CHARLES UNIVERSITY
FACULTY OF SOCIAL SCIENCES
Institute of Economic Studies

Determinants of Inequality Trends in the Czech Republic from 1992 to 2016

Bachelor's thesis

Author: Quang Linh Vo
Study program: Economics and Finance
Supervisor: Mgr. Marek Šedivý
Year of defense: 2023

Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources, literature and artificial intelligence tools in compliance with IES FSV UK, and the thesis has not been used to obtain any other academic title.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis in whole or in part and agrees with the thesis being used for study and scientific purposes.

Prague, July 24, 2023

Quang Linh Vo

Abstract

This thesis researches the determinants of income inequality trends in the Czech Republic from 1992 to 2016, focusing on the effects of socio-demographic characteristics. A mean log deviation decomposition is used to recognize the impacts of each factor on the inequality index. The variables of interest consist of the household's region, composition and the household head's demographic characteristics. It occurs that the demographic characteristics and size of a household had a relatively great between group components, nonetheless, the within effect appears to be larger. Then a linear regression on income is applied to observe the variables' effects and their significance. The results show us the significance of the highest completed education level for upper secondary and higher with positive influence. Whereas women household heads and regions except Prague had overall a negative effect at a significant level.

JEL Classification	D30, D63, P36, P46
Keywords	Inequality, Czech Republic, Decomposition, Socio-Demographic Characteristics
Title	Determinants of Inequality Trends in the Czech Republic from 1992 to 2016
Author's e-mail	84617831@fsv.cuni.cz
Supervisor's e-mail	marek.sedivy@fsv.cuni.cz

Abstrakt

Tato práce zkoumá determinanty vývoje příjmové nerovnosti v České republice v období mezi 1992 a 2016. Zaměřuje se zejména na vliv sociodemografických charakteristik. K určení dopadů jednotlivých faktorů na index nerovnosti je použita dekompozice střední logaritmické odchylky. Proměnné se skládají z regionu, struktury domácnosti a demografických charakteristik hlavy domácnosti. Ukazuje se, že demografické charakteristiky a velikost domácnosti měly poměrně velký vliv mezi složkami skupiny, nicméně vliv uvnitř skupiny se jeví jako větší. Následně je použita lineární regrese na příjem, aby bylo možné sledovat vlivy proměnných a jejich významnost. Výsledky nám ukazují významnost nejvyššího ukončeného vzdělání pro vyšší střední a výš s pozitivním vlivem. Zatímco ženy v čele domácnosti a kraje s výjimkou Prahy měly celkově negativní vliv na signifikantní úrovni.

Klasifikace JEL	D30, D63, P36, P46
Klíčová slova	Nerovnost, Česká Republika, Dekompozice, Sociodemografické charakteristiky
Název práce	Determinanty vývoje nerovnosti v České republice v letech 1992-2016
E-mail autora	84617831@fsv.cuni.cz
E-mail vedoucího práce	marek.sedivy@fsv.cuni.cz

Acknowledgments

My greatest gratitude goes to Mgr. Marek Šedivý, who was always available and gave me his support, expertise, time and patience. Then I would like to thank my family and friends for their trust, encouragement, advice and cheer during the whole period of writing this thesis. I am also grateful to the Luxembourg Income Study Cross-National Data Centre for access to the data, without which this work would not exist.

Thank You

Typeset in FSV LaTeX template with great thanks to prof. Zuzana Havrankova and prof. Tomas Havranek of Institute of Economic Studies, Faculty of Social Sciences, Charles University.

Bibliographic Record

Vo, Quang Linh: *Determinants of Inequality Trends in the Czech Republic from 1992 to 2016.* Bachelor's thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Prague. 2023, pages 81. Advisor: Mgr. Marek Šedivý

Contents

List of Tables

List of Figures

Acronyms

LIS Luxembourg Income Study

IQR Interquartile Range

CSU Czech Statistical Office

MLD Mean Log Deviation

METIS Metadata Information System

SILC Survey on Income and Living Conditions

HH Household

DHI Disposable Household Income

Chapter 1

Introduction

Income inequality is a complex and omnipresent issue affecting society globally, including the Czech Republic. There exist options on how to measure it by computing the Gini index or percentile ratios, which both indicate an increase in income inequality in Czechia from 1992 to 2016. Nonetheless, they do not show what drives the changes in income inequality. A decomposition method was introduced to analyze the impacts of different sources. In this thesis, we apply the tool to describe income inequality trends in Czechia in a period from 1992 to 2016.

The research on the impact of socio-demographic characteristics on overall income inequality in Czechia is the central purpose of our thesis. Since its establishment in 1993, the population characteristics shifted significantly in the economic and social environment in last 30 years. The country underwent a significant transformation due to a political change after the fall of communism, the transition to a market economy in the 1990s and the entrance to the EU in 2004. A rise in Czech income inequality from 1992 to 2016 is displayed by the Gini index and 90/10 ratio. The former indicates growth from 19.3 to 24.5 points according to data from the Luxembourg Income Study (2023b) database, and World Data Bank (2020) also shows a shift from 20.7 to 25.4 points. The latter compares the household incomes by taking a ratio between the top 10% and the lowest 10%, calculated from Luxembourg Income Study (2023b) data shows that the inequality rose from 2.2 to 2.9 points.

Thus for our research question, we would like to focus on households in the Czech Republic and their attributes which are their composition, size, number of members, household head's age, sex, marital status, and education. With an application of the decomposition method, we investigate how these individual factors impact income inequality.

The method of interest, decomposition, is a statistical approach that decomposes or disintegrates a measure of income inequality, such as Mean Log Deviation and Gini index, into fundamental sources like demographics, the labour market, income sources, prices, or redistribution policies. The technique identifies the inequality contributors and analyzes the importance of individual elements. To examine the effects of various factors on income we operate with a regression analysis. In our case, we work with Mean Log Deviation decomposition and linear regression.

We obtained data from the *Luxembourg Income Study* database which processed data from *Czech Statistical Office*'s *Microsensus* and *Survey on Income and Living Conditions*. The database possesses eight datasets from the period covered with sample sizes of 7 000 to 28 148 households in each year 1992, 1996, 2002, 2004, 2007, 2010, 2013, and 2016.

Our work managed to show possible contributors to inequality as education, gender, age and marital status have high between-group effects when decomposing the Mean Log Deviation. Other possible factors were the household type, similarly the number of household members and the household head's youngest child's age.

For the household level regression on income part, we examined each year separately. We observed that the variable of household heads with the highest completed education level being upper secondary and more was strongly significant with a large effect. On the other hand, we estimated that female household heads possessed less income than males.

Chapter 2 reviews the literature, specifically a summary of related works about a decomposition method to research income inequality in the Czech Republic. In addition, it mentions the results of other income inequality studies in Taiwan, Russia, Brazil, and Lithuania which were examined by decomposition methods. Chapter 3 describes a methodology, explaining a decomposition method, including its steps and computations, involving the regression analysis while defining the models. Chapter 4 outlines an overview and discussion of data, its source, the process of how they were obtained, details about the database and work usage, variables, and their description. Chapter 5 presents a description of the results of the study, tables, graphs, and primary interpretations of outcomes. Chapter 6 discusses a summary of the findings and contribution of the work, reflecting on the technique and making suggestions for future work.

Chapter 2

Literature Review

Since the establishment of the Czech Republic, researchers have created several inequality studies and examined numerous causes. Večerník (2013) looks into the effects of education on household income and in his other work Večerník (1991), where the author examines the intertemporal changes of earnings and the influence of age or gender before the Velvet Revolution. Křížková *et al.* (2010) investigates gender inequality, claiming that women earn 10% less for the same job. The within-job wage inequality exists in the local labour market and remains stable in the period covered according to the previous paper.

Švejnar (2002) suggests that the transition from a plan to a market economy supported the rising inequality by analyzing the Gini coefficient. Another paper by Matějů & Lim (1995) examines the aftermath of the end of the communist regime, which also studies the cultural and social capital, including the returns of education or the benefits from the *nomenklatura* cadre system. The relative influence of government policies such as transfer and tax changes at the beginning of the 1990s is contributed by Garner & Terrell (1998), while the redistribution policy consequences are covered by Immervoll & Richardson (2011), involving tax-benefit systems, market-income inequality or social contributions.

Besides, the number of papers using a decomposition method is not highly rich, therefore we would like to contribute to this topic. Novotný (2007) estimates regional inequality in 46 countries including the Czech Republic, using the decomposition of the Theil index. On the other hand, the Gini coefficient decomposition is analyzed thanks to Garner & Terrell (1998). Mysíková & Večerník (2018) operate with regression-based decomposition to examine the income inequality from 2004 to 2010.

2.1 Inequality studies

Initially, we take a look at possible income dispersion origins, mainly socio-demographic characteristics such as age, gender, education or housing. Besides, policy and political change in Czechia studies will also be mentioned.

Before the end of the communist regime in Czechia, Večerník (1991) explores the earning differentials developed during the normalisation era. Back then, Czechoslovakia performed as a country with a stable income inequality with a small range. The author is concerned that the income surveys were done essentially to support the ideology by showing rising living standards, meanwhile, inequality was not investigated at all. This article uses data from the Microsensus surveys but also from the LIS for cross-national comparison. The analysis discovers disparities by sex, age, education and economic sector, which might be influenced by economic and political preferences or the role of party membership. Additionally, Vecernik highlights that labour-market exposure of women and young workers could further fuel the differences, where the older generation might benefit from their economic and social capital.

Matějů & Lim (1995) conduct research to investigate the impacts of social capital gained in the normalization era on inequality in the transformation period from 1989 to 1992. Although their data consist of only middle-aged individuals, who were nearing their career high when the transformation occurred, they claim that the earnings inequality increased, despite their other finding of an intact redistribution system in 1989 from the previous era. Therefore, even with the transformation, the distribution system still featured elements of state socialism, nevertheless, the disparity widened. The authors argue that citizens that were in a party, or illegally selling scarce products were in an advantageous position as they obtained social capital during communism.

Kostelecký (2000) investigates the effect of the housing market in the transition period. He concludes that housing and social inequality cannot be compared equally, yet they both might have much in common. The housing market impact was unique, because of the change from state control and regulation to a market economy, privatization and restitution, rising diversity in regions (an enormous difference between Prague region and Ostrava region), and different sectors of the market separated into several legal and financial conditions. Since the fall of communism, the number of apartment constructions declined significantly, while newly established private firms were the reason for turning flats into offices. Also, the population rose as people arrived from abroad and children of the population boom in the 1970s were at the age of setting up families, increasing the demand for housing.

The role of education in personal and household income disparities before and after communism collapse is studied by Večerník (2013). Not only he examines the effects of education, but also of gender, age of individuals, labour market, including occupation and industry, and household composition. The author compares both periods and emphasized that before 1990, inequality was more impacted by age and gender than by education, industry and occupation. Nevertheless, after the Velvet Revolution, the age and gender influences decreased, meanwhile the individuals' market abilities contribution increased.

Vecernik explains the fact that these shifts occurred due to the modification of household income and systematic change. Both numbers of active earners and children were more important until then, however, with new opportunities and free choices the number of children in families declined, but also the number of married women that work declined. The systematic change made the role of education significant since the personal returns to education doubled, the impact of education levels of married couples increased and employment and earnings of women also rose. These results come from the analysis of data on the gender wage gap, how many children the household had and when, the women's economic activity and selective mating.

Křížková *et al.* (2010) suggest that motherhood was an important factor driving inequality in combination with policy and legislative measures for gender equality, as employers were avoiding hiring women. On the other hand, establishments more connected to the state socialism past such as state-owned firms resulted in a smaller gender wage gap. Meanwhile, the biggest gap was observed in managerial positions. However, these findings are mainly at a personal level.

2.2 Inequality Decomposition Studies in the Czech Republic

Garner & Terrell (1998) apply the Gini decomposition on inequality in Czechia, using household data. The work is inspired by Lerman & Yitzhaki (1985), Lerman & Yitzhaki (1989) and Lerman & Yitzhaki (1994), where they use the source decomposition method in the United States. Garner & Terrell (1998) use earned income, other income, government transfers and taxes as the variables of interest and studied the slightly increased inequality in the 1989-1993 period. Their goal is to determine whether government policies or market forces contributed to low overall inequality. Firstly, they claim that government policies slowed down the effect of the market system on income inequality, thanks to the tax and transfer system. In Czechia, the transfer component contributed more to reducing disparities than the tax component, as it protected the pensioners. Nonetheless, the market system supported the poor as the minimum living standard was set and that helped the moderation in inequality.

Mysíková & Večerník (2018) study personal income distribution in the Czech Republic from 2004 to 2010, covering before and after the crisis. They apply a regression-based decomposition proposed by Fiorio & Jenkins (2007), with gender, work experience, education, occupation, change of job status and city size. Their findings indicate that type of occupation impacted inequality the most, the effect of education more than doubled in the period after 2007. On the other hand, the contribution of the gender variable weakened.

2.3 Inequality Decomposition Studies in Other Countries

Furthermore, we take an insight into other analyses of other countries utilizing the decomposition methods to reveal the structure of inequality trends.

The impact of education reform in Taiwan from 1979 to 1994 on earnings distribution is researched by Bourguignon *et al.* (1998), where they use their decomposition method based on the micro-simulation technique. The outcomes of this paper indicate the increase in household income inequality due to returns of education and participation behaviour, whereas the expansion of education had a weak equalizing impact.

Černiauskas *et al.* (2021) examine a decomposition by demographics, labour market structure, returns and prices and tax-benefit system to follow the trends of income distribution in Lithuania in the 2007-2015 period. The paper's findings show an immense rise in inequality due to growing returns in the labour and capital markets, despite the tax and benefit system slowing down the disparities from the upwardly moving market income effect during the crisis from 2007 to 2011. Additionally, the marriage rates decline also contributed to the source of inequality.

Lisina & Van Kerm (2022) provide income inequality decomposition in Russia, which concentrated on the timeframe from 1994 to 2015. The authors claim that the most significant factor of inequality change was shifts in incomes. Rising public sector and pensions earnings and lessening differences in private sector incomes lead to declining inequality. The work also implies that employment structure and level, and socio-demographic characteristics alone did not notably influence inequality.

Meanwhile, Ferreira & Barros (1999) decompose the interval between 1976 and 1996 in Brazil by microsimulation as they examine occupational and education choices and labour income. The authors doubt the non-changed income distribution in the period covered. Nonetheless, their research indicates a combination of forces of those variables which consequently raised the poverty level.

Chapter 3

Methodology

We rely on two inequality measures; the Gini coefficient and the Mean Log Deviation. We cover both of them theoretically involving a method of computation and the meaning of their results. We notably refer to Ravallion (2016), who summarizes inequality measurements and the intuitions behind them. Nevertheless, we begin with data modification for inequality research.

3.1 Data modification

After we obtain the data, we apply top and bottom coding, equivalising and merging. The use of top and bottom coding helps to restrict extreme values on both sides. Neugschwender (2020) shows that not only negative values are undesirable, but also high outliers might cause biased outcomes. Although a part of the data can be lost and inequality may not be as detailed and accurate, Neugschwender (2020) highlights that such values might arrive from measurement error. Therefore, the application lowers variability in both tails, which eventually leads to reducing different distributions of measurement error of each dataset. Consequently, the top and bottom coding unifies the datasets for a more equal comparison.

We then adjust the income by CPI and PPP deflators provided by the LIS and World Bank Development Indicators. The deflators are calculated to 2017.

