Radó Márta Katalin

Tracking the Effects of Life Events on Subjective Well-being

Institute of Sociology and Social Policy

Supervisors: Tamás Bartus PhD and Judit Monostori PhD

Corvinus University of Budapest

Doctoral School of Sociology

Tracking the effects of life events on subjective well-being

Doctoral Dissertation

Radó Márta Katalin

Budapest, 2018

Table of Contents

List of Tables

List of Figures

Acknowledgements

First and foremost, I would like to thank my supervisors, Tamás Bartus and Judit Monostori, for their guidance. You made me believe in my work more, since I could trust that you would get me back on the track if I took a wrong turning. Thanks for always being accessible and ready to discuss my concerns.

Furthermore, thanks to my bosses Károly Takács and Róbert Gál. You not only helped with this work, but also inspired me to continue research beyond and after this work. Thanks to RECENS and its members for providing the best working environment one could wish for over the last three years. And thanks for JKK for helping me continue my work in academia.

I also would like to express my gratitude to colleagues, advisers and teachers who helped me to develop this work. First, I am sincerely thankful to Ágnes Neulinger who wrote several papers with me, including Chapter 6. Ági, you are the unofficial supervisor of this dissertation. Further, I am most grateful to Levente Littvay, who taught me applied methods and at the end read and commented on the dissertation. Further, I am thankful to Tamás Hajdú and György Lengyel for commenting on this work at every possible stage. I am also very grateful to Alexis Diamond for reading and commenting on my research proposal and helping me to obtain deeper knowledge of applied methods. Finally, I am also really grateful to those who commented on this work, answered my questions, and devoted time and energy to helping me finish my Ph.D., including Beáta Nagy, Gábor Király, Lilla Vicsek, Márton Csillag, Renáta Németh, Katalin Melles, Elizabeth Stuart, Bruno Arpino, Zsófia Tardos, Konrad Turek, Jaap Oude-Mulders, Christoph Bein, my PhD classmates, and many others who commented on this work at conferences or as anonymous reviewers.

I want to thank my family for their support. I am grateful to my parents, who were the first people to instil in me a love of studying. Thanks Apu, to always helping with your professional comments when there was no one else to answer my questions. Thanks Anyu for remaining my greatest source of emotional support through my life. Thanks Matyi and Hanna for being there and listening the weekly reports on how my work was developing. Finally, I would like to dedicate this work to my late grandmother, Mariska, who was physically present while I wrote the majority of this dissertation. Her all-embracing empathy and kindness paved the way for my personal and professional development.

Thanks Michael, for always having faith in me and supporting me both academically (as a co-author of the retirement paper) and beyond. You are my inspiration.

And finally, thanks to friends who put a smile on my face when I needed it.

The data for the analyses were supplied by the Hungarian Statistical Office. The author would like to thank Simon Milton for his language correction. The author received a National Talent Scholarship (NTP-NFTÖ-16-0848) from the Hungarian state, a Research Excellence Award by the Hungarian National Bank and Corvinus University of Budapest, and a Campus Mundi short-term research visit (279094) which supported the research and writing of this dissertation.

1. Introduction

1.1. Aim of the Dissertation

Who do you think is more satisfied with their life: someone who has won the lottery, or someone who became paralyzed on the same day? The vast majority who are asked this question would immediately choose the lottery winner. The answer seems to be obvious, as the lottery winner may expect more reward and less difficulty than a paralyzed person. Surprisingly, however, empirical research has shown that there is no significant difference between how people feel one year after those two dramatically different events (Brickman, Coates, & Janoff-Bulman, 1978). The reason for this counter-intuitive finding is that people eventually adopt to their new situations and, with this adaptation, their level of subjective well-being returns to its initial state. However, this type of adaptation does not occur for every event in life, since some, like unemployment, have often been found to have long-lasting effects[1] (Lucas, Clark, Georgellis, & Diener, 2004). Nevertheless, Kahneman and Krueger pointed out that this example "*challenges both everyday intuition and economic doctrine, by suggesting that in the long-run well-being is not closely related to one's circumstances and opportunities*" (Kahneman & Krueger, 2006: 16). This statement calls for further research to help challenge unsubstantiated intuition.

This dissertation is written to add to the state of the present knowledge by showing how specific life events affect subjective well-being in Hungary. There is growing demand for such research. First of all, the outcome variable of this dissertation has risen in importance in policy analysis since researchers have started arguing that the ultimate goal of politics is to produce life satisfaction, not only economic well-being (Rothstein, 2010). In contrast to financial welfare, subjective well-being captures what the stakeholder thinks is a "good life". Understanding the stakeholder's perspective is especially important when non-market outcomes are involved, such as family ties or changes in social relationships (Rothstein, 2010; Thomas & Thomas, 1928). This dissertation investigates such phenomena.

The applied theoretical framework in this dissertation is *life course theory*, which emphasizes how individuals' life trajectories can contribute to understanding macro-level changes. Based on this approach, individuals are considered the agents of demographic change, thus one needs to focus on decision formation at this level to gain deeper

[1] In other cases, life events have more complex effects; for example, they can have non-linear or heterogeneous effects (Lucas, 2007).

understanding about trends at the macro level (Elder, Johnson, & Crosnoe, 2003; Hitlin & Kirkpatrick, 2015; Kok, 2007) such as low fertility rates or widespread early retirement. Thus, the aim of this dissertation is to increase understanding of the micro mechanisms which underlie macro-level demographic changes.

Early demographic research emphasized that individuals make life-stage-related decisions based on utility maximization. For example, these early theories were influential at explaining the decrease in fertility rate, increase in number of divorces, and spreading early retirement. Based on this perspective, fertility has decreased because the economic utility obtained from having children has decreased due to the emergence of social security systems (Becker, 1981; Becker & Barro, 1988; Boldrin & Jones, 2002). This approach has also contributed to understanding the increasing divorce rate: the phenomenon was attributed to the fact that females have become less dependent on males economically due to increasing female labour market participation and income (Becker, Landes, & Michael, 1977; Hannan, Tuma, & Groeneveld, 1978; Spitze & South, 1985). Further, spreading early retirement in Hungary has also been considered to be a reaction to poor labour market opportunities or economic instability after the transition (Scharle, 2012). Despite the undeniable contribution of this early research, these arguments solely focused on objective economic well-being in explaining life-stage transitions.

Recently, some scholars have suggested that subjective indicators, such as life satisfaction, could also capture another important component of life-stage-related decisions. These authors have argued that subjective well-being has become the major engine of the current demographic changes in contemporary developed societies, because today the "*quest for happiness*" is the major concern of individuals (Billari, 2009; Caldwell & Schindlmayr, 2003; Hobcraft, 2006). Thus, they emphasize that the goal of many individuals today is not simply to maximize economic well-being, but rather to obtain a higher level of subjective well-being. For example, people have a child or retire since it makes them more satisfied, not merely because these life events are economically rational. For example, Billari (2009) formulated the hypothesis that fertility rates often decrease in contemporary developed societies since parenthood is unsatisfactory.

Moreover, observation of trends in subjective well-being can reveal those groups which are exposed to a higher level of risk at certain stages of life (Ferraro & Shippee, 2009). The effect of specific life events may significantly vary across different social groups since life trajectories have become less stable, less pre-determined, more unpredictable, and de-standardized (Kohli, 2007; Macmillan, 2005). Furthermore, one's life course depends on earlier stages, therefore the same event might have different consequences for different individuals based on the advantages and disadvantages they have accumulated (Kok, 2007; Kohli, 2007). Thus, the research in this dissertation was designed to capture the heterogeneous effects of life events on various social groups. For example, single parents and involuntary retirees are distinguished. Through observation of these groups, the dissertation seeks to contribute to understanding the inequalities that arise over a life course and to examine how social institutions can mitigate them.

Up-to-date, state-of-the-art research about the effect of life events on subjective well-being has mostly been restricted to Western countries, and little is known from Central-Eastern Europe (Baranowska & Matysiak, 2011; Sironi & Billari, 2013). However, observation of this phenomenon is especially important in this region since transition countries historically suffer from significantly lower life satisfaction than that of Western European countries, even more than 20 years after the major political changes. They are still divided by an iron curtain of unhappiness (Guriev & Zhuravskaya, 2009). Furthermore, the individuals of these countries have a lower standard of living than those in western countries, thus in this region the options for the "*quest for happiness*" are limited. Moreover, welfare regimes differ significantly from those in better studied western countries (Draxler & Van Vliet, 2010; Manning, 2004; Polese, Morris, Kovács, & Harboe, 2014). Finally, international research findings are often mixed about the effect of life-course events on subjective well-being, thus there is no obvious universal effect which would allow us to apply the results obtained in other countries to the Central Eastern European context. Thus, this dissertation responds to the growing demand for research in the Central Eastern European region.

More specifically, the issue is here re-examined in the Hungarian context. Hungary in particular represents an interesting case because it has some of the lowest levels of subjective well-being in the region, even compared to neighbouring countries. Moreover, in most cases economic development brings about higher average subjective well-being; however, in Hungary, economic growth between 1980 and 2004 was accompanied by a decrease in subjective well-being (Sacks, Stevenson, & Wolfers, 2010). Further, the permanently low fertility rate and high level of involuntary retirement call for research into the effect of parenthood and retirement. Moreover, in Hungary life transitions are embedded in a distinct economic, cultural and social context (See more in Chapters 4.2.4 and 5.2.3). In Hungary until now the research has mostly been restricted to analyses of associations using cross-sectional data. However, with the development of computational power, new methods have been invented that provide better estimations of causality.

1.2. Research Questions

To achieve the aims described above, three research questions were posed. First, the question '*How does parenthood affect overall subjective well-being in Hungary?'* was addressed. Overall subjective well-being is measured here as "*life satisfaction*". This question was tested on a longitudinal dataset. This chapter included more sub-questions; more specifically, it estimated the effect of *(1/a)* overall parenthood, *(1/b)* motherhood, *(1/c)* fatherhood, *(1/d)* having a first child, and *(1/e)* having a second child.

Second, the research described in this dissertation also investigated *how retirement affects overall subjective well-being in Hungary*. To address this issue, the same longitudinal dataset was again used as for the previous question. Here, two sub-questions were tested. First, (2/a) how retirement in general changes subjective well-being was observed. Then, (2/b), the difference between the effect of voluntary retirement and involuntary retirement on subjective well-being was estimated.

Finally, this dissertation also aimed to address *how household life-cycle status affects domain-specific subjective well-being in Hungary*. In contrast to the previous two research questions, this question was tested on a cross-sectional dataset which also enabled us to observe domain-specific subjective well-being. The effects of the following stages of life were estimated; *(3/a)* being young and childless, *(3/b)* being a parent with a young child, *(3/c)* being a parent with an older child, *(3/d)* being a single parent, *(3/e)* being a middle-aged childless person, *(3/f)* being an older childless person with a partner, and *(3/g)* being an older childless person without a partner. Correspondingly, the observed outcomes were satisfaction with life-course, future opportunities, quality of standard of living, family relations, health, work/job, housing, place of residence, income, and life as a whole (i.e. overall subjective well-being).

1.3. Causality and Methods

The present study intended to examine as closely as possible the causal relationship between certain statuses (belonging to a specified life-stage group) and subjective well-being, and adopted the potential outcome framework for this purpose (Diamond & Sekhon, 2013; DuGoff, Schuler, & Stuart, 2014; Fisher, 1925; Ho, Imai, King, & Stuart, 2011; Holland, 1986; Imai, King, & Stuart, 2007; Imai & Van Dyk, 2004; Imbens & Rubin, 2015; Neyman, 1923; Rosenbaum, 2002; Rosenbaum & Rubin, 1983, 1984, 1985a, 1985b; Rubin, 1974, 1978; Stuart, 2010). In this framework, the key independent variable – here, life-stage status – is called the treatment variable and the dependent variable –here, subjective well-being – is referred to as an outcome variable (the convention is used from now on).

A causal conclusion would theoretically require a comparison of the outcome variable for the case in which a given individual receives treatment and the case in which this individual does not receive treatment. This method is, however, not feasible with respect to the current research topic, as an individual may only belong to a single life-stage group at any specific time. This problem is referred to as a *fundamental problem of causal inference* (Holland, 1986) or *identification problem* (Imbens & Wooldridge, 2009; Kézdi, 2004).

Randomized experiments overcome this obstacle by introducing a control group that does not receive the treatment, allowing its attributes to be compared with those of the treatment group. Comparability results from the individuals are assigned to either the control or the treatment group in a random way; therefore, the treatment and control groups only differ from each other due to chance (apart from the treatment itself, naturally). Thus, the method enables a comparison to be made between individuals who belong to the specified life-stage group and those who belong to the control group that is similar in every possible way (observed or unobserved) other than their life-stage group membership (Ho et al., 2007). However, the research topic at hand does not allow a random experiment to be conducted because the researcher cannot arbitrarily decide which life-stage group the person under consideration should belong to.

In the case of the present research, I needed to rely on observational data; however, regarding such observational data, the control and the treatment group exhibited systematic differences (Rosenbaum, 2002). In other words, the members of a given life-stage group do not differ from other members of the population in terms of their current life stage only, but also in terms of a number of other variables. For instance, older generations are typically characterized by lower levels of education than younger ones, thus any differences in subjective well-being between the different life-stage groups may also stem from their education, and not from life-stage group membership alone.

To estimate causal relationships in the absence of the required experimental arrangement, this dissertation uses statistical methods to derive causal inferences from observational data: namely, matching, regression adjustment, and longitudinal analysis. These methods are aimed at estimating causality between the treatment and outcome variables by controlling for the common causes of the treatment and outcome variables, but omitting the common outcomes of these two key variables (Elwert & Winship, 2014; Rosenbaum, 1984).

First, a matching method was used to establish a quasi-experimental arrangement. The essence of this method is to assign each member of the treatment group a non-treated person(s) who is (are) as similar as possible to the latter regarding every observed variable other than the treatment itself (Diamond & Sekhon, 2013; DuGoff, Schuler, & Stuart, 2014; Holland, 1986; Rubin, 1974, 1978; Stuart, 2010). Thus, this method aims to replicate the experimental design.

After performing matching, regression adjustment is used to increase the similarity of the treatment and control groups. Nevertheless, the literature on causal inference emphasizes that performing a matching procedure prior to running a regression model is indispensable, as regression alone tends to perform poorly (results could be exposed to interpolation and extrapolation bias) unless there is sufficient overlap between the control and treatment groups (DuGoff et al., 2014; Ho et al., 2011; King & Zeng, 2006; Kuo, 2001).

Finally, where longitudinal data were available, than a longitudinal extension of the matching method was used. This extension allowed us to control for time-invariant unobserved variables such as personality traits. More specifically, I used (in Chapters 4 and 5) the matching method combined with a longitudinal method; the regressor variable method (Allison, 1990).

Finally, sensitivity analysis was used to test how sensitive the results are to the model specifications. Matching and regression are able to control for observed variables only, whereas a longitudinal design also rules out the effect of time-invariant unobserved variables. However, even when these methods are used together, unobserved time-variant variables are not controlled for. Therefore, it is essential to test the sensitivity of the estimates to these omitted unobserved time-varying variables (Rosenbaum, 2002). However, even after the application of sensitivity analysis one cannot identify causal conclusions for sure. One can argue at most that the omitted variables probably do not modify the results to a large degree. Thus, the research described herein admittedly does not draw causal conclusions, but only estimates causality.

1.4. Structure of the Dissertation

This dissertation proceeds in the following steps to address the above-mentioned research questions: Chapter 2 outlines the theoretical background of this dissertation. It first contains a review of the literature about subjective well-being, and then presents the main theoretical framework – namely, life course theory. Through this review, the main concepts and definitions are introduced which are essential for understanding the subsequent parts.

Chapter 3 details the methods that were applied and elaborates the potential outcome framework which was used to estimate causal inference. More specifically, it provides details about matching methods, regression adjustment, longitudinal analysis and sensitivity analysis. The present dissertation also contributes to the methodological literature by presenting the different types of matching procedure in a sample dataset and revealing the computations underlying these complex algorithms. Furthermore, it also illustrates on a small sample dataset why regression adjustment without matching often fails to estimate causality (in a similar manner to Ho et al. [2011]).

The three empirical studies are presented in Chapters 4, 5, and 6. These studies further justify the research questions by discussing existing theories, and reviewing earlier empirical findings. Chapters 4 and 5 focus on the effect of two specific life events on overall subjective well-being using a longitudinal dataset. More specifically, the first estimates the effect of parenthood and the second the effect of retirement. Meanwhile, Chapter 6 shows how life-stage status affects domain-specific subjective well-being using a cross-sectional dataset.

Chapters 4 and 5 use a different approach to Chapter 6 to observe how life events affect subjective well-being and are based on a longitudinal dataset (*Turning Points of Life Course*), whereas Chapter 6 uses only a cross-sectional dataset. The longitudinal dataset enables more accurate causal estimation. However, the drawback of the longitudinal dataset is that it contains only limited information about domain-specific subjective well-being (i.e. does not contain information about satisfaction with income). Thus Chapters 4 and 5 do not distinguish domain-specific subjective well-being, but only observe overall subjective well-being. However, Chapter 6 contains domain-specific subjective well-being measures as well. Thus, the cross-sectional design has less statistical power, but the dataset enables us to gain deeper understanding of the pluralizing effect of specific transitions by disentangling their effects regarding specific domains. Furthermore, this later study also observes hard-to-reach groups that could not be investigated in the representative longitudinal sample (i.e. single parents) due to the small sample size. To sum up, the two types of dataset supplement each other and together provide a bigger picture of the topic under observation.

At the end of the dissertation, findings are discussed (Chapter 7). This part summarizes the contribution of this dissertation to the state of the knowledge, explores its limitations, and make suggestions for further research.

2. Theoretical Framework

2.1. Conceptualizing Well-being

This dissertation uses well-being as an outcome variable. There are different ways to conceptualize well-being, but basically all of them elaborate the somewhat vague idea of a "good life." We can distinguish two kinds of well-being: objective well-being, and subjective well-being. The former refers to universal needs, while the latter recognizes psychological factors and the role of personal evaluation. In the 1970s a heated debate arose about the relationship between subjective and objective well-being. During this debate, it was realized that a relationship between subjective and objective measures exists, although these relationships are often weak. Cummins (2000) also points out that the correlations within subjective indicators and within objective indicators are stronger than between subjective and objective indicators. Furthermore, the Easterlin paradox also captures the difference between objective and subjective indicators: the theory says that among the developed countries there is no significant correlation between income and self-reported level of happiness (Easterlin, 1974).

There have been several efforts to explain the difference between objective and subjective indicators. Diener, Suh, and Oishi (1997) stress that the discrepancy between subjective and objective indicators can be attributed to the fact that individuals compare themselves to a reference group, and the group of relevant others varies among life conditions and statuses. This theory explains why people living under favourable conditions usually underestimate their position. Cummins (2000) argues that people's normal state is happiness, because this helps them to deal with daily challenges. He argues that people tend to distort reality using misperceptions in order to maintain their happiness. Thus, they tend to overestimate the positive nature of their lives and accomplishments. Finally, Michalos (1985) argues that subjective indicators depend on the will, evaluation and experience of the individual, and not only on their possessions.

There has been a long debate about whether objective or subjective indicators should be used in policy analysis. We can primarily distinguish between two kinds of approaches: the *Scandinavian welfare approach*, and *the American quality-of-life approach*. The first evaluates level of welfare with the help of objective measures. In other words, this approach defines well-being as "*individuals' command over, under given determinants, mobilizable resources, with whose help he/she can control and consciously direct his/her living conditions*" (Erikson, 1974: 275). Resources can include, for example, money, property, knowledge, mental and physical energy, social relations, and security. So this approach focuses on objective indicators of living standard. In contrast, the *American quality-of-life approach* appraises welfare using subjective indicators (Erikson, 1974). Thomas and Thomas argued that subjective indicators are important, since "*if men define situations as real, they are real in their consequences*" (Thomas and Thomas, 1928: 571-572). This approach emphasizes that the individual is best positioned to evaluate their own quality of life. This dissertation applies the *American quality-of-life approach* by using subjective indicators as outcome variables.

Several typologies of social indicators have been developed which come in useful in understanding how subjective well-being is related to other social indicators. Schulz (2000) has distinguished four types of variables (see Table 1), which range from objective to increasingly more subjective measures. These set of variables are defined as *social structure* (Group A), *resources and behaviour* (Group B), *evaluation of living conditions* (Group C), and *subjective quality of life* (Group D). Groups C and D are considered to contain the variables in which evaluation and cognition are clearly predominant, thus one can consider these variables as subjective indicators.

Table 1. Schulz typology of social indicators

Group A Social structure	Group B Resources and behaviour (living conditions)	Group C Evaluation of living conditions	Group D Quality of life
Socio-demographic (e.g. sex, age)	Standard of living (e.g. housing, health)	Domain Satisfaction, importance of life domains, perceived need, and fulfilment	Well-being, satisfaction, and happiness

Moreover, subjective indicators are also often categorized. Hegedűs (2002) has distinguished three kinds of subjective indicators. The first describes subjective well-being in an indirect way. For example, a variable which measures one's preferred way of spending taxes belongs to this category. The second type of subjective indicator measures well-being in a direct way. Hegedűs argues that an individual's perception of their income belongs to this category. Finally, the third type goes beyond a simple description of well-being and also incorporates an evaluation process. For example, this third type contains measurements of satisfaction with life and happiness.

Extended research has been conducted into subjective well-being,[2] but nevertheless, there are still conflicting ways to define it. First and foremost, the concept emerged in the work of Aristotle, who considered subjective well-being as people's perceptions of meaning, purpose and growth. This definition goes beyond the concept of happiness in its observation of how people realize their human potential (Graham & Nikolova, 2015). More recently, Diener (1984) defined subjective well-being as a general evaluation of a person's life. The author considers happiness and life satisfaction to be the components of this concept. In contrast, Ahuvia and Friedman (1998) have said that subjective well-being is a general and long-term state that consists of cognitive and affective components. Thus, they replaced the concept of happiness with life satisfaction and affect. In a recent article, Diener et al. (2016: 3) reviewed earlier research on subjective well-being and concluded that it is a "*broad umbrella term that refers to all different forms of evaluating one's life or emotional experience*". In this dissertation, this latter definition is used.

Within subjective well-being we can distinguish further categories. Kahneman, Diener and Schwarz (1999) have also distinguished three categories of subjective well-being: life evaluation, hedonic well-being, and eudemonic well-being. First, *life evaluation* captures people's thoughts about the quality of their lives and their overall life satisfaction. For example, the Cantril ladder is a typical way of measuring this kind of subjective well-being, whereby individuals are asked to place themselves on eleven-point scale, in which the lowest number represents the worst possible life and the highest number the best possible life. Second, *hedonic well-being* captures an individual's mood or feelings (e.g. experiences of happiness, sadness, and anger). The easiest way to measure this concept is to ask the respondent to rate how much they experience these feelings. To measure this aspect of subjective well-being both positive and negative feelings should be on the list for evaluation. Finally, *eudemonic well-being* focuses on individuals' opinions about the meaning and purpose of their life. From these three forms of subjective well-being, this dissertation focuses on life satisfaction.

The concept of subjective well-being can also be studied from a life-domain perspective. This approach measures well-being with multiple questions instead of just observing it

[2] Based on Diener et al. (2016), the topic of subjective well-being was investigated in 14,000 publications in 2015.

through a single question. The single-question approach focuses on satisfaction with whole-of-life, while the use of a series of questions captures different life domains as the components of satisfaction with life as a whole. The single-question approach usually applies the following question: "*Taken all together, how satisfied are you with your life?*" (Easterlin, 2001; Pavot & Diener, 1993; Veenhoven, 2015). Measuring subjective well-being with multiple questions has gained popularity in recent years (Cummins, Eckersley, Pallant, Van Vugt, & Misajon, 2003; Ganglmair-Wooliscroft & Lawson, 2011), and the *Satisfaction with Life Scale* is one of the most often used tools (Diener, Emmons, Larsen, & Griffin, 1985). Sirgy (2012) says the life domain satisfaction usually includes satisfaction with material well-being, work well-being, social/family well-being, leisure well-being, and residential well-being. Furthermore, others have also added satisfaction with past life course and satisfaction with future opportunities to this list (Diener & Seligman, 2002; Heckhausen, Dixon, & Baltes, 1989; Lachman, Röcke, Rosnick, & Ryff, 2008; Röcke & Lachman, 2008).

The difference between overall subjective well-being and domain-specific subjective well-being is that the first approach is top-down, whereas the second is bottom-up. The top-down perspective considers overall subjective well-being a function of personality and other stable factors. Thus, based on this approach, overall subjective well-being mediates between stable factors and satisfaction with life domains. In contrast, the bottom-up perspective assumes that overall life satisfaction is given as a function of respondents' satisfaction with many concrete domains of life (Diener, 1984; Heller, Watson, & Ilies, 2004). Although these two approaches are different, they are not contradictory, and both can work at the same time (Erdogan, Bauer, Truxillo, & Mansfield, 2012).

Numerous critiques have been articulated about the measurement of subjective well-being. Ryff (1989), for instance, criticized this measure because he found it to be too atheoretical. Tversky and Kahneman (1974) have argued that individuals tend to make fast decisions and use cognitive short cuts (so called *cognitive heuristics*) when responding to abstract questions such as "*How satisfied are you with your life?*" Others have questioned the reliability and validity of subjective well-being by arguing that this measurement mostly captures how the individual feels at the time of the questionnaire. Some researchers have documented that the circumstances of interviews can considerably influence subjective well-being. For example, people tend to change their responses if they have found money before filling out the questionnaire; furthermore, the current weather, and even the order of questions might influence answers (Schwarz, 2013).

Despite all the critiques, subjective well-being is widely accepted as a reliable and valid measurement. Several pieces of research have underpinned the claim that subjective well-being indeed measures the underlying concept as intended, and repeated tests can produce similar results (Diener, Sandvik, Seidlitz, & Diener, 1993; Ito, Sagara, Ikeda, & Kawaura, 2003; Krueger & Schkade, 2008; Larsen, Diener, & Emmons, 1985; Lepper, 1998; Lucas, Diener, & Suh, 1996). Krueger and Schkade (2008) have pointed out that although subjective well-being tends to be less reliable than objective measures (i.e. income and education), the reliability of the former is "*probably sufficiently high to support much of the research that is currently being undertaken on subjective well-being, particularly in cases where group means are being compared*" (Krueger & Schkade, 2008: 23). Other studies have encouraged the usage of subjective measures by pointing out that they predict observable events well. People with higher subjective well-being are less likely to commit suicide (Koivumaa-Honkanen et al., 2001), have better health, die later (Palmore, 1969; Sales & House, 1971), and even smile more (Ekman, Davidson, & Friesen, 1990). Some subjective measures even do better than their corresponding objective measure at predicting life events. For example, subjective health status has been found to be a better predictor of mortality than doctoral opinions (Idler & Benyamini, 1997).

Generally speaking, subjective well-being is influenced by individual characteristics and the effect of these covariates are quite stable, even in international comparison. Individual circumstances such as health status, income, education level, gender, and employment have a considerable impact on subjective well-being (Hegedűs, 2001; Hegedűs & Lengyel, 2002; Molnár & Kapitány, 2013; Spéder & Kapitány, 2002; Wang & Hesketh, 2012). Furthermore, personality has been found to be a major factor in subjective well-being (Steel, Schmidt, & Shultz, 2008), although this covariate is more significant in the case of happiness (Schimmack, Oishi, Furr, & Funder, 2004; Steel et al., 2008).

This dissertation uses life satisfaction to measure well-being. In Chapters 4 and 5 the single question approach is used, since the measurement of life-domain satisfaction is not well developed in the given dataset. In Chapter 6, life domain satisfaction is already distinguished.

2.2. Life Course Approach

This dissertation estimates the effects of life events on subjective well-being in the light of the *life course* paradigm. The concept of life course can be defined as "*the age-graded,* *socially-embedded sequence of roles that connect the phases of life*" (Mortimer & Shanahan, 2007: XI). Furthermore, life course can be also considered a paradigm. This paradigm has been defined as "*an imaginative framework comprised of a set of interrelated presuppositions, concepts, and methods that are used to study these age-graded, socially embedded roles*" (Mortimer & Shanahan, 2007: XI).

The life course approach has fundamentally changed the agenda of demographic research due to its multi-level nature. Similarly to the previous demographic approaches, it recognizes societal-level factors such as cultures, structures, institutions, political factors, and economic conditions which may determine the course of an individual life. However, it also highlights individual-level factors, thus considering the individual to be the agent of demographic action. As a consequence, in the 1970s demographic research shifted from macro- to micro-level analysis (Kok, 2007).

This approach not only shows the individual's life in snapshots, but also as a process in which every moment is determined by previous stages. In other words, previous experiences and decisions create both opportunities and obstacles for the individual. Since every life event leaves a mark on the whole life course, one cannot understand a single life event without considering the previous stages (Elder et al., 2003; Kok, 2007). Thus, early life events can have an effect on later-life outcomes. For example, childhood disadvantages can have an effect on subsequent body mass index (Elsenburg, Smidt, & Liefbroer, 2017), adult labour market income (Bartley et al., 1994; Black, Devereux, & Salvanes, 2007), educational attainment (Black, Devereux, & Salvanes, 2007), and mortality (Oreopoulos, Stabile, Walld, & Roos, 2008). Furthermore, certain life events, such as unemployment, have been found to have long-lasting effects on subjective well-being (Clark, Diener, Georgellis, & Lucas, 2008). Finally, based on this approach, gender also plays a key role in understanding life events because females and males have different experiences (Krüger & Baldus, 1999). Thus, this theoretical framework calls for studies which recognize the time-related interdependence between various life events.

Despite the individual focus of life course research, the approach also incorporates the institutions that shape individuals' opportunities and limitations. In this approach, one of the influential institutions is the family within which the individual lives. Although "family" and "household" are not synonyms, the two concepts often overlap. Thus, along with recognition of the importance of the family context, observation of the household has also gained in importance in this paradigm (Uhlenberg & Mueller, 2003). Consequently, in this dissertation the household context is also considered to play a key role in understanding transitions. The third empirical study in Chapter 6 has a specific focus on household life cycles.

Furthermore, society as a whole shapes the life course of its citizens. Every society has norms concerning life courses. These norms provide information about the "desirable" sequence of life events. For example, norms influence thoughts about at what age members of society should have a child or retire. This is what Kohli (1993) calls the institutionalization of life course. The former author pointed out that "*institutions may start as purposive social constructions but gradually become self-evident as they turn into second nature, be it in terms of a shared belief system or of a taken-for-granted structural reality*" (Kohli, 2007: 257).

Furthermore, life course can be directly influenced by policies or programs (such as education, parental leave, or retirement system reforms) that are introduced by the government. Such policies can contribute to the regulation or standardization of individuals' life courses (Mayer & Schoepflin, 1989). Accordingly, welfare states influence how the individual experiences transitions such as parenthood or retirement (Leisering, 2003). Consequently, this dissertation places high importance on describing the context (namely, the Hungarian context) in which the observed transitions occur.

In the last few decades, life courses have become less standardized, thus life events also have a pluralized effect. Brükner and Mayer (2005:28) argues that the idea that "*lives have become less predictable, less collectively determined, less stable, less orderly, more flexible, and more individualized has become one of the most commonly accepted perceptions of advanced societies.*" Nevertheless, the greater complexity and diversity of life paths means that life events such as fertility and retirement might have very different consequences for different social groups. Therefore, in this dissertation the conditions that may modify the effects of life events are also investigated. For example, the first empirical study not only estimates the effect of parenthood, but also evaluates how the number of children and gender modify this effect. Furthermore, the second empirical study not only focuses on the effect of retirement, but also observes whether voluntary and involuntary retirees experience retirement differently.

Finally, the life course approach also provides a methodological framework. Based on this framework, researchers should rely on (1) individual-level data, in order to grasp micro mechanisms, and (2) longitudinal data, which provides a bigger picture of the individuals' life course and supports the observation of interdependency between life stages (Elder et al., 2003). Thus this methodological framework calls for an individual-level longitudinal dataset such as the *Turning Points of Life Course*, which is used in the first and the second empirical study.

For the observation of life course, one needs to distinguish between ageing effects, cohort effects, and stage effects. The ageing effect highlights how certain characteristics of individuals change over their life course, whereas the cohort effect compares people from different cohorts. The observation of ageing requires longitudinal data, while in the case of cohort effect cross-sectional data is sufficient (Frijters & Beatton, 2012; Mason & Lee, 2006). Finally, the stage effect has to be distinguished from the previous two phenomena. This reflects on the effect of certain transitions such as fertility or retirement (Frijters & Beatton, 2012). The first and second studies in this dissertation focus on the latter concept; however, even in these cases one needs to have prior knowledge about the age effect, since ageing provides a context for any demographic life events. The third study combines an estimation of ageing effect and life stage effect, as in this study life cycles are defined by age, parenthood status, and partnership status.

3. Analytical Strategy

3.1. Association and Causality

Suppose we have a population of N individuals[3], in which each individual is indexed by i ($i = \{1, 2, 3...N\}$). Let Y_i denote the outcome for a given individual i and J_i denote a binary variable which takes a value of 1 if the individual receives a treatment and 0 if they receive the control. In the applied framework, the key explanatory variable is referred to as the *treatment variable,* therefore, this term will be used in the following parts of the dissertation as well. This dissertation deals with a simple treatment-and-control setting[4], thus the J_i treatment variable is a binary variable.

$$J = \begin{cases} 1 \ \textit{if i individual recieves the treatment} \\ 0 \ \textit{if i individual recieves the control} \end{cases}$$

The $N_{(j=1)}$ number of individuals who receive the treatment $(j = 1)$ belong to the K *treatment group,* whereas the $N_{(j=0)}$ number of individuals who do not receive the treatment $(j = 0)$ belong to the L *control group*. Let us assume that every i individual either receives the treatment or does not ($i \in K$ or $i \in L$). K contains those k individuals who actually receive the treatment $(j = 1)$ and L contains the l individuals who do not $(j = 0)$.

One of the goals of this dissertation is to estimate causality between life stages and subjective well-being. Regarding causal inference I seek to measure what would happen to Y_i outcome as a result of a J_i treatment variable. In this case, let Y_i outcome variable denote subjective well-being for a given individual i, and J_i denote a binary variable corresponding to life-stage groups which takes a value of 0 if an individual is not at the observed life stage and 1 if the individual is. To understand the relationship between these two variables, two concepts need to be distinguished: association and causality. Association refers to dependence between two variables but does not necessary imply causality. For example, regarding the topic of this dissertation, one might find that there

[3] The unit of analysis can be something other than an individual (e.g. a household, company, or country), but this dissertation uses individuals as the unit of analysis, so for the sake of simplicity, this chapter illustrates the methods using this unit.

[4] Alternatively, one may observe a continuous treatment effect or a multiple treatment effect; however, this subject is outside the scope of this dissertation.

is an association between a life stage and subjective well-being, but this does not mean that entering into that given life stage actually causes a change in subjective well-being.

To better understand the difference between the two concepts, let me review the types of associations. Variable J_i and variable Y_i can be associated with each other in three ways (Elwert & Winship, 2014):

1. Either of them affects the other directly or indirectly
2. They share a common cause
3. They are conditioned on a common outcome (or are descendants of a common outcome)

The first type of association implies causality, but the second and the third types do not imply causality but only association. Estimating causality requires that the latter two types of association be ruled out.

The second type of association occurs, for example, when one is interested in the effect of parenthood on subjective well-being, but the treatment and control groups differ in their subjective well-being even before the birth of a child. Imagine two hypothetical situations: In the first situation, members of the treatment group (those who had a child) have significantly higher subjective well-being than the control group (those who did not have a child) even before exposure to the treatment (childbirth). This situation occurs if people with higher subjective well-being have a higher probability of having children. This situation is illustrated in Figure 1, which shows the distribution of the outcome variable in the treatment and control groups before exposure to the treatment. In this case, the difference in post-birth subjective well-being between the treatment and control groups may be attributed to pre-birth differences, not necessarily to the birth of a child. Thus, simple comparison of the post-treatment outcome variable might indicate a positive association or an insignificant association, even in the case of a negative causal effect.

Similarly to the first hypothetical situation, one can imagine a situation in which people with lower subjective well-being have a higher probability of having children. The distribution of the outcome variable in the treatment and control groups before exposure to the treatment is illustrated in Figure 2. In this case, comparison of the post-treatment outcome can also be misleading when estimating causality: it is possible that parenthood indeed has a positive effect on subjective well-being, but the comparison of the mean shows an insignificant or even negative association between the treatment and outcome variables due to initial differences.

Figure 1. Distribution of subjective well-being before treatment in hypothetical situation 1[5]

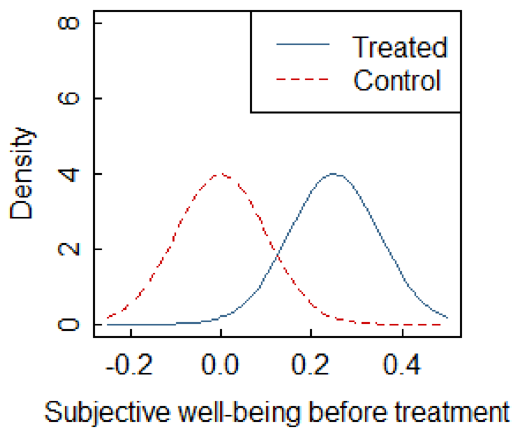
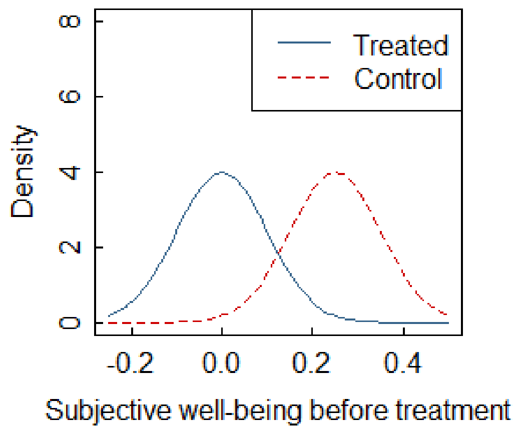


Figure 2. Distribution of subjective well-being before treatment in hypothetical situation 2[6]



Further, the treated and control groups can differ from each other not only in terms of the outcome variable, but other covariates as well. For example, it is possible that richer people have a higher level of subjective well-being, and are also more likely to have children. In this case, the reason subjective well-being is high may not be that the given individual has a child, but rather that they are rich. These common causes are referred to as *confounding variables* in the statistical literature. Upon estimating causality, failing to account for these variables creates *selection bias*.

