1-e)N + \tilde{K}}{\tilde{K}(1-\alpha)\beta - \alpha(1-e)N}$$

$$\begin{aligned}\frac{\partial K/P}{\partial N} = & \frac{\left(\tilde{K}(1-\alpha)\beta - \alpha(1-e)N\right)\left(\alpha(1-e) + \frac{\partial \tilde{K}}{\partial N}\right)}{\left(\tilde{K}(1-\alpha)\beta - \alpha(1-e)N\right)^2} - \\ & - \frac{\left(\alpha(1-e)N + \tilde{K}\right)\left(\frac{\partial \tilde{K}}{\partial N}(1-\alpha)\beta - \alpha(1-e)\right)}{\left(\tilde{K}(1-\alpha)\beta - \alpha(1-e)N\right)^2} \\ & = \frac{\alpha(1-e)(1+(1-\alpha)\beta)\left(\tilde{K} - \frac{\partial \tilde{K}}{\partial N} \cdot N\right)}{\left(\tilde{K}(1-\alpha)\beta - \alpha(1-e)N\right)^2}\end{aligned}$$

Note, that $\tilde{k} = const$ on the balanced growth path implies that $\tilde{K}$ and N grow at the same rate, which is given by n. Therefore, $\frac{\partial \tilde{K}}{\partial N} = 1$. This results in $\frac{\partial K/P}{\partial N} > 0 \iff \tilde{K} > N$. There is no reason why the opposite should be true, such that it is an innocuous assumption for $\tilde{K} > N$ to be true. In that case, $\frac{\partial K/P}{\partial N} > 0$ holds even when allowing for $\frac{\partial \tilde{K}}{\partial N} \neq 0$.

By taking the derivative of $\frac{\partial (K/P)}{\partial N}$ with respect to e, the role the share of skilled labor in the labor force plays can be determined.

$$\begin{aligned}\frac{\partial^2 K/P}{\partial N \partial e} = & \frac{\left(\tilde{K}(1-\alpha)\beta - \alpha(1-e)N\right)^2 \alpha(1+(1-\alpha)\beta(-1)}{\left(\tilde{K}(1-\alpha)\beta - \alpha(1-e)N\right)^4} - \\ & - \frac{\alpha(1-e)(1+(1-\alpha)\beta) \cdot 2\left(\tilde{K}(1-\alpha)\beta - \alpha(1-e)N\right)(\alpha N)}{\left(\tilde{K}(1-\alpha)\beta - \alpha(1-e)N\right)^4} \\ & = \frac{\alpha(1+(1-\alpha)\beta)(-1)\left(\tilde{K}(1-\alpha)\beta + (1-e)\alpha N\right)}{\left(\tilde{K}(1-\alpha)\beta - \alpha(1-e)N\right)^3} < 0\end{aligned}$$

The denominator of the cross derivative is positive since it is the denominator of $\frac{K}{P}$, which is positive, to the power of three. The numerator is negative, since $\alpha(1+(1-\alpha)\beta)(-1) < 0$.

Thus the cross derivative is negative.

Since $\frac{\partial(K/P)}{\partial N} > 0$, the effect of the population size N on the ratio $\frac{K}{P}$ is positive. For the limit case of $e = 1$, the ratio of $\frac{K}{P}$ is fully determined by the parameters α and β and thus independent of N. Therefore, in the limit case of $e = 1$, $\frac{\partial(K/P)}{\partial N} = 0$. The negative sign of the cross derivative demonstrates that as e increases, the effect of N on $\frac{K}{P}$ decreases. This is equivalent to the results derived in the main part of the paper, which neglects the effect of N on $\tilde{K}$.

C.1.4 Limit Cases of Per Capita Production Functions

To better understand the effect of population growth on per capita output and per capita capital, which is made up of traditional and automation capital, consider the limit cases of the per-capita production function discussed in Section 3.2.4.