The author implies a standard for LIS, setting the bottom line to be cubed below the interquartile range (IQR) of disposable household income, whereas the top line is cubed above it. The IQR benefits are in hand with the reasons for the coding as it is very insensitive to extreme values and it can be standardized to other data.

$$\text{botline} = \frac{Q_1}{\left(\frac{Q_3}{Q_1}\right)^3}$$

$$\text{topline} = \frac{Q_3}{\left(\frac{Q_3}{Q_1}\right)^3} \tag{3.1}$$

Where Q_1 is a 25$^{\text{th}}$ percentile of disposable household income, Q_3 a 75$^{\text{th}}$ and (Q_3/Q_1) is IQR, describing a relative size of Q_3 compared to Q_1. The IQR is of power of 3 as we are cubing it.

LIS recommends using one of the adjustment tools, equavalised income, allowing us to deal with possible economies of scale within a household and to observe more accurate outcomes as it allows for comparison between households of different sizes and compositions. For example, a single household with the same income as a household with a couple and two children might have a different standard of living. Therefore, we divide the top- and bottom-coded disposable household income by a square root of the number of household members. The square root serves as scaling or weighting since more members in a household may share their expenditures.

$$\text{equivalised disposable household income} = \frac{\text{top and bottom coded disposable household income}}{\sqrt{\text{number of household members}}} \tag{3.2}$$

Lastly, we modified our data by merging the personal level dataset into the household one, as we desire to have the socio-demographic characteristics of household heads. We filtered only heads of households, selected our chosen personal variables and assigned them by household identifier.

3.2 Mean Log Deviation

Mean Log Deviation (MLD) belongs to inequality measures with properties such as being more sensitive to the lower-tail values or having a scale where equalling zero means no inequality, while the upper bound is unlimited. The data of our distribution undergo a logarithmic transformation and then we use that as an input to our MLD, to observe how the values deviate from the mean. Then we decompose it by our socio-demographic characteristics to discover the inequality trends.

For MLD, we assume income distribution $y_1, y_2, ..., y_n$, where every $y > 0$

and have a mean $\overline{y}$, then:

$$\text{MLD} = \frac{\sum_{i=1}^{n} \ln\left(\frac{\overline{y}}{y_i}\right)}{n}$$

Afterwards, we show its decomposability by population subgroups as for MLD we consider n individuals allocated to N mutually exclusive groups. Let y_{ij} be individual i's income, where i belongs to group j. Group j consists of n_j individuals and has a population proportion denoted as p_j, then:

$$\text{MLD} = \ln \overline{y} - \sum_{j=1}^{N} p_j \sum_{i=1}^{n} \ln y_{ij} \tag{3.3}$$

This equation can be divided into within- and between-groups parts:

$$\text{MLD} = \overbrace{\sum_{j=1}^{N} p_j(\ln \overline{y} - \sum_{i=1}^{n_j} \ln y_{ij})}^{\text{Within}} + \overbrace{\ln \overline{y} - \sum_{j=1}^{N} p_j \ln \overline{y}_j}^{\text{Between}} \tag{3.4}$$

where $\overline{y}_j$ stands for the group j's earnings mean. Therefore, the interpretation of whole inequality is the sum, weighted by population, of inequality within and inequality between groups.

3.3 Gini coefficient

The Gini coefficient relates closely to the Lorenz curve. The Lorenz curve presents a graphical illustration of income disribution in the population. The share of total income going to the poorest x percent of the population lies on the y-axis, while the share of the population ranked by income is on the x-axis. The upward sloping curve with an increasing slope then shows how the income is distributed. For illustration, a perfect diagonal line means perfect equality, conversely, the more curved the line is, the more inequality occurs. If the richest person in the population possesses all of the income, the curve would be right-angled, covering the whole area below the diagonal line.

According to Ravallion (2016), doubling the space between the Lorenz curve and the 45-degree line equals to the Gini coefficient. In more detail, the Gini coefficient is defined as the average absolute difference between all couples of incomes in the population divided by the mean multiplied by two, as shown below.

$$Gini = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}|y_i - y_j|}{2n^2\overline{y}} \tag{3.5}$$

To obtain this equation, we have to begin with the average absolute difference. For the average absolute difference, we assume our disposable household income distribution of $y_1, y_2, ..., y_n$ with a mean of $\overline{y}$ and individuals i and j. Thereafter $|y_i - y_j|$ is an absolute difference between those individuals' earnings and with all the n^2 couples of earnings, the average absolute difference would be:

$$\text{average absolute difference} = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}|y_i - y_j| \tag{3.6}$$

To not have values over one, we consider the highest possible value where the richest individual possesses 100% of the earnings ($n\overline{y}$). Therefore, we adjust the average absolute difference. Ravallion (2016) emphasizes that would require:

$$\text{Gini Coefficient} = \frac{\text{average absolute difference}}{2\overline{y}} \tag{3.7}$$

as the index is $1 - (\frac{1}{n})$ when the individual has all of the earnings and in the $\lim_{n\to\infty}$ the index would achieve 1. Therefore, the interval of the index lies between 0 and 1, where 0 means perfect equality and 1 perfect inequality.

3.4 Multiple Regression

As we want to study the contributing effects of socio-demographic characteristics to income, we are using the ordinary least squares method on our multiple regression model.

To study the relationship between a dependent variable and independent variables, the method of OLS estimates the coefficients of a linear regression model. The estimating of unknown estimates is done by minimizing the sum of the squared residuals between the explained variable and the explanatory variables.

We consider the following variables: equivalised disposable income (**eqdhi**), region of the household (**region_c**), number of household members (equalised for (**nhhmem**), number of household members older than 65 (**nhhmem65**), number of household members younger than 13 (**nhhmem13**), age of the youngest child living in the household (**ageyoch**), highest completed education level (**educlev**), age in years (**age**) and (**age squared**), gender (**sex**) and marital status (**marital**). Out of the previously listed, the following are dummy variables: **region_c**, **educlev**, **sex** and **marital**.

For **region_c** in 1992 and 1996 data we have **South Bohemia**, **West Bohemia**, **North Bohemia**, **East Bohemia**, **South Moravia**, **North Moravia** and the intercept corresponding to **Prague**. **region_c** dummies for models covering period from 2002 to 2016 are **Stredocesky**, **Jihocesky**, **Plzensky**, **Karlovarsky Ustecky**, **Liberecky**, **Kralovehradecky**, **Pardubicky**, **Vysocina**, **Jihomoravsky**, **Olomoucky**, **Zlinsky**, **Moravskoslezsky**, where the intercept corresponds to region **Prague**.

For 1992 and 2002 models, the **educlev** dummy consists of **lower secondary**, **upper secondary**, **high, tertiary** and **less than primary** defined as 1. For 1996 the **educlev** is almost the same, only **high, tertiary** is replaced and we unroll it into **bachelor or equivalent**, **master or equivalent** and **doctorate or equivalent**. The year 2004 has **lower secondary** as 1, **upper secondary**, **post-secondary non-tertiary**, **bachelor or equivalent**, **doctorate or equivalent**.

Model for 2007 we define the dummy with values **primary**, set to 1, **lower secondary**, **upper secondary**, **post-secondary non-tertiary**, **short-cycle tertiary**, **bachelor or equivalent**, **master or equivalent** and **doctorate or equivalent**. The year 2010 is identical, but additionally, we have **never attended**, which is set to 1. On the other hand, the data from 2013 has **lower secondary** corresponding to the intercept, and then **upper secondary**, **post-secondary non-tertiary**, **short-cycle tertiary**, **bachelor or equivalent**, **master or equivalent** and **doctorate or equivalent**. In the last year, 2016, we have similar values to the previous model, but without **post-secondary non-tertiary** option. We decide to not group the categories for **educlev**, because we are interested in the contributions of the tertiary education level as detailed as possible. However, by not grouping them we are limiting ourselves to a better comparison. The next dummy variable is **sex** with **male** and **female**, where the former corresponds to the intercept. Lastly, **marital** is divided into **divorced**, **never married/not in union**, **widowed** and **married** set to 1.

We then check the assumptions of all our OLS models to maintain the quality of our outcomes. We apply the Breusch-Pagan test for homoskedasticity and the Shapiro-Wilk test for normality. In case of rejecting the null hypothesis of the tests, we use White's robust standard errors. For multicollinearity, we have the Variation Inflation Factor test.

Our model is then defined in the following figure:

$$
\begin{aligned}
\log dhi = {} & \beta_0 + \beta_1 \text{region}_\text{c} + \beta_2 \text{nhhmem} + \beta_3 \text{nhhmem65} \\
& + \beta_4 \text{nhhmem13} + \beta_5 \text{ageyoch} + \beta_6 \text{educlev} \\
& + \beta_7 \text{age} + \beta_8 \text{age}^2 + \beta_9 \text{sex} \\
& + \beta_{10} \text{marital} + \epsilon
\end{aligned} \tag{3.8}
$$

Chapter 4

Data

For all data used in this thesis, we rely on the cross-national data centre *LIS* database. It works as a data archive to research institutional factors, socio-economic outcomes and cross-national analyses. Luxembourg Income Study (2023a) consists of microdata in the last five decades period of 53 countries from Europe, North and South America, Africa, Asia, and Australasia, however, not every country has data for each decade.

We will use the *LIS* not only because it's the best available resource for the data at hand as the collection features both household-level and individual-level data of a high number of variables, including demography, employment, labour and capital income, public social benefits and pensions, private transfers, taxes and expenditures. All the accessible variables for each country and each year can be checked via *METIS, Metadata Information System*, which provides access to documentation, definitions, overview, comparisons, and also cross-comparing, for both variables and datasets from Luxembourg Income Study (2023c).

The surveys for data from the Czech Republic were contributed by the *Czech Statistical Office (CSU)* are available in the LIS for years 1992, 1996, 2002, 2004, 2007, 2010, 2013 and 2016. The first three obtained years were named Microsensus. The frequency of Microsensus was every three to five years, a coverage targeted the non-institutionalized households, the sampling procedure was a stratified random sample with two steps; the size of the locality and the number of apartments within the census tract. A personal interview was conducted to collect data, with household-level questions such as economic aspects, type and composition, individual and dwelling characteristics, like employment, or cash receipts from that year. The observation unit of an economic

Table 4.1: Sample Sizes of Datasets and their Non-Response Rate

	1992	1996	2002	2004	2007	2010	2013	2016
Number of Households	16 234	28148	7973	4351	11294	8866	8053	8701
Non-Response Rate	15.7%	23.8%	No data	35.2%	16.9%	14.8%	19.14%	20.2%

Source: Author base on LIS

household was then defined as individuals living together with a common budget. One or more economic households occupying the same housing made a sampling unit (a housing household), with a household head being a male. However, in cases where a male is not available, economic activity was the indicator.

Non-response rate is an aspect of households that did not respond because of refusal, non-contact, incapacity to participate or other reasons. These values were imputed and weights were assigned to each sample case.

Afterwards, *Survey on Income and Living Conditions (SILC)* replaced the Microsensus. The SILC more precisely obtain information for poverty and inequality research or the impact of policies, adding to the Microsensus' objective of measuring the quality of living and gathering household income data. Frequency quickened to annual, private or non-group living and non-institutional household coverage was introduced. The sampling frame was two-staged for the years 2004, 2007 and 2010 where the first stage became a unit stratification and the second worked as selecting families to cover the full country. For 2013 and 2016, the procedure in the second stage evolved into a more detailed sample choice due to the addition of administrative data. The questionnaire expanded with EU modules about the transmission of poverty between generations, financial exclusion and overindebtedness, or material deprivation among children.

The definitions also went through a change. Households are described depending on finance and expenditure sharing, acknowledgement of individuals in the unit permanently occupied plus members being temporally away for less than six months, with the exception of students dependent on household and having the same private address.

The overall numbers of households are shown in Table 4.1. The 1992 survey was carried out in 1993 with incomes of 1992, Czechia was still part of Czechoslovakia, until January 1993. Nonetheless, the data provided by the LIS are only for Czech households. The same goes for every other survey. The unit of time for wages and security transfers was monthly, everything else annually.

The sole exception was 2002 which defined everything per annum. For the year 2002, the non-responsive rate is unknown.

The thesis would study age, sex, number of own children living in the household (or also a number of household members 13 or younger or also the age of youngest own child living in the household), number of household members, household composition, marital status, education (3-category recode or highest completed education level or years of education).

Below we put tables with the evolution of each variable and various inequality indices.

4.1 Households' Distribution in Regions

Table A.1 describes the proportions of households in the Czech regions for the period covered. In 1992 and 1996 the distribution follows the NUTS2 level. From the 2002 dataset, all samples follow the NUTS3 level.

Nonetheless, only the 1992 and 1996 collection have its values in a different system. It was divided into eight regions where South and North Bohemia represented at least 19% of the sample size. North and East Bohemia account for 12% while regions Stredocesky and Prague have a little over one-tenth of the overall size. South and West Bohemia areas portray the sample population the least.

The survey data shows the socio-demographic structure and changes in the regions of the Czech Republic. We observe that the proportions shift throughout the period as an integral migration takes place. Among the causes of domestic mobility might be mainly economic reasons like labour market changes and job opportunities, then suburbanization, amenity migration, or environmental issues.

From 1996, we observe from the samples that the most answers were from the regions Moravskoslezsky, Prague, Jihomoravsky and Stredocesky. The number of respondents from Moravskoslezsky region peaks in 2010, but then falls steadily. Prague‘s number of respondents has a greater variance, for example in 1996 and 2002. The proportion of the Ustecky region declines throughout the observed period. Vysocina, Kralovehradecky and Pardubicky regions share a similar cut of people asked by the conductors of the survey. Lastly, the Karlovarsky region's respondent figure falls by half from 2004 to 2016.

Hubl (2020) claims that the suburbanization process after 2000 took place. From our data, we can assign the effect to the Prague region. Hampl & Müller

(2020) state that Prague and Stredocesky region supremacy is due to their location, quality of life and infrastructure. Jihomoravsky region benefited also from suburbanization.

Young people, of age 20 to 34, tend to migrate due to economic and cultural reasons, which according to Ivan & Tvrdý (2007) and Hubl (2020), is causing a brain drain from the region Moravskoslezsky to Prague region. Tonev *et al.* (2022) contribute with their finding that the population with a high school education or higher is keen to move, even for further distances. They also noted the impact of housing opportunities, as well as Hubl (2020).

Moravskoslezsky, Ustecky and Karlovarsky regions' decline explains Hubl (2020) by structural issues such as unemployment, changes in the industry and environmental concerns. Whereas in regions Jihocesky, Vysocina Hubl (2020) notes amenity migration due to the rural and environmental attraction. The impact was driven by the older population.

4.2 Household Composition

Another section follows household composition variables such as household type (Table A.2), number of household members overall (Table A.3), over 65 (Table A.4) and under 13 years of age (Table A.5), number of children (Table A.6) and age of the youngest child (Table A.7).

The development of most represented household types shifts enormously from 1992 to 2016 as singles households' share rapidly rises by almost 50% . Couples without children rise slightly, whereas the proportion of couples with children drops by 15 percentage point. The share of divorced or single parents follows an upward direction. Moreover, we find that two members' households increased from 29.88% to 36.11%, nonetheless, larger units generally decline.

Remiášová (2021) compares these trends in Czechia to the West, meaning an ageing population and fewer children born. Our data shows that nuclear families diminish. In addition, women desire to pursue a career, causing them to postpone having children or decide to not at all. Ondřichová (2014) further researches that partners do not wish to have more than three children, again due to economic reasons. They claim that the ideal is two, however, this case can be spoiled by health issues, unwillingness, or a divorce. Statistics from Czech Stastical Office (2011) show the reasons behind divorces in the Czech Republic in 2010. A majority comes from different opinions and personalities, thereafter for example infidelity or alcoholism.

Remiášová (2021) points out that a large proportion of pensioners live alone, especially widows, as the share of households with a 65-year-old person rises noticeably. On the other hand, the portion of households without them declines considerably. One pensioner household may be due to the passing away of a partner or living with children.