The third case is less intuitive, but also plays a crucial role in modelling causality. Assume that motherhood is marginally independent from subjective well-being. Moreover, there are studies which show that motherhood has a negative effect on productivity (Bryson, Forth, & Stokes, 2015). Furthermore, other studies have found that people with higher subjective well-being tend to be more productive (Bryson et al., 2015). Thus conditioning on productivity can create a spurious association between the two key variables, even if they are marginally independent from each other. The common outcomes of the key variables are called *collider variables*. Conditioning on a collider variable (or any variable which is affected by this collider variable) creates *endogenous selection bias*. To avoid this bias when estimating causality, it is advised not to control for post-treatment variables (Elwert & Winship, 2014; Rosenbaum, 1984).

However, it must be noted that controlling for only pre-treatment variables does not

[5] Author's own construction, produced in R.

[6] Author's own construction, produced in R.

always rule out endogenous selection bias. There are some variables which are affected by the treatment, although they appear before exposure to the treatment. For example, people might anticipate an event and its effect and their expectations about the future might influence the present. The literature refers to this phenomenon as the *anticipation effect*. Controlling for these variables can also cause endogenous selection bias. Regarding the research in this dissertation, the former situation may appear when observing the parenthood effect. Parenthood might have an effect before birth as parents prepare for the arrival of a child (Baetschmann, Staub, & Studer, 2016; Balbo & Arpino, 2016; Clark et al., 2008; Myrskylä & Margolis, 2014). As a consequence, controlling for only pre-treatment variables plays an important role, but it does not automatically guarantee that endogenous selection bias is ruled out.

To sum up, to estimate causality one needs to include all the confounding variables, but omit all the possible collider variables. The longitudinal design of the present research can handle both issues more effectively than cross-sectional research. First, longitudinal analysis can control for not only observed confounding variables, but also time-invariant unobserved variables, thereby reducing selection bias. Moreover, a longitudinal design enables us to measure control variables before exposure to the treatment, which can help with avoiding endogenous selection bias.

3.2. Potential Outcome Framework

The present dissertation applies the *potential outcome framework* to estimate causality. This approach was developed by Neyman (1923) and Fisher (1925), and further elaborated by Rubin (1974, 1978) and others (Diamond & Sekhon, 2013; DuGoff et al., 2014; Ho et al., 2011; P. W. Holland, 1986; Imai et al., 2007; Imai & Van Dyk, 2004; Imbens & Rubin, 2015; Rosenbaum, 2002; Rosenbaum & Rubin, 1983, 1985a, 1985b; Rubin, 1974, 1978; Stuart, 2010). Throughout the whole of the Analytical Strategy Chapter I refer to these authors if not otherwise indicated. This framework has been taken by numerous social scientists to estimate certain events, interventions or other treatment effects when formulating evidence-based policy or seeking to obtain a deeper understanding of phenomena. This chapter provides an overview of the potential outcome framework and the key concepts used within this framework. These concepts are applied in subsequent chapters which describe the methodology of this dissertation.

The potential outcome framework assumes that any i individual from the population has a greater than 0 probability of receiving the treatment, and a greater than 0 probability of receiving the control as well. Further, this approach assumes that for each i individual and j treatment scenario ($j = \{1,0\}$), Y_{ij} potential outcome exists, which the individual would encounter under j treatment. In other words, for each individual two potential outcomes may be defined: (1) the actual outcome that they in fact experience; and, (2) a counterfactual outcome that they would experience if they received a different treatment than that which they actually do.

Regarding this dissertation, for each individual the following two potential outcomes can be calculated:

	Those who receive treatment (j=1)
$Y_{i1} \mid J = 1$	actual level of subjective well-being that individual i has who belongs to the observed life-stage group
$Y_{i0} \mid J = 1$	level of subjective well-being that individual i, who belongs to the observed life-stage group, would have if they belonged to another life-stage group
	Those who do not receive treatment (j=0)
$Y_{i0} \mid J = 0$	actual level of subjective well-being that individual i has who does not belong to the observed life-stage group
$Y_{i1} \mid J = 0$	level of subjective well-being that individual i, who does not belong to the observed life-stage group, would have if they belonged to the observed life-stage group

Based on the potential outcome framework, any conclusion of causality would theoretically require a comparison of individual outcomes under each treatment possibility ($Y_{i1} - Y_{i0}$). With regard to the present research, a key issue is determining the extent to which the subjective well-being of an individual i from the given life-stage group ($j = 1$) would differ if they did not belong to the aforementioned life-stage group ($j = 0$), permitting investigation of the effects of life stages. For i individual, the Δ_i true causal effect can be calculated in the following way:

$$\Delta_i = Y_{i1} - Y_{i0} \qquad (1)$$

However, this type of comparison is not feasible in practice, given that the two potential outcomes cannot be observed simultaneously (one individual either receives the treatment or does not at any one time). Regarding the sample for the present dissertation, any one individual can only belong to a single life-stage group at any one time, thus, the level of subjective well-being is only observable in actual cases ($Y_{i1} | J = 1$ or $Y_{i0} | J = 0$), but counterfactual cases are missing ($Y_{i1} | J = 0$ or $Y_{i0} | J = 1$). This problem has been named the "*fundamental problem of causal inference*" by Holland (1986) and is also often referred to as the *identification problem* by economists (Imbens & Wooldridge, 2009; Kézdi, 2004). Table 2 illustrates the challenge of calculating the treatment effect in a sample dataset (See Table 27).

Table 2. Illustrating the "fundamental problem of causal inference"

i individual	J treatment	Y_{i1} outcome	Y_{i0} outcome	Treatment effect $Y_{i1} - Y_{i0}$
1	1	9	?	$9 - ? = ?$
2	1	8	?	$8 - ? = ?$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
40	0	?	8	$? - 8 = ?$
41	0	?	6	$? - 6 = ?$

The collection of Y_{obs} observed potential outcomes (Equation 2) can be calculated in the following way:

$$Y_{obs} = J_i \times Y_{i1} + (1 - J_i) \times Y_{i0} \tag{2}$$

Although Δ_i is not identified, the expected value of the related change can be identified. Basically, three ways of calculating this expected value have been created. First of all, one can observe individuals over a short period of time when both Y_{i1} and Y_{i0} outcomes appear, and compare these two outcomes. However, to draw causal conclusions this way one needs to assume (1) *temporal stability*, which means that Y_{i1} outcome is independent from the timing of the treatment assignment; and, (2) *causal transience*, which states that first having Y_{i0} outcome (receiving treatment) is transient, and does not change individual i to the extent that it also affects Y_{i1} outcome. Under these assumptions, one can state that "*each individual in panel data is his own best control*" (Hausman & Wise, 1979). The first and second empirical studies in this dissertation were conducted using a longitudinal dataset, thus this approach was also applied herein (See Chapter 3.7.6.).

The second way of identifying the missing potential outcome is to use a random experimental design. One of the greatest advantages of a randomized experiment is that, on average, control and treatment groups are only randomly different from one another across all observed and unobserved variables because the subjects of the experiment are randomly assigned to a treatment or control group. Therefore, the average measured potential outcome of the control group can be used as a good estimation of the average potential outcome that is missing for the treatment group (more detail in Chapter 3.4).

Third, when assignment of the treatment is not feasible, statistical adjustment may be used to estimate the expected value of any missing potential outcomes. Among these methods, the research for this dissertation employed regression adjustment, weighting, matching, and longitudinal analysis. Further, other methods can be mentioned here which would be useful for estimating causality, but which are not covered in this dissertation in detail. These include instrumental variables and regression discontinuity.

Finally, the assumptions made in this framework should be noted. The potential outcome framework relies on the *Stable Unit Treatment Value Assumption* (SUTVA), which involves two prerequisites (Rubin, 1978). First, it assumes that the potential individual outcome is not affected by whether other units receive the treatment. Regarding the research described herein, this means that one's subjective well-being is not affected by other people's life-stage status. However, mechanisms such as social learning, social pressure, social contagion, and social support might violate this assumption since they assume that individuals interact with each other (Bernardi & Klärner, 2014; Ateca-Amestoy, Aguilar, & Moro-Egido, 2014). One can avoid this problem by sampling a sufficiently large population in which interaction is less dense.

The second prerequisite of SUTVA is that there are no different versions of treatments. In other words, the characteristics of an individual do not modify the given life stage effect on subjective well-being. In certain cases, this situation can be avoided by further specifying the treatment. For example, I not only observe the effect of parenthood, but also the effect of single parenthood (as done in Chapter 6).

3.3. Different Types of Causal Parameters Within the Potential Outcomes Framework

In this chapter I introduce the theoretical foundation of the causal parameters under the potential outcome framework. In practise, these parameters are calculable only if the two potential outcomes are measurable at the same time. As I have argued before, this is not possible as one of the potential outcomes is always missing. Thus, the material in this subchapter is only theoretical, and the following chapters will unfold how the missing potential outcome can be estimated, which might also modify the calculation of the causal parameters.

First of all, one can estimate the *average treatment effect* (ATE), which is the expected value of the difference between the two potential outcomes. Namely, it is the expected value of the difference between the outcome that the individual would have upon receiving the treatment and the outcome that the individual would have without receiving the treatment (see Equation 3).

$$ATE = E\left[Y_{i1} - Y_{i0}\right] \tag{3}$$

ATE represents the expected effect of the treatment on any individual from the whole population, regardless of whether this person would be actually eligible to receive the treatment. Alternatively, one can estimate the population *average treatment effect for the treated* (ATT), which only observes the treatment effect on the treatment group (See Equation 4).

$$ATT = E\left[Y_{i1} - Y_{i0} \middle| J = 1\right] \tag{4}$$

Finally, one could estimate the *average treatment effect on the untreated* (ATU). This parameter measures how the treatment would affect the control group if this group received the treatment (See Equation 5).

$$ATU = E\left[Y_{i1} - Y_{i0} \middle| J = 0\right] \tag{5}$$

The choice of ATT over ATE or ATU often depends on the given treatment. For example, policies are often implemented through programs targeted at a narrow population instead of the whole population. Numerous social policies are aimed at helping only people who are in need, thus only selected people are eligible to receive the treatment. To estimate the effects of these policies, it makes more sense to estimate ATT which only observes what the effect may be for those who are actually eligible for treatment.

Regarding the present research, ATT would mean the effect of belonging to a certain life stage on subjective well-being for those people who actually belong to this life stage. More specifically, ATT involves a comparison between the level of subjective well-being that those have who belong to a given life stage with the subjective well-being that the same group of people would have if they did not belong to this life stage. In contrast, ATE would observe the effect of a life stage for the whole sample. Namely, ATE refers to the difference between the subjective well-being in the whole sample if everyone belonged to a given life stage and the subjective well-being that everyone in the sample would have if no one belonged to this life stage. Finally, ATU represents the effect of a life stage for those who do not belong to the given life stage. Meaning, ATU is a comparison of the subjective well-being of individuals who do not belong to a given life stage with that of the same individuals $(j = 0)$ assuming they did belong to the former life stage $(j = 1)$.

The association between ATE, ATT and ATU can be written as

$$ATE = ATT \times P(J = 1) + ATU \times P(J = 0), \tag{6}$$

where $P\left(J = 1\right)$ refers to the probability of receiving the treatment in the population, and $P\left(J = 0\right)$ is the probability of not receiving the treatment.

In the case of population surveys, one can distinguish between the sample and the population ATEs, ATTs and ATUs. The sample parameters are respectively denoted as SATE, SATT and SATU, whereas the population parameters are denoted as PATE, PATT and PATU. The main difference between sample and population parameters is that the former are generalizable to the survey only, whereas the latter describe the effect across the whole population.

Sample parameters are calculated as the average of the treatment effect on a sub-sample (SATT and SATU) or the whole sample (SATE), rather than estimating the expected value in the population. Equations 7, 8, and 9 specify the calculations for SATE, SATT and SATU, respectively:

$$SATE = \frac{1}{N} \sum_{i=1}^{N} (Y_{i1} - Y_{i0}) \tag{7}$$

$$SATT = \frac{1}{N_{(J=1)}} \sum_{i:J=1} (Y_{i1} - Y_{i0}) \tag{8}$$

$$SATU = \frac{1}{N_{(J=0)}} \sum_{i:J=0} (Y_{i1} - Y_{i0}), \tag{9}$$

where $N_{(j=1)}$ refers to the number of treated individuals and $N_{(j=0)}$ denotes the number of control individuals in the sample.

The population parameters can be estimated from the sample parameters. DuGoff et al. (2014) have argued that the best way to estimate population treatment effects is to calculate the sample treatment average, and take into account the sample design by using sample weighting.

3.4. Random Experiments

Although the research described in this dissertation does not use random experiments as a method, they serve as the gold standard of causality, so they serve as a reference point for other methods as well. Classic random experiments have the three following characteristics (Ho et al., 2011):

1. random selection of individuals from the whole population to the sample
2. random treatment assignment for of all members of sample
3. large sample size

The first characteristic reduces selection bias by ensuring that the sample selected from the given population is only randomly related to the potential outcomes. Thus, the potential outcome in the sample should be similar to the population-level potential outcome. The second feature further eliminates selection bias by ensuring that the treatment group and control groups only randomly differ from each other, even without controlling for any confounding variables. In other words, this characteristic produces a setting in which the treatment and control groups are balanced for all covariates. Finally, the third characteristic ensures that the probability that something will go wrong during random sample selection and treatment assignment is vanishingly small due to the law of large numbers. Due to these features, randomized experiments produce a setting in which there is no systematic difference between the control and treatment group. In other words, such experiments have the following characteristics:

$$Y_{i1}, Y_{i0} \perp J_i \tag{10}$$

With regard to the present research, use of a random experiment would imply that those who are at a specific life stage are selected randomly, as a consequence of which individuals at that given life stage ($j = 1$) are only randomly different from all other individuals who are not at that life stage ($j = 0$).

Under the circumstances described in Equation 10, one can estimate the missing counterfactual potential outcome of the treatment group from the average potential outcome measured in the control group, since the following equation holds:

$$E\left[Y_{i0}\middle|J=1\right]=E\left[Y_{l0}\middle|J=0\right] \text{ and } E\left[Y_{i1}\middle|J=1\right]=E\left[Y_{l1}\middle|J=0\right] \tag{11}$$

As a result, all of the causal parameters, including ATE, ATT and ATU, can be estimated by comparing the potential outcomes of the treatment and control groups under random experimental design (See Equation 12).

$$ATE=ATT=ATU=E\left[Y_{i1}\middle|J=1\right]-E\left[Y_{l0}\middle|J=0\right] \tag{12}$$

Despite the obvious advantages of random experiments, they are typically only ideal types since research in the social sciences almost always fails to ensure some of the above-mentioned characteristics. Very often, studies struggle with an insufficiently large sample size for financial reasons. In this case it is hard to generalize the effect of treatment since low sample sizes produce a large standard error. However, one can get around this problem by further controlling for the possible confounding variables using regression adjustment (Gelman & Hill, 2006).

Also, in certain cases the first characteristic (described above) cannot be guaranteed. For example, laboratory experiments usually use random treatment assignment, but sample selection is typically not random (these studies usually recruit participants who are university students although the general population is older and less educated, for example). These studies can make causal inferences due to random treatment assignment, but these inferences cannot be generalized to the entire population.

Another possible challenge with conducting experiments is that individuals may reject participation in a study. This problem is usually referred as the *problem of non-compliance*. The problem arises from the fact that even though the selection of an individual from the population and treatment assignment are random, we cannot assume that the sample of individuals who prefer not to participate are random. If the problem of non-compliance arises, than one needs to find another way to control for confounding variables.

Moreover, it is often unfeasible to conduct randomized experiments due to ethical or financial considerations. Research may also not permit the use of random treatment variables because researchers must not arbitrarily decide who is at a certain life stage. For example, the researcher cannot decide who should retire or have a child and who should not; thus, treatment cannot be randomized. Therefore, experiments were not used in this dissertation but rather observational data only (Ho et al., 2011).

3.5. Estimating Causality in Observational Data

Observational studies are defined by Rosenboum (2002:VII) in the following way: "*An observational study is an empirical investigation of treatments, policies, or exposures and the effects they cause, but it differs from an experiment in that the investigator cannot control the assignment of treatments to subjects.*"

As regards observational data, the treatment and control groups fundamentally differ from one another; that is, J_i is a non-random variable; individuals who are at a certain life stage ($j = 1$) systematically differ from individuals who are not at this life stage ($j = 0$). I will demonstrate how this arises in the case of my research questions. With regard to the effect of retirement, retirees differ from non-retirees (i.e. may have a different level of education or age), thus the difference between their subjective well-being is not solely determined by their activity status. Similarly, regarding the effect of parenthood subjective well-being is not only affected by the number of children – those who have a child differ from those who do not. In this case, for example, satisfaction with a partner can be a confounding variable. Those who have a child may be more satisfied with their partner before the baby is born, and satisfaction with a partner affects positively life satisfaction as a whole. Thus, having a child might not be the reason for changes in subjective well-being, but instead it is higher satisfaction with a partner that might cause this change.

As those who receive treatment and those who do not receive it systematically differ from each other in the case of observational data, the assumption of independence in Equation 10 is not valid for this type of research design. Since the potential outcomes are not independent of the treatment, the counterfactual of the treatment is still missing. Thus, for observational data, simply comparing the mean outcome of the control group with that of the treatment group would not identify the treatment effect, but only the association between the treatment and the outcome. To estimate causality in observational data, one needs to eliminate the confounding variables.

The present research eliminates confounding variables with the help of multivariate regression, matching, and regressor variable methods (longitudinal analysis). Use of these methods involves two main assumptions: First of all, the methods estimate missing values by relying on the *unconfoundedness assumption*. Unconfoundedness denotes the assumption that biases found when comparing the treatment and control groups may be removed by holding certain covariates constant. Let me denote the fixed set of properties of individual i as X_i. The essence of the unconfoundedness is that treatment group membership is conditionally independent of the individual's Y_i outcome response, given X_i covariates.

$$Y_i \perp J_i \mid X_i \tag{13}$$

In other words, the outcomes expected of the treatment group in the case of nonparticipation are a good reproduction of the outcomes for those control individuals who have the same characteristics (Imbens & Wooldridge, 2009). As far as the analysis in the present dissertation is concerned, this expectation means that after controlling for X_i we can use subjective well-being measured in the control group to determine the level of subjective well-being of the treatment group that they would have if they were not at this life stage.

The regression and matching methods only control for observed X_i confounding variables and assume that unobserved variables do not modify the relationship between the treatment variable and the outcome. The extension of these methods to longitudinal design can handle time-invariant unobserved variables (See Chapter 3.7.6.). However, even with this extension, time-variant unobserved variables are not controlled for. Thus, one needs to conduct sensitivity analysis to measure how unobserved effects modify the treatment effect (see more in Chapter 3.7.7.).

Second, these methods require overlap between the treatment and control groups, which can be formalized in the following way:

$$0 < P(J_i = 1 \mid X_i) < 1 \tag{14}$$

In other words, every individual has a non-zero probability of receiving the treatment, given their characteristics. Without this assumption it could happen that for a treated individual there is no control peer, as there is no sufficient overlap between the treatment and control groups. The assumption that both Equation 13 and Equation 14 are valid is often referred to as *strong ignorability*.

There are other methods which have been developed to estimate causal effects, but which do not rely on the unconfoundedness assumption, such as indicator variables and regression discontinuity. However, these methods make other assumptions. Instrumental variables rely on strong assumptions about how unobserved and observed variables are linked to each other to rule out the effect of unobserved covariates (Angrist, Imbens, &

Rubin, 1996; Imbens & Angrist, 1994). Regression discontinuity, meanwhile, uses a specific research design, which is not available in the case of the present research (Imbens & Lemieux, 2008).

3.6. Regression Adjustment

One of the most popular ways of controlling for confounding variables and obtaining causal estimates is applying regression adjustment. This section reviews only causal estimations obtained from ordinary least square regression; however, the following statement could be extended to other types of regressions as well. The difference in regression methods and their assumptions are not reviewed here, but are introduced in foundation books on econometrics and statistics (Wooldridge, 2015).

Let me demonstrate what estimating causality through regression in the potential outcome framework looks like. In the case of OLS regression, the two potential outcomes can be written as two parallel linear regression lines, which differ in the intercept by γ :

$$Y_{i0} = \alpha + \beta \times X_i + \varepsilon_i \tag{15}$$

$$Y_{i1} = \alpha + \beta \times X_i + \gamma + \varepsilon_i \tag{16}$$

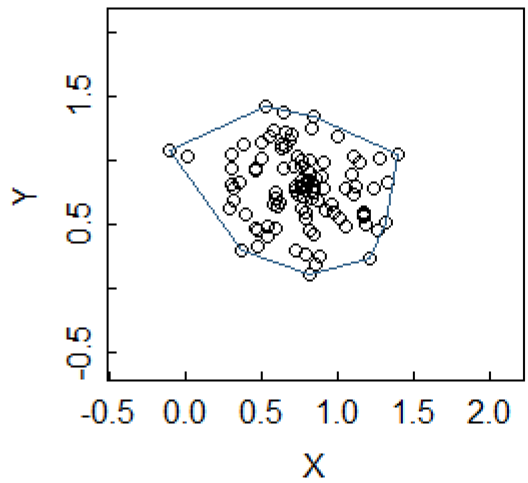
Using Equation 15 and 16, the two regression lines can be combined into one regression line with the help of Equation 2, which estimates all the observed potential outcomes together:

$$Y_{obs} = J_i \times (\alpha + \beta \times X_i + \gamma + \varepsilon_i) + (1 - J_i) \times (\alpha + \beta \times X_i + \varepsilon_i) = \alpha + \beta \times X_i + \gamma \times J_i + \varepsilon \tag{17}$$

In Equations 16 and 17, γ indicates the ATE. Further, control variables should be involved as described in Chapter 3.1.1. In other words, those X_i should be (1) involved which are the common causes of both the Y outcome variable and J treatment variable; further, those X_i should be (2) omitted which are the common outcomes of Y and J key variables. The first stipulation rules out selection bias, whereas the second guarantees that endogenous selection bias is avoided.

However, even if the necessary control variables are correctly identified, regression adjustment can still encounter two types of additional biases. These two types of biases can appear due to (1) interpolation, or (2) extrapolation. To distinguish between these two concepts one needs to define the *convex hull* that is the "*smallest convex set that contains the data*" (King & Zeng 2006: 139). Figure 3 illustrates the convex hull for a hypothetical dataset. In this figure, the dots are observations while around the dots there is a blue line which is the border of the convex hull. Inferences about observations within the convex hull are considered interpolation and outside this area extrapolation (King & Zeng, 2006; Kuo, 2001).

Figure 3. Example of a convex hull[7]



Interpolation bias is the result of improper adjustment of the control variables included within the convex hull. In the case of parametric methods, this bias can arise from using the wrong functional form (King & Zeng, 2006). For instance, OLS regression assumes a linear relationship, but sometimes this assumption is wrong (i.e. income follows an inverse U-shaped curve with aging, so quadratic regression would be more appropriate for observing the relationship between these two variables). These assumptions can be relaxed by using non-parametric methods such as a kernel-estimator or matching. However, non-parametric methods can be inappropriate for use when the sample size is small and one needs to involve several covariates (Arpino & Aassve, 2013).

Extrapolation bias occurs when the distribution of any confounding variable is different in the treatment and control groups. In other words, there is not sufficient overlap between the two groups (Ho et al., 2011; King & Zeng, 2006). It is possible that certain values of X covariates take on a positive probability in the control group, but no members of the treatment group possess this value. In this case, regression extrapolates the observations for the treatment group as well, which can create bias.

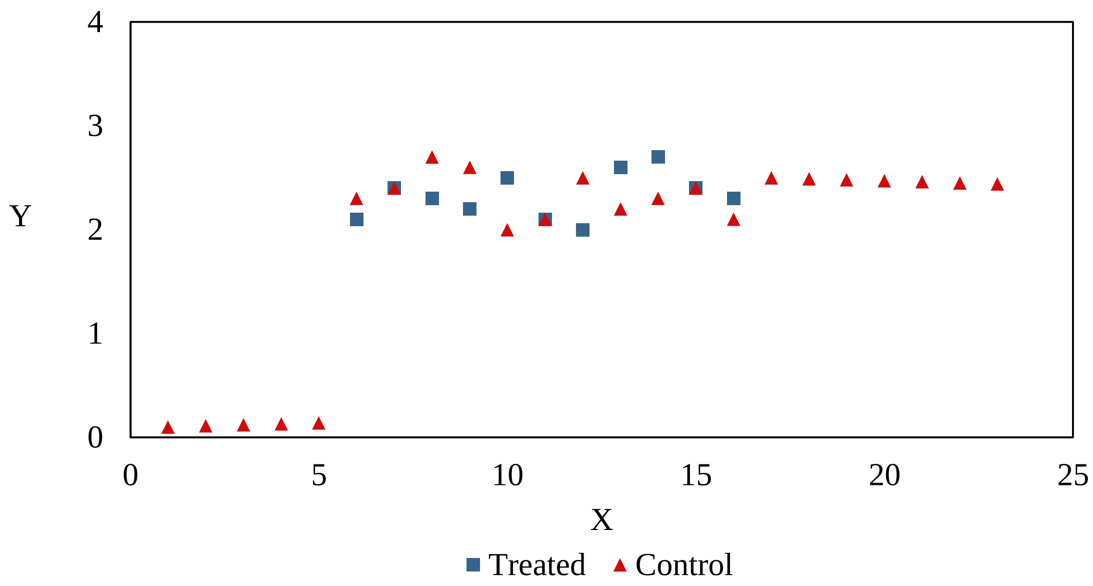
As regards the present research, X_i could be the age of i individual. Age is an important confounding variable as both life-stage status and subjective well-being depend on it.

[7] Author's own construction, produced in R.

However, there might not be sufficient overlap in age between those who are at a certain stage of life and those who are not. For example, after a certain age parenthood is very unlikely, especially for females[8]. Also, people do not usually retire at a very early age. Therefore, even though we have control observations for all adult age groups[9], the treatment group is limited to a certain age group both in the case of parenthood and retirement.

I demonstrate the problems with extrapolation in a hypothetical dataset in a similar manner to Ho et al. (2011). This hypothetical dataset can be seen in the Appendix in Table 26. The data are shown in Figure 4, in which the observations from the control group are represented with red triangles and the observations from the treatment group blue squares. As can be seen on this graph, there is no observation in the treatment group for $X \leq 4$ and $X \geq 17$, whereas the control group contains observation on a scale of $0 \leq X \leq 25$. Thus, the hypothetical dataset is not balanced in terms of X.

Figure 4. Lack of sufficient balance between the treatment and control groups (Hypothetical data)



Regression adjustment solves the problem of unbalanced datasets using extrapolation. In the case of the hypothetical data, this means that even for those X values which do not contain any treated observation, the regression calculation predicts values by extrapolating the trends. With regard to the analysis described herein, extrapolation would

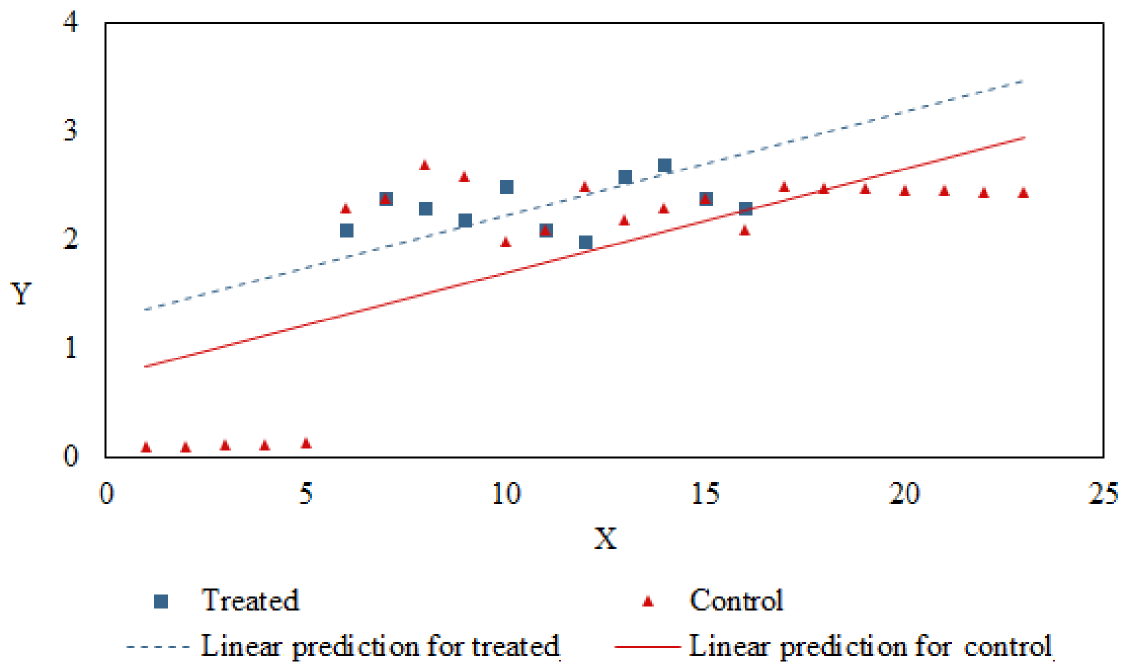
[8] The typical reproductive age range for females is widely acknowledged as being between 15 - 49 years (World Health Organization, 2006).

[9] For example, national samples are used in this dissertation

mean that we assume that even an 80 year-old might give birth, or an 18 year-old might retire.

I carried out regression on the hypothetical dataset according to Equation 17. The predicted values of this linear regression can be seen in Table 26 for both the treatment and control groups. Moreover, these predicted values are also shown in Figure 5. In this figure the regression line for the treated group is marked with a blue dashed line, whereas the regression line for the control group is marked with a solid red line. Based on this analysis, the γ treatment effect would be 0.53 and significant at a 0.002 level. However, Figure 5 shows that the regression line is extrapolated for the treated group even for age categories for which there is no actual treated observation. Thus the treatment effect could be attributed to extrapolation bias.

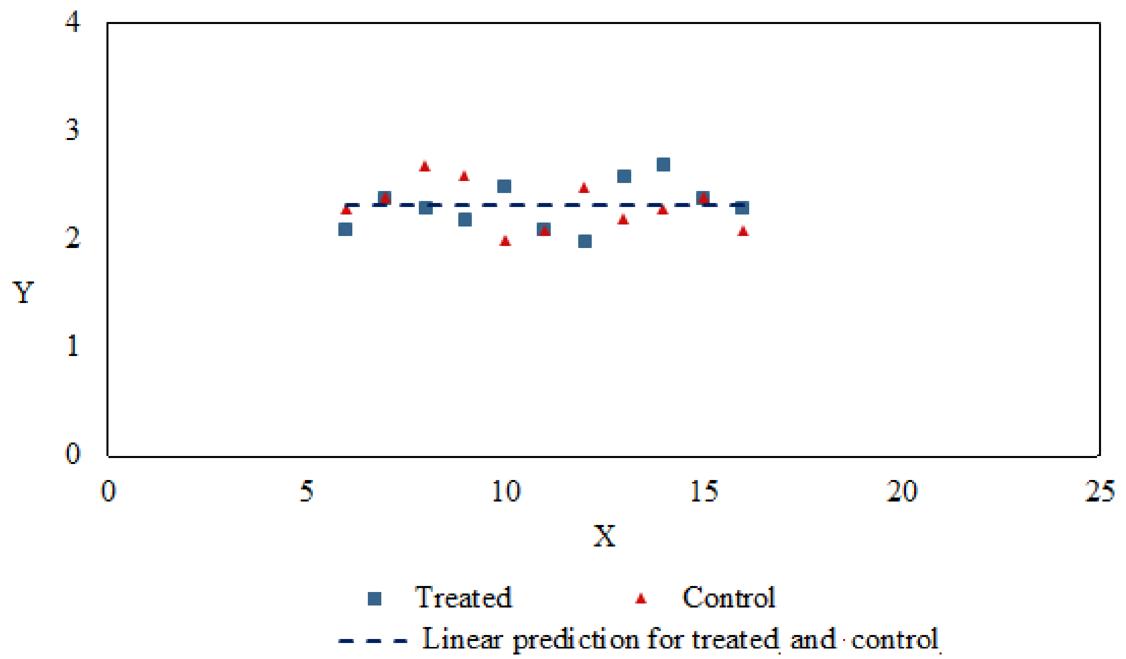
Figure 5. Extrapolation by regression (Hypothetical data)



To highlight that the treatment effect in Figure 5 can be attributed to the extrapolation bias, I reduced the dataset to the area inside the convex hull, which contains observations from both the treatment group and control groups. Thus, only those observations were kept which fall into the area $5 \leq X \leq 16$ (see the reduced observations in Figure 6). In this reduced dataset, I again performed the OLS regression (as detailed in Equation 17). The regression on the reduced dataset did not require extrapolation, in contrast to the regression run on the initial dataset. In the reduced dataset the treatment effect is already 0 and non-significant. Thus, the positive significant treatment effect which was observed for the initial dataset can only be attributed to extrapolation.

The above-described example is often called outlier bias (Wooldridge, 2015). In such cases, the solution is simple; one needs to remove the outliers. However, in certain cases the extreme points are not so easily detected. For example, the situation is less simple when an imbalance is caused not by simple outliers but the treatment tends to associate some measures with a lower probability, causing imbalance between the treatment and control groups. In this case, one might consider using an algorithm, such as matching (see the next subchapter), which creates a more balanced dataset.

Figure 6. Regression in the reduced dataset eliminates extrapolation bias



To sum up, the above-described example has shown that in certain cases regression tends to perform poorly. The problems with regression adjustment arise when there is no sufficient overlap between the treatment and control groups. In such cases, regression relies on a mechanism which works very well for prediction: it extrapolates existing data to those cases where there are otherwise none. However, this mechanism can create biased causal estimations (Ho et al., 2011; King & Zeng, 2006). Thus, although data reduction in general can increase bias, I argue that the proper process of data reduction can also eliminate extrapolation bias. The initial dataset needs to be reduced in a way that the control observations, which are responsible for the imbalance between the treated and control groups, are omitted.

3.7. Matching Methods

3.7.1. General Introduction to Matching

In the previous chapter I showed that one needs to rule out all the possible confounding variables in order to draw causal conclusions. I further argued that this goal can be achieved by using random experimental design; however, this is often not feasible. In contrast, regression methods are more readily available, but these methods often fail to produce unbiased estimations in the absence of sufficient overlap between the treated and control groups. Therefore, I suggested circumventing the shortcomings of regression adjustment by reducing the dataset. However, one needs to bear in mind that data reduction can lead to biased estimations if not conducted properly, and there is often no clear-cut way to remove outliers. For instance, in the case of measuring the parenthood effect one should avoid including in the control group an 80-year-old woman who obviously could not have a child; however, there is no clear limit to the age at which men can have children. In this chapter I introduce a method which has been developed to systematically select control observations in the best possible way to estimate causality: the so-called matching method.

Despite the long history of matching, this method is still the state-of-the-art as it has been continuously developed alongside improvements in computational capacity. Matching was first used in 1945 by Greenwood (1945) although the theoretical foundations were only developed in the second half of the century (Cochran & Rubin, 1973; Dehejia & Wahba, 2002; Heckman et al., 1997; Rosenbaum & Rubin, 1983, 1984, 1985a, 1985b; Rubin, 1973a, 1973b, 1974). Even today, the matching method still receives a great deal of attention as newer types of matching methods have also recently been invented (Abadie & Imbens, 2006; Arpino & Mealli, 2011; Chabé-Ferret, 2015; Crump, Hotz, Imbens, & Mitnik, 2009; Diamond & Sekhon, 2013; Ho et al., 2011; King & Nielsen, 2016; Lee, Lessler, & Stuart, 2010; Lunceford & Davidian, 2004).

In short, matching methods are designed to find a missing potential outcome (counterfactual) by producing the best possible approximation of the experimental arrangement using observational data. The method entails matching i treated individual(s) with one or more non-treated individual(s) who is/are as similar as possible to the given treated individual in all X_i regards, except for the J treatment itself.

The method relies on the unconfoundedness assumption (Equation 13) for calculating the counterfactual. In other words, the essence of the matching method is that the treatment is conditionally independent of the individual's response, given X_i. Thus, the expected value of the potential outcome that i treated individual would have if this i individual did not receive the treatment can be replaced with the potential outcome that those have who did not receive the treatment but have the same X properties as the given i treated person (Imbens & Wooldridge, 2009). This unconfoundedness assumption is formalized in Equation 18. Regarding this dissertation, the equation means that we can use the control group matched with individuals from an observed life stage $(j = 1)$ to determine the level of subjective well-being that they would have if they were not at this life stage.

$$E\left[Y_{i0} \middle| J = 1, X_i\right] = E\left[Y_{io} \middle| J = 0, X_i\right] \tag{18}$$

As a consequence, the average treatment effect (ATT) can be calculated in the following way by using the matching method:

$$ATT = E\left[Y_{i1} - Y_{i0} \middle| J = 1\right] = E\left[Y_{i1} \middle| J = 1, X_i\right] - E\left[Y_{i0} \middle| J = 0, X_i\right] \tag{19}$$

Based on Stuart (2010), the implementation of the matching methods involves four major steps. I have added two additional steps to these initially described four steps: more specifically, (Step V.) the extension of matching to longitudinal design (based on Imbens & Wooldridge, 2009; Balbo & Arpino, 2016; Allison, 1990; Chabé-Ferret, 2015), and (Step VI.) sensitivity analysis (based on Rosembaum, 2002).