If the labor force is unskilled ($e = 0$), per capita output is given by

$$y = k^{\alpha}(1 + p)^{1-\alpha}.$$

Deriving the marginal product of traditional and automation capital per capita:

$$r^{trad} = \frac{\partial y}{\partial k} = \alpha \frac{y}{k},$$
$$r^{auto} = \frac{\partial y}{\partial p} = (1 - \alpha)\frac{y}{(1 + p)}.$$

Setting the two marginal products equal due to capital mobility and plugging in $k = \tilde{k} - p$ and $p = \tilde{k} - k$, the optimal values of traditional capital per capita k^* and automation capital per capita p^* can be derived.

$$r^{trad} = r^{auto}$$
$$k^* = \alpha\tilde{k} + \alpha$$
$$p^* = (1 - \alpha)\tilde{k} - \alpha$$

Note, that $\tilde{k} = \frac{\tilde{K}}{N}$. If the growth rate of N, denoted by n, deviates from its balanced growth path value, this affects $\tilde{k}$. Specifically, if n decreases, $\tilde{k}$ increases. This also affects the optimal ratio of k and p, which decreases. Thus, if the population decreases, the optimal ratio of k^*/p^* decreases.

$$\frac{k^*}{p^*} = \frac{\alpha\tilde{k} + \alpha}{(1 - \alpha)\tilde{k} - \alpha},$$
$$\frac{\partial(k^*/p^*)}{\partial\tilde{k}} = \frac{-\alpha}{\left((1 - \alpha)\tilde{k} - \alpha\right)^2} < 0.$$

If the labor force is completely skilled $(e = 1)$, output per capita is given by

$$y = k^{\alpha} p^{\beta(1-\alpha)}.$$

Deriving the marginal effect of traditional and automation capital per capita:

$$r^{trad} = \frac{\partial y}{\partial k} = \alpha \frac{y}{k},$$
$$r^{auto} = \frac{\partial y}{\partial p} = \beta(1-\alpha)\frac{y}{p}.$$

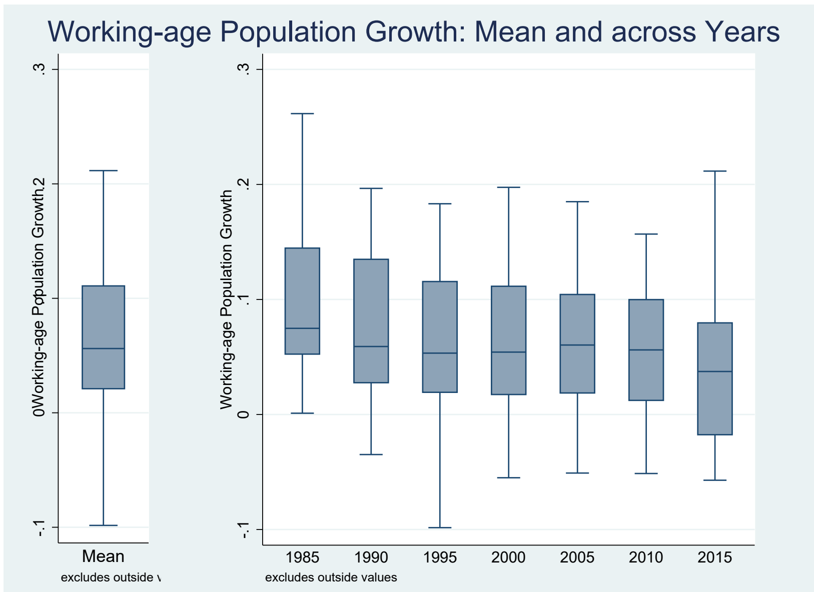
Setting equal due to capital mobility and reformulating yields

$$r^{trad} = r^{auto},$$
$$k^{*} = \frac{\alpha}{\alpha + (1-\alpha)\beta} \cdot \tilde{k},$$
$$p^{*} = \frac{(1-\alpha)\beta}{\alpha + (1-\alpha)\beta} \cdot \tilde{k},$$
$$\frac{k^{*}}{p^{*}} = \frac{\alpha}{\beta(1-\alpha)}.$$

If the workforce is fully educated, the two types of capital k and p are both allocated a constant share of total capital per capita $\tilde{k}$. Thus, the ratio $\frac{k}{p}$ is unaffected by population growth.