We observe in children's statistics a similar indication: the volume of households without a child increases, meanwhile with a child declines, measurably. This development could stem from an ageing population, as better healthcare and quality of life are achieved, therefore prolonging life expectancy.

Czech people postpone or decide to not have a child. No children in a household also can indicate that older people live alone in a household and more citizens live longer, enlarging the ratio.

Table A.7 provides an insight into the rising age mean, which could be explained by the ageing distribution among people or that children could leave the households at a later age.

4.3 Education

Table A.9 reveals the evolution of education. In general, from what we observe, the rising highest education levels of the population are a positive sign. First, the school system left communism and its principles by going through liberalization, decentralization and building upon democratic values, according to Krejčová (2012).

Secondly, multi-year grammar schools were reinstated, therefore simplifying the way to the college. Nonetheless, Krejčová (2012) notes that those grammar schools pull away pupils from elementary schools, causing a decrease in their quality. This issue is pointed out even by OECD. Moreover, during the 90s, there was no strategy or clear education system. In 2004 the Czech Republic entered the European Union, which might have caused the brain drain of bachelors.

Another enormous benefit is generally free education, including the tertiary level. The high school leaving exam was a subject of discussion since the beginning of our period, however, the reform in 2004 still did not specify the exam, which occurred seven years later.

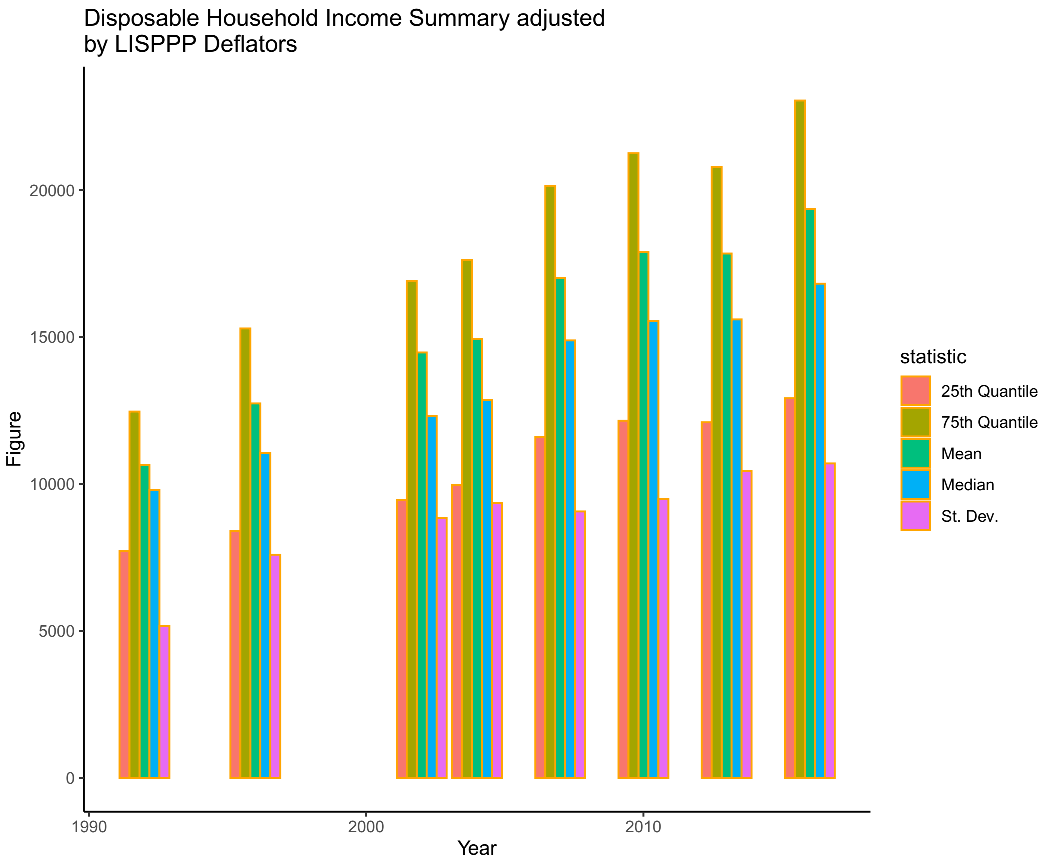
4.4 Demographic Characteristics of Individuals

We observe demographic characteristics in Table A.10, Table A.12 and Table A.13. The table about age again indicates a longer life expectancy of household heads as both the mean and the median are rising. In original datasets, women are represented with a small majority. However, for household heads, the men are represented more, which can be seen in Table A.12. Nevertheless, the women's portion is rising during the period.

We look into marital status distribution in Table A.13 and see obvious trends. For the first two datasets 1992 and 1996 option *missing values* exists, we assume that from 2002 it merges with *never married/not in union*.

Married people are slowly declining while almost double the size is the portion of divorced people. As we previously mentioned, divorce is at an increasing rate with reasons for example incompatibility, infidelity or alcohol issues. The lower rate of married people can be caused by divorce or lately, partners just do not get married due to economic reasons or not seeing the status to be much important. Widowed people could be an indicator of longer life expectancy of women than men.

Figure 4.1: Evolution of Disposable Household Income Adjusted by LISPPP



4.5 Disposable Household Income

The LIS define the Disposable Household Income as a sum of income sources, transfers, taxes and social contributions. There are monetary and non-monetary labour income and monetary capital income (non-monetary excluded due to the unavailability of the microdata and different methods of calculation), monetary and non-monetary private transfers, monetary social security transfers (work-related insurance universal and assistance transfers included), non-monetary social assistance transfers, minus the income taxes and social contributions paid.

The DHI in Czechia rises across the period if we observe all the measures. Moreover, right-skewness occurs as the mean is higher than the median, 25$^{\text{th}}$ percentile is closer to the median than the 75$^{\text{th}}$, and the standard deviation may indicate the expanding gap in the population.

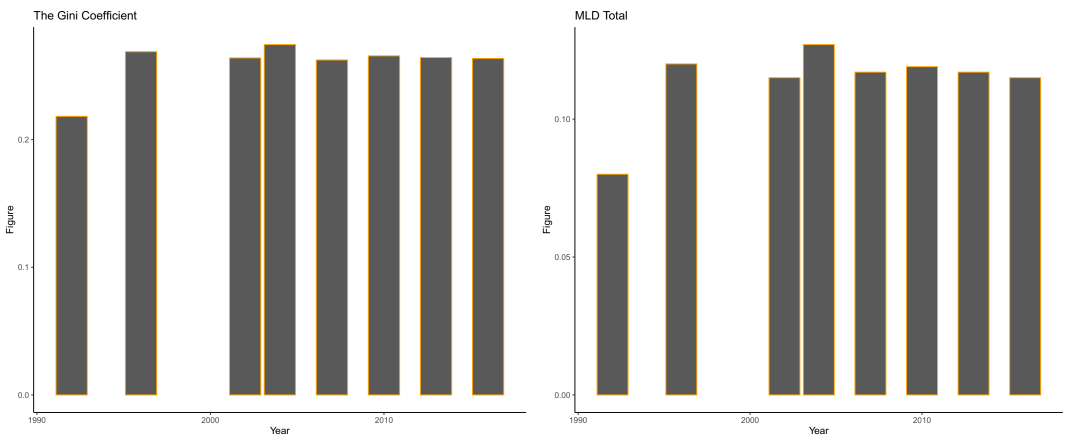
Chapter 5

Results

5.1 Inequality Indices

We are using income inequality indices Gini coefficient and Mean Log Deviation. Table 5.1 presents the evolution of both, with a computation with data that have been modified to equivalised disposable household income.

Figure 5.1: Graphs of Indices Evolution during the Period



In table 5.1, we discover that all of them follow a similar trend. In 1992

Table 5.1: Evolution of Inequality Indices for each Sample Year

	1992	1996	2002	2004	2007	2010	2013	2016
Gini	0.2181931	0.2688502	0.2639712	0.2744088	0.2623652	0.2656035	0.2642512	0.2635628
MLD	0.07988607	0.1207589	0.1151544	0.1270232	0.1171603	0.118562	0.1168964	0.1157936

Source: Author base on LIS

the values are at their lowest, then a steep increase in 1996 occurs with values similar until 2016. The MLD index for Czechia is quite moderate, with similar trends in comparison to the other measure.

Figure 5.2: Mean Log Deviation Decomposition Graphs A

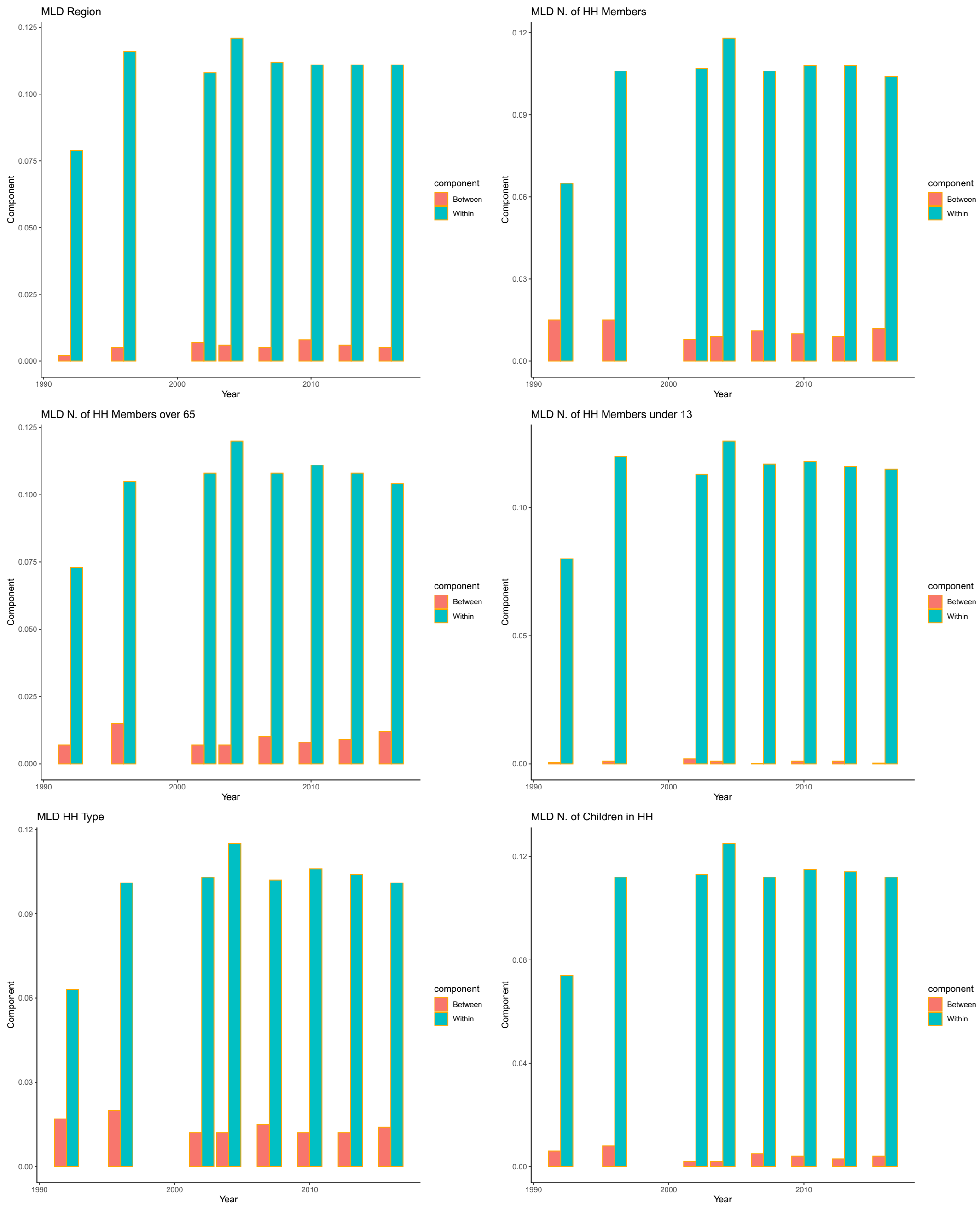
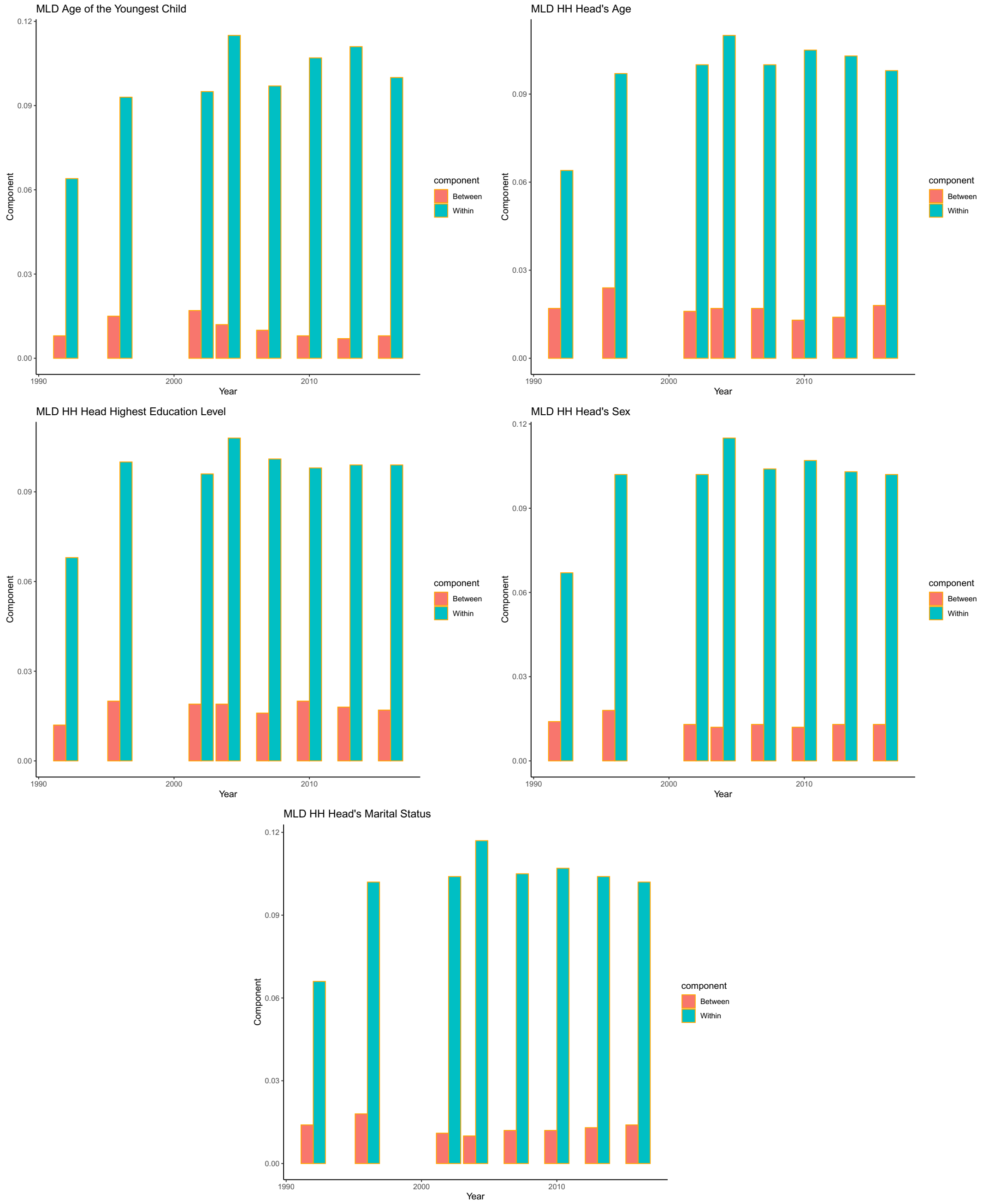


Figure 5.3: Mean Log Deviation Decomposition Graphs B



5.2 Mean Log Deviation Decomposition

Mean Log Deviation is very suitable for income inequality decomposition, as it can even be broken down into within- and between-group parts. In this section we take a look into the within- and between-group effects of each variable.