Step I. Choosing a distance definition

The $D(k,l)$ measure of distance defines how similar k treated individual ($k \in j$) and l control individuals ($l \notin j$) are based on their X properties. There are several types of distance measures which are introduced in Chapter 3.7.2. in detail. One can even combine the different distance measures, as genetic matching does, which is the primary method chosen for use in the research described in this dissertation.

Step II. Conducting matching given the chosen distance measure

Once the distance measure has been chosen, matching can be conducted. However, even with one given distance measure one can identify pairs in several ways. Thus, we can not only distinguish different distances, but also there are several types of algorithms for matching based on the selected distance measure. Different matching algorithms are introduced in Chapter 3.7.3.

Step III. Evaluating the balance in the matched dataset

After matching, one needs to assess the performance of the procedure. If it has failed to produce a balanced dataset, then the researcher should repeat Steps I and II. One can evaluate the balance based on descriptive statistics, or a display of the distribution of the propensity score in the treatment and the control group before and after matching. See more about this in Chapter 3.7.4.

Step IV. Analysis of the treatment effect on the matched data

Matching aims only to balance the observed covariates between the treated and control groups, and does not provide an estimation of the treatment effect itself. Thus after balancing the dataset by matching one needs to conduct a t-test or regression to obtain the causal parameters (See more in Chapter 3.7.5).

Step V. Extension of matching to longitudinal design

The matching method can be also extended to longitudinal design, which enables the researcher to rule out time-invariant confounding variables as well. The longitudinal extension of matching is detailed in Chapter 5.7.6. This dissertation applies genetic matching using longitudinal data.

Step VI. Conducting sensitivity analysis

Sensitivity analysis aims to estimate the sensitivity of the estimation to the omitted confounding variables. This dissertation applies Rosenbaum's (2002) sensitivity analysis for this purpose (See more in Chapter 3.7.7).

In the following subchapters, I describe these steps in more detail.

3.7.2. Types of Distance Measures

There are several ways to define the distance between k treated individual ($k \in j$) and l control individual ($l \notin j$) (Deza & Deza, 2009). Here I introduce the most common ways which are particularly noteworthy; namely, exact distance, propensity score, Euclidean distance, normalized Euclidean distance, Mahalanobis distance, and genetic matching distance.[10] I demonstrate in the following part of this chapter the calculation of the different distance measures for a small sample dataset (See Table 27), which is a reduced version of the dataset used for estimating the effect of parenthood. In this dataset the J treatment variable is parenthood status, the Y outcome is subjective well-being; further,

[10] MatchIt package also contains several other distance measures. See more about them in Ho et al. (2011).

three matching variables are used. These variables are income (X_1), satisfaction with housing (X_2), and age (X_3).

I. Exact distance:

Exact matching is the most intuitive type of matching as it allows for zero distance between the treatment group and control groups in terms of all the v number of matching variables. So this method matches a k treated individual to all of the control individuals with exactly the same X_v properties.

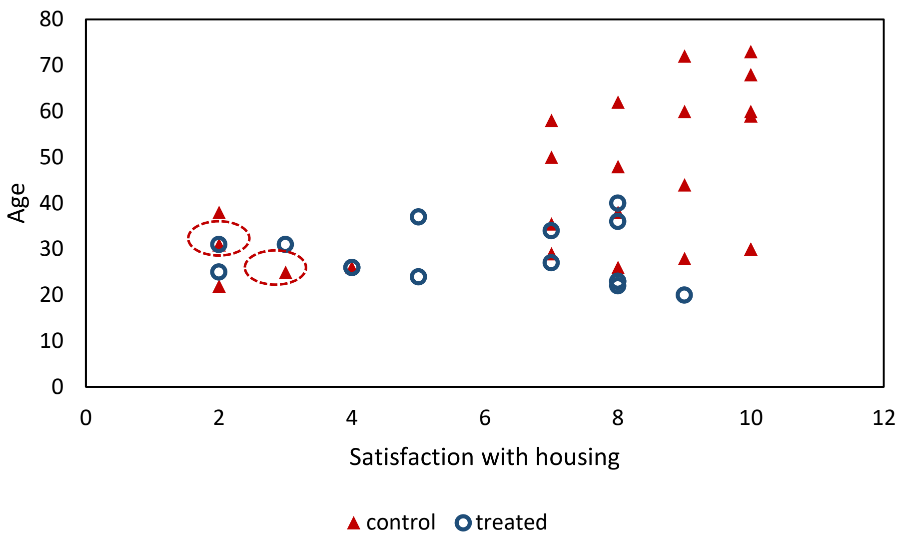
$$D^{exact}(k,l) = \begin{cases} 0 \ \ if \ \ X_{vk} = X_{vl} \\ \infty \ \ if \ \ X_{vk} \neq X_{vl} \end{cases} \quad for \ \forall \ v \tag{20}$$

This distance measure is an ideal solution for matching as it can produce unbiased results. However, exact matching is not feasible most of the time using real data. Let us assume that we have a high dimensional dataset with X_1, X_2, X_3 ... X_v covariates which also contains several continuous variables. In this case, there is a very small chance that for k treated individual $(k \in j)$ we will find a l control individual $(l \notin j)$ who is similar in all of the v aspects to the i individual. As a result, several treated individuals will not be matched. This problem is often referred to in the literature as the *curse-of-dimensionality* (Blackwell, Iacus, King, & Porro, 2009).

To illustrate this problem, let us assume that one desires to obtain better balance on the sample dataset based on satisfaction with housing and age. Figure 7 shows that only two pairs are found by exact matching[11] (the matches are indicated with a dashed red circle). In practice, this would mean that only two pairs from the fifteen are matched. However, if we wished to match using the third variable as well, then even these two pairs would not meet the requirements of exact matching. As a result, in the case of exact matching often the vast majority of observations remain unmatched, which creates biased estimations.

[11] These matches are Units 1 and 29, and Units 2 and 41.

Figure 7. Sample dataset by age and satisfaction with housing



In order to deal with the *curse-of-dimensionality* of exact matching, *Coarsened Exact Matching* (CEM) has been developed. This involves, prior to running exact matching, coarsening the continuous variables to meaningful categorical variables, and after matching returning them to the original (uncoarsened) variables (Blackwell et al., 2009).

II. Propensity score:

Another potential way to measure distance is to reduce the multivariate dimensions to a one-dimensional scalar. The most commonly used method for this purpose is *propensity score matching* (King & Nielsen, 2016; Rosenbaum, 2002; Rosenbaum & Rubin, 1983, 1985a, 1985b; Rubin, 1973a). A propensity score estimates the probability of receiving a treatment given the i subject's X_{vi} observable properties. It can be formally defined in the following way:

$$e_i\left(X_{vi}\right) \equiv P\left(J_i = 1 \middle| X_{vi}\right) \tag{21}$$

Although $e_i(X_{vi})$ is unknown in the sample, it can be estimated by logistic regression of the Y treatment on the X_v covariates:

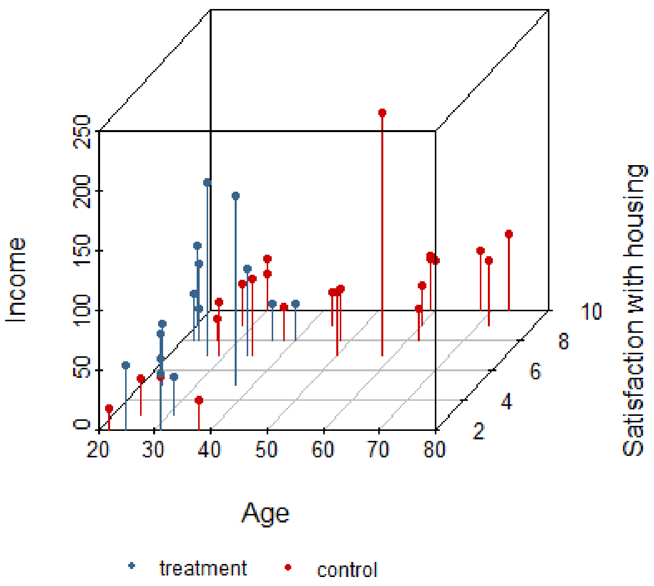
$$\ln\frac{\hat{e}_i(X_{vi})}{1-\hat{e}_i(X_{vi})} = \beta_0 + \beta_v \times X_{vi} + \varepsilon \tag{22}$$

Thus the propensity score is the reduction of the v dimensional X_{vi} vector to a one-dimensional $e_i(X_{vi})$ scalar. This score asymptotically balances the X_{vi} confounding variables. Once this scalar has been calculated for each observation, the distance metrics can be defined between any given two units in the following way:

$$D^{propensity\ score}(k,l) = \left| e_k(X_{vk}) - e_l(X_{vl}) \right| \tag{23}$$

Similarly to the illustration of exact matching, I also illustrate how to calculate propensity scores in the sample dataset (data is available in Table 27). See the joint distributions of the three matching variables (age, satisfaction with housing, and income) in Figure 8. The goal of the propensity score is to reduce the information provided by these three variables to a one-dimensional scalar.

Figure 8. The joint distribution of the three confounding variables in the sample dataset



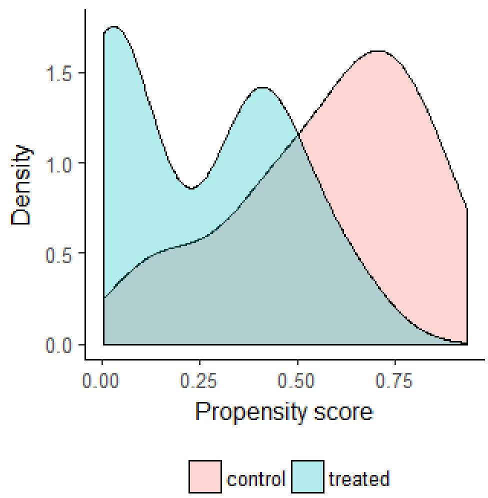
To calculate the propensity score, a logit model can be applied in which J treatment variable is regressed on income (X_1), satisfaction with housing (X_2), and age (X_3).

$$\ln\left(\frac{P(J=1)}{P(J=0)}\right) = \beta_0 + \beta_1 \times X_1 + \beta_2 \times X_2 + \beta_3 \times X_3 + \varepsilon \tag{24}$$

The predicted probability of this model gives the propensity scores (See Equation 22).

Regarding this sample dataset, the distribution of the propensity scores can be seen in Figure 9; further, Table 27 contains the propensity score of each observation calculated based on Equation 24.

Figure 9. Density of propensity scores in the sample dataset



As a result, the distance between two units can be easily calculated by taking the absolute value of the difference between their propensity scores according to Equation 23. For example, the distance between Unit 1 (whose propensity score is 0.66) and Unit 16 (whose propensity score is 0.40) is 0.26.

$$D^{Propensity\ Score}(1,16) = |e_1(X_{v1}) - e_{16}(X_{v16})| = |0.66 - 0.40| = 0.26 \tag{25}$$

Consequently, for other units the distance can be calculated in a similar manner. See the distance metrics, defined based on propensity scores, between the treated units and the control units in Table 3.

Table 3. Distance between treated and control units based on their propensity score using the sample dataset

$ID_j \setminus ID_i$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
16	0.26	0.04	0.31	0.18	0.15	0.31	0.54	0.30	0.05	0.05	0.29	0.45	0.38	0.42	0.23
17	0.62	0.40	0.67	0.54	0.50	0.67	0.89	0.06	0.30	0.40	0.65	0.81	0.74	0.77	0.13
18	0.62	0.40	0.67	0.54	0.51	0.67	0.89	0.06	0.30	0.40	0.65	0.81	0.74	0.77	0.13
19	0.54	0.31	0.58	0.46	0.42	0.58	0.81	0.02	0.22	0.32	0.57	0.73	0.65	0.69	0.04
20	0.24	0.01	0.28	0.16	0.12	0.28	0.51	0.33	0.08	0.02	0.26	0.42	0.35	0.39	0.26
21	0.26	0.04	0.31	0.18	0.15	0.31	0.54	0.30	0.05	0.05	0.29	0.45	0.38	0.42	0.23
22	0.25	0.02	0.29	0.17	0.13	0.29	0.52	0.32	0.07	0.03	0.28	0.44	0.36	0.40	0.25
23	0.34	0.12	0.39	0.26	0.23	0.39	0.61	0.22	0.02	0.12	0.37	0.53	0.46	0.49	0.15
24	0.65	0.43	0.70	0.57	0.54	0.70	0.93	0.09	0.34	0.44	0.68	0.84	0.77	0.81	0.16
25	0.66	0.43	0.70	0.58	0.54	0.70	0.93	0.10	0.34	0.44	0.69	0.85	0.77	0.81	0.17
26	0.66	0.44	0.71	0.58	0.54	0.71	0.93	0.10	0.34	0.44	0.69	0.85	0.78	0.81	0.17
27	0.66	0.43	0.70	0.58	0.54	0.70	0.93	0.09	0.34	0.44	0.68	0.84	0.77	0.81	0.16
28	0.61	0.39	0.66	0.53	0.49	0.66	0.88	0.05	0.29	0.39	0.64	0.80	0.73	0.76	0.12
29	0.00	0.22	0.04	0.08	0.12	0.04	0.27	0.56	0.32	0.22	0.03	0.19	0.12	0.15	0.49
30	0.53	0.31	0.57	0.45	0.41	0.57	0.80	0.03	0.21	0.31	0.56	0.72	0.65	0.68	0.04
31	0.31	0.08	0.35	0.23	0.19	0.35	0.58	0.26	0.01	0.09	0.33	0.50	0.42	0.46	0.19
32	0.66	0.43	0.70	0.57	0.54	0.70	0.93	0.09	0.34	0.44	0.68	0.84	0.77	0.81	0.16
33	0.23	0.01	0.28	0.15	0.11	0.28	0.50	0.33	0.09	0.01	0.26	0.42	0.35	0.38	0.26
34	0.10	0.13	0.14	0.02	0.02	0.14	0.37	0.47	0.22	0.12	0.13	0.29	0.21	0.25	0.40
35	0.33	0.11	0.38	0.25	0.22	0.38	0.60	0.23	0.01	0.11	0.36	0.52	0.45	0.48	0.16
36	0.66	0.44	0.71	0.58	0.54	0.71	0.93	0.10	0.34	0.44	0.69	0.85	0.78	0.81	0.17
37	0.66	0.43	0.70	0.58	0.54	0.70	0.93	0.09	0.34	0.44	0.68	0.85	0.77	0.81	0.16
38	0.20	0.03	0.24	0.12	0.08	0.24	0.47	0.37	0.12	0.02	0.23	0.39	0.31	0.35	0.30
39	0.06	0.17	0.10	0.03	0.06	0.10	0.33	0.51	0.26	0.16	0.08	0.24	0.17	0.21	0.44
40	0.66	0.44	0.71	0.58	0.54	0.70	0.93	0.10	0.34	0.44	0.69	0.85	0.78	0.81	0.17
41	0.22	0.00	0.27	0.14	0.11	0.27	0.50	0.34	0.09	0.01	0.25	0.41	0.34	0.38	0.27

Despite the advantages of propensity scores, this distance method has also been the subject of several critiques since the dimension reduction can leave crucial information unincorporated. Most recently, King and Nielsen (2016) have shown that propensity score matching generates higher levels of bias and model dependence than other matching methods. However, they have also argued that propensity scores can come in useful when combined with multivariate matching methods (such as genetic matching). Moreover, propensity scores are a widely used method of displaying improvements in balance (See Chapter 3.7.4.).

III. Multidimensional distances:

Other distance measures do not reduce dimensions, but are rather intended to measure the distance in a more dimensional space. One of the most basic multidimensional distance metrics is *Euclidean distance* (Deza & Deza, 2009) which is the geometric distance between two units in a multidimensional space. More specifically, this distance can be formulated as:

$$D^{Euclidean}(k,l) = \sqrt{\sum_{l}^{v}(X_{vk} - X_{vl})^2}\ , \tag{26}$$

where v is the number of variables used for calculating distance. Regarding the sample dataset (See Table 27), the Euclidean distance between Unit 1 and Unit 16 (based on values for age, satisfaction with housing and income) can be calculated in the following way:

$$D^{Euclidean}(1,16) = \sqrt{(4-10)^2 + (26-30)^2 + (55.24-42.58)^2} = 14.6 \tag{27}$$

Consequently, the Euclidean distance can be calculated between the other treated and control units (see Table 4.).

Table 4. Euclidean distance between treated and control units in the sample dataset

$ID_j \backslash ID_i$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
16	14.6	9.3	15.8	17.6	10.2	12.3	102.0	15.7	12.7	30.1	8.9	25.5	22.7	116.1	13.9
17	25.3	18.4	26.6	30.1	23.7	25.7	103.3	15.0	21.0	32.1	26.3	36.0	32.5	115.6	17.8
18	24.3	20.5	25.6	37.9	30.0	26.1	94.3	24.5	28.6	25.0	31.2	32.7	29.2	106.2	26.8
19	29.8	21.0	30.5	15.1	14.5	27.7	116.8	3.1	9.4	44.2	19.4	41.8	38.8	130.1	2.8
20	11.5	2.0	11.8	21.1	11.2	10.3	99.6	18.1	13.1	27.8	12.3	24.6	21.5	113.5	16.9
21	14.6	9.3	15.8	17.6	10.2	12.3	102.0	15.7	12.7	30.1	8.9	25.5	22.7	116.1	13.9
22	37.5	30.2	37.7	9.0	17.8	34.2	126.5	18.9	15.7	54.8	21.7	48.4	46.2	140.8	15.8
23	150.5	157.6	150.8	179.3	170.0	154.1	65.6	172.6	172.4	132.3	166.9	141.5	142.6	48.7	173.5
24	34.7	29.1	36.1	41.1	35.3	35.8	103.4	24.7	32.2	36.2	37.7	43.7	40.2	114.4	28.1
25	46.2	37.7	47.2	39.0	37.3	45.9	123.1	22.4	31.9	53.6	42.0	57.7	54.2	134.3	26.3
26	46.3	42.3	47.5	56.6	50.2	48.2	100.4	40.1	47.3	41.9	52.4	53.2	49.9	109.4	43.6
27	37.3	30.7	38.6	39.8	35.0	38.0	108.6	22.6	31.2	40.7	38.1	47.2	43.6	119.7	26.4
28	32.8	24.0	33.7	21.1	20.0	31.3	117.5	4.8	14.8	45.1	24.6	44.7	41.5	130.4	8.3
29	0.0	9.7	2.2	29.2	19.9	3.9	89.3	28.6	23.7	19.0	16.9	13.3	10.2	103.7	27.3
30	32.4	23.0	32.7	16.2	15.8	30.4	119.9	8.3	9.8	47.5	22.2	45.1	42.1	133.2	8.0
31	12.6	17.6	13.9	39.0	29.8	16.1	81.3	33.3	32.0	8.7	27.6	15.3	12.1	94.8	33.4
32	35.6	29.4	36.9	39.8	34.5	36.4	106.1	22.9	31.0	38.3	37.3	45.1	41.6	117.2	26.6
33	21.5	14.8	22.1	9.5	5.5	18.4	110.1	12.6	7.1	38.3	7.8	32.5	30.1	124.3	9.1
34	24.4	17.3	24.3	7.0	4.6	21.2	113.7	15.8	6.1	42.5	10.2	35.9	33.6	128.0	12.1
35	25.6	18.3	26.3	8.4	8.6	22.7	113.9	10.2	7.2	41.9	12.0	36.7	34.2	128.0	6.3
36	48.1	45.8	49.4	62.3	55.2	50.6	93.1	46.7	53.0	40.1	56.5	53.1	50.0	101.5	49.9
37	40.9	33.0	42.0	37.6	34.4	40.8	116.2	20.2	29.6	47.0	38.5	51.8	48.3	127.4	24.2
38	11.8	6.1	12.8	18.7	9.8	9.5	100.3	17.5	12.9	28.5	8.6	23.7	20.8	114.5	15.7
39	37.4	30.4	37.3	10.4	17.8	34.1	126.6	22.8	16.6	55.7	22.0	48.4	46.4	141.1	19.4
40	42.8	37.9	44.1	50.7	44.8	44.3	103.3	33.9	41.6	40.9	47.2	50.7	47.3	113.1	37.5
41	9.7	0.0	10.1	22.9	13.0	9.0	97.6	19.7	15.1	25.8	13.4	22.9	19.7	111.4	18.7

However, as the control variables do not usually use the same scale, their contribution to the Euclidean distance measure is also different. Also, in my example dataset the three control variables are measured on different scales. Therefore, the Euclidean distance would be developed primary on the basis of income since the scale of this variable is much wider than the scaling of the other two control variables. To overcome this challenge, one can standardize these variables. This *normalized Euclidean distance* can be written in the following way:

$$D^{Normalized\ Euclidean}(k,l)=\sqrt{\sum_{l}^{v}\frac{(X_{vk}-X_{vl})^2}{\delta_v}}\ , \qquad (28)$$

where δ_v is the variance of v variable. Equation 28 can be also written in the following way:

$$D^{Normalized\ Euclidean}(k,l)=\sqrt{(X_k - X_l)^T \times D^{-1} \times (X_k - X_l)}\ , \tag{29}$$

where D is a diagonal matrix containing the variances ($D = diag(\delta_1, \delta_2, \delta_3\ ...\ \delta_v)$), X_i matrix contains all the v covariates for i individual, and X^T is the transpose of the matrix X. Regarding the example dataset, one could calculate the normalized Euclidean distance between Unit 1 and Unit 16 in the following way:

$$D^{Normalized\ Euclidean}(1,16)=\sqrt{\frac{(4-10)^2}{7.91}+\frac{(26-30)^2}{237.61}+\frac{(55.24-42.58)^2}{1323.62}}=2.18 \tag{30}$$

Respectively, Table 5 displays the normalized Euclidian distance between each treated and control observation in the sample dataset.

Table 5. Normalized Euclidean distance between treated and control units in the sample dataset

$ID_j \backslash ID_i$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
16	2.18	2.85	2.88	0.95	2.16	1.84	3.00	1.02	2.51	1.37	0.88	0.98	1.03	3.67	0.88
17	2.04	2.40	2.62	1.69	2.03	1.90	3.12	0.63	2.11	1.26	1.69	1.95	1.72	3.41	0.86
18	1.89	2.17	2.41	1.93	1.95	1.83	2.92	0.96	1.97	1.16	1.89	2.10	1.81	3.10	1.16
19	1.78	2.24	2.41	0.97	1.63	1.55	3.29	0.15	1.84	1.29	1.08	1.60	1.38	3.73	0.14
20	0.83	0.06	0.48	2.25	0.83	1.17	3.27	2.25	0.51	1.94	2.22	2.65	2.26	3.31	2.20
21	2.18	2.85	2.88	0.95	2.16	1.84	3.00	1.02	2.51	1.37	0.88	0.98	1.03	3.67	0.88
22	1.75	2.30	2.37	0.30	1.50	1.42	3.50	0.97	1.85	1.62	0.64	1.42	1.28	4.07	0.73
23	4.67	4.94	4.91	5.35	5.15	4.74	2.56	4.87	5.20	3.90	5.06	4.54	4.44	1.95	4.95
24	3.03	3.37	3.61	2.50	3.04	2.89	3.57	1.49	3.11	2.07	2.51	2.61	2.49	3.83	1.71
25	2.85	2.99	3.31	2.53	2.74	2.78	3.97	1.43	2.69	2.24	2.62	2.95	2.73	4.11	1.69
26	3.47	3.65	3.94	3.29	3.51	3.42	3.89	2.21	3.47	2.61	3.29	3.39	3.21	3.91	2.46
27	3.09	3.41	3.66	2.53	3.07	2.95	3.72	1.51	3.13	2.17	2.57	2.71	2.58	3.97	1.74
28	2.25	2.68	2.87	1.41	2.14	2.03	3.45	0.44	2.30	1.55	1.50	1.87	1.72	3.88	0.63
29	0.00	0.81	0.71	1.64	0.55	0.39	2.68	1.82	0.80	1.27	1.51	1.84	1.46	2.94	1.71
30	1.34	0.75	1.18	2.34	1.09	1.58	3.80	2.14	0.61	2.21	2.40	2.97	2.58	3.81	2.14
31	1.24	1.85	1.91	1.33	1.44	1.06	2.28	1.02	1.69	0.25	1.13	1.21	0.86	2.70	0.99
32	3.04	3.38	3.61	2.48	3.03	2.89	3.63	1.46	3.10	2.10	2.51	2.63	2.51	3.90	1.69
33	1.87	2.52	2.56	0.53	1.78	1.52	3.11	0.86	2.14	1.32	0.54	1.01	0.95	3.74	0.64
34	0.76	0.69	0.76	1.79	0.38	0.92	3.43	2.03	0.39	1.91	1.80	2.36	2.00	3.66	1.92
35	2.25	2.88	2.94	0.85	2.15	1.91	3.31	0.96	2.49	1.58	0.91	1.22	1.25	3.96	0.81
36	3.73	3.96	4.22	3.47	3.80	3.66	3.87	2.43	3.79	2.76	3.45	3.46	3.32	3.92	2.67
37	2.90	3.14	3.42	2.43	2.83	2.78	3.80	1.35	2.84	2.12	2.50	2.75	2.57	4.01	1.60
38	1.13	1.78	1.82	0.72	1.11	0.81	2.76	0.88	1.47	0.84	0.59	1.10	0.76	3.26	0.69
39	1.27	0.99	1.04	2.15	0.89	1.42	3.92	2.46	0.78	2.45	2.21	2.82	2.48	4.12	2.34
40	3.46	3.72	3.99	3.07	3.48	3.36	3.87	2.02	3.49	2.53	3.08	3.17	3.03	4.02	2.26
41	0.81	0.00	0.45	2.27	0.85	1.17	3.22	2.26	0.55	1.92	2.22	2.64	2.24	3.26	2.21

However, even if the variables use the same scale, one should take into account the collinearity between them. If the two (or more) variables are correlated with each other, then the same information would be taken into account multiple times when calculating the distance. For example, in the sample dataset age and satisfaction with housing are correlated. See the covariance matrix in Table 6.

Table 6. Covariance matrix in the sample dataset

	Satisfaction with housing	Age	Income
Satisfaction with housing	7.91	20.67	-1.67
Age	20.67	237.67	67.27
Income	-1.67	67.27	1323.62

As a result, one needs to take into account the correlation between the variables, not just the variance of each variable. For this the *Mahalanobis distance* can be used, which is defined in the following way:

$$D^{Mahalanobis}(k,l) = \sqrt{(X_k - X_l)^T \times S^{-1} \times (X_k - X_l)}\,, \qquad (31)$$

where S is the sample covariance matrix of X. Thus, the normalized Euclidean distance is equal to the Mahalanobis distance when there is no collinearity between the variables (Equation 31 and 29 differ only in this aspect).

Equation 32 demonstrates how to calculate the Mahalanobis distance between Unit 1 and Unit 16, while the rest of the distances can be consequently calculated (See Table 7.).

$$D^{Mahalanobis}(1,16) = \sqrt{\begin{bmatrix} 4-10 \\ 26-30 \\ 55.24-42 \end{bmatrix} \begin{bmatrix} 7.91 & 20.67 & -1.67 \\ 20.67 & 237.67 & 67.27 \\ -1.67 & 67.27 & 1323.62 \end{bmatrix}^{-1} \begin{bmatrix} 4-10 & 26-30 & 55.24-42 \end{bmatrix}} = 2.31 \qquad (32)$$

Table 7. Mahalanobis distance between treated and control units using the sample dataset

$ID_j \backslash ID_i$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
16	2.31	3.28	3.08	0.85	2.33	1.86	2.98	1.42	2.92	1.54	0.74	0.97	0.94	3.77	1.19
17	1.71	2.14	2.23	1.86	1.66	1.65	3.32	0.65	1.83	1.30	1.93	2.48	2.03	3.39	0.92
18	1.61	1.84	1.99	2.25	1.63	1.70	3.27	1.11	1.64	1.38	2.29	2.80	2.30	3.14	1.35
19	1.61	2.26	2.25	1.11	1.44	1.36	3.39	0.15	1.83	1.29	1.27	2.02	1.60	3.72	0.17
20	1.11	0.06	0.56	2.77	1.05	1.57	3.69	2.21	0.52	2.12	2.81	3.43	2.88	3.36	2.29
21	2.31	3.28	3.08	0.85	2.33	1.86	2.98	1.42	2.92	1.54	0.74	0.97	0.94	3.77	1.19
22	1.85	2.68	2.54	0.35	1.65	1.44	3.50	1.05	2.23	1.63	0.70	1.61	1.33	4.07	0.77
23	4.37	4.67	4.53	5.27	4.86	4.48	2.62	4.80	4.94	3.79	5.01	4.65	4.42	1.73	4.88
24	2.52	2.89	3.02	2.39	2.49	2.43	3.63	1.28	2.61	1.88	2.45	2.87	2.52	3.71	1.52
25	2.59	2.51	2.83	2.91	2.41	2.67	4.36	1.66	2.24	2.45	3.06	3.67	3.23	4.17	1.96
26	3.03	3.00	3.28	3.45	3.02	3.13	4.20	2.23	2.86	2.65	3.52	3.94	3.52	3.88	2.51
27	2.59	2.91	3.07	2.45	2.53	2.51	3.79	1.31	2.62	2.01	2.54	2.99	2.64	3.85	1.56
28	1.97	2.55	2.59	1.41	1.83	1.75	3.53	0.38	2.14	1.49	1.56	2.20	1.84	3.84	0.54
29	0.00	1.09	0.78	1.84	0.55	0.49	2.83	1.61	0.99	1.17	1.80	2.33	1.78	2.87	1.57
30	1.79	0.86	1.37	3.14	1.53	2.18	4.39	2.38	0.85	2.66	3.26	3.98	3.43	4.01	2.53
31	1.10	1.99	1.82	1.43	1.33	0.87	2.41	1.02	1.79	0.26	1.30	1.68	1.15	2.69	0.98
32	2.53	2.90	3.03	2.38	2.49	2.45	3.69	1.26	2.60	1.92	2.46	2.89	2.54	3.77	1.50
33	2.00	2.94	2.76	0.44	1.96	1.55	3.10	1.17	2.56	1.43	0.46	1.13	0.93	3.79	0.89
34	0.79	0.81	0.76	2.10	0.40	1.08	3.62	1.78	0.44	1.86	2.18	2.90	2.37	3.59	1.78
35	2.36	3.28	3.11	0.73	2.30	1.92	3.29	1.34	2.87	1.71	0.78	1.24	1.18	4.03	1.10
36	3.15	3.25	3.48	3.42	3.18	3.19	4.01	2.26	3.11	2.60	3.46	3.78	3.41	3.76	2.52
37	2.49	2.65	2.88	2.57	2.37	2.49	4.02	1.34	2.35	2.13	2.69	3.24	2.84	3.97	1.62
38	1.15	2.10	1.91	0.83	1.17	0.74	2.81	0.85	1.76	0.81	0.79	1.48	0.98	3.25	0.65
39	1.28	0.97	1.02	2.43	0.88	1.54	4.08	2.17	0.67	2.36	2.55	3.31	2.81	4.02	2.17
40	2.91	3.08	3.30	3.03	2.89	2.92	4.00	1.85	2.86	2.38	3.09	3.50	3.13	3.88	2.12
41	1.09	0.00	0.53	2.77	1.06	1.56	3.64	2.22	0.56	2.10	2.81	3.41	2.86	3.31	2.30

Just like the previous distance measures, the Mahalanobis distance has also some limitations. Most of all, this metric performs poorly when the covariates have non-ellipsoidal distributions (e.g. a normal or t distribution). In this case, the covariance matrix might fail to account for collinearity because the distribution of covariates might differ more than their means and variance (e.g. momentum). In the case of non-ellipsoidal distributions, propensity score matching could be a better solution. Moreover, other scholars have found that the Mahalanobis distance, similarly to exact matching, performs poorly when the number of the covariates are too high; more specifically, more than eight (Sekhon, 2009; Stuart, 2010).

The empirical studies in this dissertation employ a *genetic matching distance* measure,[12]

[12] The analysis was performed using MatchIt software, which runs in the R environment.

which relies on a genetic search algorithm to maximize the balance between the treatment and control groups through the generalization of the propensity score and Mahalanobis distance matching (Diamond & Sekhon, 2013; Sekhon, 2008; Sekhon & Grieve, 2008). First, the Mahalanobis metric must be transformed by using Cholesky decomposition[13] and adding a weight parameter:

$$D^{Genetic}(k,l) = \sqrt{(X_k - X_l)^T \times (S^{-1/2})^T \times W \times S^{-1/2} \times (X_k - X_l)}\,, \qquad (33)$$

where $S^{-1/2}$ is the Cholesky decomposition of the S covariance matrix and W is a positive definite weight matrix which contains a set of weights for each X covariate. Besides the covariates, the propensity score can also be used in this matching mechanism, and thus can also influence the distance metrics with a given weight.

This distance measure runs a genetic search algorithm to find W matrix such that the optimal balance between the treatment and control groups is identified. A *genetic search algorithm* is a strategy for running multiple *local search algorithms* in parallel. Local search algorithms initially select a starting point randomly, keep track only of current states, and move to neighbouring states (i.e. local modification). Thus, these strategies optimize the solution by considering local modifications only. By running multiple local searches at the same time, a genetic algorithm is not only able to conduct local modifications but can use combinations of the states of different local searches, creating a better starting point for a new local search (Holland, 1992; Selman & Gomes, 2006).

More specifically, genetic matching relies on a *loss function* when searching for the optimal W matrix. This loss function can be specified by the user. The default loss function, which was used in this dissertation as well, minimizes the difference between the units based on P-values from Kolmogorov-Smirnov distributional tests[14] and paired t-tests[15]. The algorithm initially proposes a set of weights for the variables which are used iteratively to produce the subsequent set of weights. Each set of candidate weights is called a *generation*. Every subsequent generation is a better candidate for being the optimal W matrix in terms of the loss function than the previous generations (Diamond & Sekhon, 2013; Sekhon, 2008; Sekhon & Grieve, 2008).

---

[13] that is, $S = S^{1/2}\left(S^{1/2}\right)^T$, in which $S^{1/2}$ is a lower triangular matrix with positive diagonal elements

[14] for multinomial and continuous variables

[15] for dichotomous variables

As the calculation of the weight matrix requires huge computational power and long iterations, the present dissertation will not demonstrate how to obtain it, but only how to calculate the distance measure once the W matrix is obtained. Using the GenMatch package of R, the genetic matching process finds the following W matrix for the sample dataset $W_{original} = \begin{bmatrix} 550.44 & 0 & 0 \\ 0 & 14.97 & 0 \\ 0 & 0 & 902.30 \end{bmatrix}$. The matches do not change when this W matrix is divided by any positive scalar. Thus, this example divides the weight matrix by 1000 in order to obtain a more compact result (Diamond & Sekhon, 2013). As a result, $W = \begin{bmatrix} 0.55 & 0 & 0 \\ 0 & 0.02 & 0 \\ 0 & 0 & 0.90 \end{bmatrix}$. To obtain the genetic matching distance, one also needs to calculate the Cholesky decomposition of the S covariance matrix (See in Table 6), which is $S^{-1/2} = \begin{bmatrix} 0.41 & 0 & 0 \\ -0.04 & 0.07 & 0 \\ 0 & -0.01 & 0.03 \end{bmatrix}$ in the sample dataset.

Based on these measures, the distance between Unit 1 and Unit 16 can be calculated in the following way:

$$D^{Genetic}(1,16) = \sqrt{(X_1 - X_{16})^T \times (S^{-1/2})^T \times W \times S^{-1/2} \times (X_1 - X_{16})} = \tag{34}$$

$$= \sqrt{\begin{bmatrix} -6 & -4 & 12.66 \end{bmatrix} \times \begin{bmatrix} 0.41 & -0.04 & 0.00 \\ 0 & 0.07 & -0.01 \\ 0 & 0 & 0.03 \end{bmatrix} \times \begin{bmatrix} 0.55 & 0 & 0 \\ 0 & 0.02 & 0 \\ 0 & 0 & 0.90 \end{bmatrix} \times \begin{bmatrix} 0.41 & 0 & 0 \\ -0.04 & 0.07 & 0 \\ 0.00 & -0.01 & 0.03 \end{bmatrix} \times \begin{bmatrix} -6 \\ -4 \\ 12.66 \end{bmatrix}} =$$

$$= 2.93$$

Similarly to these cases, the distance between other units can be calculated too.

One advantage of this method is that genetic matching significantly reduces bias compared to pre-existing matching methods and can replicate even better the experimental benchmarks. Diamond and Sekhon (2013) contributed to the debate about labour economics by showing how to reproduce the result of an experiment (LaLonde, 1986) on observational data. The authors pointed out that previous matching methods have failed to produce reliable results due to poor balance, but genetic matching is able to replicate the experimental benchmark. Similarly, Sekhon and Grieve (2008) have also shown that genetic matching is able to replicate the results of a clinical experiment, in contrast to the previously described matching methods. Therefore, this dissertation uses the genetic matching distance technique as well.

3.7.3. Types of Matching Algorithms

After selecting the distance measure, one needs to find a method to match units based on this given metric. For example, genetic matching, which is applied in the empirical part, also can be conducted with several types of algorithm. The package I used by default use the algorithm which was described by Abadie and Imbens (2006)[16]. This subchapter summarizes the most commonly used algorithms. However, before describing the given algorithms, I discuss a few questions which arise in the case of all kinds of matching algorithms. Namely, (1) how many matches should be selected, and (2) whether matching should be with or without replacement.