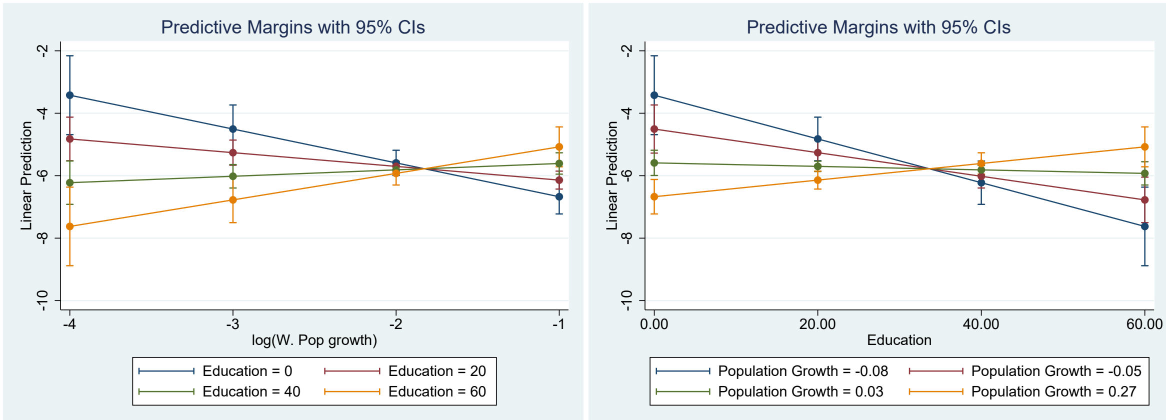
C.2 General Appendix

C.2.1 Additional Figures



Note: The figure illustrates the development of the working-age population growth rate over the years using a boxplot. In the left panel, the 25th, 50th, and 75th percentile of the distribution across years and countries is shown. The right panel shows the respective statistics across countries in 5-year intervals.

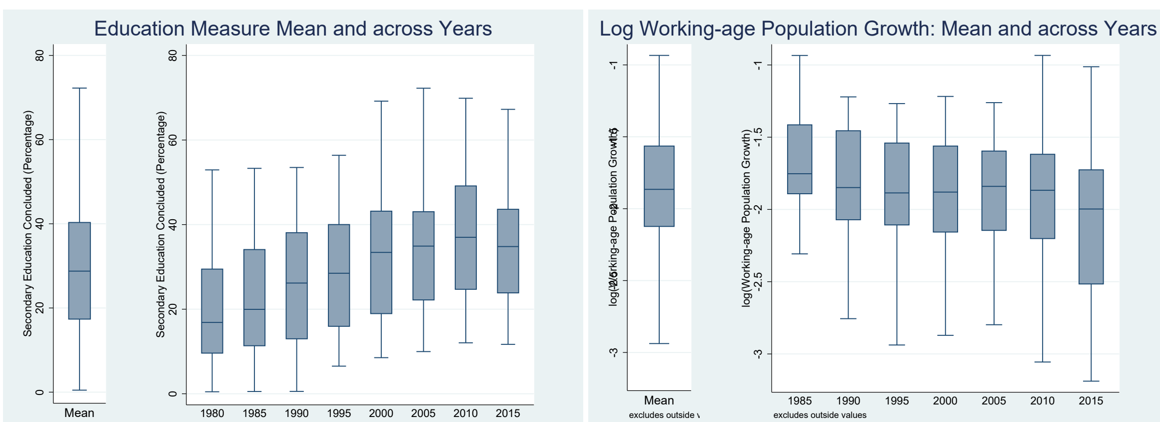
Figure C.1: *Visualization of Working-age Population Growth over Time*



(a) *Marginal Effect of Population Growth* **(b)** *Marginal Effect of Education*

Figure C.2: *Visualization of Marginal Effects from Table 3.1*

Note: The Figure illustrates the regression results reported in Table 3.1. Panel (a) shows the predicted marginal effect an increase in log working-age population growth has on automation density for different values of the education variable. Panel (b) shows the predicted marginal effect an increase in education has on automation density for different values of working-age population growth. In both cases, 95% confidence intervals of the predicted effect are reported.