In general, the within effect of all variables seem very similar, as for 1992 the value is lower by around 4 points than the other years. The year 2004 demonstrates the highest within values. Overall, the share of within impact lies in the interval from 77% to 99%.

5.2.1 Region

In regional groups, the within impact is noticeably higher than the between one, since it explains above 90% of the total MLD average. This outcome tells us that income disparities in regional groups arrive mainly within, and might stem from other socio-demographic characteristics or occupations, meanwhile, a low share of between-group effect can suggest little differences in job opportunities or resources among the regions.

5.2.2 Household Composition

Both Number of Household Members and Household Size follow very comparable within influences, however, the between effect in size is markedly higher. Nonetheless, the between component in both shares halves throughout the period.

Therefore, inequality across different household compositions exists, although, for the number of members, the effect is smaller than for the size. Mainly, the inequality stems from within the groups, as the same households might have diverse incomes or labour market and socio-demographic characteristics.

The low between effects can also be observed for variables of the number of household members over 65 and under 13. The between effects are minimal, unlike the within's shares. For under 13 groups, the share makes up more than 99% on average. This goes hand in hand with the previous claim that disparities sources originate more strongly from other factors than from the composition.

5.2.3 Socio-demographic characteristics of the Household Head

Even though the within impact is still higher, the highest education level completed, age, gender and marital status have all between-group effects share over 11% on average, where the first two reach above the 15% threshold.

The education within trend follows a pattern as the higher the level completed bigger the inequality in that group.

Except for 1992, the disparities are stronger within the gender and the working age between the mid-20s and late 50s of household heads, essentially above 0.1 value. Also marital status has similar values, besides 1992, where groups of divorced and not married people experience higher disparities than married and widowed people.

Figure 5.4: Intercept

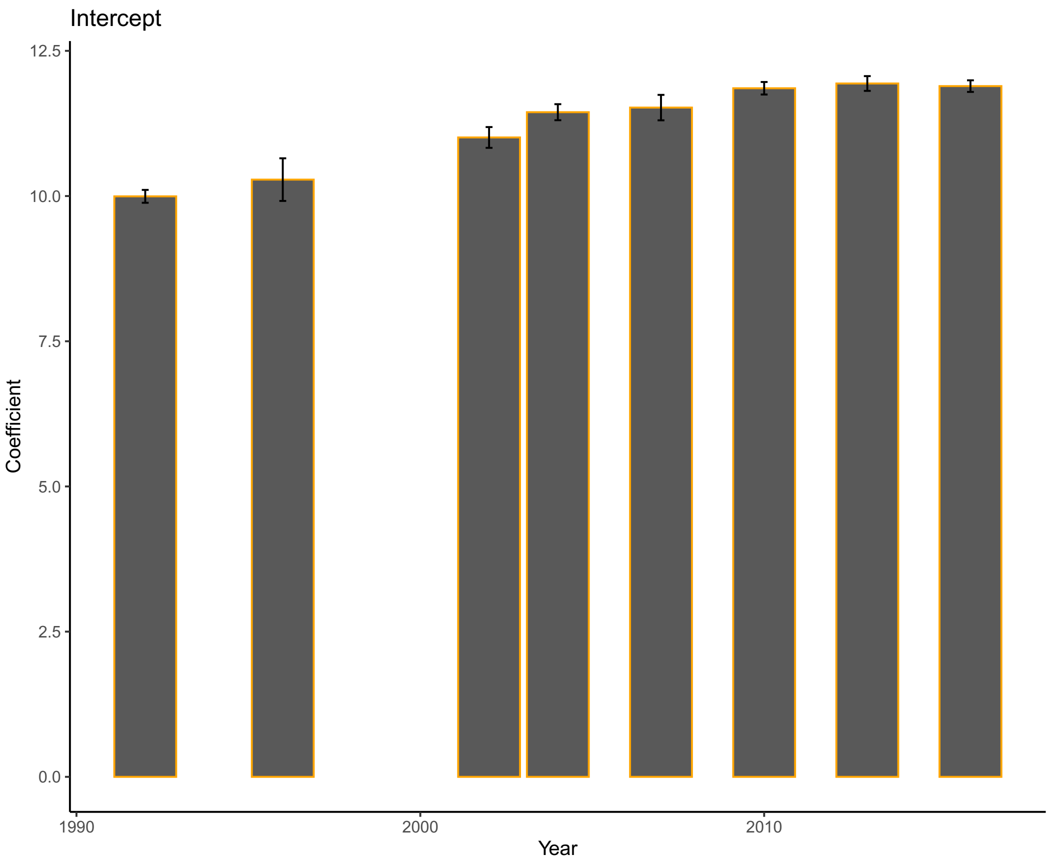


Figure 5.5: Region Coefficients 1992-1996

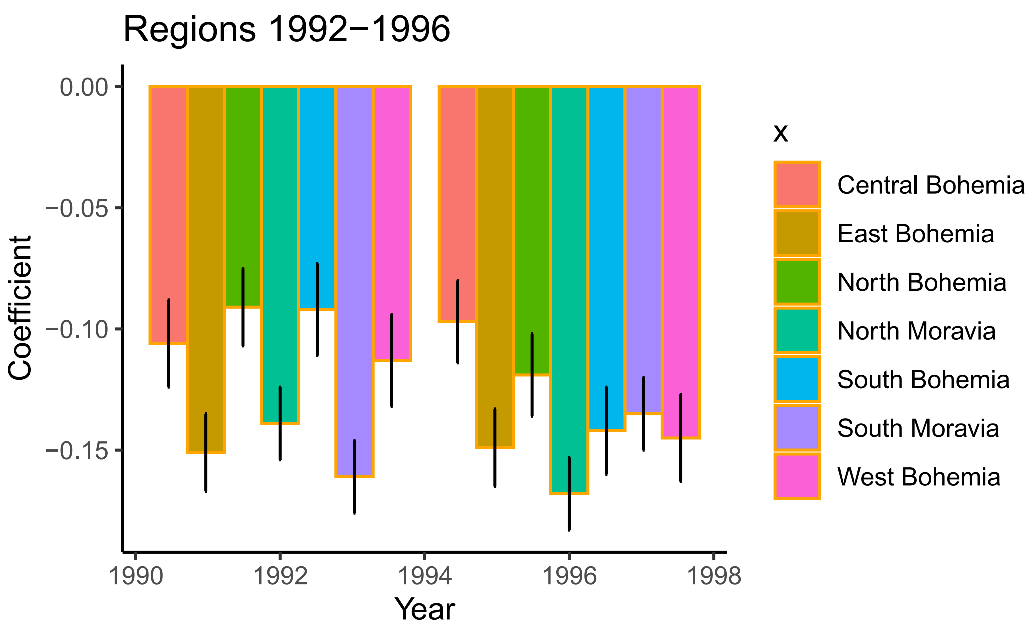
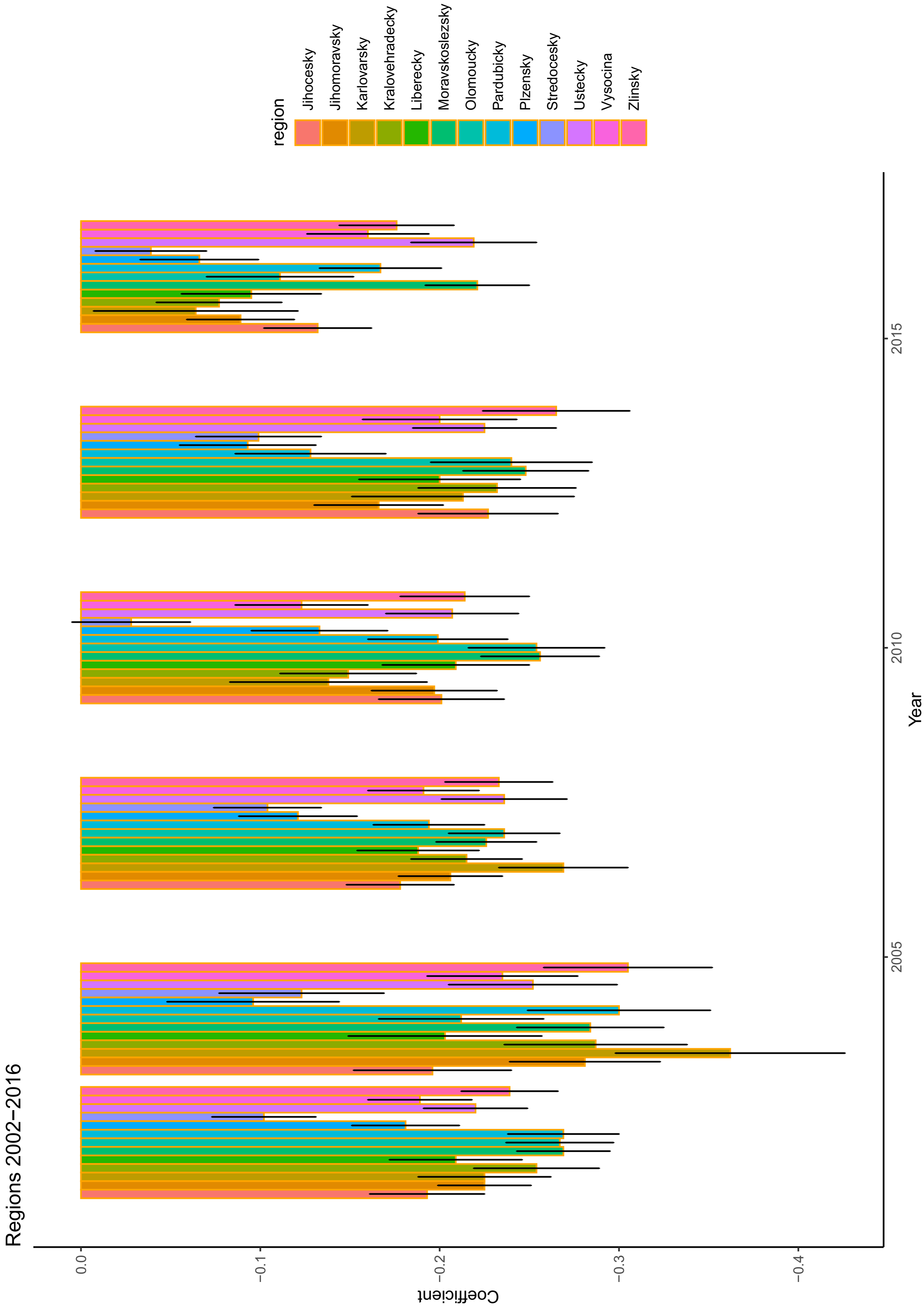


Figure 5.6: Region Coefficients 2002-2016



5.3 Socio-demographic characteristics effects on Equivalised Disposable Household Income

Overall, we regress eight models each for one year. Due to the failure of rejecting the null hypothesis of homoskedasticity, we apply White's estimators, which occupy the second column of our outcomes tables. The multicollinearity does not show any significant variables, besides age variables, which are intentionally involved. We check for normality and fail to reject the null hypothesis, however, there are more than 4 000 observations in each dataset.

At first, we desire to observe the effects of *hhtype* and *nchildren* as well. Yet, we also have *nhhmem* and *ageyoch* variables and together they may correlate and multicollinearity concern arises. We check correlations of these variables and the VIF test and both confirm our assumption therefore we do not include *hhtype* and *nchildren* into our model.

The R-squared of our models ranges between 0.274 and 0.404. We assume that since this work focuses on socio-demographic characteristics only, the absence of job-related and other income sources data contributes to lower values. Nonetheless, our F Statistics are significant and we have a large size of observations.

The results can be seen both in graphs in Figure 5.6 and tables of individual regressions in the appendix. We divide the results into subsections. First, we describe the intercept, then the regions, household composition, household head characteristics and household head highest education level.

5.3.1 Intercept

We conclude with the result that the intercept is rising with stumbles in 1996, 2010 and 2016. The intercept is significant all the time and the size of the trend occurs from 8.3 in 1992 to almost 9.3 in 2013.

5.3.2 Regions

For 1992 and 1996 the model works with different region distribution. All regions show highly significant negative coefficients except Prague. Living in one of those regions results in a decline in income of at least 8.7% (North Bohemia) to 14.8% (South Moravia) in 1992, however, the magnitude of regions' estimates ranges between Central Bohemian 9.2% and North Moravian 15.5% in 1996 in comparison to the capital city.

As suggested by the graph 5.5, during the period effects of North Moravia, South and West Bohemia are negative while those of the other regions rise.

Later during the 2002 - 2016 period the trends of regions are similar as none of them are positive, however, they are strongly significant, even at the 1% level. Among the highest, we can observe regions Stredocesky and Plzensky, where the former is located near or above the -0.1 level and the latter between -0.2 and -0.05.

The regions Moravskoslezky, Ustecky and Zlinsky occupy the bottom area close to -0.25. Even though Jihomoravsky, Karlovarsky, Pardubicky and Olomoucky regions start the period near the lowest part, they manage to rise.

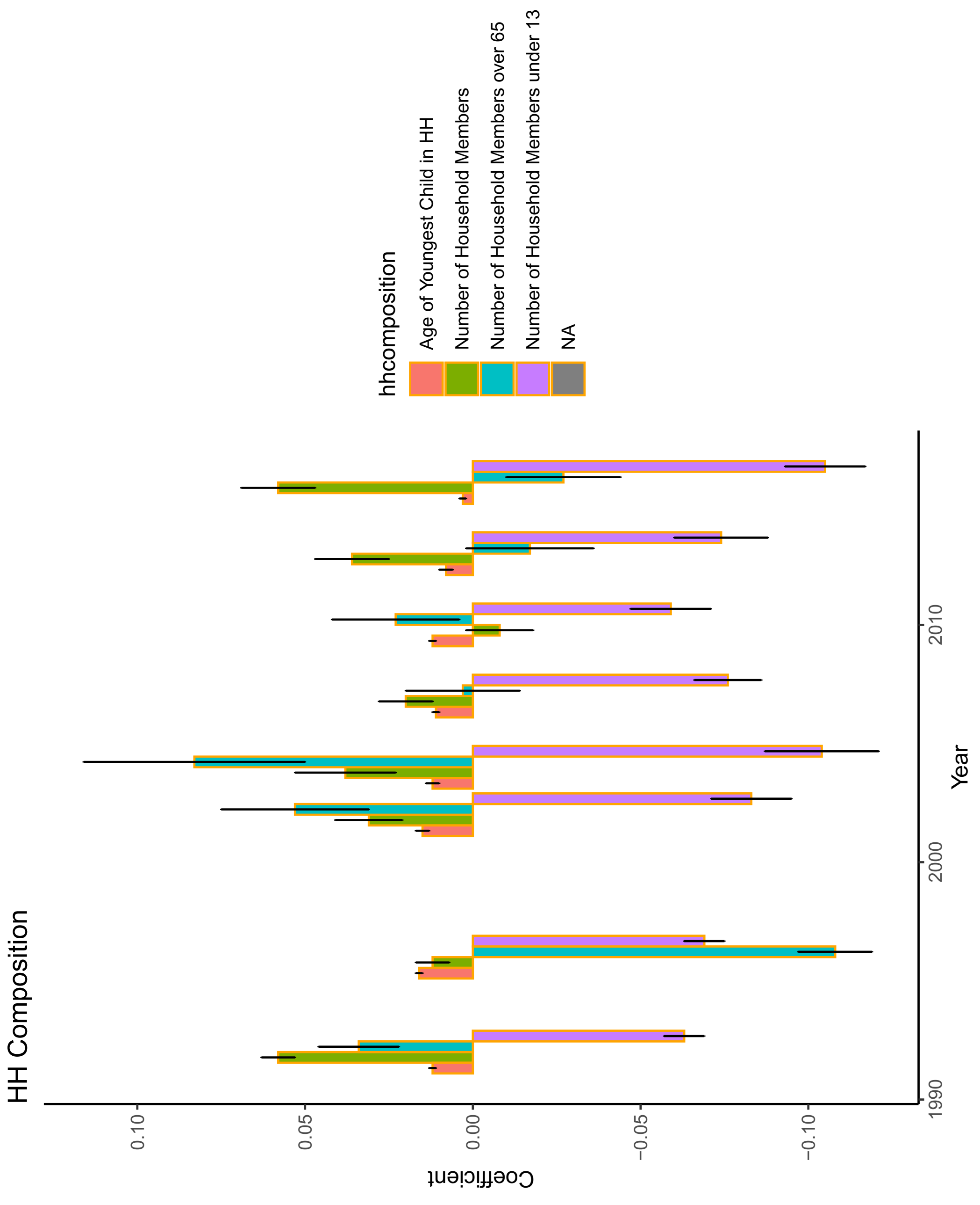
Nine of the regions record the lowest value in 2004, the minimum is given by the Karlovarsky dummy with its contribution to income by -30%, nonetheless, its figure increases afterwards.