The first question concerns the number of matches. Researchers are generally advised to match all the treated units; however, one can manipulate how many control units should be matched. A l control unit may be either matched $(m=1)$ or not $(m=0)$. In the case of a large initial dataset, one might consider selecting multiple controls for one treated unit. In this case, the treated unit is compared to the average of the selected control units. Thus, the average treatment effect introduced in Equation 3 can be reformulated in the following way:

$$ATE = E\left[Y_{i1} - \hat{Y}_{i0}\right], \tag{35}$$

where $\hat{Y}_{i0}$ is the average of the outcomes of the $N_{(j=0,m=1)}$ number of matched control units $(\hat{Y}_{i0} = \frac{1}{N_{(j=0,m=1)}} \sum_{\substack{i:\ J=0 \\ i:\ M=1}} Y_{i0})$. The other types of causal parameters can consequently be calculated by replacing Y_{i0} with $\hat{Y}_{i0}$. The number of matches obviously implies a trade-off. If $N_{(j=0,m=1)}$ is too small, the sample size is also too small, which increases the variance. However, if $N_{(j=0,m=1)}$ is large, then worse matches would be selected as well, which can distort the results. The research in this dissertation does not allow only one control unit for each treated unit, which is the default setting in the applied package.

---

[16] This algorithm applies a bias correction that renders simple nearest-neighbor matching estimators $N^{1/2}$ -consistent and asymptotically normal (Abadie and Imbens, 2006).

The second decision concerns whether matching should be conducted with or without replacement. Matching with replacement means that a control individual can be matched to more than one treated individual, whereas matching without replacements means that one control unit can be considered only once in the matching procedure. This issue also suggests a trade-off between bias and variance. The research in this dissertation uses the default setting, which allows replacement.

In the following sections, the different types of matching algorithms, such as nearest neighbour matching, optimal matching, subclassification and weighting adjustment, are discussed in detail. These algorithms are illustrated based on the propensity distance (this distance between observations can be seen in Table 3) calculated for the sample dataset (as seen in Table 27).

I. Nearest neighbour

One of the most common ways of finding matches is applying *nearest neighbour matching*. This matching algorithm chooses for each k treated individual one or more l control individuals which is/are closest to individual k based on the chosen distance measure. The $M(l)$ matched control individuals for k treated individual can be obtained by the following optimization:

$$M(l) = \arg \min_{j=0} D(k,l) \tag{36}$$

The simplest form of nearest neighbour matching is 1:1 matching which selects for each k treated individual one l control individual at the smallest distance from individual k. In contrast, one might select more control individuals for each treated individual.

The results of this matching algorithm significantly depend on the order in which treated units are matched. The R package that was used can distinguish three types of orders. First, one might match treated individuals to control individuals starting from the highest value of the distance measure to the lowest value. Second, one might match in order from the lowest distance to the highest. Finally, one might choose treated individuals in a random order and match them to the closest available control units.

This dissertation demonstrates (in Table 8) 1:1 nearest neighbour matching without replacements on the propensity score distance calculated for the sample dataset (see the calculations in the previous chapter). Matching is conducted in a way that first the smallest value of distance measures is matched and then matching continues in increasing order of distances values.

In Table 8 the smallest distance is 0 between Unit 1 and Unit 29, and also 0 between Unit 2 and Unit 41. Thus first these pairs are matched. These steps remove Units 1, 2, 29, and 41 from further consideration because each unit can be matched only once in the case of $1:1$ matching without replacements. The second largest value in the distance matrix is between Unit 10 and Unit 41 $\left(D^{propensity\ score}(10,41)=0.0055\right)$. But control Unit 41 has been selected already, hence this pairing is not realized. Similarly to this case, the third lowest distance, which is between Units 2 and 33 $\left(D^{propensity\ score}(2,33)=0.0068\right)$, also does not result in matching since Unit 2 has been already selected before. The fourth lowest distance is between Unit 9 and 31 $\left(D^{propensity\ score}(9,31)=0.0116\right)$ and, since neither of these observations has been matched before, these two units are matched. The algorithm continues in this way, selecting the matches until for each treated unit $(ID_k \leq 15)$ one control unit $(ID_l \leq 16)$ is selected. Consequently, Unit 1 is matched to Unit 29, 2 to 41, 3 to 22, 4 to 39, 5 to 34, 6 to 20, 7 to 23, 8 to 19, 9 to 31, 10 to 33, 11 to 38, 12 to 35, 13 to 21, 14 to 16, and 15 to 30. Table 8 shows the distance values for each pair (pairs which were selected are highlighted in bold and red colour).

Table 8. 1:1 nearest neighbour matching based on propensity score using the sample dataset

ID_j \ ID_i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
16	0.26	0.04	0.31	0.18	0.15	0.31	0.54	0.30	0.05	0.05	0.29	0.45	0.38	**0.42**	0.23
17	0.62	0.40	0.67	0.54	0.50	0.67	0.89	0.06	0.30	0.40	0.65	0.81	0.74	0.77	0.13
18	0.62	0.40	0.67	0.54	0.51	0.67	0.89	0.06	0.30	0.40	0.65	0.81	0.74	0.77	0.13
19	0.54	0.31	0.58	0.46	0.42	0.58	0.81	**0.02**	0.22	0.32	0.57	0.73	0.65	0.69	0.04
20	0.24	0.01	0.28	0.16	0.12	**0.28**	0.51	0.33	0.08	0.02	0.26	0.42	0.35	0.39	0.26
21	0.26	0.04	0.31	0.18	0.15	0.31	0.54	0.30	0.05	0.05	0.29	0.45	**0.38**	0.42	0.23
22	0.25	0.02	**0.29**	0.17	0.13	0.29	0.52	0.32	0.07	0.03	0.28	0.44	0.36	0.40	0.25
23	0.34	0.12	0.39	0.26	0.23	0.39	**0.61**	0.22	0.02	0.12	0.37	0.53	0.46	0.49	0.15
24	0.65	0.43	0.70	0.57	0.54	0.70	0.93	0.09	0.34	0.44	0.68	0.84	0.77	0.81	0.16
25	0.66	0.43	0.70	0.58	0.54	0.70	0.93	0.10	0.34	0.44	0.69	0.85	0.77	0.81	0.17
26	0.66	0.44	0.71	0.58	0.54	0.71	0.93	0.10	0.34	0.44	0.69	0.85	0.78	0.81	0.17
27	0.66	0.43	0.70	0.58	0.54	0.70	0.93	0.09	0.34	0.44	0.68	0.84	0.77	0.81	0.16
28	0.61	0.39	0.66	0.53	0.49	0.66	0.88	0.05	0.29	0.39	0.64	0.80	0.73	0.76	0.12
29	**0.00**	0.22	0.04	0.08	0.12	0.04	0.27	0.56	0.32	0.22	0.03	0.19	0.12	0.15	0.49
30	0.53	0.31	0.57	0.45	0.41	0.57	0.80	0.03	0.21	0.31	0.56	0.72	0.65	0.68	**0.04**
31	0.31	0.08	0.35	0.23	0.19	0.35	0.58	0.26	**0.01**	0.09	0.33	0.50	0.42	0.46	0.19
32	0.66	0.43	0.70	0.57	0.54	0.70	0.93	0.09	0.34	0.44	0.68	0.84	0.77	0.81	0.16
33	0.23	0.01	0.28	0.15	0.11	0.28	0.50	0.33	0.09	**0.01**	0.26	0.42	0.35	0.38	0.26
34	0.10	0.13	0.14	0.02	**0.02**	0.14	0.37	0.47	0.22	0.12	0.13	0.29	0.21	0.25	0.40
35	0.33	0.11	0.38	0.25	0.22	0.38	0.60	0.23	0.01	0.11	0.36	**0.52**	0.45	0.48	0.16
36	0.66	0.44	0.71	0.58	0.54	0.71	0.93	0.10	0.34	0.44	0.69	0.85	0.78	0.81	0.17
37	0.66	0.43	0.70	0.58	0.54	0.70	0.93	0.09	0.34	0.44	0.68	0.85	0.77	0.81	0.16
38	0.20	0.03	0.24	0.12	0.08	0.24	0.47	0.37	0.12	0.02	**0.23**	0.39	0.31	0.35	0.30
39	0.06	0.17	0.10	**0.03**	0.06	0.10	0.33	0.51	0.26	0.16	0.08	0.24	0.17	0.21	0.44
40	0.66	0.44	0.71	0.58	0.54	0.70	0.93	0.10	0.34	0.44	0.69	0.85	0.78	0.81	0.17
41	0.22	**0.00**	0.27	0.14	0.11	0.27	0.50	0.34	0.09	0.01	0.25	0.41	0.34	0.38	0.27

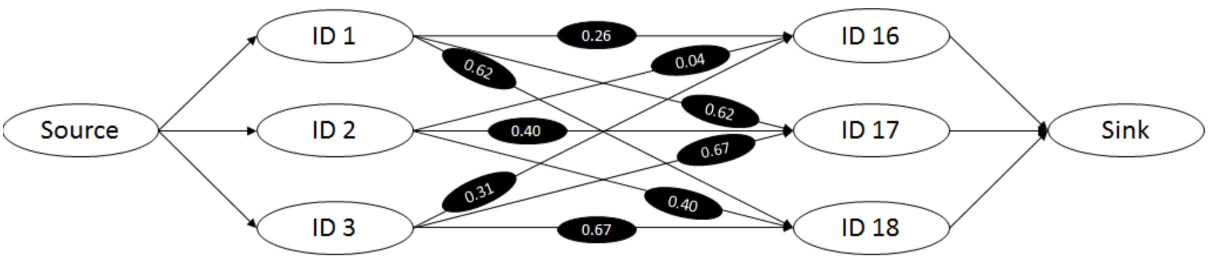
Nearest neighbour matching applies a *greedy local search* algorithm. This means that in each step moves are made which result in the greatest improvement of an objective function (here, the minimization of the distance measure). A common complaint regarding this optimization mechanism is that it leads to a local optimum, which is not necessarily the global optimum. Regarding matching, this means that this algorithm might match the closest two units in each step, but the process will not necessarily produce the minimal total distance between all the matched pairs at the end of matching. Generally, greedy nearest neighbour matching has been found to perform well when there are a sufficient amount of good matches, but it is not a good choice when there is prominent competition for good control individuals (Gu & Rosenbaum, 1993).

II. Optimal matching

In contrast to simple nearest neighbour matching, *optimal matching* relies on a *global search* algorithm. This method searches for a solution which minimizes overall distance. Thus, this method aims to find $M(1)$, $M(2)$, $M(3)$, ... $M(N_{(j=1,m=1)})$ matches in a way that minimizes $\sum_{k=1}^{N_k} D(k, M(l))$ total distance between the matched pairs (Gu & Rosenbaum, 1993; Rosenbaum, 1989; Stuart, 2010).

This algorithm uses *network flow theory* for optimization. Networks are used as a metaphor to demonstrate this solution. Based on this method, vertexes consist of N treated and control units, a source, and a sink. The arrows between the vertexes contain information about the cost of connecting a treated unit to a potential control unit, which is defined as the distance between given treated and control units. The algorithm aims to minimize the cost by finding the shortest flow between the source and the sink. The network flow is shown in Figure 10 for the first three treated and first three control units of the sample dataset. This simple form of network can be extended by including all possible pairs and looking for more than one match.

Figure 10. Demonstration of network flow theory in the first three rows and columns of the sample dataset



There are several ways of finding the *minimum cost flow* in a network through computational processes. One of the optimization methods is referred to as *the cost reducing cycle*. This method starts with a matched sample – for example, the result of greedy nearest neighbour matching – and then searches for alternative networks which have a lower total distance. The alternative networks are obtained by matching the unmatched units (*utilizing unused flow capacity*), and unmatching previously matched units and matching them to other units (*rechannelling existing flows*). See more about different ways of conducting optimal matching in Rosenbaum (1989).

Optimal matching is conducted here using the Matchit R package. Table 9 shows the propensity distance metrics and highlights those values which were chosen by the optimal matching algorithm. By optimal matching, Unit 1 is paired with Unit 23, 2 with 31, 3 with 29, 4 with 34, 5 with 16, 6 with 21, 7 with 41, 8 with 19, 9 with 35, 10 with 20, 11 with 22, 12 with 33, 13 with 38, 14 with 39, and 15 with 30.

Table 9. Optimal matching based on propensity score using the sample dataset

$ID_j \backslash ID_i$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
16	0.26	0.04	0.31	0.18	**0.15**	0.31	0.54	0.30	0.05	0.05	0.29	0.45	0.38	0.42	0.23
17	0.62	0.40	0.67	0.54	0.50	0.67	0.89	0.06	0.30	0.40	0.65	0.81	0.74	0.77	0.13
18	0.62	0.40	0.67	0.54	0.51	0.67	0.89	0.06	0.30	0.40	0.65	0.81	0.74	0.77	0.13
19	0.54	0.31	0.58	0.46	0.42	0.58	0.81	**0.02**	0.22	0.32	0.57	0.73	0.65	0.69	0.04
20	0.24	0.01	0.28	0.16	0.12	0.28	0.51	0.33	0.08	**0.02**	0.26	0.42	0.35	0.39	0.26
21	0.26	0.04	0.31	0.18	0.15	**0.31**	0.54	0.30	0.05	0.05	0.29	0.45	0.38	0.42	0.23
22	0.25	0.02	0.29	0.17	0.13	0.29	0.52	0.32	0.07	0.03	**0.28**	0.44	0.36	0.40	0.25
23	**0.34**	0.12	0.39	0.26	0.23	0.39	0.61	0.22	0.02	0.12	0.37	0.53	0.46	0.49	0.15
24	0.65	0.43	0.70	0.57	0.54	0.70	0.93	0.09	0.34	0.44	0.68	0.84	0.77	0.81	0.16
25	0.66	0.43	0.70	0.58	0.54	0.70	0.93	0.10	0.34	0.44	0.69	0.85	0.77	0.81	0.17
26	0.66	0.44	0.71	0.58	0.54	0.71	0.93	0.10	0.34	0.44	0.69	0.85	0.78	0.81	0.17
27	0.66	0.43	0.70	0.58	0.54	0.70	0.93	0.09	0.34	0.44	0.68	0.84	0.77	0.81	0.16
28	0.61	0.39	0.66	0.53	0.49	0.66	0.88	0.05	0.29	0.39	0.64	0.80	0.73	0.76	0.12
29	0.00	0.22	**0.04**	0.08	0.12	0.04	0.27	0.56	0.32	0.22	0.03	0.19	0.12	0.15	0.49
30	0.53	0.31	0.57	0.45	0.41	0.57	0.80	0.03	0.21	0.31	0.56	0.72	0.65	0.68	**0.04**
31	0.31	**0.08**	0.35	0.23	0.19	0.35	0.58	0.26	0.01	0.09	0.33	0.50	0.42	0.46	0.19
32	0.66	0.43	0.70	0.57	0.54	0.70	0.93	0.09	0.34	0.44	0.68	0.84	0.77	0.81	0.16
33	0.23	0.01	0.28	0.15	0.11	0.28	0.50	0.33	0.09	0.01	0.26	**0.42**	0.35	0.38	0.26
34	0.10	0.13	0.14	**0.02**	0.02	0.14	0.37	0.47	0.22	0.12	0.13	0.29	0.21	0.25	0.40
35	0.33	0.11	0.38	0.25	0.22	0.38	0.60	0.23	**0.01**	0.11	0.36	0.52	0.45	0.48	0.16
36	0.66	0.44	0.71	0.58	0.54	0.71	0.93	0.10	0.34	0.44	0.69	0.85	0.78	0.81	0.17
37	0.66	0.43	0.70	0.58	0.54	0.70	0.93	0.09	0.34	0.44	0.68	0.85	0.77	0.81	0.16
38	0.20	0.03	0.24	0.12	0.08	0.24	0.47	0.37	0.12	0.02	0.23	0.39	**0.31**	0.35	0.30
39	0.06	0.17	0.10	0.03	0.06	0.10	0.33	0.51	0.26	0.16	0.08	0.24	0.17	**0.21**	0.44
40	0.66	0.44	0.71	0.58	0.54	0.70	0.93	0.10	0.34	0.44	0.69	0.85	0.78	0.81	0.17
41	0.22	0.00	0.27	0.14	0.11	0.27	**0.50**	0.34	0.09	0.01	0.25	0.41	0.34	0.38	0.27

One way to compare how well a matching algorithm has performed is to consider the total distance between the matched pairs $\sum_{k=1}^{N_k} D(k, M(l))$. In the sample dataset this measurement is 2.78 in the case of optimal matching, and 2.86 in the case of nearest neighbour matching. This difference demonstrates that even though the nearest neighbour algorithm may be optimal in each separate step, it can be suboptimal globally. Further, Table 9 shows that, in contrast to nearest neighbour matching, optimal matching does not choose any of the two 0.00 distance values since they might be optimal in a local search algorithm but are suboptimal globally.

However, both nearest neighbour and optimal matching might leave a lot of information unincorporated. These algorithms select only a limited number of control units for each treated individual,[17] although there might be more suitable control units available. Moreover, these methods mostly enable us to estimate ATT. The methods introduced in the following chapter use all units, which enables us to estimate ATE as well.

III. Subclassification

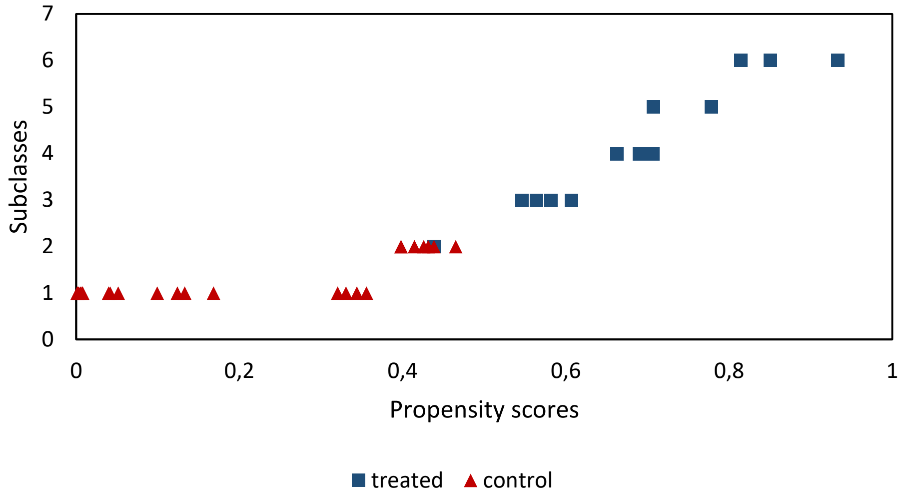
Subclassification matching divides the units into subclasses in a way that every unit within the subclasses is as similar to each other as possible (the researcher does not discard any of the units, thus all of them belong to one of the subclasses). The treatment effect is first calculated for every subclass separately, and the average of these treatment effects gives the overall treatment effect (Cochran, 1968; Lunceford & Davidian, 2004; Stuart, 2010; Yang, Imbens, Cui, Faries, & Kadziola, 2016).

The major issue in the case of subclassification is the number of subclasses. Rosenbaum and Rubin (1984) have found that the application of five subclasses based on the propensity score can remove at least 90% of the bias in estimating the treatment effect. Further, Lunceford and Davidian (2004) have argued that one only needs to use a higher number of subclasses if the number of observations is higher.

Figure 11 shows the results of subclassification matching based on propensity scores on the sample dataset. This figure also demonstrates one of the major drawbacks of this algorithm. More specifically, almost all the subclasses either include only treated or only control units (except for the third subclasses) due to pronounced initial differences between the treated and control group. The problem with this is that the calculation of the treatment effect in the subclasses requires at least one control and one treated unit. Thus, in certain subclasses the treatment effect cannot not be calculated at all, however, discarding some of the subclasses from the calculation of the overall effect can lead to biased estimations.

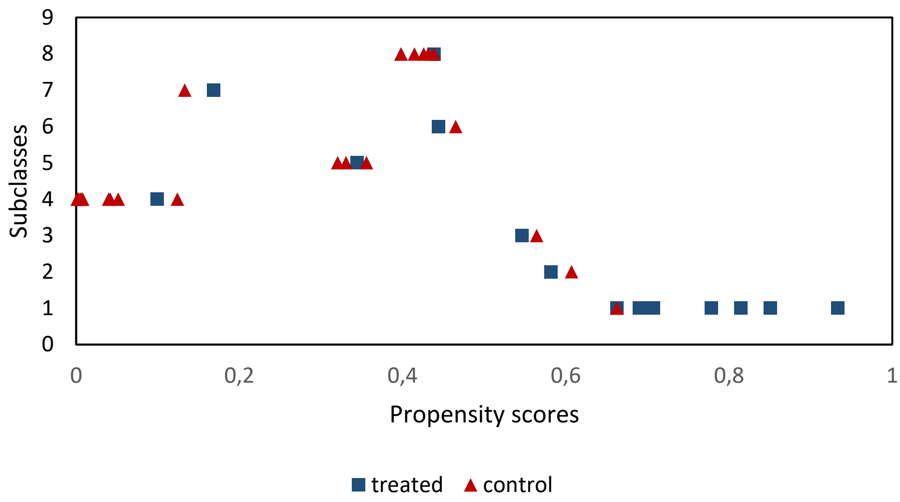
[17] For example, the method presented above selects only one control unit for each treated unit, but this number can also be increased.

Figure 11. Subclassification based on propensity scores on the sample dataset



One extension of subclassification is *full matching*, which overcomes the above-mentioned problem by creating subclasses in which there is at least one treatment unit and at least one control unit (Hansen, 2004; Stuart, 2010). Thus, comparison within the subclasses is guaranteed. Figure 12 shows the subclasses that full matching would have created based on propensity scores for the sample dataset. This figure demonstrates that full matching ensures that there is at least one treated and one control unit in every subclass. However, due to this property, this algorithm sometimes matches pairs which are quite far from each other.

Figure 12. Full matching based on propensity scores for the sample dataset



Subclassification methods are suitable for use when one aims to avoid the discarding of observations (or in other words, wants to estimate ATE). However, these methods are not the best choice if some of the units have a zero probability of receiving the treatment. In this case, one needs to estimate ATT by removing some of the control units.

IV. Weighting adjustment

Matching weights are produced in all matching mechanisms to award different levels of importance to the control units when estimating the treatment effect. 1:1 nearest neighbour and optimal matching produce a very simple weight which takes a value of 1 if the unit is matched and 0 if the unit has been discarded. Other methods might match more control units to a treated unit, but they weight them based on their importance in the matching. Often, each control unit receives a weight which is proportional to the number of treatment units to which they were matched. This weighting is also used in the empirical part of this dissertation.

Further, the *inverse probability of treatment weighting* is a particularly noteworthy method. This technique defines for each i individual a w_i weight which is $w_i = \frac{J_i}{\hat{e}_i} - \frac{1-J_i}{1-\hat{e}_i}$. Thus, the process of calculating this matching weight resonates with the Horvitz–Thompson estimator for calculating survey weights (Czajka, Hirabayashi, Little, & Rubin, 1992; Horvitz & Thompson, 1952). A disadvantage of this weighting technique is that the method might produce weights with overly large variance (Osborne, 2008). Thus, one runs the risk that the estimations are the result of the estimation procedure instead of the actual probabilities. Therefore, some scholars have suggested trimming these weights to obtain more reliable results (Scharfstein, Rotnitzky, & Robins, 1999).

Another weighting technique which is widely used is *weighting by the odds*. This technique is particularly appropriate for estimating ATT. In this case, the weights can be defined as $w_i = J_i + (1-J_i) \times \frac{\hat{e}_i}{1-\hat{e}_i}$. Alternatively, one can use also *kernel weighting*, which is mostly popular among economists. This technique applies a distance function with a bandwidth parameter which controls the smoothness of the estimate (Stuart, 2010).

3.7.4. Evaluating the Balance on the Matched Dataset

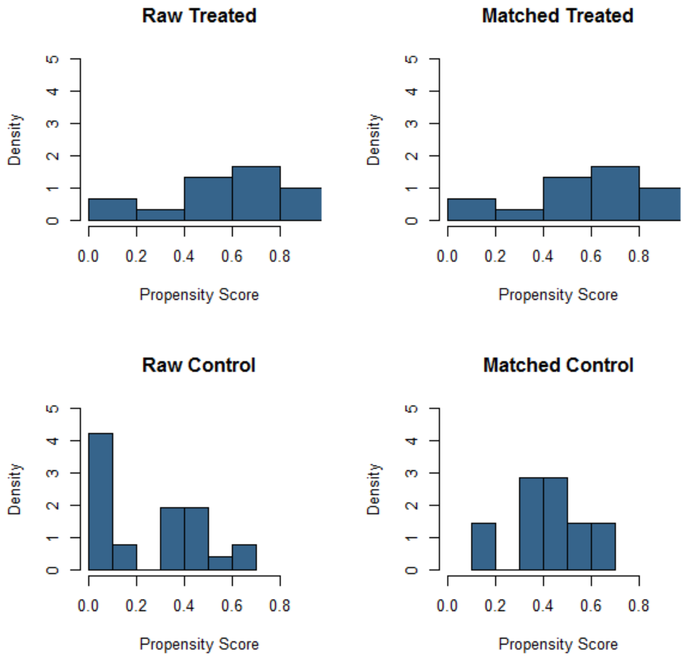
After matching is conducted, one needs to evaluate the balance between the treatment and control groups on the matched dataset. A well balanced dataset is indispensable for quantifying to what extent the difference between the control group and treatment groups can be attributed to the treatment effect rather than the difference in composition of the two groups. Good balance means that the treatment and control groups only randomly differ from each other in all covariates (Stuart, 2010). In the case of exact matching, the balance between the treatment group and control groups is guaranteed; however, this is not so obvious in the case of the other matching methods.

Ho et al. (2011) has proposed that the balance can be assessed by devising a descriptive statistic that summarizes how the mean and the standard deviation of the various variables have changed in the treatment and control groups upon matching (see Tables 28, 30, 32, 34, 36, 38, and 39).[18] Based on these statistics, we could conduct a t-test or other procedures to see whether the treatment group and control groups differ from each other significantly. But Imai et al. (2007) have pointed out that such a procedure would be highly misleading and should never be used to assess balance. The reason for this can be summarized the following way. First, balance is about the sample, not about the population. Second, hypothesis tests measure not only balance, but often statistical power as well. As a result, this dissertation uses descriptive statistics instead of hypothesis tests.

Furthermore, one-dimensional measures can also be used to capture the balance between the treatment and control groups. More specifically, an improvement in balance may also be captured by displaying a histogram of the propensity score (See Equations 21 and 22) before and after matching in the treatment and control groups. Figures 16, 17, 18, 19, 20, 21, 22, and 23 serve to illustrate this concept for different treatment conditions in the empirical research part of this dissertation. Further, Figure 13 illustrates a histogram for the genetic matching conducted on the sample dataset. In this figure raw data refer to the initial dataset before matching and matched data refer to the one after matching. Although in this dataset the number of observations is too small to see considerable differences, even here matching is able to increase the overlap between the treatment group and control groups to some extent. There is a significantly higher number of units which have a propensity score of between 0.0 and 0.1 (which signifies very low probability of receiving the treatment) in the initial control group than in the initial treatment group. Matching in this case removes these control units.

[18] In these tables, distance refers to logit propensity scores (in the case of genetic matching this occurs by default).

Figure 13. Propensity score distribution before and after matching on the sample dataset



3.7.5. Analysis of the Treatment Effect on the Matched Data

A matching method in itself can only ensure the appropriate balance between the treatment and control groups; however, it is incapable of estimating causal relationships. Therefore, a t-test or a regression analysis must be conducted for that purpose. DuGoff et al. (2014) argue that matching should be followed by multivariate regression analysis that involves the control variables that are used in matching to further improve the balance between the treatment and control groups. As a result, the research for this dissertation involved running multivariate OLS regression (and its extension to longitudinal design) after matching, controlling for the same variables which were used during matching.

Finally, an important issue is how to combine matching weights with sampling weights when estimating the treatment effect. Sample weights play an important role in the present research as it seeks to specify population-level inference. In the first and second empirical studies, sampling weights are especially important due to the longitudinal design (especially in the second one, which deals with older people who have a higher chance of dropping out during the observation period due to mortality), whereas the sampling weights are also necessary for the third study due to stratified sample design. In this later sample, certain life-stage groups were over- or underrepresented in the data (i.e. single parents were overrepresented) to achieve a sufficient sample size in each life-stage group. Thus, one needs to apply sampling weights to estimate the treatment effect at the population level. DuGoff et al. (2014) suggest that the original sampling weights should also be involved in the matching process if we are to reach conclusions pertaining to the entire population. They also advise creating a new weight variable for the purpose of the regression estimation, generated as the product of the sampling weight and the matching weight. The empirical studies in this dissertation consequently use weights.

3.7.6. Extension of Matching to Longitudinal Design

Longitudinal datasets are widely used for estimating causal inference as they represent a powerful tool for capturing missing potential outcomes, as described in Chapter 3.2. In my dissertation two empirical studies applied the matching method on a longitudinal dataset; more specifically, the first and second empirical studies. First, in this subchapter I briefly discuss how longitudinal analysis can be used to estimate the treatment effect. After that, I review the literature about the extension of the matching method to longitudinal data. This subchapter is mostly based on Allison (1990), Athey and Imbens (2006), Balbo and Arpino (2016), Chabé-Ferret (2015), and Imbens and Wooldridge (2009) (I refer to these authors if not otherwise indicated).

The longitudinal data include the pre-post treatment setting, thus each variable is measured both before and after treatment. In the present research, the treatment (parenthood in the first empirical study, and retirement in the second) occurs always between the first wave ($t=1$) and the second wave ($t=2$). The treatment effect is observed between the first wave ($t=1$) and the second wave ($t=2$) to measure the short-term change, and between the first wave ($t=1$) and fourth wave ($t=4$) to measure the long-term change. For the sake of simplicity, I only demonstrate the longitudinal data analysis for the short-term effect, although the method is applicable for analysing the changes between the first and fourth wave.

In order to estimate the γ treatment effect (i.e. ATE), one could apply a *Difference-in-Difference estimator* (DiD estimator). This method calculates the difference between the control and treatment groups over time. Let Y_{ij}^{t} denote i individual's subjective well-being at time t who has received j treatment between times 1 and 2. In order to obtain a DiD estimator, one can extend the equation for linear regression (Equation 17) with a t time dimension. In the first period, no one has received the treatment yet $\left(i \notin J \middle| t=1 \;\; for\, \forall i\right)$; however, in the second period the treated individuals receive the treatment. Consequently, the regression can be formalized for times 1 and time 2 respectively:

$$Y_i^1 = \alpha + \beta \times X_i^1 + \delta \times A_i + \varepsilon_i^1 \tag{37}$$

$$Y_i^2 = \alpha + \mu + \beta \times X_i^2 + \delta \times A_i + \gamma_i \times J_i + \varepsilon_i^2 \ , \tag{38}$$

where α is the intercept in time 1, μ is the difference in the intercept between time 2 and 1, X_i^t are the time-varying covariates and A_i the time-invariant covariates. Differencing Equations 37 and Equation 38 removes the A_i time-invariant covariates, but X_i^t time-varying confounding variables still remain in the equation. Thus, these variables need to be controlled for when estimating the treatment effect. More formally, the γ_{DID} treatment effect can be calculated in the following way:

$$Y_i^2 - Y_i^1 = \mu + \beta \times (X_i^2 - X_i^1) + \gamma_{DID} \times J_i + (\varepsilon_i^2 - \varepsilon_i^1) \tag{39}$$

DiD indeed identifies the causal effect when certain assumptions hold. First and foremost, this method posits the same assumptions as an OLS regression. Furthermore, the identification of the treatment effect by DiD requires that the *common-trend assumption* also holds. This assumption is that the difference over time in the potential non-treatment outcome should be conditionally independent from the treatment group membership given X covariates (Lechner, 2011). More formally, this assumption states:

$$\begin{aligned} & E(Y_0^2 \mid X = x, J = 1) - E(Y_0^1 \mid X = x, J = 1) = E(Y_0^2 \mid X = x, J = 0) - E(Y_0^1 \mid X = x, J = \\ & = E(Y_0^2 \mid X = x) - E(Y_0^1 \mid X = x) \quad \forall x \in X \end{aligned} \tag{40}$$

In the case of random treatment assignment, one does not necessarily need to control for time-varying confounding variables, thus any differentiation in the X_i^t covariates can be also ignored, giving the following equation:

$$Y_i^2 - Y_i^1 = \mu + \gamma_{DID} \times J_i + (\varepsilon_i^2 - \varepsilon_i^1) \tag{41}$$

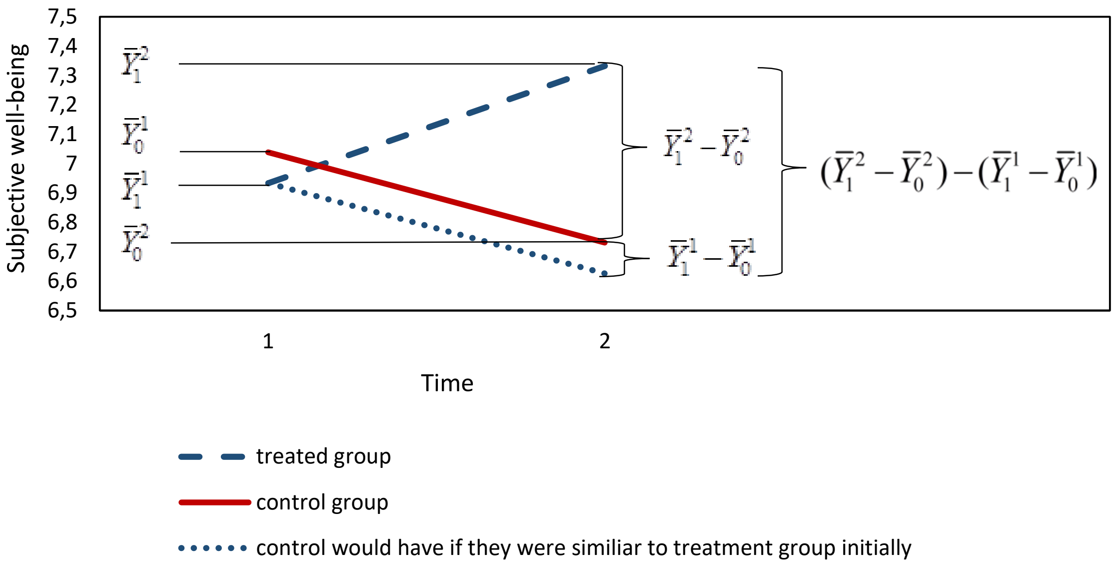
Equation 41 can be written in the following way as well:

$$\gamma_{DID} = (\overline{Y_1^2} - \overline{Y_0^2}) - (\overline{Y_1^1} - \overline{Y_0^1}) = (\overline{Y_1^2} - \overline{Y_1^1}) - (\overline{Y_0^2} - \overline{Y_0^1}) \, , \tag{42}$$

where $\bar{Y}_j^t$ is the average subjective well-being of the J treatment group at time t. Consequently, $\bar{Y}_1^1$ denotes the average subjective well-being measured in the first wave for the treatment group and $\bar{Y}_0^1$ for the control group. Furthermore, let $\bar{Y}_1^2$ denote the average subjective well-being measured in the second wave for the treatment group and $\bar{Y}_0^2$ for the control group. In this case, $\bar{Y}_1^1 - \bar{Y}_0^1$ indicates the initial difference between the treatment and control group, whereas $\bar{Y}_1^2 - \bar{Y}_0^2$ indicates the difference between the treatment group and the control group after exposure to treatment. After differentiating between treatment groups, one needs to differentiate over time as well, which gives the treatment effect (second part of Equation 42). Similarly, one can differentiate first over time and then by treatment group (third part of Equation 42). The "double difference" in this dissertation would mean that the life stage effect is given as the difference between those who belong in the second wave to the observed life-stage group and those who do not, over time.

Figure 14 illustrates the DiD approach in the case of random treatment assignment on the sample dataset. First, assume that the J_i treatment variable is random, and thus we do not need to control for any X_i^t covariates, thus Equation 41 and 42 can be used to estimate the treatment effect. In Figure 14 the dashed blue line refers to the change in the outcome among the treatment group, and the solid red line refers to the trajectory of the outcome among the control group. Based on the trends in the control group and the initial starting point of the treatment group we can predict what would happen with the treatment group in the case of not receiving the treatment. For this estimation we rely on the *common-trend assumption* of DiD, which states that the slope of the regression line would be the same for the treatment group in the case of not receiving the treatment as it was in the control group. Thus, this unobserved regression line should be perceived as a line parallel to the control regression line (red solid line) from the $\bar{Y}_1^1$ initial state of the treatment group. The dotted blue line represents the counterfactual changes in the treatment group that would occur if this group did not receive the treatment. The treatment effect is the difference between the second wave states of the actually observed treatment outcome (dashed blue line) and the sample (dotted blue line) treatment outcome.

Figure 14. Difference-in-Difference on the sample dataset (assuming random treatment assignment)



This figure also shows how to calculate the treatment effect based on Equation 42 on the right side of the graph. Under random assignment conditions, the actual treatment effect[19] on the sample dataset would be calculated by DiD in the following way:

$$\gamma_{DID} = (\overline{Y_1^2} - \overline{Y_0^2}) - (\overline{Y_1^1} - \overline{Y_0^1}) = (7.3 - 6.7) - (6.9 - 7.0) = 0.7 \qquad (43)$$

Another approach, which is similar to DiD is the *regressor variable method* (Allison, 1990). This method involves a Y_i^1 lagged outcome variable as predictor, which can be formulated in the following way[20]:

$$Y_i^2 - Y_i^1 = \mu + \beta_1 \times Y_i^1 + \beta_2 \times X_i^1 + \gamma_{regressor} \times J_i + \Delta\varepsilon_i \qquad (44)$$

This method assumes unconfoundedness given the lagged value of the outcome variable (Y_i^1) and the lagged values of the predictor variables (X_i^1). In other words, it assumes

[19] However, this example does not actually show the treatment effect as J_i treatment variable was not actually assigned randomly, thus, even though this method rules out time invariant confounders it does not control for time-variant confounders.