(a) *Education* **(b)** *Log Working-age Population Growth*

Figure C.3: *Visualization of Variable Distribution Across Time*

Note: The figure illustrates the development of the education measure and working-age population growth over the years using a boxplot. In the respective left panel, the 25th, 50th, and 75th percentile of the distribution across years and countries is shown. The right panel shows the respective statistics across countries in 5-year intervals

To interpret the effect of working-age population growth on automation, it may be helpful to look at the marginal effect at certain levels of education. The distribution of education across years is visualized in Panel (a) of Figure C.3. Fixing the share of the educated population at 0%, 20%, 40%, and 60%, the respective marginal effects of working-age population growth are -1.1, -0.4, 0.2, and 0.8. These numbers are calculated as $\frac{\partial automate}{\partial popgrowth} = \eta_1 + \eta_3 \cdot educ$, inserting the values 0, 20, 40, and 60 for education. The analysis can also be visualized, which is done in Figure C.2. It shows the estimated marginal effect of working-age population growth on automation for the same fixed shares of the educated population. For low levels of education, represented by Education = 0 and Education = 20, the marginal effect of working-age population growth is negative and for high levels of education (Education = 40 and Education = 60) it is positive.

The same analysis of marginal effects is done for the education variable at different levels of working-age population growth. In that case, the marginal effect is calculated as $\frac{\partial automate}{\partial educ} = \eta_2 + \eta_3 \cdot popgrowth$. Again, this is calculated for some meaningful values of working-age population growth. In the data, the working-age population growth rate ranges from -0.10 to 0.30. The distribution of the growth rate across years and by years is visualized in Panel (b) of Figure C.3. For the visualization in Figure C.2b, the values -0.08, -0.05, 0.03, 0.27 are used.[3] At these values, the marginal effect of an increase in education on the density of automation patents is given by -0.07, -0.04, -0.01, 0.03, which corresponds to the respective slope of the lines. The negative sign for low levels of working-age population growth is driven by the effect working-age population growth has on automation density, whose negative effect outweighs the positive effect education has on automation density.

[3]Note, that these are the values before the rescaling of the population growth variable. $\log(-4)$ in the figure corresponds to a rescaled working-age population growth rate of 0.18, which, taking into account the rescaling, is equivalent to a growth rate of -0.08.

Graphical Analysis of Error Terms

(a) *Autocorrelation* **(b)** *Partial Autocorrelation*

Figure C.4: *Graphical Analysis of Time Series Estimation Error Terms*

Note: The figure shows the autocorrelation and partial autocorrelation of error terms from the time series analysis of US data, results of which are reported in Table 3.3.

Figure C.4 shows the autocorrelation and the partial autocorrelation of the error terms resulting from time series analysis, the results of which are reported in Table 3.3. There is no distinctive pattern in the error terms for different lag times. Furthermore, none of the correlations across lags are statistically significant, indicating that there is no serial correlation of the error terms.

C.2.2 Replication Regression Results

Table C.1: *Replication of Abeliansky and Prettner (2021) using Patent data 1993-2020*

	(1)	(2)	(3)
log(Pop Growth)	0.0112	-0.1481	-0.2754
	(0.06)	(-0.68)	(-1.35)
Investment Share	-0.0014	0.0011	0.0076
	(-0.20)	(0.14)	(1.03)
Constant	-1.7019***	-1.6279***	0.0000
	(-3.31)	(-2.82)	(.)
Time FE	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
R^2	0.25	0.10	0.08
Observations	311	247	250

Note: Dependent variable in column (1), column (2) and column (3) is the box-cox transformed growth rate of automation patents per capita as defined by the *auto*1, *auto*90 and *auto*95 measure, respectively. All explanatory variables are lagged by one period. log(Pop growth) is the box-cox transformed population growth rate. Investment share refers to gross fixed capital formation in % of GDP. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics are reported in parentheses.

Tables C.1 and C.2 show regression results from replicating Abeliansky and Prettner (2021) using patent data. The outcome of similar regression results using patent data covering the same time period cannot be obtained (Table C.1). However, one of the advantages of using patent data is that it covers a much longer time period than the robot data. While reliable data on robots starts in 1993, the OECD patent data goes back to 1977. When running the same regressions using the whole time period the patent data is available, the finding of a significantly negative relationship between population growth and automation growth can be replicated (Table C.2). This is robust to using different classifications and thus measures of automation patents.[4]

Table C.3 reports results from running a dynamic corrected fixed effects regression using log growth rates of automation measures as a dependent variable as in Abeliansky

[4]Probably due to reasons detailed in Section 3.3.2, this finding is limited to an analysis using IP5 patents, while analyses using Triadic family patents or PCT patents do not result in statistically significant coefficients.