Another drop is situated in 2013, where the region Jihocesky manages a fall with the size of the effect being 20%.

Then we can see two peaks in 2010 and 2016. In 2010, regions Stredocesky and Vysocina are at their maximum. In 2016, Plzensky, Karlovarsky, Jihomoravsky, Jihocesky, Liberecky, Olomoucky, Kralovehradecky, Moravskoslezsky and Zlinsky regions document the highest figures.

By 2016, some regions and their significance weaken as well as their negative effect. Among the variables that are significant at 1% level are regions Jihocesky, Pardubicky, Vysocina, Jihomoravsky, Olomoucky, Zlinsky, Ustecky, and Moravskoslezsky, where the last two have a size over 19.7%.

Figure 5.7: HH Composition Coefficients



5.3.3 Household Composition

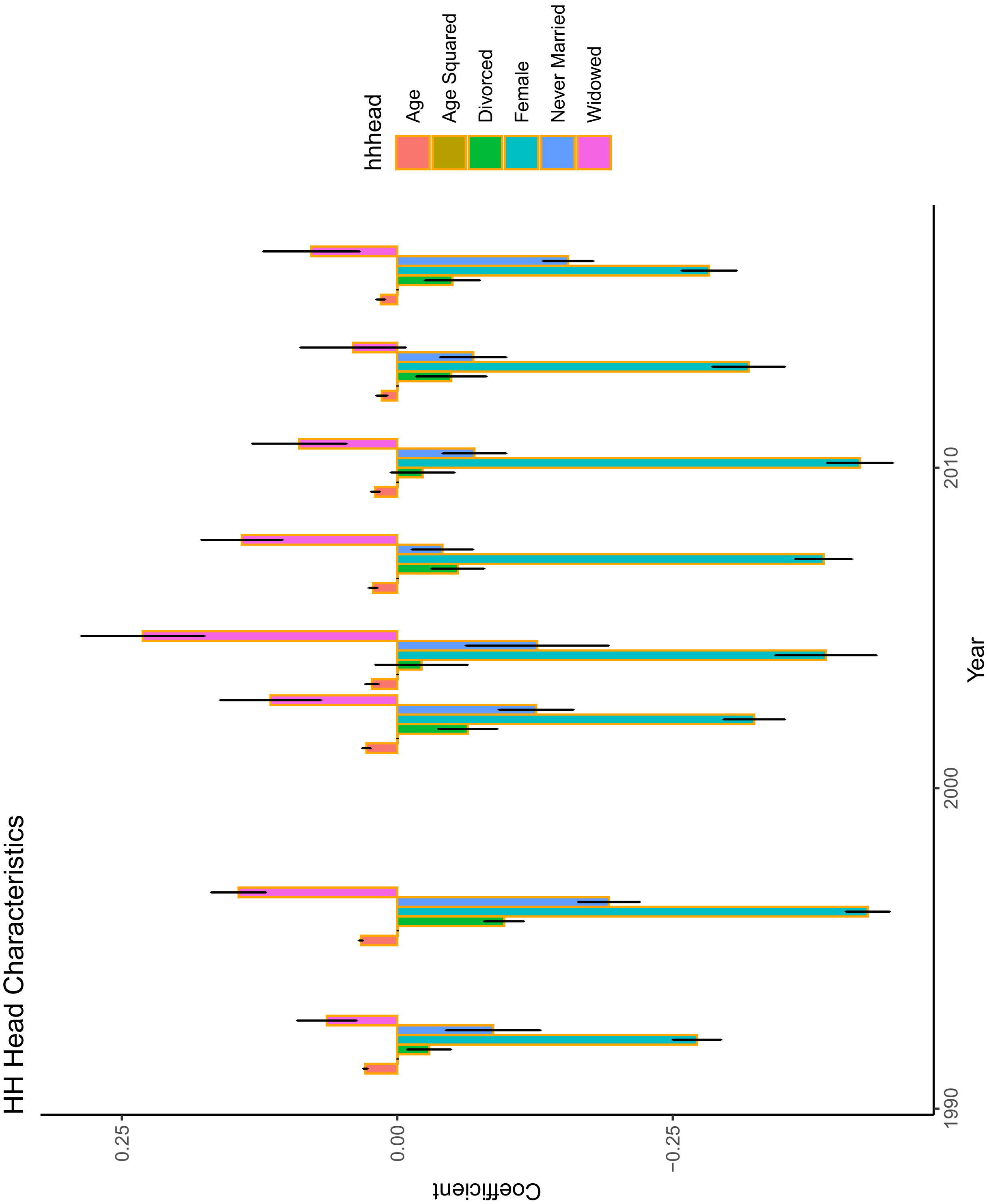
We find that all of the coefficients are statistically significant, except for the Number of HH Members over 65 from 2007 onward.

Among the positive effects, we find the Number of HH Members and the Age of Youngest Child in HH. The latter lies slightly above zero at a constant level, while the former has peaks above the 5% figure in 1992 and 2016. However, in 2010 the Number of HH Members reaches a negative value.

The Number of HH Members over 65 moves with higher fluctuation in comparison with other HH Composition coefficients, but only in the first half of the period, where it records values between -10.23% and 8.65%. From 2007 the effect holds near the zero level, ending below it.

The only full negative coefficient in this group is the Number of HH Members under 13 as it shifts between -0.05 and -0.1.

Figure 5.8: HH Head Characteristics Coefficients



5.3.4 Household Head Characteristics

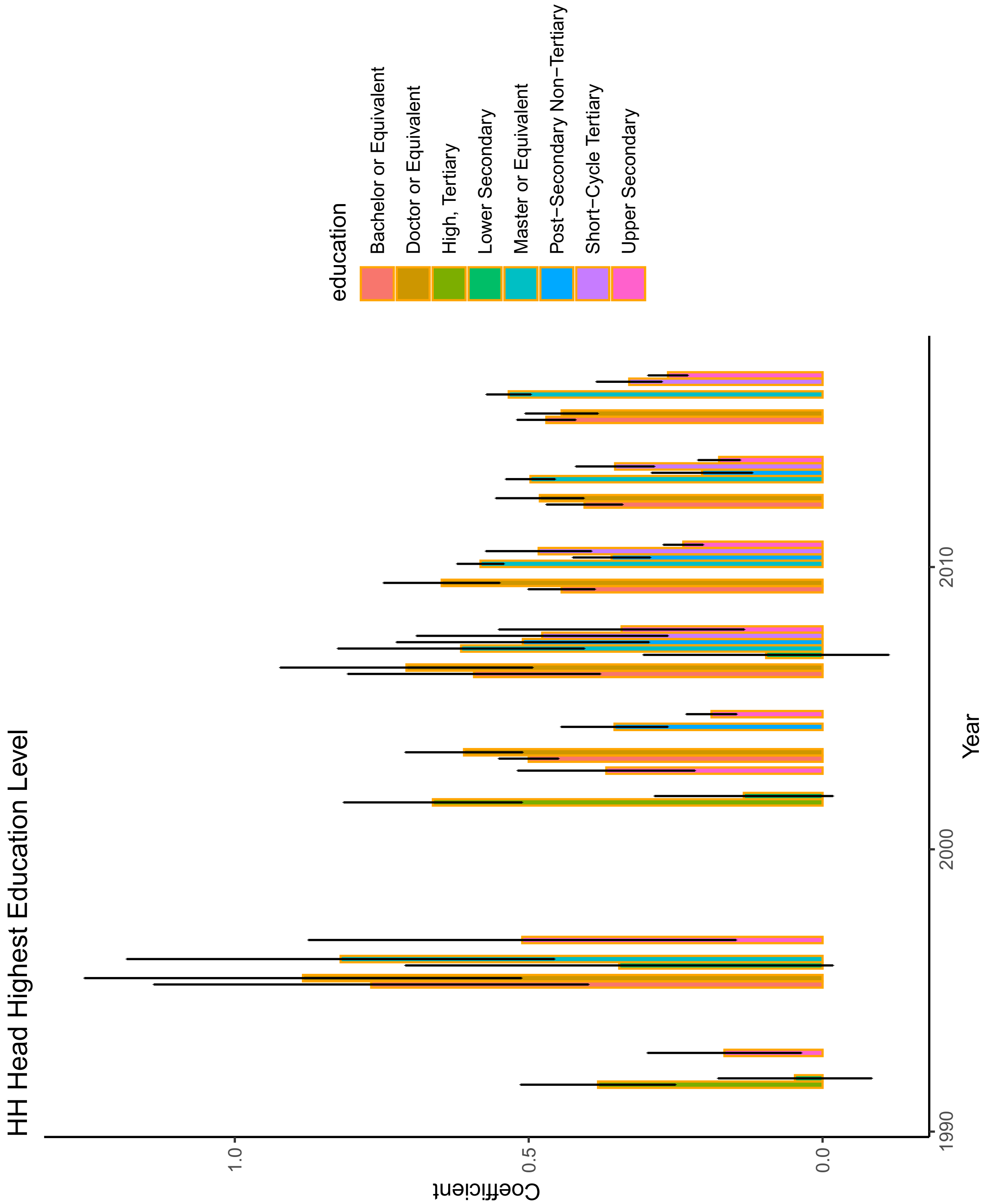
Both the Age and the Age Squared are almost constant and highly significant at the 1% level. The Age's trend is, however, decreasing from 3% to 1%. The Age Squared variable holds negative figures but with little to no effects.

The Female dummy variable possesses significant negative coefficients, with a large magnitude reaching -0.4 in most cases.

The Marital status also offers significant results. Nonetheless, the statuses of Divorced in 1992, 2004, 2010 and 2013, Never Married in 2007 and Widowed in 2013 are insignificant.

The Divorced and Never Married dummies are both negative, occurring in an area between 0 and -0.2 figures. Nevertheless, the Widowed status is strongly positive, peaking in 2004 at nearly 26%.

Figure 5.9: HH Head Highest Education Level Coefficients



5.3.5 Household Head Education Level

Despite the different distribution of the education levels throughout the period, we can observe some patterns. The figures are all positive and the significance of these results is present, except for the Lower Secondary level in 1992, 1996, 2002 and 2007 and the Upper Secondary in 1992, 1996 and 2007.

The High or Tertiary Education prevails over others, as possessing Bachelor, Master or Doctorate Education has the biggest effects. Mostly, the types peak in 1996, where the High level reaches values near 0.8, the Upper Secondary around 0.5 and the Lower Secondary just below 0.4. Another increase can be seen in 2007 and slightly at the end of the research period. The falls occur in 2004 and reach the minimum in 2013.

5.3.6 Overall trends

In general, among the strongly significant variables with negative influence are Regions except for Prague, the Number of Household Members younger than 13, Female and Age Squared. After that, the dummy of marital status Never Married.

Otherwise, the Number of Household Members, Age of Youngest Child, Age, Widowed status and Highest Completed Tertiary Education Levels show a high significance for positive influences.

The significant negative effects of regional dummies on income might stem from the supremacy of the constant, the Prague region as Hampl & Müller (2020) claims. Moreover, Ivan & Tvrdý (2007) states that young people might be attracted by economic reasons.

Křížková *et al.* (2010) and Bernard (2022) point out the gender pay gap issue, as being a female markedly decreases income. Also, we can think of a theory that single mothers as household heads would not have much income. Age squared indicates the non-linear relationship between age and our income in logarithmic form, as with increasing age, the income falls.

On the other hand, the results show us that the higher the level of education completed, the larger the positive effect on income, as more qualified workers can earn more.

Nonetheless, the intercept is also very significant, with the declines in 1996, which could have been contributed by the establishment of the Czech Republic, or in 2010 after a financial crisis.

Chapter 6

Conclusion

Our thesis focused on finding the determinants of inequality trends from 1992 to 2016 in Czechia. We obtained access to the Luxembourg Income Study database which provides eight datasets for Czechia during the period covered. We narrowed our research down to examine the effects of socio-demographic characteristics.

The variables of interest were Region, Number of Household Members, Number of Household Members Over 65, Number of Household Members Under 13, as well as the Age of the Youngest Child, Household Type, Household Head's Highest Completed Education Level, Age, Gender and their Marital Status.

After we modified our data, by top and bottom coding and equivalising income, we studied the Mean Log Deviation Index and its decomposition. The index was constantly moderate throughout the period, only between 1992 and 1996 the index rose sharply from 0.08 to 0.12. The Gini Coefficient behaved similarly, ranging from 0.22 to 0.27.

Firstly, we decomposed the MLD and discovered relatively high values of between group effects exceeding 0.01. The variables of our interest were the Household Head's Age, Education, Gender and Marital Status, Number of Household Members, Household Type and Age of the Youngest Head's Child. However, the within effect was still noticeably larger.

After the regression on the equivalised disposable household income, we observed that region estimates, in comparison to Prague, were in general significant and negative, and so was the female dummy variable. On the contrary, the highest completed education levels of upper secondary and above were significant with positive effects.

With our work, we desired to contribute to income inequality studies in Czechia by applying the decomposition method and covering the effects of socio-demographic characteristics.

Among the limitations of our research could be, for example, not using the panel data analysis, as we had many different dummy variables in each period. We could have also study a regression-based Gini decomposition like Bourguignon *et al.* (1998) or Lisina & Van Kerm (2022). Nonetheless, with the data provided, we could have also used other decomposition methods. For example, the decomposition of Theil or Gini index, even though Garner & Terrell (1998) studied the decomposition of the Gini coefficient for the Czech Republic, the research can be done with recent data. We also limited ourselves and examined only socio-demographic characteristics. Therefore, the next studies can focus on the labour market, income sources or policies.

Bibliography

BERNARD, A. (2022): "Economic Transformation and Inequality: Czech Republic, 1978-2016." Accepted: 2022-10-04T15:18:25Z Publisher: Univerzita Karlova, Fakulta sociálnĂch vÄ›d.

BOURGUIGNON, F., M. FOURNIER, & M. GURGAND (1998): "Distribution, Development and Education: Taiwan, 1979-1994."

CZECH STASTICAL OFFICE (2011): "Demografická ročenka ČR - 2011."

FERREIRA, F. H. G. & R. P. d. BARROS (1999): "The Slippery Slope: Explaining the Increase in Extreme Poverty in Urban Brazil, 1976-96."

FIORIO, C. & S. JENKINS (2007): "Regression-based inequality decomposition." *United Kingdom Stata Users' Group Meetings 2007* Number: 03 Publisher: Stata Users Group.