[20] This method is equivalent to regressing Y_i^2 on Y_i^1 instead of regressing $Y_i^2 - Y_i^1$ on Y_i^1. These two methods give the same estimate for the treatment effect (Allison, 1990; Baltagi, 2014). However, the research in this dissertation employed the later version as it is more widely used in the literature (Balbo & Arpino, 2016; Imbens & Wooldridge, 2009).

that the treatment group membership is conditionally independent of the individual's Y_i^2 response at time 2, given X_i^1 and Y_i^1 measured at time 1.

$$Y_i^2 \perp J_i \mid X_i^1, Y_i^1 \tag{45}$$

As I have shown before, one of the main advantages of the DiD approach is that it is able to rule out A_i time-invariant confounders once the common-trend assumption holds. The regressor variable method also controls for A_i time-invariant confounding if the unconfoundedness assumption holds (Equation 45). By including Y_i^1 variable as a predictor, A_i time-invariant confounders are controlled for under the unconfoundedness assumption, because the effect of A_i time-invariant confounders on Y_i^2 is already manifested in Y_i^1. For example, let us assume that we wish to observe the parenthood effect (J) on subjective well-being (Y). As in Chapter 3.1, we recognise that the common cause of both the treatment and the outcome should be ruled out. One might argue that personality traits might be such a common cause which would cause selection bias in the estimate. Personality traits are often considered to be time-invariant (Le Moglie, Mencarini, & Rapallini, 2015). This implies that personality affects subjective well-being in the same way every time. Thus, controlling for the lagged value of subjective well-being controls for the effect of personality traits at a later period as well.

Both DiD and the regressor variable method are powerful tools for estimating causality once longitudinal data is available;[21] however, these two methods substantively differ from each other in terms of their assumptions. The DiD approach requires the common-trend assumption (Equation 40). This assumption also implies that controlling for the Y_i^1 lagged outcome would create obstacles to the comparison as this lagged value might be correlated with the ε_i error term. In contrast, the regressor variable methods make the assumption of unconfoundedness (Equation 45). Imbens and Wooldridge (2009) generally favour the regressor variable method which relies on the unconfoundedness

[21] Alternatively, one could use fixed or random effect models as well but these models require two or more pre-treatment periods. Technically I could have used these methods, since I have three waves (thus I could have observed the treatment after two pre-treatment periods). However, I was interested in the short- and long-term changes that occurred after the treatment, thus I only used one period before the exposure to the treatment, and two afterwards. This setting is not suitable for the application of fixed or random effect models.

assumption. "*As a practical matter, the DID approach appears less attractive than the unconfoundedness based approach in the context of panel data. It is difficult to see how making treated and control units comparable on lagged outcomes will make the causal interpretation of their difference less credible, as suggested by the DID assumptions*." (Imbens & Wooldridge, 2009: 68) In contrast, others have argued that the DiD approach might work better in certain cases (Allison, 1990; Chabé-Ferret, 2015). The present research used the regressor variable method in line with Imbens and Wooldridge's (2009) suggestion.

Both the DiD and regressor variable method can be applied together with the matching method, which enable us to combine the advantage of both. These combinations improve the robustness of the matching method through eliminating possible time-invariant unobservable variables, but maintain the advantages of the matching method too. They can be combined with matching by first running matching on a certain set of control variables, and then applying the given longitudinal method to the same set of variables. The combination of matching and DiD approach is called Difference In Difference Matching (Chabé-Ferret, 2015), but this was not used in this dissertation. This dissertation combines matching with the regressor variable method.

In the case of a combination of the regressor variable method and matching, one conducts matching on the lagged predictor variables and the lagged outcome variable. Arpino and Aassve (2013) suggest that individuals in the treatment group should be matched with individuals in the control group that have similar values (including the outcome variable itself) before the treatment, and changes in outcomes should be compared. The first and second empirical studies employed this method. More specifically, in the present research individuals who belonged to a life stage (e.g. had a child, or retired) between the two waves were matched with individuals who did not belong to this given life stage at the end of the observational period, but who had had similar properties initially, including subjective well-being.

To sum up, a longitudinal design is suitable for eliminating selection bias (See Chapter 3.1.) by controlling for confounding variables to a higher degree than cross-sectional data would allow. Table 10 summarizes what types of confounding variables matching combined with longitudinal analysis controls for. It rules out all observed covariates and time-invariant unobserved confounders. However, it should be emphasized here that even this method cannot control for time-variant unobserved covariates (See Table 10). The next chapter introduces the sensitivity analysis that can be conducted to estimate the sensitivity of the results to these variables.

Table 10. What types of variables does matching on longitudinal data control for?

	Time-variant	Time-invariant
Observed confounding variable	Controlled	Controlled
Unobserved confounding variable	Not controlled	Controlled

Another crucial advantage of longitudinal design is that it can help to avoid endogenous selection bias (See Chapter 3.1.) which arises from controlling for collider variables. These collider variables are the common outcomes of the treatment and the outcome variable, thus they are mostly post-treatment variables. Therefore, controlling for only pre-treatment variables is considered to be the best way of avoiding having to control for collider variables (Elwert & Winship, 2014; Rosenbaum, 1984). A longitudinal design enables us to use only those X_i^1 covariates which are measured before exposure to the treatment; therefore, it is suitable for eliminating endogenous selection bias. Consequently, the research in this dissertation also controls for only the pre-treatment variables in Chapter 4 and 5 when longitudinal design was available.

3.7.7. Sensitivity Analysis

Given that the estimations of causality depend on the validity of the unconfoundedness assumption (Equations 13 and 45), capturing the sensitivity of this assumption is clearly important. This dissertation employs the Rosenbaum (2002) sensitivity analysis to assess to what degree the results are sensitive to a given quantifiable increase in uncertainty.

Let us assume again that every i individual either receives the treatment or does not $(i \in K \; or \; i \in L)$. K contains those k individuals who actually received the treatment $(j=1)$ and L contains those l individuals who did not receive the treatment $(j=0)$. Further, $m(l)$ individuals are those control individuals who were matched to a treated individual. Let $P(J=1)_i$ be the probability that i individual receives the treatment.

If there were no hidden bias in the study, than $P(J=1)_i$ would be the function of the observed covariates X_i for all i individuals. In contrast, there is a hidden bias if a k treated individual and a $m(l)$ control individual have the same covariates $X_k = X_{m(l)}$, but they have a different chance of receiving the treatment $P(J=1)_k \neq P(J=1)_{m(l)}$. The odds that k treated individual will receive the treatment is $P(J=1)_k / P(J=0)_k$ and the odds that $m(l)$ control individual who is matched to k treated individual will also receive the treatment would be $P(J=1)_{m(l)} / P(J=0)_{m(l)}$. Rosenbaum (2002) proved that the odds ratio for the odds for k treated individuals and odds for $m(l)$ matched control individuals is bounded by the Γ parameter and the reciprocal of the Γ parameter:

$$\frac{1}{\Gamma} \leq \frac{P(J=1)_k / P(J=0)_k}{P(J=1)_{m(l)} / P(J=0)_{m(l)}} \leq \Gamma, \tag{46}$$

for all k and $m(l)$.

The test relies on this Γ parameter that assumes a certain degree of departure from the unconfoundedness assumption; that is, from the random assignment of the treatment given the controlled covariates. If $\Gamma = 1$, then for every k treated individual and $m(l)$ matched control individuals with the same covariates $X_k = X_{m(l)}$ would have an equal chance of receiving the treatment $P(J=1)_k = P(J=1)_{m(l)}$; in other words, the study would be free of bias. If $\Gamma = 2$, then given k treated individual and $m(l)$ matched control individual with the same covariates $X_k = X_{m(l)}$, one could be twice as likely as the other to receive the treatment.

During the sensitivity analysis one observes how much the inferences about the treatment effects can be altered by hidden biases with various Γ parameters. If the results remain significant even for a high value of Γ, then there is a robust treatment effect even if the unconfoundedness assumption (Equations 13 and 45) do not stand entirely and some confounders were not controlled for. This dissertation reports the critical value of Γ parameters using a 90% confidence level. There is no straightforward and reliable critical Γ value which should be considered statistically valid, but DiPrete and Gangl (2004) suggest that a value of 1.5 or over should be considered a robust result in the social sciences. Sensitivity analysis was conducted using the *rbound* package, which runs in the R environment (Keele 2010).

4. Estimating the Effect of Parenthood on Subjective Well-being

4.1. Introduction

Developed countries have for decades been experiencing below-replacement level fertility. This situation has caught the attention of scholars and decision makers alike due to its implications for population ageing and associated costs. Researchers have argued that one of the reasons for this low fertility is that potential parents do not perceive that having children will sufficiently increase their subjective well-being (Aassve, Arpino, & Balbo, 2016; Billari, 2009; Le Moglie et al., 2015; Luppi, 2016; Margolis & Myrskylä, 2015; Parr, 2010). Consequently, a growing number of scientific papers have investigated whether having children actually leads to a decrease in subjective well-being. So far, most longitudinal evidence has come from western countries, finding that parenthood in general has a positive effect on subjective well-being (Balbo & Arpino, 2016; Baranowska & Matysiak, 2011; Kohler, Behrman, & Skytthe, 2005; Mikucka, 2016; Le Mogliea et al., 2015; Pollmann-Schult, 2014). However, inconsistencies remain regarding how the effect of having children changes in specific circumstances; for example, when children grow older, and according to the parity and gender of the parents (Angeles, 2010; Baetschmann et al., 2016; Balbo & Arpino, 2016; Baranowska & Matysiak, 2011; Clark & Georgellis, 2013; Frijters & Beatton, 2012; Keizer, Dykstra, & Poortman, 2010; Kohler et al., 2005; Mikucka, 2016; Myrskylä & Margolis, 2014; Pollmann-Schult, 2014).

The present chapter contributes to the current debate by examining this issue in the Hungarian context. Although the topic has received significant attention in the West, in Central-Eastern Europe – the area in which fertility is the "lowest-low" (Kohler, Billari, & Ortega, 2002) – only limited research has been undertaken (Baranowska & Matysiak, 2011; Sironi & Billari, 2013). However, in the affected region fertility decisions are embedded in a very different economic, cultural, and social context than in Western countries. Firstly, a lower standard of living in CEE countries may limit individuals' options in their quest for happiness (Szalai, 1991). Secondly, the low level of childlessness in the CEE region indicates that the above-average fertility rate is mainly attributable to a low level of second births (Miettinen & Szalma, 2014; Szalma & Takács, 2015). Finally, individuals in CEE countries have historically had a significantly lower level of life satisfaction than those in Western European countries (Guriev & Zhuravskaya, 2009). Within CEE countries, this dissertation focuses on Hungary, which is an especially interesting case since the relatively high level of child poverty, persistently low fertility rate, and low level of subjective well-being is exceptional, exceeding that of most neighbouring countries (Guriev & Zhuravskaya, 2009; Spéder & Kapitány, 2014; Szalma & Takács, 2015). Finally, in this country there has been no research to date about the effect of parenthood on subjective well-being using longitudinal data.

The hypotheses that will be tested are based on a set of four theories. First, the *value of children theory* has emphasized the positive side of having a child (Hoffmann & Hoffmann, 1973; Nauck, 2007). Second, others have claimed that parenthood is also associated with enormous *costs* (Lavee, Sharlin, & Katz, 1996; Stanca, 2012; Twenge, Campbell, & Foster, 2003; Zimmermann & Easterlin, 2006). Third, the *demand and reward theory* emphasizes that the positive and the negative effects of having a child offset each other. This latter theory also postulates that the rewards of parenthood decline with the ageing of the child; consequently, parenthood should have also a declining effect on subjective well-being (Hansen, 2012; Nomaguchi, 2012; Nomaguchi & Milkie, 2003; Umberson, Pudrovska, & Reczek, 2010). Finally, *set-point theory* argues that the effect of parenthood is only temporary, thus subjective well-being eventually returns to its pre-birth baseline level (Headey & Wearing, 1989; Kammann, 1983; Lykken & Tellegen, 1996).

The hypotheses are tested by using state-of-the-art methods. Investigation of the causal relationship between parenthood and subjective well-being poses statistical challenges, since only observational data are available. The present study applies genetic matching using longitudinal data. The advantage of the technique is due to a combination of matching and longitudinal analysis, which enables us to control for both time-invariant and observed time-variant confounders. The method of matching using longitudinal data has been used in the international literature (Baetschmann et al., 2016; Balbo & Arpino, 2016; Sironi & Billari, 2013) to analyse the effects of fertility; however, this method has never been applied to examine the relationship between fertility and subjective well-being in Hungary.

Overall, this research finds that parenthood has a positive, long-lasting effect on subjective well-being. Moreover, not only a first child but also a second child appears to increase subjective well-being. This finding raises the question why the total fertility rate has persistently remained less than two in Hungary. One answer concerns the moderating effect of gender. The research finds that both women and men benefit from having a child in the short term, but in the long run fatherhood has no positive significant effect on subjective well-being. This finding is crucial as childbearing decisions typically involve shared decision making, thus both genders should experience an increase in subjective well-being for parity progression (Aassve et al., 2016). However, the other findings described in this chapter indicate that subjective well-being is probably not the main driver of the low fertility rate in Hungary.

4.2. Background

4.2.1. General Effects of Parenthood

Parenthood has complex consequences on subjective well-being, as this life event is both simultaneously rewarding and stressful. It is not surprising, therefore, that various – sometimes conflicting – theories have emerged about the topic. This study reviews value of children theory, the cost of children approach, demand and reward theory, and set-point theory.

First of all, *value of children theory* postulates that parenthood has a positive effect on subjective well-being. This theory argues that children fulfil different parental needs. Hoffmann and Hoffmann (1973) suggested several ways in which children can modify parental satisfaction, such as by strengthening primary group ties, providing entertainment, expanding the sense of self, creating a social identity and a sense of achievement, providing economic utility, and generating an advantage in terms of social comparison. Furthermore, this theory also claims that parenthood has a persistent effect as ageing children fulfil different types of needs throughout their entire lives (Nauck, 2007).

Others have emphasized that parenthood can have a negative effect on subjective well-being by drawing attention to the *cost of children*. Hansen (2012) distinguishes between psychological cost, marital cost, financial cost, and opportunity cost in relation to childbearing. The concept of psychological cost recognizes that fertility increases financial stress (Stanca, 2012; Zimmermann & Easterlin, 2006), worsens work-family balance (Craig, 2016; Kimmel & Connelly, 2007), and reduces personal freedom (Twenge et al., 2003). Marital cost refers to child-related decreases in the quality of a partnership (Lavee et al., 1996). Child also come at a substantial financial cost (Evertsson, 2016; Reizer, 2011). Finally, parents face opportunity costs which affect their careers, education, and leisure time (Sanchez & Thomson, 1997). Further, the cost of children approach assumes that subjective well-being increases over time due to a decrease in costs. Based on this theory, childbearing-related costs generally decrease as children grow up, since parents with younger children are more likely to be exposed to financial problems (Nelson, Kushlev, & Lyubomirsky, 2014), an overwhelming amount of housework (Nomaguchi & Milkie, 2003), intense work-family conflict (Nomaguchi, 2009), marital conflict (Nelson et al., 2014; Nomaguchi & Milkie, 2003), sleep disturbance (Nelson et al., 2014), a shortage of leisure time (Claxton & Perry-Jenkins, 2008), and a loss of networking time (Munch, McPherson, & Smith-Lovin, 1997).

Further, *demand and reward theory* argues that both the positive and the negative effects of having children should be taken into account as these effects offset each other (Nomaguchi, 2012; Nomaguchi & Milkie, 2003; Umberson et al., 2010). In general, this approach emphasizes that the benefits of having a child, or more specifically, relationship satisfaction with a child, is the key to understanding changes in the effects of parenthood. Nomaguchi (2012) found that the rewards that parents obtain from their children decline as the child grows older. The author claims that emotional benefits are greatest when the child is less than five years old. Thus, this theory assumes that the effect of parenthood should decline as a child grows up (Nomaguchi, 2009, 2012; Nomaguchi & Milkie, 2003).

Furthermore, *set-point theory* also claims that this transition has a declining effect on subjective well-being. However, this theory is more restrictive than demand and reward theory as it also specifies that this effect not only declines but entirely disappears in the long run. This theory claims that people have a stable baseline level of subjective well-being which is determined by personality traits and other genetic factors (Headey & Wearing, 1989; Kammann, 1983; Lykken & Tellegen, 1996). According to this theory, after major life events, such as having a child, individuals eventually adopt to their new situation and their subjective well-being returns to the initial baseline level (Myers, 1999). The theory, thus, posits that parenthood only temporarily alters subjective well-being, but does not have a significant effect in the long run. Although this theory has been influential and supported by several research efforts (Headey & Wearing, 1989), other studies have recently questioned it by showing that adaptation to life events is not universal (Diener, Lucas, & Scollon, 2006; Headey, 2006).

Up-to-date empirical evidence about the causal relationships between parenthood in general and subjective well-being is mostly restricted to observations from western countries where longitudinal data were available. The majority of the state-of-the-art research on longitudinal data has found a positive effect for the average person (Balbo & Arpino, 2016; Baranowska & Matysiak, 2011; Kohler et al., 2005; Mikucka, 2016; Le Mogliea et al., 2015; Pollmann-Schult, 2014), while only limited evidence has shown an insignificant effect (Angeles, 2010; Keizer et al., 2010). In contrast, cross-sectional studies, which often cover non-western countries, produced mixed evidence. The majority of cross-sectional research finds a negative link between parenthood and subjective well-being (Hansen, 2012; McLanahan & Adams, 1987; Stanca, 2012; Vanassche, Swicegood, & Matthijs; 2013), although some has found a positive link (Aassve, Goisis, & Sironi, 2012).

State-of-the-art longitudinal research projects are able to observe how the effect of parenthood changes over time. All of the reviewed papers found that subjective well-being indeed declines with the ageing of the child, which finding is in line with the demand and reward approach. Further, and consistent with set-point theory, some of the former has shown that subjective well-being not only declines but also returns to the pre-birth baseline. (Balbo & Arpino, 2016; Clark et al., 2008; Frijters & Beatton, 2012; Myrskylä & Margolis, 2014). Other empirical evidence is at odds with the set-point theory and shows that this effect remains significant even in the long run (Baetschmann et al., 2016; Mikucka, 2016; Pollmann-Schult, 2014).

Moreover, parenthood not only affects subjective well-being after the birth of a child, but before the birth too as parents prepare for the new arrival. The literature refers to this phenomenon as the *anticipation effect*, whose existence is supported by several studies (Baetschmann et al., 2016; Clark et al., 2008; Frijters & Beatton, 2012; Mikucka, 2016; Myrskylä & Margolis, 2014). Although there is consensus about the existence of such an effect, there is mixed empirical evidence about how much time before the birth this phenomenon occurs. Most studies have found that the impact occurs in the year before childbirth (Baetschmann et al., 2016; Clark et al., 2008; Mikucka, 2016; Myrskylä & Margolis, 2014), although a piece of research identified the anticipation effect only 6-9 months before the birth (Frijters & Beatton, 2012), and other authors as far in time as 2–3 years before the birth, but only for females (Clark & Georgellis, 2013).

In Hungary no research project has so far evaluated the causal relationship between parenthood and subjective well-being on longitudinal data. However, there are some studies in which the impact of parenthood was not the main point of interest but where subjective well-being was regressed using a series of control variables, including the presence of children. Molnár and Kapitány (2013) found that individuals who are on parental leave have significantly higher subjective well-being than the rest of the population. However, Molnár and Kapitány (2006) previously also found that fertility is negatively associated with satisfaction with income.

Comparative cross-sectional research also often included Hungary in the study of how macro-level factors mediate the relationship between parenthood and subjective well-being. However, the goal of these studies was not to determine causality, and they often arrived at different conclusions. Aassve et al. (2012) found that in richer Western countries there exists a significant positive effect between parenthood and subjective well-being, but in Eastern European countries the relationship is rather negative. In contrast, Billari (2009) concluded in his descriptive analysis that even in Eastern European countries, including Hungary, there is a positive effect. Finally, Margolis and Myrskylä (2011) found that it is especially in the former socialist countries that a higher number of children can be associated with higher satisfaction.

4.2.2. Parity Specific Parenthood Effects

Most research focuses on examining the effect of a first child on subjective well-being, although subsequent children might have different effects than the first. First, *set-point theory* recognizes parity differences in the parenthood effect on subjective well-being. This theory assumes that the transition to parenthood associated with the first child is more novel than the birth of the subsequent children. Thus, the effect of parenthood is stronger with the first child, and weaker effects on well-being are associated with higher-order children (Aassve et al., 2012; Mikucka, 2016).

Demand and reward theory posits the moderating effect of parity based on a comparison of the marginal utility and marginal cost of having a child. First, empirical results show that the law of diminishing marginal utility also apply to the *value of children*, thus, the reward of having a child decreases with each additional child (Nauck, 2007; Nomaguchi, 2012). Second, there is inconsistency in the empirical findings about the *marginal cost of children*. On the one hand, some studies emphasize that the marginal *cost of children* diminishes – for example, marginal financial costs can decrease since items can be shared between siblings (such as furniture or clothes) (Thévenon, 2010). On the other hand, other research finds that marginal costs actually increase with each additional child - some studies have shown that households with more children are exposed to a proportionately greater risk of poverty and higher opportunity costs than families with fewer children (Troske & Voicu, 2009). Considering marginal utility and marginal costs together, several authors have argued that additional children are less ‘beneficial’, thus, elicit less subjective well-being (Aassve et al., 2012; Mikucka, 2016).

The vast majority of the state-of-the-art longitudinal studies have found that first children have a positive effect on subjective well-being (Baetschmann et al., 2016; Balbo &

Arpino, 2016; Baranowska & Matysiak, 2011; Clark et al., 2008; Clark & Georgellis, 2013; Frijters, Johnston, & Shields, 2011; Kohler et al., 2005; Matysiak, Mencarini, & Vignoli, 2016; Mikucka, 2016; Myrskylä & Margolis, 2014), but a few research projects have not supported this finding (Angeles, 2010; Pedersen & Schmidt, 2014). Further, several empirical studies have supported set-point theory and demand and reward theory by finding that the effect of parenthood gradually decreases for higher-order children. These studies have found that second children have a non-significant (Angeles, 2010; Baranowska & Matysiak, 2011; Kohler, et al., 2005; Pollmann-Schult, 2014) or only a temporary effect (Balbo & Arpino, 2016; Matysiak et al., 2016; Myrskylä & Margolis, 2014). However, Mikucka (2016) found that second children have a strong and lasting effect; furthermore, first children elicit less subjective well-being than second children for women.

4.2.3. Gender Specific Parenthood Effects

Men and women experience the transition to parenthood differently as they face distinct child-related opportunities and restrictions. However, the moderating effect of gender is especially important as couples generally make fertility decisions together (Bauer & Kneip, 2012). Thus, if only one of the genders derives well-being from parenthood, this may create obstacles to parity progression. For example, Aassve et al. (2016) found, based on British data, a multiplicative effect for partners' subjective well-being on parity progression. Further, they found that females' subjective well-being matters more in decisions about the first child, whereas males' subjective well-being has more influence over decision-making about higher-order children.

Various theories have been developed to understand how gender differences are reflected in the parenthood effect on subjective well-being. Firstly, the *value of children theory* argues that, on average, females benefit more from having children than males since women more often claim that children strengthen primary group ties, provide fun, expand the self, and create social identity. However, this theory also claims that parenthood is rewarding for both genders, since both parents are more liable to claim than childless people that children provide fun, bring love and companionship, and expand the self (Hoffman & Hoffman, 1973; Hoffman, Thornton, & Manis, 1978).

In contrast to value of children theory, the *cost of children approach* emphasizes that women might be more heavily burdened during parenthood, indicating that women experience more negative effects on their subjective well-being after the arrival of children. During parenthood women tend to experience higher stress (Nomaguchi &

Milkie, 2003), greater work-life conflict (Goldsteen & Ross, 1989), more household and parenthood duties (Baxter, Hewitt, & Haynes, 2008; Bianchi, Robinson, & Milke, 2006; Bird, 1999), worse career opportunities (Bianchi, 2000), and have less leisure time (Craig & Mullan, 2013; Mattingly & Bianchi, 2003). Moreover, several studies have emphasized the gender inequality in childrearing activities. Men tend to engage in more enjoyable activities, such as playing, whereas women undertake more time-inflexible, day-to-day care, or management activities (Musick, Meier, & Flood, 2016; Raley, Bianchi, & Wang, 2012; Sayer, Bianchi, & Robinson, 2004). However, in general financial-related stress is more pronounced for men than women as males experience more pressure to provide for their families (Pollmann-Schult, 2014).

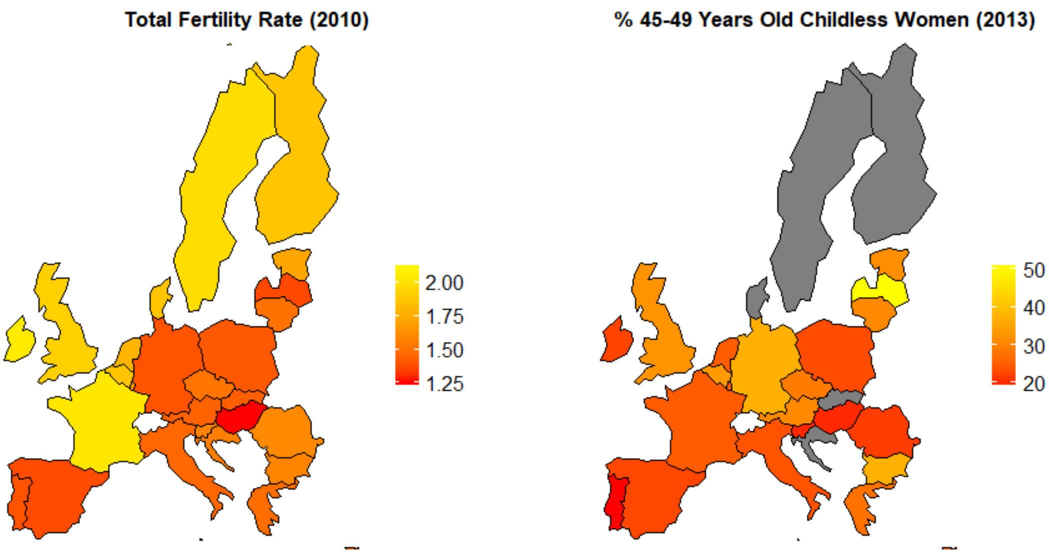
In contrast to these theories, the cost of children approach emphasizes that women might be more heavily burdened during parenthood, indicating that women experience more negative effects on their subjective well-being after the arrival of children. During parenthood women tend to experience higher stress (Nomaguchi & Milkie, 2003), greater work-life conflict (Goldsteen & Ross, 1989), more household and parenthood duties (Baxter et al., 2008; Bianchi et a., 2006; Bird, 1999), worse career opportunities (Bianchi, 2000), and have less leisure time (Craig & Mullan, 2013; Mattingly & Blanchi, 2003). Moreover, several studies have emphasized the gender inequality in childbearing activities. Men tend to engage in more enjoyable activities, such as playing, whereas women undertake more time-inflexible, day-to-day care, or management activities (Musick et al., 2016; Raley et al., 2012; Sayer et al., 2004; Yeung et al., 2001). However, in general financial-related stress is more pronounced for men than women as males experience more pressure to provide for their families (Pollmann-Schult, 2014).

Empirical findings are also inconsistent about the moderating effect of gender. First, most studies have found that women have higher subjective well-being after having a child than men (Angeles, 2010; Baranowska & Matysiak, 2011; Baetschmann, et al., 2016; Clark & Georgellis, 2013; Kohler et al., 2005; Myrskylä & Margolis, 2014; Sironi & Billari, 2013). Furthermore, some of these studies have found that fatherhood has non-significant (Sironi & Billari, 2013) or only a temporary effect (Baranowska & Matysiak, 2011). Second, other research has found that both genders equally benefit from having a child (Pollmann-Schult, 2014). Finally, some studies have even supported the idea that men benefit more from having children in terms of subjective well-being (Aassve, et al., 2012; Balbo & Arpino, 2016; Nelson, Kushlev, English, Dunn, & Lyubomirsky, 2013), while some research claims that only fathers benefit significantly from parenthood (mothers do not) (Aassve et al., 2012; Nelson, et al., 2013).

4.2.4. The Hungarian Context

The focus of this dissertation is Hungary, where the fertility rate has been decreasing for 40 years, births per women have remained under 1.5 for 25 years, and between 2009 and 2011 this country had the lowest fertility rate among all EU countries (See the fertility rate in the European Union in 2010 on the left side of Figure 15). There are multiple reasons for the relatively low fertility level in Hungary, including poor economic conditions, the dismantling of institutions, value shifts, and social anomie (Spéder & Kapitány, 2014). Further, empirical studies suggest that the low number of children can mostly be attributed to the fact that second children are not being born, whereas childlessness still plays a relatively minor role in the low levels of fertility (Miettinen & Szalma, 2014; Szalma & Takács, 2015). Figure 15 (right-hand side) illustrates the low level of childlessness in Hungary, displaying the proportion of woman living without their own children in the European Union. Moreover, Hungarians generally have strong intentions of having children, since there is a considerable gap between the ideal and actual number of children (Molnár, 2009; Kapitány & Spéder, 2015). The low level of childlessness and high ideal number of children together indicate that in Hungary the value of children is still relatively high, despite the low fertility rate.

Figure 15. Fertility statistics in the European Union[22]



[22] The figure is the author's own work based on Ewen Gallic's blog (Access: http://egallic.fr/en/european-map-using-r/). It was produced in R. Source of fertility rate (left-hand side): Eurostat, Source of childlessness (right-hand side): OECD (http://www.oecd.org/els/family/database.htm#structure)

Institutional arrangements are able to influence the effect of fertility on subjective well-being to a large degree. Hungary is an exceptional case due to the generosity of the family support system. By law, parents are eligible to 24 weeks' leave at 70% of average pre-birth salary, with no ceiling on payments. Moreover, after this period parents receive flat-rate benefits until the child's third birthday (Makay, 2015). This is one of the longest terms of parental leave in Europe, and, as a consequence, its generosity leads one to expect that the short-term cost of children would be relatively low, and the short-term positive impact on subjective well-being relatively strong.

However, the long period of paid parental leave is paired with low opportunities for flexible work, such as part-time jobs, and poor access to child-care (Radó, Nagy, & Király, 2016). These factors encourage parents to exhaust the three-year period of parental leave, thereby increasing the opportunity cost of children in the long run. As it is usually females who take parental leave, they are penalized more (Aassve et al., 2012, Bartus, Murinkó, Szalma, & Szél, 2013). However, the high opportunity cost for women can also create spillover effects on males' subjective well-being, since in male-breadwinner households it is typically men who experience financial-related stress (Pollmann-Schult, 2014). Taken as a whole, the accumulation of costs of childbearing generated by the Hungarian social security system may also reflect on subjective well-being. One might expect that long-term effect of parenthood on subjective well-being would be relatively worse in Hungary than in western countries.

4.3. Hypotheses

This chapter employs three research questions, defined as the following: how does the effect of parenthood on subjective well-being change (1) with the ageing of the child, (2) by parity of children, and (3) by the gender of the parents? To investigate these questions, often competing hypotheses are formulated based on the state-of-the-art empirical evidence, the theoretical background, and the Hungarian context.

Hypothesis 1A. The parenthood effect decreases with the age of the child. This hypothesis is supported by all of the reviewed research projects which analysed longitudinal data. The decreasing effect is consistent with demand and reward theory and set-point theory. Further, the tendency for wellbeing to decrease over time is also consistent with Hungarian social policy context, which supports childbearing in the short run, but creates opportunity costs in the long run (Bartus et al., 2013; Makay, 2015).

Hypothesis 1B. Parenthood has a positive long-lasting effect on subjective well-

being, even in the long run. This expectation is consistent with some of the empirical evidence (Baetschmann et al., 2016; Mikucka, 2016; Pollmann-Schult, 2014). Further, it is in line with value of children theory.

Hypothesis 1C. Parenthood has only a temporary positive effect on subjective well-being or no effect at all. This hypothesis is also in line with much international evidence (Balbo & Arpino, 2016; Clark et al., 2008; Frijters & Beatton, 2012; Myrskylä & Margolis, 2014). Further, such a temporary effect is consistent with set-point theory.

Hypothesis 2A. The first child increases subjective well-being to a higher degree than the second child. This claim is in line with the majority of the state-of-the-art longitudinal studies (Balbo & Arpino, 2016; Baranowska & Matysiak, 2011; Kohler et al., 2005; Matysiak et al., 2016; Myrskylä & Margolis, 2014; Pollmann-Schult, 2014). Also, it supports set-point theory and demand and reward theory.

Hypothesis 2B. Second children either do not have an effect, or have only a temporary effect on subjective well-being. This hypothesis is based on findings from the majority of state-of-the-art international research projects (Balbo & Arpino, 2016; Baranowska & Matysiak, 2011; Matysiak et al., 2016; Myrskylä & Margolis, 2014; Kohler et al., 2005; Pollmann-Schult, 2014). Further, this expectation is consistent with set-point theory. This proposition explains the low incidence of second children in Hungary (Miettinen & Szalma, 2014; Szalma & Takács, 2015).

Hypothesis 2C. Second children elicit long-lasting positive changes in subjective well-being. Only limited research has supported this claim (Mikucka, 2016) although it is in line with previous studies which found that the value of children is high in Hungary (Molnár, 2009). Furthermore, there is evidence that second children are sometimes associated with lower marginal costs than first children (Thévenon, 2010), which could explain this expectation as well.

Hypothesis 3A. Parenthood has a long-lasting effect on both mothers and fathers. This expectation is based on some of the empirical evidence (Mikucka, 2016; Pollmann-Schult, 2014) and is in line with value of children theory, which emphasizes that parenthood can be rewarding for both genders.

Hypothesis 3B. Mothers experience only a temporary effect or no effect at all. This hypothesis is also in line with some of the empirical studies (Aassve et al., 2012; Balbo & Arpino, 2016; Clark & Georgellis, 2013; Keizer et al., 2010; Myrskylä & Margolis, 2014; Nelson et al., 2013). It is also consistent with set-point theory. Furthermore, Hungarian mothers face particularly high opportunity costs in the long run due to long parental leave (Aassve et al., 2012, Bartus et al., 2013).

Hypothesis 3C. Fathers experience only a temporary effect or no effect at all. This expectation is also supported by international research projects (Baranowska & Matysiak, 2011; Clark & Georgellis, 2013; Keizer et al., 2010; Myrskylä & Margolis, 2014; 2016; Sironi & Billari, 2013). Further, this hypothesis is also consistent with set-point theory. Moreover, it is also in line with nature of the Hungarian social support system, since the long period of parental leave in Hungary creates a higher level of stress for males as their partners are outside the labour force, and thus they become the main breadwinners (Pollmann-Schult, 2014).

4.4. Data

The empirical foundation of the first (Chapter 4) and second (Chapter 5) empirical studies in this dissertation was the *Turning Points of Life Course* survey (also known as *Hungarian Generations and Gender Survey*), a longitudinal piece of research carried out by the Hungarian Demographic Research Institute. The first wave of data for this undertaking was collected between November 2001 and January 2002 (hereafter, 2001/2002), the second between November 2004 and January 2005 (hereafter, 2004/2005), the third between November 2008 and January 2009 (hereafter, 2008/2009), and the fourth between November 2012 and January 2013 (hereafter, 2004/2005). The research described in Chapter 4 and 5 uses data from those waves in which subjective well-being was measured (the first, the second and the fourth).

The first wave was representative of Hungarian residents aged between 18-75 years, which was gradually extended with a sample to replace young people. Respectively, the last wave contained respondents between 18-86 years. Initially, the raw dataset (i.e. the dataset before matching) was not reduced by age or any other control variable (see more in the chapter on Analytical Strategy). However, the matched dataset which is used for estimating causality reduced the initial dataset to create a balance between the treatment and control groups.

Longitudinal data are never free of sample attrition. In the last wave, 8103 people were addressed, whereas twice as many (16363) participated in the first wave. The most frequent reason for dropping out from the study was refusal to participate. Between the first and the second wave 6% of the initial sample refused to answer, whereas this percentage reached 11% in the fourth wave. However, the major advantage of this research was that less than 8% of the initial sample dropped out due to moving to an unknown destination throughout the whole period of research. The high drop-out rate, just like the other missing data, might cause biased estimations. This problem was handled with longitudinal weighting (see more about the weighting of the given dataset in Bartus [2015]).

In the following subchapters, the applied variables are introduced. Concerning these variables, there are several similarities between this chapter and the next chapter. First, the majority of the matching variables are the same in these two studies. Second, in both cases matching variables and control variables coincide, therefore the same set of covariates are used for matching and estimating the treatment effect using the regressor variable method (See Equation 44). Finally, the outcome variables are also defined in the same way.

4.4.1. Treatment Variables

Table 11 summarizes the treatment variables in this study, which are (1) parenthood in general, (2) motherhood, (3) fatherhood, (4) having a first child, and (5) having a second child. This table also details the compositions of the treatment and control groups. In all cases, those whose child(ren) was (were) born between the first wave (2001/2002) and 2003 were omitted from the analysis to eliminate the anticipation effect.[23]

23 The claim to a one-year anticipation effect is in line with the vast majority of the literature (Baetschmann et al., 2016; Balbo & Arpino, 2016; Clark et al., 2008; Frijters & Beatton, 2012; Myrskylä & Margolis, 2014; Pollmann-Schult, 2014). To my knowledge, only one study has found that this impact appears 2-3 years before birth for women (Clark & Georgellis, 2013). Thus, I also tried to match individuals two years before childbirth, but this generated similar results to the scenario that I had obtained assuming a one-year anticipation effect. As a result, I narrowed down the anticipation effect to one year, which permits a higher number of observations than a longer anticipation period would.