Table C.2: *Replication of Abeliansky and Prettner (2021) using Patent data 1977-2020*

	(1)	(2)	(3)
log(Pop Growth)	-0.2419*	-0.2987**	-0.3484***
	(-1.94)	(-2.20)	(-2.62)
Investment Share	0.0057	0.0038	0.0097
	(1.07)	(0.63)	(1.64)
Constant	-0.5952*	-0.4360	-0.5800*
	(-1.86)	(-1.26)	(-1.70)
Time FE	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
R^2	0.59	0.56	0.59
Observations	429	336	336

Note: Dependent variable in column (1), column (2) and column (3) is the box-cox transformed growth rate of automation patents per capita as defined by the *auto*1, *auto*90 and *auto*95 measure, respectively. All explanatory variables are lagged by one period. log(Pop growth) is the box-cox transformed population growth rate. Investment share refers to gross fixed capital formation in % of GDP. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics are reported in parentheses.

and Prettner (2021). The autocorrelation is insignificant. The estimated coefficient of log population growth though is smaller and no longer significant. However, that may be due to the much-reduced sample size.

Tabel C.4 uses the same dependent variable as column (1) in Tables C.1 and C.2. In columns (2) and (3) education and an interaction term of education with the population growth variable are added. The results show clearly that including education is important, especially the interaction term. The coefficient of both education and the interaction term is statistically significant and positive. Additionally, the point estimate for the population growth variable has a higher significance and a higher magnitude.

Table C.3: *Corrected FE Estimates 1977-2020*

	(1)	(2)	(3)
$\log(\Delta auto1)_{t-1}$	0.114		
	(1.37)		
$\log(\Delta auto90)_{t-1}$		0.002	
		(0.02)	
$\log(\Delta auto95)_{t-1}$			0.059
			(0.36)
log(Pop Growth)	-0.124	0.099	-0.091
	(-0.63)	(0.27)	(-0.32)
Investment Share	0.003	0.002	0.014
	(0.39)	(0.20)	(1.36)
Time FE	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
N	289	141	144

Note: Dependent variable in column (1), column (2), and column (3) is the log growth rate of automation patents per capita as defined by the *auto*1, *auto*90, and *auto*95 measure, respectively. A three-year lag of the dependent variable is included as an explanatory variable in each regression. All other explanatory variables are lagged by one period. log(Pop growth) is the log of population growth. Investment share refers to gross fixed capital formation in % of GDP. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics are reported in parentheses.

Table C.4: *Replication of Abeliansky and Prettner (2021) adding Education data 1977-2020*

	(1)	(2)	(3)
log(Pop Growth)	-0.2419*	-0.2381*	-0.3663***
	(-1.94)	(-1.92)	(-2.61)
Education		0.0029	0.0151**
		(0.71)	(1.99)
log(Pop growth) $\times$ Education			0.0062*
			(1.91)
Investment Share	0.0057	0.0082	0.0083
	(1.07)	(1.50)	(1.52)
Time FE	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
R^2	0.59	0.59	0.60
Observations	429	426	426

Note: Dependent variable in all columns is the box-cox transformed growth rate of automation patents per capita as defined by the *auto*1. All explanatory variables are lagged by one period. log(Pop growth) is the box-cox transformed population growth rate. Education measures the share of the population with some tertiary education. Investment share refers to gross fixed capital formation in % of GDP. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics are reported in parentheses.

C.2.3 Additional Regression Tables

Using Different Education Measures

Table C.5: *Using Some Tertiary Education as Education Measure*

	(1)	(2)	(3)
log(W. Pop growth)	-0.1646	-0.1551	-0.5227**
	(-1.14)	(-1.07)	(-2.20)
Some Tertiary Education		-0.0053	0.0433
		(-0.56)	(1.63)
log(W. Pop growth) × Education			0.0259*
			(1.95)
Investment Share	0.0362***	0.0351***	0.0352***
	(3.47)	(3.29)	(3.32)
Time FE	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
R^2	0.60	0.60	0.61
Observations	328	325	325

Note: Dependent variable is the log of the automation measure *auto*1, constructed from patent data reported by the OECD and divided by working-age population to arrive at a per-capita measure. All explanatory variables are lagged by one period. log(W. Pop growth) is the log of working-age population growth. Education measures the share of the working-age population with at least some tertiary education as reported by Barro and Lee. Investment share refers to gross fixed capital formation as a share of GDP. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics are reported in parentheses.