GARNER, T. I. & K. TERRELL (1998): "A Gini decomposition analysis of inequality in the Czech and Slovak Republics during the transition1." *Economics of Transition* **6(1)**: pp. 23–46. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-0351.1998.tb00035.x.

HAMPL, M. & J. MÜLLER (2020): "Development of geographical distribution of the population in Czechia after 2000: post-transformation tendencies?" *Geografie* **124(4)**: pp. 385–409. Publisher: The Czech Geographical Society.

HLAVAC, M. (2022): *stargazer: Well-Formatted Regression and Summary Statistics Tables.* Social Policy Institute, Bratislava, Slovakia. R package version 5.2.3.

HUBL, R. (2020): *Vnitřní migrace obyvatelstva ve východní části České republiky: regionálně-geografická analýza.* Ph.D. thesis, Masarykova univerzita, Přírodovědecká fakulta.

Immervoll, H. & L. Richardson (2011): "Redistribution Policy and Inequality Reduction in OECD Countries: What Has Changed in Two Decades?"

Ivan, I. & L. Tvrdý (2007): "The changes in spatial mobility of the people in Moravian-Silesian Region." In "Území, znalosti a rozvoj na počátku 21. století. Sborník příspěvků sekce konference Nové výzvy pro rozvoj regionů, států a mezinárodních trhů.", pp. 1–20. Ostrava: Ekonomická fakulta VŠB-TU Ostrava.

Kostelecký, T. (2000): "Housing and Its Influence on the Development of Social Inequalities in the Post-Communist Czech Republic." *Czech Sociological Review* **8(2)**: pp. 177–193. Publisher: Institute of Sociology of the Academy of Sciences of the Czech Republic.

Krejčová, S. (2012): *Debata o reformě českého školství v letech 1989 - 2004*. Ph.D. thesis, Masarykova univerzita, Fakulta sociálních studií.

Křížková, A., A. M. Penner, & T. Petersen (2010): "The Legacy of Equality and the Weakness of Law: Within-job Gender Wage Inequality in the Czech Republic." *European Sociological Review* **26(1)**: pp. 83–95.

Lerman, R. I. & S. Yitzhaki (1985): "Income Inequality Effects by Income Source: A New Approach and Applications to the United States." *The Review of Economics and Statistics* **67(1)**: pp. 151–156. Publisher: The MIT Press.

Lerman, R. I. & S. Yitzhaki (1989): "Improving the accuracy of estimates of Gini coefficients." *Journal of Econometrics* **42(1)**: pp. 43–47.

Lerman, R. I. & S. Yitzhaki (1994): "Effect of Marginal Changes in Income Sources On U.S. Income Inequality." *Public Finance Quarterly* **22(4)**: pp. 403–417. Publisher: SAGE Publications.

Lisina, A. & P. Van Kerm (2022): "Understanding Twenty Years of Inequality and Poverty Trends in Russia." *Review of Income and Wealth* **68(S1)**: pp. S108–S130. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/roiw.12574.

Luxembourg Income Study (2023a): "LIS Cross-National Data Center in Luxembourg - About LIS."

LUXEMBOURG INCOME STUDY (2023b): "LIS Cross-National Data Center in Luxembourg - Our Data."

LUXEMBOURG INCOME STUDY (2023c): "METIS."

MATĚJŮ, P. & N. LIM (1995): "Who Has Gotten Ahead After the Fall of Communism? The Case of the Czech Republic." *Czech Sociological Review* **3(2)**: pp. 117–136. Publisher: Institute of Sociology of the Academy of Sciences of the Czech Republic.

MYSÍKOVÁ, M. & J. VEČERNÍK (2018): "Personal Earnings Inequality and Polarization: The Czech Republic in Comparison with Austria and Poland." *Eastern European Economics* **56(1)**: pp. 57–80. Publisher: Routledge _eprint: https://doi.org/10.1080/00128775.2017.1402685.

NEUGSCHWENDER, J. o. (2020): "LIS Cross-National Data Center in Luxembourg."

NOVOTNÝ, J. (2007): "On the measurement of regional inequality: does spatial dimension of income inequality matter?" *The Annals of Regional Science* **41(3)**: pp. 563–580.

ONDŘICHOVÁ, M. (2014): "Je správné mít v dnešní době více dětí? Historie, současný význam vícedětných rodin a jejich perspektivy se zvláštním ohledem na situaci v ČR." Accepted: 2021-11-19T12:54:28Z Publisher: Jihočeská univerzita.

RAVALLION, M. (2016): *The economics of poverty: history, measurement, and policy.* New York: Oxford University Press.

REMIÁŠOVÁ, R. (2021): "Sňatečnost a rozvodovost v České republice z regionálního pohledu v letech 1970-2011." Accepted: 2022-04-11T09:14:17Z Publisher: Univerzita Karlova, Přírodovědecká fakulta.

TONEV, P., M. HALÁS, P. KLAPKA, & S. KRAFT (2022): *Vybrané prostorové aspekty migrace specifických skupin obyvatelstva v ČR.* Masarykova univerzita.

ČERNIAUSKAS, N., D. SOLOGON, C. ODONOGHUE, & L. TARASONIS (2021): "Income Inequality and Redistribution in Lithuania: The Role of Policy, Labor Market, Income, and Demographics." *Review of Income and Wealth* **68**.

Večerník, J. (1991): "Earnings Distribution in Czechoslovakia: Intertemporal Changes and International Comparison." *European Sociological Review* **7(3)**: pp. 237–252. Publisher: Oxford University Press.

Večerník, J. (2013): "The changing role of education in the distribution of earnings and household income: The Czech Republic, 1988-2009." *Economics of Transition* **21**.

Švejnar, J. (2002): "Transition Economies: Performance and Challenges." *Journal of Economic Perspectives* **16(1)**: pp. 3–28.

World Data Bank (2020): "Gini index - Czechia | Data."

Appendix A

Tables

Table A.1: Evolution of Household Distribution in Czech Regions for each Sample Year

	1992	1996	2002	2004	2007	2010	2013	2016
Praha	10,34%	7,28%	19,90%	10,78%	8,42%	9,82%	11,57%	12,32%
Stredocesky	11,20%	11,17%	8,60%	10,55%	10,38%	11,31%	11,41%	10,94%
Jihocesky	-	6,64%	5,63%	5,72%	6,64%	7,11%	7,02%	6,82%
Plzensky	-	5,71%	5,18%	6,32%	5,60%	5,37%	6,05%	6,07%
Karlovarsky	-	2,93%	4,44%	2,71%	3,34%	2,41%	2,02%	2,18%
Ustecky	-	9,38%	7,94%	8,32%	8,25%	8,12%	7,77%	7,54%
Liberecky	-	4,31%	3,69%	4,00%	4,12%	3,93%	4,25%	4,10%
Kralovehradecky	-	5,56%	4,08%	5,26%	5,15%	5,19%	5,04%	5,08%
Pardubicky	-	5,06%	4,89%	4,76%	5,22%	4,73%	4,87%	5,06%
Vysocina	-	4,71%	4,88%	5,36%	5,52%	5,48%	5,24%	5,05%
Jihomoravsky	-	11,98%	8,98%	9,77%	10,19%	10,77%	10,67%	11,40%
Olomoucky	-	6,28%	5,57%	7,08%	6,65%	6,05%	6,46%	5,78%
Zlinsky	-	6,09%	5,21%	5,54%	6,25%	6,05%	5,63%	5,21%
Moravskoslezsky	-	12,91%	11,01%	13,84%	14,27%	13,67%	12,01%	12,46%
South Bohemia	6,86%	7,38%	-	-	-	-	-	-
West Bohemia	8,46%	8,64%	-	-	-	-	-	-
North Bohemia	12,03%	12,96%	-	-	-	-	-	-
East Bohemia	12,25%	12,24%	-	-	-	-	-	-
South Moravia	19,87%	20,93%	-	-	-	-	-	-
North Moravia	19,00%	19,41%	-	-	-	-	-	-

Source: Author base on LIS

Special thanks to Hlavac (2022) for Stargazer package.

Table A.2: Evolution of Household Types Distribution for each Sample Year

	1992	1996	2002	2004	2007	2010	2013	2016
one person household	21,48%	24,26%	27,81%	28,22%	28,14%	28,61%	30,92%	32,65%
couple without children	24,15%	24,07%	26,09%	26,32%	26,46%	28,15%	28,52%	29,59%
couple with children	40,73%	37,68%	31,39%	30,59%	30,57%	28,65%	26,92%	25,68%
one parent with children	5,70%	6,37%	7,78%	7,33%	7,25%	7,71%	7,07%	9,41%
couple without children and relatives	0,92%	1,05%	0,98%	1,15%	0,98%	0,91%	0,78%	0,91%
couple with children and relatives	3,07%	2,43%	1,69%	1,91%	2,27%	1,83%	1,65%	1,76%
one parent with children and relatives	0,67%	0,63%	0,82%	0,90%	0,59%	0,60%	0,62%	-
relatives living together (no family nucleus)	2,47%	2,68%	2,36%	2,44%	2,86%	2,86%	2,72%	-
head and other	0,39%	0,43%	0,66%	0,69%	0,52%	0,37%	0,37%	-
couple without children and other	0,12%	0,11%	0,13%	0,18%	0,04%	0,05%	0,07%	-
couple with children and other	0,23%	0,19%	0,20%	0,23%	0,24%	0,20%	0,29%	-
one parent with children and other	0,06%	0,10%	0,10%	0,05%	0,07%	0,05%	0,06%	-

Table A.3: Evolution of Number of Household Members Distribution for each Sample Year

	1992	1996	2002	2004	2007	2010	2013	2016
1 hh member	21,48%	24,26%	27,81%	28,22%	28,14%	28,61%	30,92%	32,65%
2 hh members	29,88%	30,43%	33,55%	33,76%	33,69%	35,60%	35,89%	36,11%
3 hh members	19,51%	19,32%	17,46%	16,80%	16,60%	16,67%	15,41%	14,33%
4 hh members	21,43%	19,76%	16,62%	16,39%	16,37%	14,60%	13,39%	12,84%
5 hh members	5,96%	4,84%	3,60%	3,59%	3,97%	3,36%	3,27%	3,20%
6 hh members	1,36%	1,04%	0,69%	0,94%	0,92%	0,70%	0,79%	0,62%
7 hh members	0,27%	0,28%	0,24%	0,18%	0,25%	0,33%	0,22%	0,22%
8 hh members	0,08%	0,06%	0,04%	0,11%	0,02%	0,07%	0,04%	0,03%
9 hh members	0,02%	0,01%	-	-	0,03%	0,06%	0,05%	-
10 hh members	0,01%	0,01%	-	-	0,01%	-	0,02%	-
11 hh members	0,01%	0,01%	-	-	0,01%	-	-	-
12 hh members	0,01%	-	-	-	-	0,01%	-	-

Source: Author base on LIS

Table A.4: Evolution of Number of Household Members older than 65 Distribution for each Sample Year

	1992	1996	2002	2004	2007	2010	2013	2016
no hh members 65 or older	72,63%	69,44%	70,94%	69,04%	66,04%	65,82%	60,81%	57,80%
1 hh member 65 or older	20,39%	21,91%	20,95%	22,66%	23,86%	23,43%	26,91%	27,71%
2 hh members 65 or older	6,94%	8,58%	8,03%	8,16%	10,03%	10,70%	12,22%	14,31%
3 hh members 65 or older	0,04%	0,07%	0,09%	0,14%	0,07%	0,05%	0,06%	0,18%

Table A.5: Evolution of Number of Household Members younger than 13 Distribution for each Sample Year

	1992	1996	2002	2004	2007	2010	2013	2016
no hh members 13 or younger	69,26%	74,55%	76,76%	77,98%	79,14%	80,25%	81,09%	80,75%
1 hh member 13 or younger	16,03%	14,11%	13,46%	12,04%	11,88%	11,37%	10,51%	10,42%
2 hh member 13 or younger	12,50%	9,89%	8,47%	8,85%	7,69%	6,90%	6,98%	7,30%
3 hh member 13 or younger	1,87%	1,19%	1,10%	0,97%	1,14%	1,27%	1,15%	1,28%
4 hh member 13 or younger	0,28%	0,20%	0,15%	0,16%	0,14%	0,15%	0,19%	0,22%
5 hh member 13 or younger	0,04%	0,05%	0,05%	-	0,01%	0,03%	0,06%	0,03%
6 hh member 13 or younger	0,01%	0,01%	0,01%	-	-	-	0,01%	-
7 hh member 13 or younger	-	-	-	-	-	0,02%	0,01%	-
10 hh member 13 or younger	0,01%	-	-	-	-	-	-	-

Source: Author base on LIS

Table A.6: Evolution of Number of Households Living with own Children Distribution for each Sample Year

	1992	1996	2002	2004	2007	2010	2013	2016
not living with own children	44,16%	47,02%	53,10%	66,29%	66,15%	67,45%	68,53%	68,64%
living with 1 own child	22,66%	22,61%	21,21%	16,28%	15,98%	16,26%	15,89%	16,14%
living with 2 own children	26,33%	24,54%	21,24%	14,28%	14,61%	13,04%	12,42%	12,16%
living with 3 own children	5,82%	4,87%	3,65%	2,57%	2,71%	2,69%	2,60%	2,53%
living with 4 own children	0,87%	0,76%	0,60%	0,45%	0,46%	0,37%	0,38%	0,45%
living with 5 own children	0,12%	0,15%	0,17%	0,08%	0,07%	0,12%	0,13%	0,08%
living with 6 own children	0,02%	0,03%	0,02%	0,04%	0,01%	0,03%	0,01%	0,01%
living with 7 own children	0,01%	-	-	-	0,01%	0,02%	0,02%	-
living with 8 own children	-	-	-	-	-	0,00%	0,02%	-
living with 9 own children	0,01%	0,00%	-	-	0,01%	0,01%	-	-
living with 10 own children	0,01%	-	-	-	-	-	-	-

Table A.7: Evolution of Youngest Child's Age Distribution for each Sample Year

	1992	1996	2002	2004	2007	2010	2013	2016
Mean	12	13	13	15	16	16	17	18
Weighted Mean	12	13	14	17	17	17	18	19

Source: Author base on LIS

Table A.8: Evolution of Education Distribution for each Sample Year

	1992	1996	2002	2004	2007	2010	2013	2016
less than primary	0,24%	0,10%	0,17%	0,30%	0,07%	0,09%	0,11%	0,10%
never attended	0,20%	-	-	-	-	-	-	-
primary	-	-	-	-	0,27%	0,15%	0,07%	0,07%
lower secondary	22,58%	21,33%	15,20%	15,09%	14,51%	13,65%	11,54%	10,63%
upper secondary	51,11%	53,96%	59,17%	58,12%	59,31%	59,10%	59,95%	58,96%
post-secondary non-tertiary	-	-	-	1,84%	1,14%	0,95%	0,70%	-
high, tertiary	6,03%	-	8,16%	-	-	-	-	-
short-cycle tertiary	-	-	-	-	0,74%	0,72%	0,93%	1,12%
bachelor or equivalent	-	0,31%	-	7,46%	1,05%	1,76%	2,06%	2,26%
master or equivalent	-	5,89%	-	-	6,66%	8,59%	9,71%	11,18%
doctorate or equivalent	-	0,09%	-	0,69%	0,72%	0,36%	0,33%	0,47%
missing	19,84%	18,32%	17,31%	16,50%	15,52%	14,63%	14,60%	15,20%

Table A.9: Evolution of Household Head's Highest Completed Education Level Distribution for each Sample Year