Table 11. Description of the observed treatments in Chapter 4

Observed phenomenon	Treatment group	Control group
General effect of parenthood	Those whose child(ren) was (were) born between 2003 and the second wave (2004/05), but to whom no children were born between the first wave (2001/02) and 2003	Those to whom no children were born between the first (2001/02) and second wave (2004/05)
The effect of motherhood	Women whose child(ren) was (were) born between 2003 and the second wave (2004/05), but to whom no children were born between the first wave (2001/02) and 2003	Women to whom no children were born between the first (2001/02) and second wave (2004/05)
The effect of fatherhood	Men whose child(ren) was (were) born between 2003 and the second wave (2004/05), but to whom no children were born between the first wave (2001/02) and 2003	Men to whom no children were born between the first (2001/02) and second wave (2004/05)
The effect of the first child	Those who had their first child between 2003 and the second wave (2004/05)	Those who remained childless between 2003 and the second wave (2004/05)
The effect of the second child	Those who had their second child between 2003 and the second wave (2004/05), but to whom no children were born between the first wave (2001/02) and 2003	Those who had a child before the first wave (2001/02), but to whom no children were born between the first wave (2001/02) and 2003

For estimating the overall effect of parenthood, information about the birth of children was required. In the second wave of the research, respondents were asked to list details about all of their children, including year of birth. This question served to measure the treatment variable, which took a value of 1 if a respondent's child(ren) was (were) born between 2003 and the second wave (2004/05) and 0 if they did not.

Further, the parity effect was distinguished by reducing the initial dataset for (1) those who had had no children before the observation period, and (2) those who had had only one child before the observation period. Firstly, the dataset which contained those who were initially childless was used to distinguish the effect of the first child. Here, the treatment variable was awarded a value of 1 if the respondent had had a first child between 2003 and the time of the second wave (2004/2005), and a value of 0 if this person had remained childless. Secondly, the dataset including those who initially had one child was used to estimate the effect of the second child. In these cases, the treatment variable took a value of 1 if the respondent had had an additional child between 2003 and the time of the second wave (2004/2005), and a value of 0 if they had not.

Finally, the distinct effect of motherhood and fatherhood was observed by splitting the initial dataset between women and men. Afterwards, the same treatment variables which had been developed to measure the overall effects of parenthood were applied to these two subgroups.

4.4.2. Outcome Variables

In contrast to the cross-sectional data, *Turning Points of Life Course* research does not include domain-specific subjective well-being in detail. Thus, in this case I needed to focus only on overall subjective well-being. Subjective well-being was measured in both waves with the following question: "*On an eleven-point scale, how satisfied are you with the trajectory of your life?*" This variable used a value of 0 to mean 'not satisfied at all', and a value of 10 for 'completely satisfied'. The outcome variable was the change in subjective well-being before and after exposure to treatment. More specifically, the change in subjective well-being was calculated in terms of the change in life satisfaction between the first and second waves to measure short-term change[24], and between the first and fourth waves for estimating long-term change[25]. Subjective well-being was treated as an interval variable as other studies have found that it makes little difference treating this variable as an ordinal (Ferrer-i-Carbonell & Frijters, 2004).

4.4.3. Matching Variables

As argued in Chapter 3.1., those variables should be involved as covariates that have an effect on the treatment and the outcome, although they should not be affected by the treatment variable (Elwert & Winship, 2014; Heckman et al., 1997). Thus, one should involve only pre-treatment covariates which are less likely to be affected by the treatment (Elwert & Winship, 2014; Rosenbaum, 1984). Therefore, all the matching variables were measured in the first wave (2001/2002) at least one year before exposure to treatment. Previous research in this topic has identified the confounding variables (e.g. Balbo &

[24] Children were born between 2003 and 2004/2005, whereas we observed a change in subjective well-being between 2001/2002 and 2004/2005, thus the effect of having a zero- to two-year-old child is observed here. With children of this age, parents are entitled to parental leave in Hungary. Therefore, the generosity of parental leave directly influences the short-term effect.

[25] Children were born between 2003 and 2004/2005, whereas we observed a change in subjective well-being between 2001/2002 and 2012/2013, thus the effect of having a seven- to ten-year-old child is observed here.

Arpino, 2016; Baranowska & Matysiak, 2011; Myrskylä & Margolis, 2014; Sironi & Billari, 2013) which were included in this research too.

Matching variables included demographic and socio-economic variables. Education was categorized as primary or less, secondary vocational, general secondary, and tertiary. Four categories of residence were distinguished: villages, smaller cities, bigger cities, and the capital city. Age (given in years) and equivalent household income (thousand HUF) were measured as a continuous variable. Gender was also controlled for. Further, subjective health status was measured using the following question: "*On an eleven-point scale, how satisfied are you with your health*?" Higher scores indicated higher satisfaction with health. Satisfaction with housing was measured in a similar way as satisfaction with health. Further, perceived well-being was also included, as measured with the following question: "*How would you rate your standard of living?*" This variable took a value of 0 to mean the worse living standard imaginable, and 10 to indicate the best standard of living. Finally, the first wave value for subjective well-being, which was measured similarly to the outcome variable, was also involved.

I also controlled for labour-market-related characteristics. In general, the analysis incorporated data relating to whether the respondent had ever experienced unemployment (0 if respondent had experienced unemployment; 1 if never). An attitude variable was also included to measure whether respondents enjoyed working. Individuals were asked to rate the validity of the statement "*I usually do not enjoy working*" using a four-point scale (value of 1: respondent considered statement not valid at all; value 4: completely valid). Moreover, labour market status was categorized as employed, entrepreneur or self-employed owner, unemployed, and other non-working. Furthermore, type of work was also classified as blue-collar and white-collar. The former group was used to refer to those who undertook manual work[26] and the latter one to professional jobs[27]. Finally, data about whether the respondents' jobs were private or public was collected.

Family-related characteristics were also controlled for. Marital status was recorded as single, married living together, married living apart, divorced, and widow/er. Satisfaction with partner was categorized as does not have a partner, dissatisfied, neutral, rather satisfied, very satisfied, and no answer. Partner activity status was measured similarly to

[26] such as farmers, traders, manufacturers, service providers, skilled workers, unskilled workers, semi-skilled workers, and temporary workers

[27] such as production managers, top managers, middle managers, lower-level managers, subordinate non-manual workers with degrees, subordinate graduate non-manual workers, subordinate non-graduate non-manual workers, and temporary non-manual workers

the respondent's own labour market status (employed, entrepreneur or self-employed owner, retired, unemployed, and other non-working). Moreover, the analysis recorded how many children the respondent had.

To estimate the subgroup effect of parenthood, certain matching variables were omitted since they played a role in conceptualizing the treatment variable. For example, gender was not controlled for when estimating the effect of mother or fatherhood. Similarly, number of children was not used as a matching variable when estimating the effect of the first or second children, but played an important role in estimating the general and gender-specific effect of parenthood. In these latter cases, the number of children guaranteed that the parity effect was controlled for.

See the distribution of the matching variables in Tables 28, 30, 32, 34, and 36 for different subgroups. These tables also shows the balance improvement after matching[28].

4.5. Results

4.5.1. General Effects of Parenthood

First, the correlation between subjective well-being and the birth of a child was analysed. Results are displayed in Table 12. In all waves, people whose child was born during the observation period had higher subjective well-being than those to whom no children were born. This finding is in line with earlier studies which found a positive correlation between childbearing and subjective well-being in Hungary (Billari, 2009; Margolis & Myrskylä, 2011; Molnár & Kapitány, 2013). Moreover, the longitudinal design allows observation of the change in subjective well-being over time. This analysis reveals that people whose child was born in the observation period experienced a significantly greater increase in subjective well-being between 2001/2002 and 2004/2005 compared to those to whom no children were born. However, there is no significant difference between the two groups in terms of the change between 2001/2002 and 2012/2013. Thus, the arrival of a new child appears to be associated with an increase in subjective well-being in the short term, but not the long term.

[28] The values displayed in these tables are not weighted but the sample weights are instead involved as a matching variable. This is due to the special requirement of using sample weights in matching (See more about this in Chapter 3.7.5.).

Table 12. Difference in subjective well-being between those to whom a child was born between 2003 and 2004/2005 and those to whom no children were born in this period (mean, standard deviation, and level of significance)

	Child was born during the observation period		Child was not born during the observation period		Sig.
	Mean	SD	Mean	SD	
Subjective well-being measured in 2001/2002 (1-3 years before treatment group had a child)	7.01	0.11	6.53	0.03	0.01
Subjective well-being measured in 2004/2005 (0-2 years after treatment group had a child)	7.38	0.11	6.55	0.03	0.01
Subjective well-being measured in 2012/2013 (7-10 years after treatment group had a child)	7.44	0.11	6.95	0.03	0.01
Change in subjective well-being between 2001/2002 and 2004/2005	0.37	0.12	0.02	0.03	0.01
Change in subjective well-being between 2001/2002 and 2012/2013	0.43	0.13	0.42	0.03	0.93

The mean differences in subjective well-being might be attributable to a factor other than the birth of the child. For example, people whose child was born in the observation period were on average 28 years old in the first wave, whereas those to whom no children were born were on average 46 years old. This difference is crucial, as ageing itself can have an effect on subjective well-being, regardless of parenthood status (Blanchflower & Oswald, 2008; Gwozdz & Sousa-Poza, 2010). In order to rule out confounding variables, those to whom a child was born in the observation period (i.e. treatment group) were matched to those to whom no children were born in this period (i.e. control group) using genetic matching. This process reduced the number in the control group from 6331 to 259 but caused no change in the number in the treated group (307). More about the improvement in balance can be found in the Appendix in Table 28 and Figure 16.

In contrast to the simple correlation analysis (see in Table 12), the multivariate analysis (see Table 13) reveals that parenthood has a large positive significant effect both in the short and the long run, even though the effect decreases over time. The estimated treatment effect in this research is 0.56 in the short run and 0.39 in the long run, which is higher than what was found, for example, in Great Britain (0.20 in the short run and 0.04

in the long run) when similar methodology was applied (Balbo & Arpino, 2016).

The multivariate result is consistent with Hypothesis 1A, which predicted a decreasing effect on subjective well-being with the ageing of the child. This result is in line with demand and reward theory. Whereas the Hungarian social security system might also underpin the decrease in effect through modifying the cost of children in two ways: firstly, the generous parental leave diminishes the short-term costs of having a child, and secondly, the system of long periods of parental leave is associated with low access to child-care facilities and flexible job arrangements, which has a negative long-term effect on parents' employment. Moreover, the findings are also in line with Hypothesis 1B, which predicts that the arrival of a new child has a long-lasting positive significant effect on subjective well-being, based on value of children theory. However, this finding contests Hypothesis 1C by suggesting that the subjective well-being of parents does not return to the pre-birth baseline level even seven to ten years after the birth. Thus, the empirical evidence presented in this research contradicts set-point theory. Also, findings from the Hungarian case appear to be similar to those from other countries where parenthood is persistently linked to higher subjective well-being (Baetschmann et al., 2016; Mikucka, 2016; Pollmann-Schult, 2014).

Table 13. Parenthood status in regression models after matching (regression coefficient, and level of significance)

	Parenthood status
Short-term change (between 2001/2002 and 2004/2005)	0.56***
Long-term change (between 2001/2002 and 2012/2013)	0.39**

Note: This table contains only the treatment variable; the entire analysis can be seen in Table 29.
P-values: ***<0.001, **<0.05, *<0.1

I used Rosenbaum's (2002) sensitivity analysis to bound the treatment effect estimates. The $\Gamma_{general}$ parameter is 1.45 in terms of the short-term effect, and 1.32 for the long-term effect. The parameter for the short-term effect is close to the 1.5 threshold, which suggests the robustness of the estimates and indicates that it is very unlikely that an unobserved difference in covariates would change the inference of the short-term effect. However, the $\Gamma_{general}^{long-term}$ parameter for the long-term effect on parenthood is more significantly below this threshold. Thus, parenthood also has a long-term positive effect at a level, but it is more sensitive to unobserved confounders.

4.5.2. Parenthood Effects According to Parity

Next, the analysis estimated separately (1) the first child effect, and (2) the second child effect. In the first case, the treatment group contained those whose first child was born between 2003 and 2004/2005, whereas the control group contained respondents who remained childless during this period. In the second case, the treatment group referred to those whose second child was born between 2003 and 2004/2005, and the control group to those who had no additional children in this period, but had had one before the first wave of data collection.

First, the correlation between parenthood and subjective well-being is analysed in terms of parity. Table 14 displays the results. Having a child is associated with higher subjective well-being both before and after birth, regardless of whether the child is a first or second one. However, only the first child appears to increase subjective well-being significantly in the short term.

Table 14. Difference in subjective well-being between those to whom a child was born between 2003 and 2004/2005 and to whom no children were born by parity of children (mean and level of significance)

	First child			Second child		
	First child was born during the observation period	First child was not born during the observation period	Sig.	Second child was born during the observation period	Second child was not born during the observation period	Sig.
Subjective well-being measured in 2001/2002 (1-3 years before treatment group had a child)	6.85	6.27	0.01	7.33	6.31	0.01
Subjective well-being measured in 2004/2005 (0-2 years after treatment group had a child)	7.33	6.25	0.01	7.59	6.45	0.01
Subjective well-being measured in 2012/2013 (7-10 years after treatment group had a child)	7.16	6.40	0.01	7.79	6.74	0.01
Change in subjective well-being between 2001/2002 and 2004/2005	0.48	-0.02	0.02	0.30	0.14	0.59
Change in subjective well-being between 2001/2002 and 2012/2013	0.31	0.13	0.63	0.47	0.43	0.86

Again, matching was undertaken to obtain sufficient balance between the treatment and control groups, a process which is necessary for making unbiased causal inference. In terms of measuring the first child effect, matching reduced the control group from 726 to 89 responses but did not modify the size of the treatment group (134). In the case of the second child, matching reduced the dataset from 1274 control individuals and 82 treated individuals to 63 control and 82 treated. The improvement in balance can be seen in Tables 30 and 32, and in Figures 17 and 18.

Matching using longitudinal data (see in Table 15) revealed that both the first child and the second child had a positive effect both in the short and long run. This evidence is at odds with Hypothesis 2B and supports Hypothesis 2C. The fact that even second children have a lasting positive effect on subjective well-being is especially interesting in light of the relatively low number of second children in Hungary.

Nevertheless, this model unexpectedly revealed that the effect of the first child is greater than that of the second child in the short term, but the opposite is true in the long term. Thus, Hypothesis 2A stands only in the case of a short-term effect and is not supported in case of a longer one. Application of the demand and rewards approach (Nomaguchi, 2012; Nomaguchi & Milkie, 2003; Umberson et al., 2010) would suggest that this finding is due to changes in marginal utility and cost (i.e. in the short term the marginal utility of having a child decreases faster than the marginal cost, however, in the long run marginal utility increases faster than marginal cost). Further, the short-term moderating effect of parity may be also explained by set-point theory (Headey & Wearing, 1989; Kammann, 1983; Lykken & Tellegen, 1996), which argues that novelty is higher with the first child and adaptation to parenthood is stronger in the case of the higher-order children, thus, first children should elicit greater changes than second children. However, this theory is clearly at odds with the observed long-term effects.

Table 15. Parenthood status in the regression models before and after matching according to parity of children (regression coefficient and level of significance)

	Fist child	Second child
Short-term change (between 2001/2002 and 2004/2005)	0.81***	0.51**
Long-term change (between 2001/2002 and 2012/2013)	0.54**	0.69**

Note: This table contains only the treatment variable. The whole regression can be seen in Table 31 (for the first child effect) and Table 33 (for the second child effect). P-values: ***<0.001, **<0.05, *<0.1

Again, Rosenbaum's (2002) sensitivity analysis was conducted to measure the robustness of the estimations. The estimations for the first child effect were fairly robust both in terms of short-term changes ($\Gamma_{first}^{short-term}$ parameter: 1.55) and long-term changes ($\Gamma_{first}^{long-term}$ parameter: 1.47). The estimation of second children effects was highly robust in terms of the measurement of long-term changes ($\Gamma_{second}^{short-term}$ parameter: 1.62); however, the estimation of short-term effects of the second child was slightly more sensitive to unobserved confounders ($\Gamma_{second}^{long-term}$ parameter: 1.42).

4.5.3. Parenthood Effects in Terms of Gender

In this section, the effect of parenthood is described separately by gender. In case of motherhood effect, the treatment group contained women whose child was born between 2003 and 2004/2005, while those women to whom no children were born in this period belong to the control group. In case of fatherhood effect, the treatment group referred to men whose child was born in the observation period, and the control group to those men to whom no children were born in this period.

The results of correlation analysis is displayed in Table 16. Among women those whose child was born in the observation period had higher level of subjective well-being than women to whom no children were born in this period, both before and after the childbirth. In contrast, men whose child was born between 2003 and 2004/2005 did not differ from other men to whom no children were born during this period either before birth (2001/2002) or long after it (2012/2013). However, even fathers reported to having higher subjective well-being than other men for a short time (0-2 years) after the arrival of the child. Regarding the changes between waves, motherhood was not associated with a change in subjective well-being in the short term or long term, whereas fatherhood was associated with a short-term increase in subjective well-being.

Table 16. Difference in subjective well-being between those to whom a child was born between 2003 and 2004/2005 and those to whom no children were born in this period by gender (mean and level of significance)

	Women			Men		
	Child was born during the observation period	Child was not born during the observation period	Sig.	Child was born during the observation period	Child was not born during the observation period	Sig.
Subjective well-being measured in 2001/2002 (1-3 years before treatment group had a child)	7.31	6.53	0.01	6.75	6.49	0.17
Subjective well-being measured in 2004/2005 (0-2 years after treatment group had a child)	7.61	6.56	0.01	7.15	6.50	0.01
Subjective well-being measured in 2012/2013 (7-10 years after treatment group had a child)	7.51	6.87	0.01	7.24	6.94	0.22
Change in subjective well-being between 2001/2002 and 2004/2005	0.30	0.03	0.11	0.40	0.01	0.02
Change in subjective well-being between 2001/2002 and 2012/2013	0.20	0.34	0.50	0.49	0.45	0.98

Similarly to the previous cases, causality was estimated by reducing the initial dataset to create a more balanced one. In the case of females, matching reduced the control group from 3156 to 105, and maintained the 163 members of the treatment group. In the case of males, matching decreased the number of members of the control group from 2806 to 95, while the number of individuals in the treated group stayed at 127. Tables 34 and 36 and Figures 19 and 20 show how the balance between the treatment group and control groups greatly improved upon matching.

In contrast to the correlation using raw data (see in Table 16), matching using longitudinal data (see in Table 17) shows that motherhood has a high and long-lasting positive effect. Moreover, the latter model confirms that fatherhood has a moderate positive effect in the short term, and no effect in the long term. This finding is in line with Hypothesis 3C, but it contests Hypothesis 3A and 3B. Thus, results are in line with the claims of set-point theory only for males. However, value of children theory was confirmed for females, but not for males. Furthermore, demand and reward theory can explain results concerning gender differences in terms of costs and benefits. According to this approach, the finding shows that for females the benefits outweigh the costs, but for males the costs offset benefits in the long run. Finally, this finding is in line with the results of some earlier studies conducted in other countries (Baranowska & Matysiak, 2011; 2016; Sironi & Billari, 2013).

Table 17. Parenthood status in the regression models before and after matching by sex (regression coefficient and level of significance)

	Women	Men
Short-term change (between 2001/2002 and 2004/2005)	0.64***	0.46**
Long-term change (between 2001/2002 and 2012/2013)	0.48**	0.28

Note: This table contains only the treatment variable. The whole regression can be seen in Table 35 (motherhood effect) and Table 37 (fatherhood effect). P-values: ***<0.001, **<0.05, *<0.1

Sensitivity analysis on the estimations of motherhood effect reveals that the results are fairly robust (Γ_{female} parameter: 1.64 for short-term effect; 1.58 for long-term effect). However, the short-term effect of fatherhood is less robust (Γ_{male} parameter: 1.26). Thus, although fatherhood also has a significantly positive short-term effect, this result is more sensitive to unobserved confounders. As a consequence, the gender differences in terms of the parenthood effect might be even more pronounced than this analysis has suggested.

4.6. Discussion

During the last decades, the effect of parenthood on subjective well-being has been identified as the missing link in understanding recent fertility trends (Billari, 2009). Consequently several other scholars have also argued that the phenomenon can shed light on the underlying process behind fertility decisions and the low fertility rate (Aassve et al., 2016; Le Moglie et al., 2015; Luppi, 2016; Margolis & Myrskylä, 2015; Parr, 2010). However, there is still little known about the parenthood effect on subjective well-being in CEE countries where fertility is “the lowest-low”. The goal of this chapter was to extend the research about the parenthood effect on subjective well-being to Hungary.

By combining longitudinal data analysis with the matching method, this article has made further steps to estimate causality between parenthood and subjective well-being. Similar methodology has been used to estimate this relationship in the international literature (Baetschmann et al., 2016; Balbo & Arpino, 2016; Sironi & Billari, 2013); however, no prior research has applied this technique in Hungarian data. The application of the approach is highly beneficial since by using these methods one can control for all the time-invariant unobserved confounding variables and all observed confounding variables. Furthermore, this method can eliminate the extrapolation and interpolation bias, which prevalent in regression models (Ho et al., 2011; King & Zeng, 2006).

Overall, the research described in this chapter finds that parenthood has a large and long-lasting positive effect on subjective well-being. The finding resonates with some of the state-of-the-art international studies (Baetschmann et al., 2016; Mikucka, 2016; Pollmann-Schult, 2014). Furthermore, it supports value of children theory (Hoffmann & Hoffmann, 1973; Nauck, 2007). Nevertheless, the evidence that parenthood has a slightly decreasing effect supports the demands and rewards approach. This theory suggests that the effect declines as children age due to worsening parent-children relationships (Nomaguchi, 2012). However, this result is at odds with set-point theory, which argues that major life events are able to alter subjective well-being only temporarily since people adopt to their new situations eventually (Headey & Wearing, 1989; Kammann, 1983; Lykken & Tellegen, 1996). Finally, the identification of a decreasing effect is also in line with the Hungarian social policy context which supports parents in the short run, but creates opportunity costs in the long run through provision of long parental leave (Bartus et al., 2013; Makay, 2015).

This chapter also documented the moderating effect of parity. The effect of having a first child over remaining childless, and the effect of having a second child over having only one child was estimated. It was shown that not only the arrival of the first child, but also the second permanently increases subjective well-being. These findings are exceptional in international comparison, since only in Russia has it been found that second children have such a long-term effect (Mikucka, 2016). Moreover, these results raise the question why Hungarians do not have more second children even though doing so could positively impact life satisfaction. Understanding this paradox is crucial as the low Hungarian fertility rate is mainly attributable to the low number of second children (Miettinen & Szalma, 2014; Szalma & Takács, 2015).

Moreover, the research also looked at the moderating effect of sex and revealed that only females witness a long-lasting increase in subjective well-being, whereas parenthood has no long-lasting effect on males. This finding is in line with previous research conducted in the CEE countries, and more specifically, in Poland (Baranowska & Matysiak, 2011) and in Bulgaria (Sironi & Billari, 2013). As this result was only found in the CEE countries, it might be characteristic of this region. Since couples make decisions about transitions to parenthood together, the experience of both sexes matters in fertility-related decisions (Aassve et al., 2016). Therefore, the fact that fatherhood does not have a long-lasting positive effect may be contributing to the low fertility rate in Hungary and in other countries in the region. This finding suggests that further research in which sample size allows for the further elaboration of the effect of fatherhood by also incorporating the parity effect is needed.

To sum up, the argument that life satisfaction matters in understanding fertility trends makes only a limited contribution to the discussion about why the fertility rate is persistently low in Hungary. The only trend with subjective well-being which could contribute to such a low fertility rate is the fact that fatherhood does not cause a significant, long-term effect on life satisfaction. However, every other subgroup reported to experiencing positive changes upon the arrival of children. As a result, in Hungary one needs to go beyond observing subjective well-being to understand the low fertility rate.

The major limitation of this research is that even though unobserved time-invariant and all observed confounding variables were controlled for, however, unobserved time-variant confounding variables could not be controlled. First, pre-birth expectations about how children affect subjective well-being were not taken into account, since this variable was neither measured nor can it be assumed that it is time-invariant. However, Kravdal (2014) argues that it is important to take into account this factor, suggesting that it can significantly influence parenthood-related choices (as those who expect to have an enjoyable parenthood are more likely to have children than those who fear an undesirable parenthood). Similarly to most of the pre-existing studies in this field, the findings of the present study might be subject to this selection bias. Second, this chapter did not control for events which occur after treatment took place (i.e. post-birth divorce). In general, statisticians emphasize that one should not control for such variables when estimating causality, since these variables are most of the time affected by the treatment (i.e. they can be collider variables, as described in Chapter 3.1.) and controlling for them can create endogenous selection bias (Elwert & Winship, 2014; Rosenbaum, 1984). Although it is unlikely, we note that this practice might create selection bias if the given post-treatment variable were indeed not influenced by the treatment. Third, intention to have a further child was not controlled for, as this variable might be also affected by the treatment (i.e. it might also be a collider variable). Intention not only affects parity progression but can also be affected by future parenthood status since parenthood might be anticipated upon formulating this intention. Thus, intention is similar to post-treatment variables in the sense that (1) controlling for them can create endogenous selection bias if they were affected by the treatment, but (2) omitting them might create selection bias if they were indeed not affected by the treatment. I deal with intention to have a further child in a similar manner as I do for other possible collider variables, not controlling for them to avoid endogenous selection bias. To sum up, several variables were not controlled for due to lack of data access or statistical considerations which might have created selection bias. The present research has made an attempt to assess how sensitive the results are to a certain amount of selection bias using sensitivity analysis (Rosenbaum, 2002), finding that most of the results are fairly robust in terms of the estimation of the short-term effect, although weaker for the long-term overall effect and the short-term effect of second children and fatherhood. Future research could also test how robust the results are in terms of the moderating effects of certain post-treatment variables (such as divorce, unemployment), poor post-birth work-life balance, or intention to have further children (Matysiak et al., 2016).

To obtain deeper understanding about the changes in subjective well-being upon parenthood, this dissertation also devotes a chapter to observing which life domain is affected the most by parenthood. Thus, Chapter 6 not only observes overall life satisfaction, but also analyses the effect of domain-specific subjective well-being. However, prior to providing details about this, the next chapter (Chapter 5) shows how retirement affects overall subjective well-being.

5. Estimating the Effect of Retirement on Subjective Well-being

5.1. Introduction

Understanding the meaning of retirement in the life course of individuals has become a highly relevant policy field. Knowing the impact of retirement on well-being might help us to evaluate the effectiveness of the pension system; especially in ageing societies where the sustainability of the pension system has been endangered, and governmental spending on the pension system has grown. Despite the relevance of this topic, the previous empirical results produced mixed evidences (Barrett & Kecmanovic, 2013; Bonsang & Klein, 2012; Charles, 2002; Fonseca, Kapteyn, Lee, & Zamarro, 2017; Henning, Lindwall, & Johansson, 2016; Heybroek, Haynes, & Baxter, 2015; Kesavayuth, Rosenman, & Zikos, 2016; Latif, 2011; Luhmann, Hofmann, Eid, & Lucas, 2012).

This lack of empirical evidence is not surprising since there are several theoretical perspectives that can be employed to investigate the impact of retirement on subjective well-being. *Role theory* assumes positive changes in subjective well-being upon retirement, while in contrast, *role-strain theory* predicts negative changes, whereas *continuity theory* postulates insignificant changes. Further, *set-point theory* assumes that retirement should not have a long-term effect (Headey & Wearing, 1989; Kammann, 1983; Lykken & Tellegen, 1996). Finally, the *resource-based dynamic perspective* emphasises that an individual's resources (such as education or whether retired voluntarily or involuntarily) will determine whether retirement is a positive or negative experience for the given person (Wang, Henkens, & van Solinge, 2011).

Although this topic has received great attention in western countries, there is still little known about retirement effects on subjective well-being in Central-Eastern Europe (CEE). Such a research is important as the CEE countries are facing with similar challenges due to population ageing as the western social security systems (Simonovits, 2009), thus, the estimation of the effects of retirement might shed a light on underlying processes which needs to be addressed by policy measures. A positive relationship between retirement and subjective well-being would provide further explanation about the early labour market exit in this country, which is especially widespread in the CEE countries. Whereas, a negative relationship between retirement and subjective well-being would call for further active ageing policies.

The present chapter extends the scope of the evidences to Hungary. Observing the

Hungarian case is important as in this country the retirement system is embedded in a different context than in the well-observed western countries. In absolute terms, the Hungarian elderly is poor in European comparison (Szalai, 1991; Zaidi, 2011). However, this can be mainly attributed to the fact that Hungarians are in general poorer than their western peers. In relative terms, the Hungarian elderly is doing well compared to the rest of society as the Hungarian pension system has high pension replacement rate (Monostori, 2015; OECD, 2015; Zaidi, 2009). Thus, in the one hand, Hungarian elderlies might face with limited opportunities and more restrictions around retirement than their western peers. But on the other hand, the transition from work to retirement cannot be considered a financial shock in this country. Furthermore, the absolute deprivation and unfavourable labour market opportunities of the Hungarian elderlies also generated exceptionally high incident rates of involuntary retirement (Dorn & Sousa-Poza, 2010; Kohli, 2014; Szalai, 1991). Thus, this dimension requires further investigation. Finally, the observation of subjective well-being is utmost important since in Hungary the level of subjective well-being is significantly lower than in the Western European countries (Guriev & Zhuravskaya, 2009).

The research question at the hand requires the estimation of causality on observational data. The present study applies the genetic matching on longitudinal data, which method controls for both observed confounding variables and unobserved time-invariant variables. Thus, I am able to control for pre-retirement characteristics to a large degree, which helps to distinguish accumulated disadvantages from transition related risks. Most recently this method has been used for estimating the effects of various life events such as parenthood (Baetschmann, et al., 2016; Balbo & Arpino, 2016; Sironi & Billari, 2013), but it has never been applied for the observation of retirement effect on subjective well-being. Similarly to the previous study (Chapter 4), the empirical foundation of this research was the Turning Points of Life Course survey (Hungarian Generations and Gender Survey).

Based on my estimations, retirement does not have a statistically significant effect on subjective well-being. In fact, this result provides evidence that the Hungarian pension system is able to facilitate the transition from work to retirement. However, I also show that voluntary retirement does have an effect on subjective well-being: voluntary retirees achieve higher subjective well-being than their involuntary peers when previously existing differences are controlled for.

5.2. Background

5.2.1. The Effect of Retirement on Subjective Well-being in General

The effect of retirement on subjective well-being is not obvious as it is a complex life event which can bring about multiple changes in individuals' lives. There are five main theories that provide an explanation for how well-being changes upon retirement. *Role theory* argues that employment is a significant source of identity (Ballweg, 1967; Ellison, 1968; George & Maddox, 1977; Carter & Cook, 1995). Since the individual loses his or her important role after retirement, one can assume that subjective well-being will decline. In contrast, advocates of the *role-strain theory* argue that retirement, in effect, relieves individuals from expectations, which leads to an increase in their subjective well-being. Furthermore, retirees can devote more time to their families or to leisure, which might further increase well-being (Shultz, Morton, & Weckerle, 1998).

Further, *continuity theory* argues that subjective well-being is not affected by retirement since individuals try to maintain their standard of living, their self-esteem, and their values over time (Atchley 1971, 1989; Kim & Moen, 2002). More specifically, Atchley (1971) argued that employment is not as important as has been suggested by role theory and role-strain theory, claiming that identity is based on multiple sources such as family and social networks, not solely on employment status. Furthermore, Atchley also argued that even after retirement people tend to consider themselves as still belonging to their prior occupational group. Thus, they can transfer some aspects of their former occupations (such as skills) to their retired life.

Further, *set-point theory* can be mentioned here as well (see a description of this theory in more detail in Chapter 4.5.1.). This theory argues that even if a life event has a short term effect, this effect will diminish over time as individuals tend to adopt to their new circumstances (Headey & Wearing, 1989; Kammann, 1983; Lykken & Tellegen, 1996). Thus, this theory argues that retirement does not affect subjective well-being in the long term.

Finally, Wang et al. (2011) have developed a new theoretical framework; namely, the *resource-based dynamic perspective*, which integrates the previous approaches. It aims to understand how retirement affects well-being by building on the life course approach and the resource perspective. Based on this approach, certain factors provide resources which can foster retirement adjustment processes. These factors include individual attributes (such as good health), pre-retirement job-related characteristics (such as unemployment before retirement, or work stress), family-related characteristics (such as marital status), transition-related characteristics (such as voluntary retirement) and finally, post-retirement activities (such as leisure activities). This theory does not provide a universal answer to how subjective well-being changes upon retirement but rather focuses on the interindividual differences in experiencing the transition from work to retirement.

International research has produced mixed evidence about the effects of retirement on subjective well-being. There are some research projects which have found a positive relationship (Barrett & Kecmanovic, 2013; Charles, 2002; Gall, Evans, & Howard, 1997, 1997; Johnston & Lee, 2009; Kesavayuth et al., 2016; Montizaan & Vendrik, 2014). Others have suggested that life satisfaction diminishes after retirement (Bossé, Aldwin, Levenson, & Ekerdt, 1987; Coursolle, Sweeney, Raymo, & Ho, 2010; Nikolova & Graham, 2014; Szinovacz & Davey, 2004). Furthermore, some others have shown that retirement has no significant effect on subjective well-being at all (Baker, Gruber, & Milligan, 2009; Bonsang & Klein, 2012; Börsch-Supan & Jürges, 2006; Clark & Fawaz, 2009; Davis, 2012; Fonseca et al., 2017; Heybroek et al., 2015; Latif, 2011; Luhmann et al., 2012).

To summarize these results, Luhmann et al. (2012) have carried out a meta-analysis on the effect of retirement on subjective well-being. They have found that retirement is typically a neutral life event which has costs as well as benefits. However, this analysis does not attempt to capture differences between welfare states, countries, time of observation, and the applied methods. Thus, their meta-analysis may disguise significant relationships by aggregating conflicting results. However, Henning et al. (2016) also undertook a short review of the topic and found no systematic difference between countries, the time of observation, and methods. This review paper also concludes that the effect of retirement is mostly neutral, and sometimes even a positive life event.

To date, no serious effort was made to examine the impact of retirement on subjective well-being in Hungary. Previous studies provided only descriptive evidence. Molnár (2004) has shown that 60-75-year old citizens have the same subjective well-being as younger Hungarians. However, the analysis of a retrospective question showed that, the 60-75-year old population experiences more decline in their life conditions than younger people do. Furthermore, the most unsatisfied people are those who have just retired. Another line of research was by Monostori (2008), who already had longitudinal data. However, she has not estimated causality either, rather, she has provided a detailed analysis of retired people's stratification by subjective well-being. Monostori has found that retirees' subjective well-being can be explained mostly "*by the health status, the equivalent household income and whether somebody has experienced unemployment during his or her life course*" (Monostori 2008: 109).

5.2.2. The Effect of Voluntary Retirement on Subjective Well-being

There are multiple ways to define voluntariness of retirement. Desmet, Jousten, and Perelman (2005), for instance, split the population of retirees into two subgroups. They referred to those people who retire due to bad labour market conditions as the *true unemployed*, while they used the term *optimizers* for those who actually have an opportunity to follow individual utility maximization. Dorn and Sousa-Poza (2010) considered *voluntary retirees* those retirees who prefer retirement over the continuation of their job. Based on their definition, *involuntary retirees* were those who retire due to labour market constraints. They stressed that involuntary retirement is usually caused by unexpected constraints such as sickness or unemployment. This dissertation uses Dorn and Sousa-Poza's definition of voluntary and involuntary retirement.

Three sets of theories which were detailed in the previous subchapter also pose expectations about how retirement effects differ between those who retire voluntarily and those who do so involuntarily. First, *continuity theory* does not differentiate based on the form of retirement and suggests that retirement should not have an effect on anyone. Therefore, based on this theory we may expect that voluntary and involuntary retirees will not differ in terms of subjective well-being trends around retirement once we take into account pre-retirement differences. Second, the *resource-based dynamic perspective* assumes that retiring voluntarily instead of retiring involuntarily provides resources for coping with this transition (Wang et al., 2011). Thus, this theory predicts that, even when pre-retirement characteristics are taken into account, the level of subjective well-being of voluntary retirees increases to a greater extent than that of involuntary retirees. Finally, *set-point theory* suggests that life events may have an important short-term impact on subjective well-being, although convergence to original levels will occur as people adapt to their new situations (Headey & Wearing, 1989; Kammann, 1983; Lykken & Tellegen, 1996). As a consequence, potential differences between voluntary and involuntary retirees regarding how their subjective well-being changes upon retirement should disappear over time.

International research also produced mixed results concerning the effects of voluntary and involuntary retirement. Bonsang and Klein (2012) have found that voluntary retirement is the best predictor of retirement-related change in subjective well-being. Based on their work, involuntary retirement has a negative effect on subjective well-being mainly because people experience a higher decrease in satisfaction with the household income than what they gain with the increase in satisfaction with leisure. Bonsang and Klein (2012) have also shown that involuntary retirement is similar to unemployment, because individuals do not work; although, they in fact intend to do so. Consequently, several studies supported that involuntary retirement worsen the subjective well-being (Barrett & Kecmanovic, 2013; Bender, 2012; Shultz et al., 1998). In contrast, Abolhassani and Alessie (2013) have suggested that neither voluntary nor involuntary retirement have any significant effect on current life satisfaction.

5.2.3. The Hungarian Context

After the transition as the economy has changed from centrally planned state economy to market economy several people lost their job and unemployment peeked, which was mostly absorbed by the early retirement system (Szalai, 1991). However, in the last 20 years, various policy measures have entered into force in this country, all aiming to extend working careers and by that mitigate the negative consequences of ageing (Monostori, 2015). But still employment rates of men aged 60-64 was the lowest in Hungary in the observation period (2001 and 2010) among the OECD countries (Ebbinghaus & Hofäcker, 2013).