Tables C.5 and C.6 report results from estimating the main regressions but using some tertiary education or completed tertiary education as an education measure. The coefficient on working-age population growth decreases both in magnitude and significance. The coefficient of the interaction term also varies in significance and magnitude. The overall findings however can be replicated quite well.

Table C.6: *Using Completed Tertiary Education as Education Measure*

	(1)	(2)	(3)
log(W. Pop growth)	-0.1646	-0.1660	-0.5611**
	(-1.14)	(-1.15)	(-2.41)
Completed Tertiary Education		0.0054	0.0881**
		(0.41)	(2.18)
log(W. Pop growth) × Education			0.0441**
			(2.16)
Investment Share	0.0362***	0.0368***	0.0371***
	(3.47)	(3.47)	(3.52)
Time FE	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
R^2	0.60	0.60	0.61
Observations	328	325	325

Note: Dependent variable is the log of the automation measure *auto*1, constructed from patent data reported by the OECD and divided by working-age population to arrive at a per-capita measure. All explanatory variables are lagged by one period. log(W. Pop growth) is the log of working-age population growth. Education measures the share of the working-age population with at least completed tertiary education as reported by Barro and Lee. Investment share refers to gross fixed capital formation as a share of GDP. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics are reported in parentheses.

Using Total Population Growth

Table C.7: *Total Population Growth and Automation Density*

	(1)	(2)	(3)
log(Pop growth)	-0.7332***	-0.7563***	-2.0402***
	(-3.23)	(-3.40)	(-5.11)
Education		-0.0008	0.0661***
		(-0.13)	(3.58)
log(Pop growth) $\times$ Education			0.0344***
			(3.83)
Investment Share	0.0390***	0.0381***	0.0467***
	(3.72)	(3.69)	(4.53)
Time FE	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
R^2	0.61	0.63	0.65
Observations	341	334	334

Note: Dependent variable is the log of the automation measure *auto*1, constructed from patent data reported by the OECD and divided by population to arrive at a per-capita measure. All explanatory variables are lagged by one period. log(Pop growth) is the log of total population growth. Education measures the share of the total population with at least completed secondary education as reported by Barro and Lee. Investment share refers to gross fixed capital formation as a share of GDP. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics are reported in parentheses.

Table C.7 reports results from estimating the main regressions but using total population growth instead of working-age population growth. All coefficients remain significant and even slightly increase in magnitude.

Using PTC and Triadic Patent Data

Table C.8: *Main Regression using Triadic and PCT Patents*

	Triadic			PCT		
	(1)	(2)	(3)	(4)	(5)	(6)
log(W. Pop growth)	-0.2155	-0.2149	-1.5596***	-0.4690**	-0.4747**	-2.3359***
	(-1.10)	(-1.09)	(-4.21)	(-1.97)	(-1.99)	(-5.31)
Education		0.0007	0.0859***		0.0125	0.1241***
		(0.08)	(3.93)		(1.28)	(5.08)
Interaction			0.0471***			0.0617***
			(4.23)			(4.94)
Investment Share	0.0217	0.0217	0.0189	0.0193	0.0183	0.0164
	(1.52)	(1.50)	(1.35)	(1.21)	(1.15)	(1.07)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.43	0.43	0.47	0.74	0.74	0.77
Observations	324	321	321	323	320	320

Note: Dependent variable is the log of automation patents per capita as reported by the OECD. In columns (1), (2), and (3), only automation patents registered with the EPO, the JPO, and the USPTO are used for the analysis. In columns (4), (5), and (6), only automation patents filed with the PCT are used for analysis. All explanatory variables are lagged by one period. log(W. Pop growth) is the log of working-age population growth. Education measures the share of the working-age population with at least completed secondary education as reported by Barro and Lee. Investment share refers to gross fixed capital formation as a share of GDP. The variable Interaction is defined as follows: Interaction = log(W. Pop growth) $\times$ Education. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics are reported in parentheses.