	1992	1996	2002	2004	2007	2010	2013	2016
less than primary	-	0,0002%	0,002%	0,002%	0,005%	0,0005%	0,0003%	0,007%
never attended	0,20%	-	-	-	-	-	-	-
primary	-	-	-	-	0,22%	0,0007%	0,01%	0,001%
lower secondary	20,14%	18,89%	13,77%	11,21%	10,68%	9,36%	8,64%	7,98%
upper secondary	68,96%	69,74%	74,75%	73,91%	74,21%	73,13%	73,51%	73,31%
post-secondary non-tertiary	-	-	-	2,01%	1,19%	1,09%	0,90%	-
high, tertiary	10,69%	-	11,26%	-	-	-	-	-
short-cycle tertiary	-	-	-	-	0,80%	0,72%	1,03%	1,05%
bachelor or equivalent	-	0,005%	-	11,50%	1,14%	1,62%	1,80%	2,15%
master or equivalent	-	10,65%	-	-	10,54%	13,29%	13,42%	14,86%
doctorate or equivalent	-	0,002%	-	1,12%	1,21%	0,67%	0,59%	0,83%

Source: Author base on LIS

Table A.10: Evolution of Household Head's Age in Years Distribution for each Sample Year

	1992	1996	2002	2004	2007	2010	2013	2016
Mean	50,12	51,96	51,77	52,67	54,59	55,07	56,29	57,14
Median	48	51	51	53	55	56	58	59
Standard Deviation	15,97	16,08	16,62	16,63	16,24	16,22	16,23	16,51

Source: Author base on LIS

Table A.11: Evolution of Gender Distribution for each Sample Year

	1992	1996	2002	2004	2007	2010	2013	2016
male	47,74%	48,20%	47,49%	47,58%	47,56%	47,66%	47,44%	47,58%
female	52,26%	51,80%	52,51%	52,42%	52,44%	52,34%	52,56%	52,42%

Table A.12: Evolution of Household Head's Gender Distribution for each Sample Year

	1992	1996	2002	2004	2007	2010	2013	2016
male	78,52%	78,06%	75,31%	77,85%	76,61%	76,80%	75,15%	75,71%
female	21,47%	21,94%	24,69%	22,15%	23,39%	23,20%	24,85%	24,29%

Source: Author base on Luxembourg Income Study (2023c)

Table A.13: Evolution of Marital Status Distribution for each Sample Year

	1992	1996	2002	2004	2007	2010	2013	2016
married	51,70%	50,49%	47,10%	47,56%	47,49%	46,86%	45,80%	45,64%
never married/not in union	15,89%	19,30%	35,95%	35,61%	35,55%	35,56%	35,42%	35,20%
divorced	4,93%	5,73%	8,51%	7,74%	8,03%	8,78%	9,80%	9,94%
widowed	7,64%	8,07%	8,45%	9,09%	8,92%	8,80%	8,98%	9,21%
missing	19,84%	16,41%	-	-	-	-	-	-

Table A.14: Evolution of Household Head's Marital Status Distribution for each Sample Year

	1992	1996	2002	2004	2007	2010	2013	2016
married	67,57%	65,75%	58,35%	60,33%	57,76%	56,52%	51,95%	51,48%
never married/not in union	7,22%	8,38%	12,22%	12,73%	13,79%	15,55%	17,12%	18,56%
divorced	10,58%	15,79%	15,50%	14,07%	15,20%	15,43%	17,59%	17,09%
widowed	14,64%	9,39%	13,92%	12,87%	13,25%	12,50%	13,34%	12,87%

Source: Author base on Luxembourg Income Study (2023c)

Table A.15: Evolution of Disposable Household Income Adjusted by LISPPP

	1992	1996	2002	2004	2007	2010	2013	2016
Mean	10643,46	12742,36	14476,45	14943,35	17010,04	17897,36	17844,18	19353,83
Median	9790,48	11050,42	12312,60	12856,37	14886,90	15554,18	15601,53	16818,54
25th Quantile	7721,11	8395,18	9453,37	9971,90	11595,28	12154,53	12099,81	12917,92
75th Quantile	12464,76	15293,51	16903,74	17621,28	20149,94	21256,06	20793,16	23051,33
St. Dev.	5160,05	7593,01	8843,03	9349,21	9065,61	9497,53	10448,01	10699,82

Source: Author base on LIS

Table A.16: Mean Log Deviation Decomposition Results

		1992	1996	2002	2004	2007	2010	2013	2016
Total		0,080	0,120	0,115	0,127	0,117	0,119	0,117	0,115
Region	Within	0,079	0,116	0,108	0,121	0,112	0,111	0,111	0,111
	Between	0,002	0,005	0,007	0,006	0,005	0,008	0,006	0,005
Number of HH members	Within	0,065	0,106	0,107	0,118	0,106	0,108	0,108	0,104
	Between	0,015	0,015	0,008	0,009	0,011	0,010	0,009	0,012
Number of HH members over 65	Within	0,073	0,105	0,108	0,120	0,108	0,111	0,108	0,104
	Between	0,007	0,015	0,007	0,007	0,010	0,008	0,009	0,012
Number of HH members below 13	Within	0,080	0,120	0,113	0,126	0,117	0,118	0,116	0,115
	Between	0,0005	0,001	0,002	0,001	0,0002	0,001	0,001	0,0003
HH type	Within	0,063	0,101	0,103	0,115	0,102	0,106	0,104	0,101
	Between	0,017	0,020	0,012	0,012	0,015	0,012	0,012	0,014
HH Head Education	Within	0,068	0,100	0,096	0,108	0,101	0,098	0,099	0,099
	Between	0,012	0,020	0,019	0,019	0,016	0,020	0,018	0,017
HH Head age	Within	0,064	0,097	0,100	0,110	0,100	0,105	0,103	0,098
	Between	0,017	0,024	0,016	0,017	0,017	0,013	0,014	0,018
HH Head Gender	Within	0,067	0,102	0,102	0,115	0,104	0,107	0,103	0,102
	Between	0,014	0,018	0,013	0,012	0,013	0,012	0,013	0,013
HH Head Marital Status	Within	0,066	0,102	0,104	0,117	0,105	0,107	0,104	0,102
	Between	0,014	0,018	0,011	0,010	0,012	0,012	0,013	0,014
Number of Children in HH	Within	0,074	0,112	0,113	0,125	0,112	0,115	0,114	0,112
	Between	0,006	0,008	0,002	0,002	0,005	0,004	0,003	0,004
Age of the Youngest child	Within	0,064	0,093	0,095	0,115	0,097	0,107	0,111	0,100
	Between	0,008	0,015	0,017	0,012	0,010	0,008	0,007	0,008

Source: Author base on LIS

Table A.17: Regression of Income 1992, pt. 1

	Dependent variable:	
	Equivalised Disposable Household Income in Logarithmic Form	
	OLS	White
	(1)	(2)
Central Bohemia	−0.106***	−0.106***
	(0.015)	(0.018)
South Bohemia	−0.092***	−0.092***
	(0.017)	(0.019)
West Bohemia	−0.113***	−0.113***
	(0.016)	(0.019)
North Bohemia	−0.091***	−0.091***
	(0.015)	(0.016)
East Bohemia	−0.151***	−0.151***
	(0.015)	(0.016)
South Moravia	−0.161***	−0.161***
	(0.013)	(0.015)
North Moravia	−0.139***	−0.139***
	(0.013)	(0.015)
Number of Household Members	0.058***	0.058***
	(0.005)	(0.005)
Number of Household Members over 65	0.034***	0.034***
	(0.013)	(0.012)
Number of Household Members under 13	−0.063***	−0.063***
	(0.006)	(0.006)
Age of Youngest Child in HH	0.012***	0.012***
	(0.001)	(0.001)
Constant	8.413***	8.413***
	(0.111)	(0.141)
Observations	8,191	8,191
R^2	0.308	0.308
Adjusted R^2	0.307	0.307
Residual Std. Error (df = 8170)	0.300	0.300
F Statistic (df = 20; 8170)	182.105***	182.105***
Note:		*p<0.1; **p<0.05; ***p<0.01

Table A.18: Regression of Income 1992, pt. 2

	Dependent variable:	
	Equivalised Disposable Household Income in Logarithmic Form	
	OLS	White
	(1)	(2)
Lower Secondary	0.047	0.047
	(0.101)	(0.131)
Upper Secondary	0.167*	0.167
	(0.101)	(0.131)
High, Tertiary	0.382***	0.382***
	(0.101)	(0.132)
Age	0.029***	0.029***
	(0.002)	(0.002)
Age Squared	−0.0003***	−0.0003***
	(0.00002)	(0.00002)
Female	−0.272***	−0.272***
	(0.018)	(0.022)
Never Married/Not in Union	−0.087***	−0.087**
	(0.028)	(0.043)
Divorced	−0.029*	−0.029
	(0.017)	(0.020)
Widowed	0.064***	0.064**
	(0.024)	(0.027)
Observations	8,191	8,191
R^2	0.308	0.308
Adjusted R^2	0.307	0.307
Residual Std. Error (df = 8170)	0.300	0.300
F Statistic (df = 20; 8170)	182.105***	182.105***
Note:		*p<0.1; **p<0.05; ***p<0.01

Source: Author's own calculations base on LIS

Table A.19: Regression of Income 1996, pt. 1

	Dependent variable:	
	Equivalised Disposable Household Income in Logarithmic Form	
	OLS	White
	(1)	(2)
Central Bohemia	−0.097***	−0.097***
	(0.015)	(0.017)
East Bohemia	−0.149***	−0.149***
	(0.015)	(0.016)
North Bohemia	−0.119***	−0.119***
	(0.015)	(0.017)
North Moravia	−0.168***	−0.168***
	(0.014)	(0.015)
South Bohemia	−0.142***	−0.142***
	(0.016)	(0.018)
South Moravia	−0.135***	−0.135***
	(0.014)	(0.015)
West Bohemia	−0.145***	−0.145***
	(0.016)	(0.018)
Number of Household Members	0.012***	0.012**
	(0.005)	(0.005)
Number of Household Members over 65	−0.108***	−0.108***
	(0.012)	(0.011)
Number of Household Members under 13	−0.069***	−0.069***
	(0.006)	(0.006)
Age of Youngest Child in HH	0.016***	0.016***
	(0.001)	(0.001)
Constant	8.307***	8.307***
	(0.249)	(0.367)
Observations	13,336	13,336
R^2	0.404	0.404
Adjusted R^2	0.403	0.403
Residual Std. Error (df = 13313)	0.345	0.345
F Statistic (df = 22; 13313)	409.975***	409.975***
Note:		*p<0.1; **p<0.05; ***p<0.01

Table A.20: Regression of Income 1996, pt. 2

	Dependent variable:	
	Equivalised Disposable Household Income in Logarithmic Form	
	OLS	White
	(1)	(2)
Lower Secondary	0.346	0.346
	(0.245)	(0.364)
Upper Secondary	0.511**	0.511
	(0.244)	(0.364)
Bachelor or Equivalent	0.768***	0.768**
	(0.249)	(0.370)
Master or Equivalent	0.820***	0.820**
	(0.244)	(0.364)
Doctorate or Equivalent	0.884***	0.884**
	(0.258)	(0.372)
Age	0.033***	0.033***
	(0.002)	(0.002)
Age Squared	−0.0004***	−0.0004***
	(0.00002)	(0.00002)
Female	−0.427***	−0.427***
	(0.016)	(0.020)
Never Married/Not in Union	−0.192***	−0.192***
	(0.022)	(0.028)
Divorced	−0.097***	−0.097***
	(0.015)	(0.018)
Widowed	0.144***	0.144***
	(0.022)	(0.025)
Observations	13,336	13,336
R^2	0.404	0.404
Adjusted R^2	0.403	0.403
Residual Std. Error (df = 13313)	0.345	0.345
F Statistic (df = 22; 13313)	409.975***	409.975***
Note:		*p<0.1; **p<0.05; ***p<0.01

Source: Author's own calculations base on LIS

Table A.21: Regression of Income 2002, pt. 1

	Dependent variable:	
	Equivalised Disposable Household Income in Logarithmic Form	
	OLS	White
	(1)	(2)
Stredocesky	−0.102***	−0.102***
	(0.026)	(0.029)
Jihocesky	−0.193***	−0.193***
	(0.030)	(0.032)
Plzensky	−0.181***	−0.181***
	(0.032)	(0.030)
Karlovarsky	−0.225***	−0.225***
	(0.033)	(0.037)
Ustecky	−0.220***	−0.220***
	(0.027)	(0.029)
Liberecky	−0.209***	−0.209***
	(0.034)	(0.037)
Kralovehradecky	−0.254***	−0.254***
	(0.037)	(0.035)
Pardubicky	−0.269***	−0.269***
	(0.031)	(0.031)
Vysocina	−0.189***	−0.189***
	(0.030)	(0.029)
Jihomoravsky	−0.225***	−0.225***
	(0.026)	(0.026)
Olomoucky	−0.267***	−0.267***
	(0.030)	(0.030)
Zlinsky	−0.239***	−0.239***
	(0.029)	(0.027)
Moravskoslezsky	−0.269***	−0.269***
	(0.024)	(0.026)
Number of Household Members	0.031***	0.031***
	(0.010)	(0.010)
Number of Household Members over 65	0.053**	0.053**
	(0.025)	(0.022)
Observations	3,347	3,347
R^2	0.395	0.395
Adjusted R^2	0.390	0.390
Residual Std. Error (df = 3320)	0.357	0.357
F Statistic (df = 26; 3320)	83.284***	83.284***

Note: *p<0.1; **p<0.05; ***p<0.01

Table A.22: Regression of Income 2002, pt. 2

	Dependent variable:	
	Equivalised Disposable Household Income in Logarithmic Form	
	OLS	White
	(1)	(2)
Number of Household Members under 13	−0.083***	−0.083***
	(0.012)	(0.012)
Age of Youngest Child in HH	0.015***	0.015***
	(0.002)	(0.002)
Lower Secondary	0.134	0.134
	(0.164)	(0.152)
Upper Secondary	0.368**	0.368**
	(0.163)	(0.151)
High, Tertiary	0.663***	0.663***
	(0.164)	(0.152)
Age	0.028***	0.028***
	(0.004)	(0.004)
Age Squared	−0.0003***	−0.0003***
	(0.00004)	(0.00004)
Female	−0.324***	−0.324***
	(0.025)	(0.028)
Never Married/Not in Union	−0.126***	−0.126***
	(0.032)	(0.034)
Divorced	−0.064***	−0.064**
	(0.025)	(0.027)
Widowed	0.115***	0.115**
	(0.039)	(0.046)
Constant	8.664***	8.664***
	(0.183)	(0.179)
Observations	3,347	3,347
R^2	0.395	0.395
Adjusted R^2	0.390	0.390
Residual Std. Error (df = 3320)	0.357	0.357
F Statistic (df = 26; 3320)	83.284***	83.284***
Note:		*p<0.1; **p<0.05; ***p<0.01

Source: Author's own calculations base on LIS

Table A.23: Regression of Income 2004, pt. 1

	Dependent variable:	
	Equivalised Disposable Household Income in Logarithmic Form	
	OLS	White
	(1)	(2)
Stredocesky	−0.123***	−0.123***
	(0.043)	(0.046)
Jihocesky	−0.196***	−0.196***
	(0.049)	(0.044)
Plzensky	−0.096*	−0.096**
	(0.049)	(0.048)
Karlovarsky	−0.362***	−0.362***
	(0.063)	(0.064)
Ustecky	−0.252***	−0.252***
	(0.045)	(0.047)
Liberecky	−0.203***	−0.203***
	(0.053)	(0.054)
Kralovehradecky	−0.287***	−0.287***
	(0.050)	(0.051)
Pardubicky	−0.300***	−0.300***
	(0.049)	(0.051)
Vysocina	−0.235***	−0.235***
	(0.047)	(0.042)
Jihomoravsky	−0.281***	−0.281***
	(0.043)	(0.042)
Olomoucky	−0.212***	−0.212***
	(0.045)	(0.046)
Zlinsky	−0.305***	−0.305***
	(0.048)	(0.047)
Moravskoslezsky	−0.284***	−0.284***
	(0.040)	(0.041)
Number of Household Members	0.038***	0.038***
	(0.014)	(0.015)
Number of Household Members over 65	0.083**	0.083**
	(0.035)	(0.033)
Observations	1,784	1,784
R^2	0.359	0.359
Adjusted R^2	0.349	0.349
Residual Std. Error (df = 1756)	0.378	0.378
F Statistic (df = 27; 1756)	36.380***	36.380***
Note:		*p<0.1; **p<0.05; ***p<0.01