The situation of the Hungarian elderly is rather different from what people experience in western countries. In absolute terms, the Hungarian elderly face a lower standard of living than their western peers. In terms of severe material deprivation, 65+ year-old Hungarian men are ranked 22nd and Hungarian women 23rd among the 28 European member states for this risk. However, absolute poverty may be mainly attributed to the fact that Hungarians in general are not doing as well as their western peers. In comparison to the whole of society, the relative income position of the Hungarian elderly is one of the best in the European Union. A formerly median-level earner in Hungary presently receives a pension which amounts to 94% of their previous earnings (Monostori, 2015). Furthermore, as Table 18 shows, the at-risk-of-poverty ranking, which is a relative poverty measurement, for 65 year-or-older males is the 3rd lowest and for females the lowest in Hungary among all European countries. As a result, the transition to retirement cannot be considered a financial shock in this country. Finally, the high absolute poverty and low relative poverty of elderly Hungarians is not only exceptional in the European context, but exceeds that of Visegrad countries as well.

Table 18. Poverty of the Hungarian elderly (65+) compared to Visegrad and European countries in 2015[29]

		Hungarian	Visegrad 3[30]	EU28	Hungarian rank[31]
Severe material deprivation	female	16.4	8.1	6.3	22nd
	male	10.4	5.8	4.5	23rd
At-risk-of-poverty[32]	female	4.1	10.4	16.0	1st
	male	4.8	5.4	11.8	3rd

Finally, another major characteristic of the Hungarian retirement system which makes this study highly relevant is that this country has the highest rate of involuntary retirement among the OECD countries (Dorn & Sousa-Poza, 2010; Kohli, 2014). Kohli (2014) has found that 41.5% of the Hungarian males retire involuntarily, whereas this measure is 39.1% for women. Dorn and Sousa-Poza (2010) used a different definition, but they arrived at a similar conclusion in that involuntary retirement is alarmingly high in Hungary compared to other OECD countries. Based on the data from the 1997 International Social Survey Program, they have shown that the percentage of involuntary early retirees in Hungary was 62.1% of all early retirees. Previous studies in this issue claim that the high incidence rate of involuntary retirement can be attributed to the lack of demand for workers (Dorn & Sousa-Poza, 2010; Szalai, 1991), poor labour market position (Radó, 2012), alternative commitments (such as care responsibilities) (Kohli, 2014) or to health limitations (Kohli, 2014; Radó, 2012). All these factors related to the deprived status of the Hungarian elderly in absolute terms. Dorn and Souse-Poza (2010) say that the low standard of living in Eastern Europe inhibited voluntary retirement even among well-educated people, while Radó (2012) found that involuntary retirement is closely related to low pre-retirement social status and bad labour market position.

5.3. Hypotheses

The goal of this chapter is twofold: First, it aims to estimate the effect of retirement on

[29] Data: Eurostat, Róbert Iván Gál's calculation

[30] Visegrad 3: unweighted average of data from the Czech Republic, Poland and Slovakia

[31] HU rank: position of Hungary among the EU28 countries

[32] Threshold: 60% of median equalized income after social transfer

subjective well-being. Second, it also contributes to understanding the pluralized effect of retirement on subjective well-being by comparing how this life event affects voluntary and involuntary retirees. I formulate the following hypotheses based on the international empirical evidence and the Hungarian context.

Hypothesis 1. Retirement does not affect subjective well-being significantly either in the short or long term. This hypothesis is consistent with continuity theory (Atchley 1971, 1989; Kim & Moen, 2002) and the majority of international findings (Baker et al., 2009; Bonsang & Klein, 2012; Börsch-Supan & Jürges, 2006; Clark & Fawaz, 2009; Davis, 2012; Fonseca et al., 2017; Heybroek et al., 2015; Latif, 2011; Luhmann et al., 2012). On the one hand, retirement should not have a negative effect since, on average, this life event is not a financial shock, as reflected in the low level of relative poverty among the Hungarian elderly compared to the rest of society. On the other hand, neither should retirement have a positive effect in Hungary since the Hungarian elderly are also faced with a very low standard of living in absolute terms which limits the opportunities to increase subjective well-being and enjoy retired life.

Hypothesis 2. Retirement does not have significantly different effects on voluntary and involuntary retirees when their pre-retirement social and economic status are taken into account. Previous Hungarian research has shown that voluntary and involuntary retirees differ according to their pre-retirement status to a large degree (Radó, 2012). These initial dissimilarities might explain away all differences after retirement. Further, this hypothesis is in line with some of the international findings about this topic (Abolhassani & Alessie, 2013). Finally, this expectation is also in line with continuity theory (Atchley 1971, 1989; Kim & Moen, 2002).

5.4. Data

Similarly to the previous chapter, this study also relies on the Turning Points of Life Course survey. Consequently, the outcome variables and most of the matching variables are the same in these two chapters, therefore these variables will be only briefly mentioned here (See more about them and about the dataset in general in Chapter 4.4.). This subchapter will only detail those variables which are new compared to the previous study. Furthermore, the treatment effect is estimated the same way as in the previous chapter (through matching and the application of the regressor variable method based on Equation 44).

5.4.1. Treatment Variables

This empirical study was designed to observe two treatment variables; (1) retirement, and (2) voluntary retirement. Table 19 describes these treatment variables and also highlights their compositions.

Table 19. Description of the observed treatments in Chapter 5

Observed phenomenon	Treatment group	Control group
Retirement effect	Those who retired between the first (2001/02) and second waves (2004/05)	Those who did not retire between the first (2001/02) and second waves (2004/05), and were not retired yet in the first (2001/02)
Voluntariness of retirement effect	Those who voluntarily retired between the first (2001/02) and second waves (2004/05)	Those who involuntarily retired between the first (2001/02) and second waves (2004/05)

Firstly, retirement was measured in the second wave of the research (2004), which contains information about who retired between 2001 and 2004. Retirement is a binary variable in this dissertation. It takes a value of 1 if someone retired between 2001 and 2004, and a value of 0 if they did not retire[33]. Those people who retired before 2001 are not included in the analysis. After deletion of the missing values, the sample consisted of 306 people who retired between 2001 and 2004, and 4209 respondents who did not.

Second, whether somebody retired voluntarily or involuntarily was also observed in the second wave (2004). This variable was measured based on three questions: (i) whether it was the respondent's decision to retire, (ii) whether the interviewee is satisfied with the timing of retirement, and (iii) whether fear of unemployment played a role in the decision. This variable takes two values: voluntary or involuntary retirement. Those people are considered voluntary retirees who (1) did not want to retire later, (2) made the decision about their own retirement, and (3) did not make this decision due to fear of unemployment. All the rest are considered involuntary retirees. Descriptive statistics shows that only 51.63% of respondents (158) retired voluntarily while 48.37% of them

[33] There is empirical evidence that retirement, similarly to parenthood, can have an effect even one year before the transition occurs (Kesavayuth et al., 2016). This phenomenon is referred to as the anticipation effect in the literature. I eliminated the anticipation effect by omitting from the analysis those who retired between the first wave (2001/2002) and 2003. However, the treatment effect was similar with and without consideration of the anticipation effect. Since the sample size is higher without considering anticipation, I display only these results.

(148) were involuntary retirees.

5.4.2. Outcome Variables

Subjective well-being was measured in the same way as in the previous study (See more in Chapter 4.4.2.)

5.4.3. Matching Variables

Similarly to the previous study, matching variables included demographic and socio-economic variables such as education, residence, age, subjective health status, income, and the first wave value for subjective well-being. Further I also controlled for labour-market-related characteristics such as unemployment history, enjoyed working or not, labour market status, types of work, private or public jobs. Family-related characteristics were also controlled for such as marital status, satisfaction with partner, partner's labour market status, number of children, and number of grandchildren.

Although most of these matching variables were defined the same way as in the previous chapter (see Chapter 4.4.) there are some differences. In this study, partner's labour market status was categorized as employed, entrepreneur or self-employed owner, unemployed, retired, and other non-working. Thus, retired people were also distinguished in this study who had previously belonged to the 'other non-working' category. Further, in this study the number of children were distinguished based on their gender. This was necessary as previous studies found that male and female children contribute to their parents' subjective well-being to the same extent when they are not yet adults (Margolis & Myrskyla, 2016). However, adult male and female children might elicit different subjective well-being in old age since they differ from each other in terms of keeping contact with the elderly parents (Dykstra & Fokkema, 2011; Silverstein & Bengtson, 1997; Suitor & Pillemer, 2006), and providing intergenerational transfers (Dwyer & Coward, 1991; Horowitz 1985; Örkény, Koltai, & Székelyi; 2011; Stone et al., 1987). Finally, this study also incorporated the number of grandchildren besides the number of children. This was necessary, as some empirical evidence has found that grandchildren have an effect on subjective well-being (Powdthavee, 2011).

5.5. Results

5.5.1. The Effect of Retirement on Subjective Well-being in General

As a first step, the effect of retirement on subjective well-being was calculated by comparing the means of those who retired between 2001 and 2004 and those who were non-retirees in 2004 (see Table 20). In all waves, non-retired people have a significantly higher level of subjective well-being. However, there is no significant difference between the two groups in terms of change in subjective well-being either in the short or long term. This means that retirement is not accompanied by a change in subjective well-being.

Table 20. Difference in subjective well-being between retired and non-retired people (mean, standard deviation, and level of significance)

	Retired between 2001 and 2004		Non-retired in 2004		ANOVA P-value
	Mean	SD	Mean	SD	
Subjective well-being measured in 2001 (0-3 years before the treatment group retire)	6.37	2.06	6.81	1.89	0.01
Subjective well-being measured in 2004 (0-3 years after the treatment group retire)	6.47	1.91	6.82	1.86	0.01
Subjective well-being measured in 2012 (8-11 years after the treatment group retire)	6.84	2.21	7.12	2.05	0.01
Change in subjective well-being between 2001 and 2004	0.10	2.20	0.01	1.98	0.39
Change in subjective well-being between 2001 and 2012	0.47	2.43	0.31	2.26	0.17

A comparison of means, however, provides insight only into correlation but not causality. Since retired and non-retired groups systematically differ from each other, there may be confounding variables which modify the relationship between retirement and change in subjective well-being (i.e. the level of subjective well-being prior to retirement). Thus, one needs to rule out the confounding variables.

First, I conducted regression adjustment on the initial dataset (before matching) by using the regressor variable method (See Equation 44). The estimated results of this model can be seen in Table 21 (Model 1). This table highlights the estimations for key variables (namely, their coefficients and the significance level) whereas the coefficients of the rest of the control variables are provided in the Appendix (Tables 40 and 41). The retirement effect in this model is not significant either in the short or long term. However, as I argued in Chapter 3, regression adjustment often fails to create an unbiased estimation when the balance is poor between the treatment groups. As Table 38 and Figure 21 in the Appendix show, the balance in this case is rather poor, thus further investigation is needed to rule out extrapolation and interpolation bias.

In order to increase the balance between the treatment and control groups, a matching method was applied. Due to this matching, the number of respondents in the treated group (those who retired in the observation period) remained at 339, while the number of respondents in the control group (those who did not retire during the observation period) decreased from 4209 to 250. Figure 21 and Table 38 also show the improvement in balance. The treatment effects were again obtained by using the regressor variable method, but this time on the matched dataset. See the results in Table 21, Model 2. The retirement effect on subjective well-being is still insignificant even after eliminating the possibility of extrapolation and interpolation bias. This result supports Hypothesis 1.

Table 21. The variable retirement in the regression models before and after matching

	Model 1: Raw data regression adjustment	Model 2: Matched data regression adjustment
Short-term change (0-3 years after retirement)	0.07	-0.07
Long-term change (8-11 years after retirement)	-0.01	-0.16

Note: This table contains only the treatment variable; the entire analysis can be seen in Tables 40 and 41.
P-values: ***<0.001, **<0.05, *<0.1

5.5.2. The Effect of Voluntary Retirement

Although retirement may have no significant effect on subjective well-being, it can have different consequences for those who retire voluntarily and those who do so involuntary. Firstly, a simple mean difference was observed in subjective well-being (see Table 22). In particular, this dissertation has found that voluntary retirees have a significantly higher level of subjective well-being than involuntary retirees, both before and after retirement. However, the two groups do not differ significantly from each other in change in subjective well-being, neither in the short run, nor in the long run.

Table 22. Difference in subjective well-being between voluntary retired and involuntary retired people (mean, standard deviation, and level of significance)

	Voluntary Retirees		Involuntary Retirees		ANOVA P-value
	Mean	SD	Mean	SD	
Subjective well-being measured in 2001 (0-3 years before retirement)	6.72	2.05	6.01	2.10	0.01
Subjective well-being measured in 2004 (0-3 years after retirement)	6.81	1.91	6.29	1.94	0.01
Subjective well-being measured in 2012 (8-11 years after retirement)	7.23	2.22	6.53	2.89	0.01
Change in subjective well-being between 2001 and 2004	0.09	2.08	0.28	2.45	0.41
Change in subjective well-being between 2001 and 2012	0.52	2.42	0.52	2.44	0.98

This finding suggests that voluntary and involuntary retirees experience similar changes at the time of their retirement. However, one should make certain whether this non-significant correlation actually stems from the voluntariness of retirement or is simply attributable to initial differences in the compositions of voluntary and involuntary retirees. For this purpose, confounding variables are again ruled out by using the regressor variable method (See Equation 44) with matching (Model 1 in Table 23) and without matching (Model 2 in Table 23).

Model 1 (See Table 23) observes the relationship between voluntary retirement and subjective well-being on the raw data (before matching). This model shows that voluntary retirement has no significant relationship with change in subjective well-being either in the short or long term. However, as argued in the Analytical Strategy chapter, this regression adjustment often fails to produce an unbiased estimation when there is no sufficient overlap between the treated and control groups, while Table 39 and Figure 22 show that the balance between voluntary and involuntary retirees is indeed rather poor in the raw dataset. Although they used different samples, earlier studies that applied regression adjustment might have also been subject to such bias by omitting to improve the balance between the two subgroups (Abolhassani and Alessie, 2013; Barrett and Kecmanovic, 2013; Bonsang and Klein, 2012).

Matching was performed to improve the balance. The matching procedure identified a total of 225 retirees (158 voluntary) who satisfied the matching criteria. Thus, matching significantly reduced the sample size compared to the pre-matching state (306 observations). However, in the case of matching, only the raw dataset needs to be sufficiently large. Unbiased estimates can also be obtained using an even smaller set of matched data provided that this is well balanced. The improvement in balance in all covariates can be seen in Table 39 and Figure 22 in the Appendix. These figures show a significant improvement in balance compared to the raw dataset, thus reducing the exposure to extrapolation and interpolation bias. As a result, Model 2, using the matched data, provides a less biased causal estimation than Model 1, which uses raw data.

In contrast to Model 1, Model 2, using matched data (See Table 23), shows that retirement induces higher subjective well-being among voluntary retirees than among involuntary retirees when the composition of these two groups is taken into account. Furthermore, this effect remains significant 8 - 11 years after retirement, thus contradicting set-point theory. The short-term effect is similar or higher than what previous research suggested[34], whereas the long-term effect is higher here than what was found in previous studies. However, this long-term effect is less than it is 0 - 3 years after retirement, thus some adaptation to retirement might indeed occur. Further, the positive short-term and long-term effects of the voluntariness of retirement are also at odds with continuity theory, a finding which was predicted in Hypothesis 2.

[34] Albohassani and Alessie (2013) have found a very small and insignificant effect of voluntary retirement (-0.038) and involuntary retirement (-0.145), whereas Bonsang and Klein (2012) found that voluntary retirement increases subjective well-being by 0.147 unit and involuntary retirement significantly decreases it by -0.526 unit.

Table 23. The variable voluntary retirement in the regression models before and after matching

	Model 1: Raw data regression adjustment	Model 2: Matched data regression adjustment	Γ
Short-term change (0-3 years after retirement)	0.31	0.82***	1.85
Long-term change (8-11 years after retirement)	-0.12	0.58**	1.32

Note: This table contains only the treatment variable; the entire analysis can be seen in Tables 42 and 43. P-values: ***<0.001, **<0.05, *<0.1

As was mentioned in the methodology section, Rosenbaum's (2002) sensitivity analysis should be used for bounding the treatment effect estimates. This Γ parameter is 1.56 for estimating the short-term effect while 1.13 for estimating the long-term effect. The parameter of the short-term effect is above 1.5, set as a threshold, which provides evidence for the robustness of the estimates. This suggests that it is very unlikely that an unobserved difference in covariates would change the inference in the case of the short-term effect. However, the Γ parameter of the long-term effect is below this threshold. Thus, although voluntary retirement has also a positive long-term effect, this result is more sensitive to unobserved confounders.

5.6. Discussion

The aim of this study was to extend the research about the effect of retirement on subjective well-being to Hungary. A growing number of international research projects have produced inconsistent results concerning how they answer this research question (Henning et al. 2016; Luhmann et al. 2012; Wang et al. 2011), whereas no research has been done in the Hungarian context so far. However, more evidence in this area could contribute to the policy agenda in terms of reforming the Hungarian social security system which is endangered by population ageing. This chapter has argued that evidence from better observed western countries cannot be applied in the present context since Hungary is an extreme case. It has a higher rate of absolute poverty among the elderly in comparison to western countries, although the retirement system is able to maintain the pre-retirement earning profiles to a large degree.

I have found that retirement has no significant effect on subjective well-being in Hungary.

The insignificant effect resonates with some of the previous international studies (Baker et al., 2009; Bonsang & Klein, 2012; Börsch-Supan & Jürges, 2006; Clark & Fawaz, 2009; Davis, 2012; Fonseca et al., 2017; Heybroek et al., 2015; Latif, 2011; Luhmann et al., 2012). This finding hence reflects the fact that the Hungarian social security system can successfully facilitate the transition from work to retirement. Despite the fact that the Hungarian elderly are not doing as well as their western peers, the Hungarian retirement system is able to mitigate this disadvantage due to a high net pension replacement rate. Furthermore, the lack of a positive relationship shows that the widespread practice of early labour market exit cannot be explained by increasing subjective well-being upon retirement. Although there is growing interest in the demographic literature which claims that subjective well-being may be the “missing link” in understanding decisions about life events (Billari, 2009), this argument makes no contribution to understanding early retirement in Hungary. Finally, this finding supports so-called continuity theory (and contests both role theory and role-strain theory), which assumes that individuals’ well-being is not influenced by retirement, since individuals try to maintain their standard of living, their self-esteem, and their values during their entire life course.

Besides the overall effect of retirement, the present chapter also observed how voluntary retirement can facilitate the transition from work to retirement. It is particularly interesting in Hungary, because the country has the highest involuntary retirement rate among developed countries. The present study suggests that voluntary retirees experience greater shifts in subjective well-being than involuntary retirees do after controlling for previously existing differences between the two groups. This result implies that even when lifelong accumulated advantages are taken into account (such as pre-retirement income or health), the transition has distinct effect on voluntary and involuntary retirees. This result is at odds with continuity theory and supports what the resource-based dynamic perspective assumes. Further, this finding is in line with some of the previous research findings in this area (Barrett & Kecmanovic, 2013; Bender, 2012; Shultz et al., 1998). As a result, the high incidence of involuntary retirement raises equity-related questions as well. Thus, this chapter shows that decreasing involuntary retirement by enabling the elderly to stay in the labour market as long as they wish is of utmost importance in Hungary. However, this must not be done at the expense of higher unemployment among older people, as it is already high in Hungary (Micheel, Roloff, & Wickenheiser, 2011).

Moreover, the effect of the voluntariness of retirement is persistent over time and only slightly diminishes a decade after retirement. These results contribute to the existing body of research that tested set-point theory following different life events. Authors such as (Lucas et al., 2004; Clark 2008) have suggested that although set-point theory seems to make correct predictions in many cases, there are some events – like unemployment – that seem to permanently alter the level for the set point. According to the present results, it seems that involuntary retirement belongs to this category. Clearly, the empirical evidence found in the research for this study does not support the claim that people eventually entirely adapt to their new situations. However, adaptation does seem to occur to a certain extent as the difference in subjective well-being between voluntary and involuntary retirees declines over time.

The previous chapter and the present chapter have focused on how certain life events influence overall subjective well-being. The following chapter extends the scope to include an investigation of the effects of the whole life course on several life domains. Thus, this last empirical chapter aims to provide a more in-depth analysis of the topic described above.

6. Estimating the Effects of Household Life-cycle on Overall and Domain-specific Subjective Well-being

6.1. Introduction

The previous two chapters documented the effects of parenthood and retirement on overall subjective well-being. The questions arose whether (1) other kinds of life events have effects too, and (2) whether life events have a distinct effect on specific satisfaction measures in different domains. Therefore, this study estimates the effect of household life cycles on domain satisfaction.

Household life-cycle theory argues that households follow a specific life path, and changes in their demographics are accompanied by adjustments in their lifestyles and well-being. The theoretical approach has had undiminished popularity since its early application in sociology; however, the operationalization of the concept has changed continuously and has been criticized due to the inherent flaws of the life cycle concept (see for example Derrick and Lehfeld [1980]). One of the most-frequently cited of the early models was developed by Wells and Gubar (1966). Their model consists of nine life-cycle stages that are defined according to the age, marital status and employment status of the household head, and the age of the youngest child in the family. However, the greatest deficiency of the early models is that they take into account only classic family types, unlike modern life cycle models. An example of the latter is Gilly and Enis's (1982) model, which incorporates modern types of cohabitation and, hence, addresses single-parent households separately.

Life stages can have complex consequences on overall subjective well-being as they can influence the specific domains in a distinct way. For example, motherhood can be rewarding, as females often spend more time with the family after giving birth (Nomaguchi, 2012; Nomaguchi & Milkie, 2003), although this life event may decrease opportunities to be engaged in gainful employment (Sanchez & Thomson, 1997). This part of the dissertation aims to capture such complex effects by observing domain-specific subjective well-being. Most of the previous studies have focused only on a certain life event, or only on one specific sub-domain. Only a few studies have observed the effect of all life stages on domain-specific subjective well-being, and most of these were either admittedly explanatory or observed only the effect of ageing instead of the effect of life stages (Bardo, 2017; Schafer, Mustillo, & Ferraro, 2013).

The present study, similarly to the previous two studies, intends to approximate as closely as possible the causal relationship between belonging to a life-cycle group and subjective well-being. Again, to achieve this goal a matching method was used. However, in contrast to the previous empirical chapters, for this research only cross-sectional data were available.

6.2. Background

6.2.1. Ageing and Overall Subjective Well-being

The life-course perspective emphasizes understanding how subjective well-being changes across the entire lifespan. Accordingly, in the literature the change in subjective well-being with ageing is often studied controlling for the influence of variables such as income and education. Even after taking into account these stable demographic variables, ageing can be influential in understanding satisfaction with life.

International empirical results with regard to ageing effects on overall subjective well-being are highly controversial. For Western European countries, it has been shown severally that subjective well-being follows a U-shaped curve with ageing, which means that a typical individual's subjective well-being reaches its minimum in middle age. Accordingly, U-shaped curves have been found across many studies, including Clark and Oswald (1994, 2006), based on the British Household Panel Survey and a General Health Questionnaire; McAdams, Lucas, and Donnellan (2012), based on the British Household Panel; Lang, Llewellyn, Hubbard, Langa, and Melzer (2011), based on a Health Survey for England; and Van Landegham (2008, 2012), based on a German Socio-Economic Panel. Moreover, Blanchflower and Oswald (2008) observed 72 countries and showed that a U-shape curve is valid for each of them. On the contrary, some studies have found an inverted U-shape curve for the US (Mroczek & Spiro 2005, Easterlin 2006, Easterlin & Sawangfa 2007). Some authors have found that the U-shaped curve vanishes after controlling for socio-demographic variables (Gwozdz & Sousa–Poza 2010, Frijters & Beatton 2012, Kassenboehmer & Haisken-DeNew 2012), and some studies have found a decline in subjective well-being with ageing (Gwozdz & Sousa–Poza 2010), or non-significant effect (Costa et al., 1987). Thus, ageing can be seen as having different effects in different countries and under different circumstances. It is important to note that Steptoe, Deaton, and Stone (2015), and Lengyel and Janky (2002) have found that the association between subjective well-being and ageing is independent from demographic variables.

The Hungarian results about the correlation between ageing and subjective well-being are also controversial. Graham and Pozuelo (2017) found that subjective well-being across individual life-cycles is U-shaped in Hungary. However, the majority of scholars argue that a U-shaped relationship is valid for only rich English-speaking countries, whereas in middle-income countries, and especially in former Soviet countries, the elderly seem to be the most dissatisfied age group (Deaton, 2008; Neulinger & Simon, 2011; Steptoe et al., 2015).

Furthermore, there have been studies even in Hungary which observed how ageing affects subjective well-being by controlling for social economic variables. The majority of these studies have found a U-shaped relationship between the two key variables (Blanchflower & Oswald, 2008; Graham & Pozuelo, 2017; Hajdu & Hajdu, 2013; Molnár & Kapitány, 2006; Spéder & Kapitány, 2002). More specifically, they found that the relatively low subjective well-being of the elderly can be explained by their lower level of income and declining health status. Up to now, only limited research has found a non-significant relationship (Murinkó, 2007), or negative relationship[35] (Lengyel & Janky, 2002).

6.2.2. Life Events and Overall Subjective Well-being

Besides the observation of ageing, there is growing interest in how specific life events change overall subjective well-being. Chapter 4 detailed the parenthood effect and Chapter 5 detailed the retirement effect. Both of these chapters argued that these life events have complex effects by introducing the conflicting theoretical background. In general, state-of-the-art longitudinal studies have found that parenthood has an initially positive effect on overall subjective well-being, which effect disappears or at least decreases when the child grows up (Balbo & Arpino, 2016; Baranowska & Matysiak, 2011; Kohler et al., 2005; Mikucka, 2016; Le Mogliea et al., 2015; Pollmann-Schult, 2014). There is mixed evidence about the effect of retirement, but meta-analyses have found that this life event is mostly neutral or even positive (Henning et al., 2016; Luhmann et al., 2012).

This chapter also differentiates life cycles based on partnership status. Based on previous empirical studies, we know that divorce and widowhood have generally negative effects on subjective well-being (Amato & Hohmann-Marriott, 2007; Bierman, Fazio, & Milkie, 2006; Giczi, 2008; Hohmann-Marriott & Amato, 2008; Hughes & Waite, 2009; Lorenz, Wickrama, Conger, & Elder Jr., 2006), while marriage and partnership seem to have a

[35] In this study, the age variable was included but the square of age was not, which could have actually shown the non-linear effect of age.

positive effect (Graham & Pozuelo, 2017; Molnár & Kapitány, 2006; Murinkó, 2007).

Finally, this chapter also observes how partnership status modifies the relationship between parenthood and subjective well-being.[36] There is consistent evidence that single parents are less satisfied with their lives than non-single parents. The prior group has less access to economic resources and less time to spend with their children, thus childbearing is more demanding for them (Hansen, 2012). However, there is mixed evidence about the effect of single parenthood compared to remaining childless. Some scholars have argued that childbearing outside of a union is also becoming a viable strategy, therefore single parents are more satisfied than childless people (Kohler et al., 2005; Margolis & Myrskylä, 2015). Others have found that the disadvantage of being a single parent is so large that such individuals have even worse subjective well-being than their childless peers (Nelson et al., 2014; Nomaguchi & Milkie, 2003). Finally, some scholars have found that the difference in the level of subjective well-being between single parents and childless adults is not significant (Aassve, Goisis, & Sironi, 2012; Baranowska & Matysiak, 2011).

6.2.3. Ageing and Domain-specific Subjective Well-being

Ageing can have a complex effect on subjective well-being, which can be captured by observing the distinct effects on different domain-specific subjective well-being measures. In this section I review earlier findings about the ageing effect on domain-specific subjective well-being, while the subsequent chapter summarizes the literature about the effect of life events on domain-specific subjective well-being.

Certain domain-specific indicators consistently found to decline with ageing include satisfaction with health status (Easterlin, 2006) and satisfaction with future opportunities. The first indicator declines as objective health status also decreases with age. Furthermore, the latter measurement decreases over time since younger people tend to evaluate their future opportunities more optimistically than the elderly, who tend to accept that their situations might not change (Heckhausen et al., 1989; Lachman et al., 2008; Röcke & Lachman, 2008).

Other domain-specific subjective well-being measurements tend to increase with ageing. First, income and standard of living tend to increase over time (Diener & Eunkook Suh, 1997; Easterlin, 2006; George, 1992; Plagnol, 2011; Plagnol & Easterlin, 2008). One

[36] This issue was not observed in Chapter 4 due to the insufficiently low number of single parents in the longitudinal sample.

might expect that these measures would change according to the actual earning profile, which follows an inverted U-shaped curve, with a peak at midlife. However, empirical studies have shown otherwise. There are several theories about why subjective financial indicators move independently from objective measures, including those which recognize decreasing needs and expectations, increasing assets, and a reduction in the dependency burden as children grow up and parents die (Diener & Eunkook Suh, 1997; Easterlin, 2006; George, 1992; Plagnol, 2011; Plagnol & Easterlin, 2008). Second, satisfaction with life-course often also increases since the elderly recall positive events more often than younger people (Heckhausen et al., 1989; Lachman et al., 2008; Röcke & Lachman, 2008). Third, satisfaction with housing also tends to increase over the life course (Adams, 1992; Fernández-Carro, Módenes, & Spijker, 2015; Russell, 2008). This measurement changes with the needs and characteristics of living arrangements. On the one hand, housing-related needs change as the number of people living together varies over the life course (Miller & Crader, 1979), and with ageing new needs arise such as the requirement to live near to certain services (Fernández-Carro et al., 2015; Russell, 2008). On the other hand, the characteristics of housing also change over the life course (e.g. owning or renting accommodation). Needs usually fluctuate with ageing (Fernández-Carro et al., 2015; Miller & Crader, 1979; Russell, 2008), but the characteristics generally improve (Andrews & Sánchez, 2011).

Finally, empirical results have produced mixed evidence about the effect of ageing on satisfaction with family relationships and satisfaction with jobs. First, satisfaction with family relationships is mostly influenced by satisfaction with children and partner. Satisfaction with children usually decreases (Nomaguchi, 2012), whereas satisfaction with partner either also decreases (Blood & Wolfe, 1960) or follows a U-shaped curve (Rollins & Cannon, 1974; VanLaningham, Johnson, & Amato, 2001). Second, studies are inconclusive about how job satisfaction changes over time. Most show that job satisfaction follows a U-shaped curve with ageing (Clark, Oswald, & Warr, 1996; Gazioglu & Tansel, 2006; Hochwarter, Ferris, Perrewe, Witt, & Kiewitz, 2001; Zacher, Jimmieson, & Bordia, 2014), but some scholars have found an inverted U-shaped curve (Easterlin, 2006) or a positive relationship (Besen, Matz-Costa, Brown, Smyer, & Pitt-Catsouphes, 2013; Ng & Feldman, 2010).[37]

In Hungary only limited research has observed how domain-specific subjective well-

[37] This positive relationship resonates with the results of early studies on this topic (Glenn, Taylor, & Weaver, 1977; Lindström, 1988; O'brien & Dowling, 1981; Weaver, 1980; White & Spector, 1987).

being changes over the life course, and these research projects were not usually intended to estimate causality. Murinkó (2007) observed the correlation between age and domain-specific satisfaction. She found that satisfaction with health decreases with age, whereas satisfaction with income stagnates over the life course. Further, she found that the trajectory of satisfaction with work and family differs among females and males. In the case of females, these two measures are initially high in the twenties, decrease in the thirties, peak in the forties, and gradually decrease thereafter. In the case of males, satisfaction with family does not vary over the life course, although satisfaction with work decreases with age.

Other Hungarian studies have controlled for stable demographic variables using multivariate regression, but they only focused on one domain. There is evidence that satisfaction with work increases with age (Medgyesi & Róbert, 2000), whereas satisfaction with economic status follows an inverted U-shaped curve, taking into account demographic variables (Molnár & Kapitány, 2006).

6.2.4. Life Events and Domain-specific Subjective Well-being

Most studies which observe the effect of life events on domain-specific subjective well-being, consider only one life event or only one domain at one time. In the following section I summarize these studies.

First of all, previous research has found that marriage and partnership have positive effects on satisfaction with health (DeMaris, 2017), family relationships (Ubesekera & Luo, 2008), and financial satisfaction (Stevenson & Wolfers, 2009). Some research projects have also found that these life events have positive (Bowen, Radhakrishna, & Keyser, 1994; Khaleque & Rahman, 1987), or insignificant (Scott, Swortzel, & Taylor, 2005) effects on job satisfaction.

Second, parenthood has also been found to have distinct effects on life domains. Financial satisfaction, job satisfaction, and health satisfaction is expected to decrease since having children involves significant objective cost (Hansen 2012; Reizer, 2011). In fact, there is evidence that job satisfaction decreases upon parenthood (Holtzman & Glass, 1999), although using longitudinal data Bernardi, Bollmann, Potarca, and Rossier (2017) found that this indicator decreases significantly only for females. Further, previous studies have shown that parenthood has a negative effect on satisfaction with health (Newman, 2008) and satisfaction with housing (Miller & Crader, 1979). In contrast, parenthood influences satisfaction with family relationships rather positively, although this event has dissimilar effects on relationships with different family members. On the one hand, this may decrease since satisfaction with the partner tends to decrease after childbirth (Lawrence, Rothman, Cobb, Rothman, & Bradbury, 2008; Pollmann-Schult, 2013; Rollins & Cannon, 1974; Twenge et al., 2003; VanLaningham et al., 2001). On the other hand, satisfaction with family relationships is expected to increase as satisfaction with a child can compensate for worsening partnership satisfaction (Nomaguchi, 2012).

This study also considers how single motherhood affects domain-specific subjective well-being. Previous studies in this field showed that single parents experience lower satisfaction with their life-course, and satisfaction with the future. Further, single parents might experience high levels of work-family conflict which can have negative effects on their job satisfaction (Burden, 1986; Matjeke, 2016). However, their level of satisfaction with health (Herbst, 2012) and satisfaction with their financial situation (Medgyesi & Zólyomi, 2016) do not differ significantly from the rest of the population.

6.3. Hypotheses

In this study, the following hypotheses are tested, which are consistent with the state-of-the-art empirical research on this topic:

Hypothesis 1: Overall subjective well-being follows a U-shaped curve with ageing, once confounding variables are controlled for.

Hypothesis 2: The elderly are more satisfied than the rest of society with their financial situation, standard of living, housing, and previous life course, when confounding variables are taken into account.

Hypothesis 3: The elderly are less satisfied than the rest of the society with their future opportunities, health status, and family relationships, when confounding variables are taken into account.

Hypothesis 4: Parenthood has an initially positive effect on overall subjective well-being, but this effect disappears as the child grows up.

Hypothesis 5: Parenthood has a positive effect on satisfaction with family relationships.

Hypothesis 6: Parenthood has a negative effect on satisfaction with standard of living, satisfaction with job, satisfaction with income, satisfaction with health and satisfaction with housing.

6.4. Data

This study uses cross-sectional data which were collected with the help of face-to-face interviews by Ipsos from a national sample of 1000 respondents (final respondent number after refusal) using random sampling in Hungary on March 2014. The questionnaire was developed by Ágnes Neulinger, Katalin Melles, and Márta Radó.

All respondents were the primary shoppers in the given household, where the primary shopper was identified by means of the following question: "*Which member of your family does the shopping (compiles the shopping list) most frequently; who makes the purchase decisions?*" This filter question was necessary because the treatment variable of this analysis is household life cycle, whereas the outcome – subjective well-being – cannot be measured at the household level. Thus, the household had to be narrowed down to one person who could report his or her level of subjective well-being. We chose the primary shopper for this purpose, as this member of the family is most aware of their household income and expenditure, and the household situation in general.[38]

Data collection was stratified according to ten life-cycle stages (See description below), and 100 persons were selected from each group. This approach was intended to ensure the appropriate number of responses from each life stage. Due to the stratification sampling method, the sample was originally not representative of the country's entire population; therefore, a weight variable was created to draw conclusions at the national level as well. This weight was created by the Hungarian Ipsos.

The following sub-chapters introduce the treatment, outcome and matching variables. The estimation of the treatment effect is calculated by matching and then regression adjustment (See Equation 17).

6.4.1. Treatment Variables

Life-cycle stages constitute the treatment variable in the analysis. This variable included information about the individual's gender, age, partnership status, their partner's age (where appropriate), and the number and ages of any children in the household (where appropriate). These life-cycle stages were mutually exclusive and collectively exhaustive; that is, every individual was certain to fit into one, but only one, of the categories. Based on these variables, the following ten life-cycle stages can be distinguished:

1. *Young alone*: Those respondents, who were 34 years old or younger and were childless, but did not live together with a partner.

[38] Furthermore, in the paper, which was published in the International Journal of Consumer Studies, we also observed the direct effect of life cycle on subjective well-being controlled for expenditure structure. Expenditure structure is a variable which is hard to measure and the primary shopper can recall this best.

2. *Young with a partner*: Those respondents, who were 34 years old or younger (if the respondent was male than the partner was 34 years or younger), and they were childless, but they lived together with their partner.
3. *Full nest 1*: Those respondents, who were cohabiting with their partner and lived together with a child aged below 6 (if there were more children than the youngest is aged below 6).
4. *Full nest 2*: Those respondents, who were cohabiting with their partner and lived together with a child aged between 6 and 18 (if there were more children than the youngest is aged between 6 and 18).
5. *Full nest 3*: Those respondents, who were cohabiting with their partner and lived together with a child aged above 18 (if there were more children than the youngest is aged above 18).
6. *Single parent 1*: Those respondents, who did not live together with their partner (or had no partner at all), but lived together with a child aged below 18 (if there were more children than the youngest is aged below 18).
7. *Single parent 2*: Those respondents, who did not live together with their partner (or had no partner at all), but lived together with a child aged above 18 (if there were more children than the youngest is aged above 18).
8. *Middle aged without child*: Those people, who were between 35 and 64 (if the respondent was male and had a partner than the partner was between 35 and 64 years), and did not have a child, but might or might not have a partner living together.
9. *Empty nest cohabiting*: Those people, who were 65 years or older (if the respondent was male than his partner was 65 years old or older) they were cohabiting with their partner, but they were not together with a child.
10. *Empty nest alone*: Those respondent, who were aged above 65 and they did not have neither partner nor child living together with them.