Table C.8 reports regression results from running the baseline specification but using Triadic patent data and PCT patent data instead. The estimated coefficients are very similar in magnitude and significance.

Using auto90 and auto95 Data

Table C.9: *Main Regression using auto90 and auto95*

	auto90			auto95		
	(1)	(2)	(3)	(4)	(5)	(6)
log(W. Pop growth)	-0.2070	-0.2097	-1.0128***	-0.5406*	-0.5432*	-2.0826***
	(-1.16)	(-1.17)	(-2.94)	(-1.83)	(-1.83)	(-3.88)
Education		0.0045	0.0594***		0.0067	0.0855***
		(0.58)	(2.76)		(0.86)	(3.52)
Interaction			0.0306***			0.0439***
			(2.72)			(3.42)
Investment Share	0.0383***	0.0378***	0.0380***	0.0359***	0.0352**	0.0407***
	(2.94)	(2.88)	(2.93)	(2.58)	(2.52)	(2.96)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.32	0.32	0.35	0.31	0.31	0.34
Observations	304	301	301	300	297	297

Note: Dependent variable is the log of automation patents per capita as reported by the OECD. In columns (1), (2), and (3), the dependent variable is calculated using automation patents as measured by the variable *auto*90. In columns (4), (5), and (6), the dependent variable is calculated using automation patents as measured by the variable *auto*95. All explanatory variables are lagged by one period. log(W. Pop growth) is the log of working-age population growth. Education measures the share of the working-age population with at least completed secondary education as reported by Barro and Lee. Investment share refers to gross fixed capital formation as a share of GDP. The variable Interaction is defined as follows: Interaction = log(W. Pop growth) $\times$ Education. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics are reported in parentheses.

Table C.9 reports regression results from running the baseline specification but using *auto*90 and *auto*95 patent data as dependent variables instead. The pattern of the different estimation specifications is similar to the baseline case. The estimated coefficients however are smaller and have a lower significance level. This is likely due to higher noise in the automation measures *auto*90 and *auto*95, compared to the preferred measure of *auto*1.

Subsample of G20 Member Countries

Table C.10: *Main Regression using only G20 Countries*

	(1)	(2)	(3)
log(W. Pop growth)	-0.6435	-0.6764	-2.3878***
	(-1.47)	(-1.56)	(-2.96)
Education		-0.0140	0.0615*
		(-1.57)	(1.95)
log(W. Pop growth) × Education			0.0408**
			(2.49)
Investment Share	0.1467***	0.1464***	0.1343***
	(7.29)	(7.34)	(6.72)
Time FE	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
R^2	0.71	0.72	0.74
Observations	113	113	113

Note: Dependent variable is the log of automation patents per capita as reported by the OECD. Only the subsample of G20 member states is used for regression. All explanatory variables are lagged by one period. log(W. Pop growth) is the log of working age population growth. Education measures the share of the working-age population with at least completed secondary education as reported by Barro and Lee. Investment share refers to gross fixed capital formation as a share of GDP. Significance stars are defined as follows: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. t-statistics are reported in parentheses.

Table C.10 reports regression results from repeating the regressions from Table 3.1 in the sub-sample of G20 member states. The pattern of results is the same as in the full sample.

C.2.4 List of Countries

There are 59 countries for which all variables necessary for estimating the main specification are available. They are:

Argentina, Australia, Austria, Belgium, Brazil, Bulgaria, Canada, Chile, China, Colombia, Croatia, Czech Republic, Denmark, Egypt, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, India, Indonesia, Ireland, Israel, Italy, Japan, Korea, Latvia, Lithuania, Luxembourg, Malaysia, Mexico, Morocco, Netherlands, New Zealand, Norway, Pakistan, Peru, Philippines, Poland, Portugal, Romania, Russia, Saudi Arabia, Singapore, Slovenia, South Africa, Spain, Sweden, Switzerland, Thailand, Tunisia, Turkiye, Ukraine, United Kingdom, United States.