Table A.24: Regression of Income 2004, pt. 2

	Dependent variable:	
	Equivalised Disposable Household Income in Logarithmic Form	
	OLS	White
	(1)	(2)
Number of Household Members under 13	−0.104***	−0.104***
	(0.017)	(0.017)
Age of Youngest Child in HH	0.012***	0.012***
	(0.002)	(0.002)
Upper Secondary	0.189***	0.189***
	(0.037)	(0.043)
Post-Secondary Non-Tertiary	0.354***	0.354***
	(0.077)	(0.091)
Bachelor or Equivalent	0.500***	0.500***
	(0.045)	(0.051)
Doctorate or Equivalent	0.610***	0.610***
	(0.097)	(0.100)
Age	0.023***	0.023***
	(0.006)	(0.006)
Age Squared	−0.0003***	−0.0003***
	(0.0001)	(0.0001)
Female	−0.389***	−0.389***
	(0.042)	(0.046)
Never Married/Not in Union	−0.127***	−0.127**
	(0.047)	(0.065)
Divorced	−0.022	−0.022
	(0.040)	(0.042)
Widowed	0.231***	0.231***
	(0.057)	(0.056)
Constant	9.071***	9.071***
	(0.143)	(0.138)
Observations	1,784	1,784
R^2	0.359	0.359
Adjusted R^2	0.349	0.349
Residual Std. Error (df = 1756)	0.378	0.378
F Statistic (df = 27; 1756)	36.380***	36.380***
Note:		*p<0.1; **p<0.05; ***p<0.01

Source: Author's own calculations base on LIS

Table A.25: Regression of Income 2007, pt. 1

	Dependent variable:	
	Equivalised Disposable Household Income in Logarithmic Form	
	OLS	White
	(1)	(2)
Stredocesky	−0.104***	−0.104***
	(0.026)	(0.030)
Jihocesky	−0.178***	−0.178***
	(0.028)	(0.030)
Plzensky	−0.121***	−0.121***
	(0.030)	(0.033)
Karlovarsky	−0.269***	−0.269***
	(0.035)	(0.036)
Ustecky	−0.236***	−0.236***
	(0.028)	(0.035)
Liberecky	−0.188***	−0.188***
	(0.032)	(0.034)
Kralovehradecky	−0.215***	−0.215***
	(0.030)	(0.031)
Pardubicky	−0.194***	−0.194***
	(0.030)	(0.031)
Vysocina	−0.191***	−0.191***
	(0.029)	(0.031)
Jihomoravsky	−0.206***	−0.206***
	(0.026)	(0.029)
Olomoucky	−0.236***	−0.236***
	(0.029)	(0.031)
Zlinsky	−0.233***	−0.233***
	(0.028)	(0.030)
Moravskoslezsky	−0.226***	−0.226***
	(0.025)	(0.028)
Number of Household Members	0.020***	0.020**
	(0.008)	(0.008)
Number of Household Members over 65	0.003	0.003
	(0.020)	(0.017)
Observations	4,630	4,630
R^2	0.343	0.343
Adjusted R^2	0.338	0.338
Residual Std. Error (df = 4599)	0.355	0.355
F Statistic (df = 30; 4599)	79.942***	79.942***
Note:		*p<0.1; **p<0.05; ***p<0.01

Table A.26: Regression of Income 2007, pt. 2

	Dependent variable:	
	Equivalised Disposable Household Income in Logarithmic Form	
	OLS	White
	(1)	(2)
Number of Household Members under 13	−0.076***	−0.076***
	(0.010)	(0.010)
Age of Youngest Child in HH	0.011***	0.011***
	(0.001)	(0.001)
Lower Secondary	0.096	0.096
	(0.162)	(0.209)
Upper Secondary	0.342**	0.342
	(0.161)	(0.209)
Post-Secondary Non-Tertiary	0.510***	0.510**
	(0.171)	(0.215)
Short-Cycle Tertiary	0.477***	0.477**
	(0.172)	(0.214)
Bachelor or Equivalent	0.593***	0.593***
	(0.168)	(0.215)
Master or Equivalent	0.615***	0.615***
	(0.162)	(0.210)
Doctorate or Equivalent	0.708***	0.708***
	(0.168)	(0.215)
Age	0.022***	0.022***
	(0.004)	(0.004)
Age Squared	−0.0003***	−0.0003***
	(0.00004)	(0.00004)
Female	−0.387***	−0.387***
	(0.023)	(0.026)
Never Married/Not in Union	−0.041	−0.041
	(0.025)	(0.028)
Divorced	−0.055**	−0.055**
	(0.022)	(0.024)
Widowed	0.141***	0.141***
	(0.034)	(0.037)
Constant	9.079***	9.079***
	(0.180)	(0.219)
Observations	4,630	4,630
R^2	0.343	0.343
Adjusted R^2	0.338	0.338
Residual Std. Error (df = 4599)	0.355	0.355
F Statistic (df = 30; 4599)	79.942***	79.942***
Note:		*p<0.1; **p<0.05; ***p<0.01

Source: Author's own calculations base on LIS

Table A.27: Regression on Income 2010, pt. 1

	Dependent variable:	
	Equivalised Disposable Household Income in Logarithmic Form	
	OLS	White
	(1)	(2)
Stredocesky	−0.028	−0.028
	(0.031)	(0.033)
Jihocesky	−0.201***	−0.201***
	(0.034)	(0.035)
Plzensky	−0.133***	−0.133***
	(0.038)	(0.038)
Karlovarsky	−0.138***	−0.138**
	(0.050)	(0.055)
Ustecky	−0.207***	−0.207***
	(0.034)	(0.037)
Liberecky	−0.209***	−0.209***
	(0.041)	(0.041)
Kralovehradecky	−0.149***	−0.149***
	(0.035)	(0.038)
Pardubicky	−0.199***	−0.199***
	(0.038)	(0.039)
Vysocina	−0.123***	−0.123***
	(0.035)	(0.037)
Jihomoravsky	−0.197***	−0.197***
	(0.031)	(0.035)
Olomoucky	−0.254***	−0.254***
	(0.037)	(0.038)
Zlinsky	−0.214***	−0.214***
	(0.035)	(0.036)
Moravskoslezsky	−0.256***	−0.256***
	(0.030)	(0.033)
Number of Household Members	−0.008	−0.008
	(0.010)	(0.010)
Number of Household Members Over 65	0.023	0.023
	(0.023)	(0.019)
Observations	3,459	3,459
R^2	0.325	0.325
Adjusted R^2	0.319	0.319
Residual Std. Error (df = 3429)	0.383	0.383
F Statistic (df = 29; 3429)	56.938***	56.938***
Note:		*p<0.1; **p<0.05; ***p<0.01

Table A.28: Regression on Income 2010, pt. 2

	Dependent variable:	
	Equivalised Disposable Household Income in Logarithmic Form	
	OLS	White
	(1)	(2)
Number of Household Members Under 13	−0.059***	−0.059***
	(0.013)	(0.012)
Age of The Youngest Child	0.012***	0.012***
	(0.001)	(0.001)
Upper Secondary	0.237***	0.237***
	(0.030)	(0.034)
Post-Secondary Non-Tertiary	0.359***	0.359***
	(0.069)	(0.066)
Short-Cycle Tertiary	0.483***	0.483***
	(0.079)	(0.090)
Bachelor or Equivalent	0.444***	0.444***
	(0.059)	(0.057)
Master or Equivalent	0.582***	0.582***
	(0.035)	(0.040)
Doctorate or Equivalent	0.648***	0.648***
	(0.094)	(0.099)
Age	0.020***	0.020***
	(0.004)	(0.004)
Age Squared	−0.0003***	−0.0003***
	(0.00005)	(0.00004)
Female	−0.420***	−0.420***
	(0.026)	(0.030)
Never Married/Not in Union	−0.070***	−0.070**
	(0.027)	(0.029)
Divorced	−0.023	−0.023
	(0.025)	(0.029)
Widowed	0.089**	0.089**
	(0.040)	(0.043)
Constant	8.745***	8.745***
	(0.112)	(0.108)
Observations	3,459	3,459
R^2	0.325	0.325
Adjusted R^2	0.319	0.319
Residual Std. Error (df = 3429)	0.383	0.383
F Statistic (df = 29; 3429)	56.938***	56.938***
Note:		*p<0.1; **p<0.05; ***p<0.01

Table A.29: Regression of Income 2013, pt. 1

	Dependent variable:	
	Equivalised Disposable Household Income in Logarithmic Form	
	OLS	White
	(1)	(2)
Stredocesky	−0.099***	−0.099***
	(0.032)	(0.035)
Jihocesky	−0.227***	−0.227***
	(0.036)	(0.039)
Plzensky	−0.093**	−0.093**
	(0.039)	(0.038)
Karlovarsky	−0.213***	−0.213***
	(0.061)	(0.062)
Ustecky	−0.225***	−0.225***
	(0.038)	(0.040)
Liberecky	−0.200***	−0.200***
	(0.044)	(0.045)
Kralovehradecky	−0.232***	−0.232***
	(0.040)	(0.044)
Pardubicky	−0.128***	−0.128***
	(0.041)	(0.042)
Vysocina	−0.200***	−0.200***
	(0.039)	(0.043)
Jihomoravsky	−0.166***	−0.166***
	(0.032)	(0.036)
Olomoucky	−0.240***	−0.240***
	(0.039)	(0.045)
Zlinsky	−0.256***	−0.256***
	(0.039)	(0.041)
Moravskoslezsky	−0.248***	−0.248***
	(0.033)	(0.035)
Number of Household Members	0.036***	0.036***
	(0.011)	(0.011)
Number of Household Members over 65	−0.017	−0.017
	(0.024)	(0.019)
Observations	2,948	2,948
R^2	0.274	0.274
Adjusted R^2	0.266	0.266
Residual Std. Error (df = 2918)	0.391	0.391
F Statistic (df = 29; 2918)	37.913***	37.913***
Note:		*p<0.1; **p<0.05; ***p<0.01

Table A.30: Regression of Income 2013, pt. 2

	Dependent variable:	
	Equivalised Disposable Household Income in Logarithmic Form	
	OLS	White
	(1)	(2)
Number of Household Members under 13	−0.074***	−0.074***
	(0.013)	(0.014)
Age of Youngest Child in HH	0.008***	0.008***
	(0.001)	(0.002)
Upper Secondary	0.176***	0.176***
	(0.034)	(0.036)
Post-Secondary Non-Tertiary	0.205**	0.205**
	(0.101)	(0.086)
Short-Cycle Tertiary	0.353***	0.353***
	(0.075)	(0.067)
Bachelor or Equivalent	0.405***	0.405***
	(0.063)	(0.065)
Master or Equivalent	0.497***	0.497***
	(0.039)	(0.042)
Doctorate or Equivalent	0.481***	0.481***
	(0.107)	(0.075)
Age	0.014***	0.014***
	(0.005)	(0.005)
Age Squared	−0.0002***	−0.0002***
	(0.00005)	(0.00005)
Female	−0.319***	−0.319***
	(0.029)	(0.033)
Never Married/Not in Union	−0.069**	−0.069**
	(0.028)	(0.030)
Divorced	−0.049*	−0.049
	(0.027)	(0.032)
Widowed	0.040	0.040
	(0.045)	(0.048)
Constant	9.342***	9.342***
	(0.121)	(0.127)
Observations	2,948	2,948
R^2	0.274	0.274
Adjusted R^2	0.266	0.266
Residual Std. Error (df = 2918)	0.391	0.391
F Statistic (df = 29; 2918)	37.913***	37.913***
Note:		*p<0.1; **p<0.05; ***p<0.01

Source: Author's own calculations base on LIS

Table A.31: Regression of Income 2016, pt. 1

	Dependent variable:	
	Equivalised Disposable Household Income in Logarithmic Form	
	OLS	White
	(1)	(2)
Stredocesky	−0.039	−0.039
	(0.029)	(0.031)
Jihocesky	−0.132***	−0.132***
	(0.032)	(0.030)
Plzensky	−0.066*	−0.066**
	(0.034)	(0.033)
Karlovarsky	−0.064	−0.064
	(0.057)	(0.057)
Ustecky	−0.219***	−0.219***
	(0.033)	(0.035)
Liberecky	−0.095**	−0.095**
	(0.038)	(0.039)
Kralovehradecky	−0.077**	−0.077**
	(0.036)	(0.035)
Pardubicky	−0.167***	−0.167***
	(0.035)	(0.034)
Vysocina	−0.160***	−0.160***
	(0.036)	(0.034)
Jihomoravsky	−0.089***	−0.089***
	(0.028)	(0.030)
Olomoucky	−0.111***	−0.111***
	(0.035)	(0.041)
Zlinsky	−0.176***	−0.176***
	(0.035)	(0.032)
Moravskoslezsky	−0.221***	−0.221***
	(0.028)	(0.029)
Number of Household Members	0.058***	0.058***
	(0.011)	(0.011)
Number of Household Members over 65	−0.027	−0.027
	(0.020)	(0.017)
Observations	3,206	3,206
R^2	0.306	0.306
Adjusted R^2	0.300	0.300
Residual Std. Error (df = 3177)	0.371	0.371
F Statistic (df = 28; 3177)	49.971***	49.971***
Note:		*p<0.1; **p<0.05; ***p<0.01

Table A.32: Regression of Income 2016, pt. 2

	Dependent variable:	
	Equivalised Disposable Household Income in Logarithmic Form	
	OLS	White
	(1)	(2)
Number of Household Members under 13	−0.105***	−0.105***
	(0.011)	(0.012)
Age of Youngest Child in HH	0.003***	0.003***
	(0.001)	(0.001)
Upper Secondary	0.263***	0.263***
	(0.030)	(0.034)
Short-Cycle Tertiary	0.329***	0.329***
	(0.074)	(0.056)
Bachelor or Equivalent	0.470***	0.470***
	(0.052)	(0.050)
Master or Equivalent	0.534***	0.534***
	(0.034)	(0.038)
Doctorate or Equivalent	0.444***	0.444***
	(0.083)	(0.062)
Age	0.015***	0.015***
	(0.004)	(0.004)
Age Squared	−0.0002***	−0.0002***
	(0.00004)	(0.00004)
Female	−0.283***	−0.283***
	(0.024)	(0.025)
Never Married/Not in Union	−0.155***	−0.155***
	(0.022)	(0.023)
Divorced	−0.050**	−0.050**
	(0.024)	(0.025)
Widowed	0.078*	0.078*
	(0.042)	(0.044)
Constant	9.283***	9.283***
	(0.101)	(0.100)
Observations	3,206	3,206
R^2	0.306	0.306
Adjusted R^2	0.300	0.300
Residual Std. Error (df = 3177)	0.371	0.371
F Statistic (df = 28; 3177)	49.971***	49.971***
Note:		*p<0.1; **p<0.05; ***p<0.01

Source: Author's own calculations base on LIS

Appendix B

Content of Enclosed ZIP file

We enclosed a ZIP file to this thesis, named *Determinants of Inequality Trends in the Czech Republic from 1992 to 2016*. Open file **READ_ME.pdf** first to navigate yourself.