Due to the requirements of the modelling only seven life-cycle stages were used as I needed to recode Stage 1 and 2 (*Young alone* and *Young with a partner*), Stage 4 and 5 (*Full nest 2* and *Full nest 3*), and Stage 6 and 7 (*Single parent 1* and *Single parent 2*). These recorded life stages were called respectively *Young childless, Full nest 2,* and *Single parent*.

For each life-cycle stage a binary variable was created that took a value of 1 if the respondent was at the given life stage, and 0 otherwise. Thus, in this study I always compare whether someone belongs to a certain life stage against belonging to the rest of the population. In other words, the control group is always constituted by all of those people who do not belong to the observed life stage.

6.4.2. Outcome Variables

Both a single question on whole-of-life satisfaction and multiple questions about different life subdomains were applied using a 10-point Likert scale. These were measured with the following question: "*Please indicate how satisfied you are with the following things. If you are not satisfied at all, say 0; if completely satisfied, say 10. How satisfied are you with...?*"

- Your life-course
- Your future opportunities
- Your quality of living standard
- Your family relations
- Your health
- Your work/job
- Your housing
- Your place of residence
- Your income
- Your life as a whole

The last variable, satisfaction with whole life, is referred to as overall subjective well-being from now on. Other elements will be referred to as domain-specific subjective well-being.

6.4.3. Matching Variables

Subjective well-being is influenced by a large number of factors other than life stages. The variables that might have concealed or explained the relationship between the two variables were controlled for; in other words, stable socio-demographic variables, and life course-related factors. The analyses were controlled for education, settlement type, gender, income, nature of employment (whether the respondent is employed in the public sector, and whether she/he has ever been unemployed), number of wage-earners in the household, and perceived health status (this later variable was not included upon estimation of satisfaction with health). However, the variables that determined the respondents' membership in the given life-cycle stage, such as age or household size, were not controlled for. In selecting the variables to be controlled, all of the relevant research findings concerning demographic trends, fertility and parenthood, and subjective well-being were considered; see for example the works of Balbo et al. (2016). The variable mentioned here was also used in the regression adjustment; in other words, the matching variables coincide with the control variables.

6.5. Results

In line with previous studies, the present results suggest that perception of overall subjective well-being varies across life-cycle stages. The last row of Table 24 shows the correlation between life stages and overall subjective well-being. According to this table, young childless people and couples with a young child have the highest life satisfaction, while single parents, empty nest couples, and empty nest singles were found to have the lowest level of subjective well-being.

The analysis of domain-specific subjective well-being shows that satisfaction with life-course, future opportunities, quality of living standard, health, and housing tend to decrease with age. However, satisfaction with family relationships peaks at mid-life, except for with single parents who experience below-average family satisfaction. Based on these results, most of the domain-specific indicators are lowest when the child leaves the parental house or someone raise a child without a partner, which is also reflected in the significantly below average overall subjective well-being.

Table 24. Correlation between subjective well-being and life-cycle stages (regression coefficient and significance)

	Young childless	Full nest 1	Full nest 2	Single parent	Middle aged without a child	Empty nest cohabiting	Empty nest alone
Life-course	0.04	0.05	0.04	-0.12***	0.06	-0.07**	-0.08**
Future opportunities	0.12***	0.05	0.04	-0.09***	-0.02	-0.16***	-0.06
Quality of living standard	0.07**	0.03	0.06	-0.11***	0.04	-0.10***	-0.07**
Family relationships	-0.03	0.07**	0.09***	-0.08***	0.10***	-0.15***	-0.05
Health	0.18***	0.12***	0.04	0.04	-0.21***	-0.29***	-0.10***
Work/job	0.04	0.04	0.01	-0.06	0.00	-0.02	-0.06
Housing	0.02	0.02	0.05	-0.05	0.04	-0.09***	-0.03
Place of residence	0.03	0.05	0.02	-0.07**	-0.01	-0.02	0.01
Income	0.04	0.01	0.05	-0.08**	0.02	-0.05	-0.06**
Overall subjective well-being	0.07**	0.07**	0.04	-0.09***	0.01	-0.12***	-0.07**

Note: P-values: ***<0.001, **<0.05, *<0.1

Table 24 shows only correlation. To estimate causality, a matching method is combined with regression adjustment (See Equation 17) here. Without matching, a regression adjustment might fail to produce unbiased results. More specifically, matching is necessary before regression adjustment if the dataset is not well-balanced, which is the case here since the treatment and control groups considerably differ from each other. Figure 23 in the Appendix illustrates this by displaying the propensity score distribution before and after matching in the treatment and control groups for one life-cycle group, while the overlap is similar in the case of the other life-cycle groups.

The effect of belonging to a certain life-cycle on overall subjective well-being is shown in Table 25. These findings suggest that having a young child and when children leave the parental home are indeed the two most satisfactory periods of life, whereas there is no life stage which significantly differs in a negative direction from the other life stages. This result is at odds with Hypothesis 1, which postulated that overall subjective well-being should follow a U-shaped curve. For a U-shaped curve to exist, young childless people and empty nest singles should be more satisfied than this study found. However, this study supports the claim that individuals in one of the elderly life stages — more specifically, empty nest families — are more satisfied than the average. Finally, Hypothesis 4, which stated that parenthood has an initially positive effect, is supported for those parents who have a partner. However, single parents were not differentiated according to the age of the child, thus, the initial period of parenthood – which was satisfactory for couples – cannot be traced for them. But the insignificant negative coefficients suggests that single parents do not benefit from having a child.

Table 25. Estimating causality between subjective well-being and life-cycle stages by matching and multivariate regression (regression coefficient and significance)

	Young childless	Full nest 1	Full nest 2	Single parent	Middle aged without a child	Empty nest cohabiting	Empty nest alone
N before matching (control/treatment)	500/207	638/69	568/139	561/146	630/77	633/74	635/72
N after matching (control/treatment)	140/207	52/69	86/139	111/146	73/77	40/74	32/72
Life-course	-0.08	0.33	0.10	-0.37	-0.21	1.07**	-0.21
Future opportunities	-0.01	0.16	-0.11	-0.42	-0.27	0.75**	-0.49
Quality of living standard	-0.09	0.01	0.15	-0.24	-0.11	0.51	-0.20
Family relationships	-0.34	0.80**	0.81***	-0.68**	0.24	0.72	-0.96**
Health	0.98**	-0.37	-0.12	0.10	0.23	-1.67***	-1.02**
Work/job	-0.36	-0.11	0.38	-0.07	-0.21	-	-
Housing	-0.49**	0.60	0.55**	0.10	-0.09	0.21	-0.54
Place of residence	-0.26	0.28	0.13	-0.04	-0.11	0.06	-0.31
Income	-0.24	0.13	-0.32	-0.21	-0.36	0.27	0.89**
Overall subjective well-being	-0.15	0.60**	0.01	-0.19	0.02	0.63**	-0.34

Note: This table contains only the treatment variables; the entire analysis can be requested from the author of this dissertation. P-values: ***<0.001, **<0.05, *<0.1

The results of the matching procedure suggest that some of the domain-specific subjective well-being measures change in old age (See Table 25). First, I found that certain measures are above average for the elderly. Both satisfaction with the life-course and future opportunities peak in old age when the respondent is cohabiting with his or her partner. The first finding supports Hypothesis 2, which states that the elderly tend to recall their past more positively than younger people. But the second finding is at odds with Hypothesis 3, which postulated that young people should be more optimistic about their future opportunities. Unexpectedly, the elderlies are not more satisfied with their housing than the rest of the society, however, aging might have a positive effect on this domain as young childless people below average satisfied with it. Further, satisfaction with income is also above average for those elderly who live without a partner and a child, which is line with Hypothesis 2.

Other domain-specific measures decrease with age. Unsurprisingly, satisfaction with health gradually decreases over the life course. Further, satisfaction with family relationships is significantly below average in old age when the respondents do not live together with their partners, or their children. These results were predicted in Hypothesis 3.

Parenthood elicits changes in certain measures of domain-specific well-being. Satisfaction with family relationships peaks when the respondents live together with their child and partner. This finding is in line with Hypothesis 5. However, single parents face significantly below average family satisfaction, which is inconsistent with Hypothesis 5. In other words, the birth of a child can compensate for the often reported decline in satisfaction with the partner, but it cannot compensate for the lack of a partner. Furthermore, satisfaction with housing also peaks when couples have an older child, which was not expected based on the previous studies. Surprisingly, this research has found that living together with a child when the partner is also present does not have any negative effect on any domain-specific measures. This finding is at odds with Hypothesis 6, which expected that the costs of children would be manifested in changes in domain-specific subjective well-being.

Finally, satisfaction with certain life domains does not vary across the life course, which is unexpected based on the literature. More specifically, satisfaction with quality of living standard, place of residence, and job are not influenced by life stages after the potentially confounding variables are controlled for. This means that all the differences that one can see in the correlation analysis can be attributed to the fact that individuals of certain life stages differ in their socioeconomic status.

6.6. Discussion

The present analysis employed a quasi-experimental design to assess the household life-cycle effects on subjective well-being. This approach provides a deeper understanding of subjective well-being throughout the whole life course and allowed for the differentiation between satisfactions with domains. Previous studies in this field have mainly focused on overall subjective well-being instead of domain-specific well-being (Balbo & Arpino, 2016; Clark et al., 2008; Clark & Oswald, 2006; Fonseca et al., 2017; Pollmann-Schult, 2014), even though such an approach is widely recognized. However, those studies which observed domain-specific subjective well-being over the life course either narrowed their focus to only one specific life event (Bernardi et al., 2017; Lapa, 2013) or one specific domain (Clark et al., 1996; Glenn et al., 1977; Plagnol, 2011; Rollins & Cannon, 1974; Zacher et al., 2014) or they observed ageing effects on subjective well-being instead of life-cycle effect (Bardo, 2017; Schafer et al., 2013). Thus, there is only limited research on this topic (Easterlin, 2006), especially in Hungary.

Although the vast majority of earlier research found that subjective well-being follows a U-shaped curve with age after controlling for confounding variables (Blanchflower & Oswald, 2008; Graham & Pozuelo, 2017; Hajdu & Hajdu, 2013; Molnár & Kapitány, 2006; Spéder & Kapitány, 2002), the research behind this paper does not support this relationship entirely. A U-shaped curve would imply that young people and the elderly have above average life satisfaction. On the one hand, contrary to expectations, this study showed that young childless people do not differ from the average significantly, which might be attributed to their below average satisfaction with housing. On the other hand, some of the elderly indeed reported above average life satisfaction. More specifically, those who live together with their partner are more satisfied with their life than average, but those who live alone do not differ from the average. The difference between the elderly living with or without their partners can be also captured in their different levels of domain-specific subjective well-being. The presence of a partner makes the elderly more satisfied with several life domains (namely, life course, future opportunities, and family relationships) than the average, which was not found for those elderly people who live alone. However, elderly people living alone reported above average satisfaction with their income, in contrast to the elderly cohabiting with their partners.

This study found that parenthood has an initially positive effect on life satisfaction when the partner is present. This result is in line with the findings of the majority of the international literature (Angeles, 2010; Balbo & Arpino, 2016; Baranowska & Matysiak, 2011; Kohler et al., 2005; Mikucka, 2016). Further, the observation of domain-specific well-being helps create a deeper understanding of why couples with young children report above average life satisfaction. The presence of a child and a partner is associated with significantly higher satisfaction with family relationships. So, this domain could be responsible for the increase in overall subjective well-being upon the arrival of the child.

However, the presence of a child does not necessarily lead to an increased level of subjective well-being under every circumstance. Parenthood only has a positive effect on overall subjective well-being in the case that a partner is present and the youngest child is under six years old (Full Nest 1 life cycle), whereas single parents and parents with partners and older children do not differ significantly from the rest of the population. Furthermore, the results about the empty-nest life stages revealed that the subjective well-being of cohabiting is above the population average, signifying that children leaving the parental home might be a positive life event.

This research is original in that the pairing of individuals is performed using genetic matching, which appeared as a novel method in the international literature only a few years previously, while its merits have been recognized in pre-existing research (Balbo & Arpino, 2016; Diamond & Sekhon, 2013; Sironi & Billari, 2013). However, one limitation of this analysis is the cross-sectional design. Further, the relatively small sample size may have had an effect on the outcome of the analysis. Finally, the Hungarian context also restricts the generalizability of the research findings and calls for future research.

7. Summary and Conclusion

7.1. The Objectives of the Dissertation

This dissertation has documented the relationship between belonging to a certain life stage and subjective well-being. A growing number of studies have emphasized that individuals engage in certain life changes if they benefit from those given transitions in terms of subjective well-being (Billari, 2009; Caldwell & Schindlmayr, 2003; Hobcraft, 2006). Therefore, observation of this topic can contribute to understanding macro-level demographic changes. For example, if parenthood did not make people more satisfied with their lives, this could trigger a decrease in the fertility rate. The dissertation applies the life course paradigm as a theoretical framework, which also emphasizes that one needs to observe individuals' life trajectories in order to understand macro-level demographic changes (Elder et al., 2003; Kok, 2007).

Furthermore, the research topic of this dissertation also reveals inequalities which emerge over the life course (Ferraro et al., 2009). This dissertation observes marginalized groups such as single parents and involuntary retirees. Thus, it also sheds light on groups which should be better supported by the social security system.

Although a growing number of international studies have examined the effects of life transitions on subjective well-being, there is still little known about this topic in Hungary. This is unfortunate, as this country permanently has one of the lowest levels of life satisfaction among the OECD countries (Guriev & Zhuravskaya, 2009). Furthermore, Hungarians also have a very low standard of living, which limits opportunities to increase subjective well-being. The economic, social, and cultural differences (Draxler & Van Vliet, 2010; Manning, 2004; Polese et al., 2014) between Hungary and better observed western countries may also modify the relationship between life-stage status and subjective well-being. Thus, the present dissertation discusses in detail the Hungarian context of parenthood and retirement.

This dissertation was designed to extend prior knowledge by estimating causal relationships using the state-of-the-art causal inference approach (Diamond & Sekhon, 2013; DuGoff et al., 2014; Fisher, 1925; Ho et al., 2011; Holland, 1986; Imai et al., 2007; Imai & Van Dyk, 2004; Imbens & Rubin, 2015; Neyman, 1923; Rosenbaum, 2002; Rosenbaum & Rubin, 1983, 1984; Rubin, 1974, 1978; Stuart, 2010). Causal relationships are here estimated by using statistical methods, more specifically, the matching method (see Chapter 6) and its extension to longitudinal data (Chapters 4 and 5). These methods come in useful when randomized experiments are not feasible. This is the case with this research, since we cannot arbitrarily choose who should experience a life transition (e.g. have a child), thus the treatment variable cannot be manipulated and, as a consequence, cannot be randomized. In the Analytical Strategy chapter I unfold the advantages and the limitations of different types of matching methods.

This work contains three empirical studies which observe from different perspectives how subjective well-being changes over the life course. The first and the second studies focus on the effects of specific life transitions on overall subjective well-being using a longitudinal dataset. More specifically, parenthood and retirement transitions are investigated in these chapters, whereas the third study estimates how domain-specific subjective well-being measures change over the whole life course using a cross-sectional dataset. The third empirical study and the other two studies complement each other since the third study details domain-specific subjective well-being, but the other two have more power to estimate causality due to the longitudinal data that is available. These studies cannot be fully compared, as the populations were somewhat different (the third study observed only the primary shoppers, who are mostly women, whereas the first and the second studies used a nationally representative sample), but in the following subchapters I synthesize the related findings.

7.2. Summary of Empirical Findings

The present dissertation has found that subjective well-being varies across the life course. Most earlier studies suggested that overall subjective well-being follows a U-shaped curve with ageing in Hungary, after controlling for social economic variables (Blanchflower & Oswald, 2008; Graham & Pozuelo, 2017; Hajdu & Hajdu, 2013; Molnár & Kapitány, 2006; Spéder & Kapitány, 2002). This finding implies that young people and the elderly are the two most satisfied age groups. But the findings of the third empirical study (Chapter 6) in this dissertation does not entirely support this link. It shows that young childless people and the elderly who live alone are not more satisfied with their lives than the rest of society. However, the elderly who live with their partners and who are without children indeed seem to enjoy above-average life satisfaction.[39]

The observation of domain-specific subjective well-being contributes to understanding why, contrary to expectations, overall subjective well-being does not follow a U-shaped

[39] Further, those who live together with their children were not distinguished based on their age, therefore this study cannot conclude anything about the domain-specific well-being of the elderly who live together with offspring.

curve. First of all, the unexpectedly low overall subjective well-being of the young and childless might be attributed to their poor housing opportunities, since these young people report significantly below average satisfaction with housing. The difference in domain-specific subjective well-being among the elderly with or without a partner might also explain why the prior group reports significantly above average subjective well-being, whereas the latter one does not differ from the average. This study found that the presence of a partner and the absence of a child make the elderly significantly more satisfied with their life-course and future opportunities than the rest of society. In contrast, the absence of both a partner and a child leads the oldest respondents to be less satisfied with their family relationships, and they do not differ from the rest of the population in terms of their satisfaction with their life course and future opportunities. However, members of this latter life-cycle group are more satisfied with their income than average, which is not true for the elderly who have a cohabiting partner.

Further, this dissertation estimated the effect of parenthood on subjective well-being using both cross-sectional data (Chapter 6) and longitudinal data (Chapter 4). Both the longitudinal and the cross-sectional analyses support the claim that parenthood in general has an initially positive effect on overall subjective well-being. This result is in line with the vast majority of international studies that used longitudinal data (Balbo & Arpino, 2016; Baranowska & Matysiak, 2011; Kohler et al., 2005; Mikucka, 2016; Pollmann-Schult, 2014). Moreover, this finding supports the value of children theory, which predicts a positive effect based on the argument that children fulfil different parental needs (Hoffmann & Hoffmann, 1973; Nauck, 2007).

Both the longitudinal (Chapter 4) and the cross-sectional analyses (Chapter 6) support the claim that parenthood in general has an initially positive effect on overall subjective well-being. However, these studies do not agree concerning the longer-term effect. On the one hand, the cross-sectional analysis found that parenthood has no effect in the long term, since those parents whose (youngest) child is over six years of age are not more satisfied with their life than other members of society. On the other hand, the longitudinal study in Chapter 4 found that parenthood has a long-lasting effect on subjective well-being that is significant even 7-10 years after childbirth. Thus, the cross-sectional study supported set-point theory, but the longitudinal one did not. This theory postulates that parenthood has only a temporary effect, since after this event individuals eventually adopt to their new situation and their subjective well-being returns to the initial baseline level (Headey & Wearing, 1989; Kammann, 1983; Lykken & Tellegen, 1996). Nevertheless, both studies are in line with the expectations of the demands and rewards approach, which suggests that the effect declines as children age due to worsening parent-children relationships (Nomaguchi, 2012). The cross-sectional and longitudinal studies may have arrived at different conclusions about the long-term effect of parenthood due to their different methodological approaches or their different sample composition. Moreover, the longitudinal research followed parents only until their children were 7-10 years old. Thus, it is possible that parenthood has a long-lasting effect only until children are 10, after which adaptation takes place, as reflected in the cross-sectional results.

The positive general effect of parenthood was further specified by distinguishing the parenthood effect according to (1) the gender of the parent, (2) the parity of the child, and (3) the partnership status of the parent. First, the longitudinal study (Chapter 4) observed the distinct effect of motherhood and fatherhood. The study reveals that motherhood is more rewarding than fatherhood both in the short- and the long-term. Moreover, the positive effect related to fatherhood does not have a significant effect in the long term, whereas those for motherhood are long lasting. This result is consistent with previous studies which found that motherhood is more rewarding than fatherhood (Angeles, 2010; Clark et al., 2008; Baranowska & Matysiak, 2011; Baetschmann et al., 2016; Kohler et al., 2005; Myrskylä & Margolis, 2014; Sironi & Billari, 2013), while positive effect on fatherhood is temporary (Baranowska & Matysiak, 2011). Second, the longitudinal data analysis also revealed that both the first and second children have positive long-lasting effects, a finding which is in line with some of the empirical studies (Matysiak et al., 2016; Mikucka, 2016; Myrskylä & Margolis, 2014; Pollmann-Schult, 2014). Finally, the cross-sectional analysis (Chapter 6) showed that single parenthood does not have a significant effect on overall subjective well-being, which finding is also consistent with some international results (Aassve et al., 2012; Baranowska & Matysiak, 2011).

Moreover, the third empirical study (Chapter 6) contributed to understanding the complex effect of parenthood by investigating how it influences different measures of domain-specific well-being. First, parenthood does not have any significantly negative effect on any life-domain measure when partners cohabit. This suggests that the costs of children (Hansen, 2012) are not manifested in changes in domain-specific subjective well-being. Further, parenthood significantly increases satisfaction with family relationships when partners cohabit. However, several studies have argued that this domain might suffer upon the arrival of a child since satisfaction with a partner tends to decrease at this time (Lawrence et al., 2008; Rollins & Cannon, 1974; Twenge et al., 2003; VanLaningham et al., 2001). The positive effect of parenthood on satisfaction with family relationships shows that the creation of new family ties with the child can compensate for the possible decrease in partnership satisfaction. However, the absence of a partner suppresses this otherwise positive effect, since single parents experience below average satisfaction with family relationships.

Further, Chapter 6 analyses how the retirement transition affects overall subjective well-being using longitudinal data. This study demonstrates that retirement does not have a significant effect on overall subjective well-being. This finding resonates with most of the previous international studies (Baker et al., 2009; Bonsang & Klein, 2012; Börsch-Supan & Jürges, 2006; Clark & Fawaz, 2009; Davis, 2012; Fonseca et al., 2017; Heybroek et al., 2015; Latif, 2011; Luhmann et al., 2012) and also supports continuity theory, which argues that, in the case of life transitions, people aim to maintain their standard of living, their self-esteem and their values and, as a consequence, their subjective well-being (Atchley 1971, 1989, Kim & Moen 2002). We cannot disentangle how retirement affects domain-specific subjective well-being since the cross-sectional study, which observed domain-specific measures, did not distinguish retirees, but only those who were older than 65 years and living without their children. However, the situation of the elderly living without children might also reflect on retirement effects to a certain degree, since 95% of them were retired. Based on this proposition, retirement may negatively affect satisfaction with health. Further, if a partner is present then retirement might influence positively satisfaction with life course, and satisfaction with future opportunities. However, in the absence of a partner retirees might experience below average satisfaction with family relationships but above average satisfaction with income.

Besides these findings, the research described in this dissertation was able to identify certain social groups which were exposed to a higher level of risk when encountering a given transition. First, the second empirical study (Chapter 5) found that those who retire involuntarily benefit less from this transition than those who retire voluntarily, even when pre-retirement differences are taken into account. This finding is in line with the conclusions of previous research projects which have found that it matters whether retirement is voluntary or involuntary (Barrett & Kecmanovic, 2013; Bender, 2012; Shultz et al., 1998). Further, this disadvantage of involuntary retirees is permanent over time (even over the long term). Therefore, involuntary retirement shows similarities with unemployment, which has also proven to be resistant to adjustment in the literature (Lucas et al., 2004). Second, the third empirical study (Chapter 6) highlights the disadvantage of being a single parent compared to becoming a parent with a cohabiting partner. The finding is that although parents who live together with a partner have significantly higher overall subjective well-being, this is not true of those who raise their child(ren) alone. By observing how individuals varied according to the measures of domain-specific subjective well-being, I found that the lower overall well-being among single parents is mostly attributable to their lower level of satisfaction with family relationships.

7.3. Main Contributions of the Dissertation

This dissertation contributes to our prior knowledge in several major ways. First and foremost, it extends the scope of research about the effect of life events on subjective well-being in Hungary. This topic has been the subject of growing attention in western countries, but there is still little known about the issue in Hungary, and in general in the CEE countries. Observation of this issue in different countries is of utmost importance, as in each country life events are embedded in a different economic and cultural context, thus one may expect that individuals will react in a different way to these transitions under different circumstances.

A key strength of the present studies is that they use the state-of-the-art matching method to estimate causality between belonging to a certain life stage and subjective well-being. Furthermore, when longitudinal data were available (Chapters 4 and 5), the matching method was extended to longitudinal data analysis. The application of longitudinal data can rule out unobserved time-invariant heterogeneity, whereas matching eliminates interpolation and extrapolation bias. Thus, the applied methods bring us one step closer to obtaining unbiased causal estimations (Ho et al., 2011; King & Zeng, 2006). In the international literature, matching using longitudinal data analysis has already been undertaken to estimate the effect of certain life events such as parenthood (Balbo & Arpino, 2016; Sironi & Billari, 2013). However, to my knowledge no prior research has applied this technique to estimate the effect of retirement. Furthermore, a matching method has been never used in the Hungarian context to investigate the effect of life events on subjective well-being.

The first empirical study (Chapter 4) contributes to understanding why fertility is the lowest of the low in Hungary. The hypothesis has arisen that the low fertility rate might be a consequence of unsatisfactory parenthood (Billari, 2009). Therefore, the effect of parenthood on subjective well-being was identified as the “missing link” in understanding parity progression (Aassve et al., 2016; Le Moglie et al., 2015; Luppi, 2016; Margolis & Myrskylä, 2015; Parr, 2010). As a consequence, a growing number of studies have investigated this effect. In Chapter 4 I describe how this hypothesis was tested in the Hungarian context. Overall, the research described in this chapter found that fertility has a long-lasting positive effect on subjective well-being. Moreover, not only the first child but also the second one generates an increase in subjective well-being. This finding is especially important as the low number of second children depresses the fertility rate in Hungary (Miettinen & Szalma, 2014; Szalma & Takács, 2015). The only result which might explain the low fertility rate relates to the moderating effect of gender. This dissertation shows that women benefit from having children both in the short and long term, whereas men experience only a temporary short-term increase in subjective well-being upon the arrival of a child. The effect of fatherhood is important as parenthood typically involves a joint decision, thus both genders should benefit from this life event in terms of realizing further parity progression (Aassve et al., 2016). To sum up, the research described in this chapter finds that the pre-existing theory that uses the link between parenthood and subjective well-being to explain fertility trends make a limited contribution to the discussion about the Hungarian situation. Thus, there appear to be other reasons for the low fertility rate in this country.

The second study (Chapter 5) tested whether the social security system is able to provide a smooth transition from work to retirement. This is an especially urgent question as the current retirement system needs to be reformed to be sustainable. However, before any reform can take place one needs to understand how the current system works. Overall, the related research finds no evidence that retirement affects subjective well-being. Thus, the present retirement system appears to be able to facilitate the transition from work to retirement even in a country where the standard of living is generally low. However, voluntary retirees achieve a significantly higher level of subjective well-being after retirement than those who retire involuntarily, even when the accumulated differences between the two groups are taken into account. Therefore, it appears that involuntary retirement might be creating a new form of inequality, alongside more traditional forms of disparities.

The third empirical study (Chapter 7) contributes to understanding the pluralized effect of life stages. First of all, the study aimed to capture the complex effects of life stages on different life domains by using domain-specific subjective well-being instead of only using overall life satisfaction. This is an important goal, as certain life events can have contradictory effects on different life domains (Bardo, 2017; Bernardi et al., 2017; Lapa, 2013; Schafer et al., 2013). For example, this research found that being young and childless affects satisfaction with health positively, but satisfaction with housing negatively. Furthermore, this study also observed the pluralized effect of life events by further elaborating life-stage groups. The same life events may have different effect on distinct groups (Hansen, 2012). The parenthood effect on individuals, for example, significantly depends on whether they are raising a child alone or with a cohabiting partner.

Besides the empirical contributions, the present dissertation made a further attempt to compare how different types of matching methods are calculated and when these methods should be used. The Analytical Strategy chapter (Chapter 3) illustrated the computational background of these methods using a small sample dataset. To my knowledge, no prior work has systematically compared the differences between matching methods this way.

7.4. Limitations and Suggestions for Further Research

This dissertation has some limitations. First of all, the available datasets in Hungary are either cross-sectional or do not provide a whole picture of the domain-specific measures. This is unfortunate, as state-of-the-art longitudinal data analysis may not be used to observe domain-specific effects. Future research should deal with this issue.

Furthermore, the methodology that was applied uses the Stable Unit Treatment Value Assumption (SUTVA), which presupposes that (1) the potential outcome for an individual is not affected by whether other units received the treatment or not, and (2) that treatment does not have different versions (Rubin, 1978). As I explain below, I should acknowledge that to a certain degree both of these conditions were violated, but not to the degree that they discredit the validity of the results. I do believe that future research should and could address these issues.

The implication of the first assumption of SUTVA for this research is that any given individual's subjective well-being should depend only on this person's life-stage status, without being influenced by the life-stage status of others. However, this assumption may be violated due to social learning, social pressure, social contagion, and social support mechanisms (Bernardi & Klärner, 2014; Ateca-Amestoy et al., 2014). For example, imagine that your best friend had a child with whom you interacted frequently after their birth. You saw how satisfied the child made your friend, so you decided to have your own child too. Thereafter, the birth of your own child increased your satisfaction with your life. Can you state in this situation that your subjective well-being was independent of the life-stage status of others? In reality, it was your own child that directly increased your level of satisfaction, but this child is the ‘outcome’ of your best friend’s fertility, thus your level of well-being was also affected by your friend’s life-stage status. This example illustrates how the first assumption of SUTVA is violated once respondents interact. The present dissertation tackles this problem by observing a sufficiently large sample, in which interactions are rare. However, future research on this topic should take into account the fact that individuals are embedded in social networks.

The second assumption of SUTVA implies that the effect of belonging to a given life stage should not depend on individuals’ characteristics. This prerequisite can be ensured by further specifying life events. The analysis in the first empirical study, for example, not only observes the general effect of parenthood, but also analyses the moderating effect of gender and parity. However, certain subgroups should have been further distinguished, but the sample size was too small for the dataset to be split. For example, it would have been interesting to distinguish the effects of motherhood and fatherhood by parity as well. Further, one could also observe the diverse effect of life events based on income or education level. Moreover, although single parenthood was observed in Chapter 6 using cross-sectional data, this effect could be also observed in longitudinal data to obtain more precise casual estimations. Finally, the effect of other atypical forms of parenthood, such as early or late parenthood, are also important issues for future research.

Moreover, the limited length (11 years) of the longitudinal study (*Turning Points of Life Course*) creates a barrier to following respondents long after their exposure to the treatments, although the effects of life events can change in the long term. For example, the first empirical study in this dissertation (Chapter 4) could only observe how parents benefit from having a child whose age ranged from 0-10 years old. This period of time enabled us to test whether children contribute to parental well-being as *consumption goods* (Becker & Barro, 1988). However, children can also be *investment goods* since they often provide financial and emotional support for their elderly parents when they grow up (Boldrin, De Nardi, & Jones, 2015; Caldwell, 1978; Caldwell, 1982; Leibenstein, 1957; Neher, 1971). Thus, one may only capture the entire effect of parenthood by observing this effect from the birth of the child until the end of the parents’ lives. However, this was not feasible, since there have not yet been any longitudinal studies that capture such a long period of time in Hungary. Therefore, extending the longitudinal research with further data would be beneficial for understanding this topic more in depth.

Also, the frequency of data collection in the longitudinal study created limitations. The longitudinal research at hand was repeated only every 3-4 years; furthermore, the third wave could not be used in the present analysis since at that time subjective well-being questions were not asked. However, previous research has shown that the effect of life events tends to change annually or even monthly (Clark et al., 2008; Mikucka, 2016; Pollmann-Schult, 2013). Thus, more frequent data collection in the future could also reveal unexpected patterns.

Another limitation is that we might not have been able to control for all possible confounding variables, which is otherwise essential for estimating causality. Although the research for this dissertation took steps to rule out confounding variables by applying matching on longitudinal data (see Chapters 4 and 5) which controls for all unobserved time-invariant variables and observable variables, even these studies could not rule out unobserved time-variant confounding variables. For example, Kravdal (2014) has argued that all previous estimations about the effect of parenthood were biased since none of these studies controlled for expectations about the effect of parenthood. This variable can cause selection bias since those who had expected to have an enjoyable parenthood were more likely to have a child than those who had not been looking forward to this event. Similarly, this argument can be applied to expectations about retirement which could distort estimations about the effect of retirement. Other confounding variables should also have been taken into account, but the dataset did not contain them. For example, work life balance should ideally have been controlled for to estimate the effect of parenthood on subjective well-being, since prior research has found that this variable influences the effect of parenthood (Bernardi et al., 2017; Matysiak et al., 2016). In order to tackle the problem of selection bias, this dissertation conducted sensitivity analyses (Rosenbaum, 2002). The majority of these indicated fairly robust findings, which means that the results would have been unlikely to change if further matching variables had been involved.

The problem of post-treatment variables also needs to be mentioned here. Imagine that we are interested in the effect of the first child using the available longitudinal dataset. As the observational period in this dataset was quite long (0-3 years in case of the short-term effect, but 7-11 years in case of the long-term effect), it is possible that in this period not only was a first child born, but the parents may also have divorced. In this case, observed changes in subjective well-being might have occurred due to the divorce (Amato, 2004; Johnson & Wu, 2002). On the one hand, divorce can be a consequence of having a child. For example, parenthood often decreases satisfaction with a partner, sometimes leading to divorce (Lawrence et al., 2008; Pollmann-Schult, 2013; Rollins & Cannon, 1974; Twenge et al., 2003; VanLaningham et al., 2001). In this case, divorce would be a collider variable, and controlling for it would cause endogenous selection bias. On the other hand, divorce can be also independent from having a child even if it occurs after childbirth. In this case, not controlling for divorce can cause selection bias. In general, statisticians have suggested not controlling for post-treatment variables as they are typically influenced by the treatment (Elwert & Winship, 2014; Rosenbaum & Rubin, 1984). Thus, in the research for this dissertation I followed statisticians' advice to control only for pre-treatment variables. However, I acknowledge the problem. Conducting a sensitivity analysis was also important due to these variables. Further, future research should also observe the sensitivity of these effects by controlling for post-treatment events as well.

Finally, this dissertation was not designed to capture all life events. In the future, other life events could be investigated with the same methodology. For example, future research could observe the effect of leaving the parental home on domain-specific subjective well-being. This research could help identify policies which should be created to improve young adults' lives.

Appendix

Table 26. Hypothetical situation to illustrate extrapolation

X explanatory variable	Y_{i1} outcome for treated units	Prediction for treated units	Y_{i0} outcome for control units	Prediction for control units
1		1.37237	0.1	0.84048
2		1.46786	0.11	0.93597
3		1.56335	0.12	1.03146
4		1.65884	0.13	1.12695
5		1.75433	0.14	1.22244
6	2.10	1.84982	2.30	1.31793
7	2.40	1.94531	2.40	1.41342
8	2.30	2.0408	2.70	1.50891
9	2.20	2.13629	2.60	1.6044
10	2.50	2.23178	2.00	1.69989
11	2.10	2.32727	2.10	1.79538
12	2.00	2.42276	2.50	1.89087
13	2.60	2.51825	2.20	1.98636
14	2.70	2.61374	2.30	2.08185
15	2.40	2.70923	2.40	2.17734
16	2.30	2.80472	2.10	2.27283
17		2.90021	2.5	2.36832
18		2.9957	2.49	2.46381
19		3.09119	2.48	2.5593
20		3.18668	2.47	2.65479
21		3.28217	2.46	2.75028
22		3.37766	2.45	2.84577
23		3.47316	2.44	2.94126

Table 27. Sample dataset to illustrate the “fundamental problem of causal inference”, distance matrices, matching and longitudinal analysis

ID	Satisfaction with housing	Age	Income	Treatment	Propensity score	Life satisfaction 1. wave	Life satisfaction 2. wave
1	4	26	55.24	1	0.66260	8	9
2	2	31	47.14	1	0.43815	7	8
3	2	25	55.24	1	0.70724	5	5
4	8	23	26.51	1	0.58176	7	7
5	4	26	35.35	1	0.54626	7	9
6	5	24	52.02	1	0.70686	10	9
7	7	27	144.51	1	0.93326	9	7
8	8	40	30.66	1	0.09913	7	7
9	3	31	32.05	1	0.34393	7	7
10	7	34	72.25	1	0.44364	7	8
11	8	22	39.31	1	0.69000	7	7
12	9	20	65.99	1	0.85081	7	5
13	8	23	64.10	1	0.77812	5	7
14	5	37	158.38	1	0.81436	6	8
15	8	36	30.17	1	0.16826	5	7
16	10	30	42.58	0	0.39814	10	9
17	8	48	43.35	0	0.04162	5	6
18	7	50	53.03	0	0.03981	5	2
19	8	38	28.28	0	0.12401	7	5
20	2	31	45.12	0	0.42596	5	6
21	10	30	42.58	0	0.39814	10	9
22	8	26	18.00	0	0.41458	7	8
23	7	58	202.31	0	0.32026	8	8
24	10	59	46.24	0	0.00796	7	8
25	8	62	26.51	0	0.00326	6	6
26	9	72	54.80	0	0.00135	10	8
27	10	60	41.04	0	0.00601	5	5
28	9	44	28.28	0	0.05155	6	6
29	4	26	55.24	0	0.66260	9	8
30	2	38	25.21	0	0.13299	3	3
31	7	35	63.58	0	0.35550	7	7
32	10	59	43.35	0	0.00742	8	8
33	9	28	34.47	0	0.43136	7	8
34	3	25	30.93	0	0.56415	10	8
35	10	30	30.66	0	0.33038	8	8
36	10	73	63.58	0	0.00140	10	10
37	9	60	33.14	0	0.00508	5	5
38	7	29	44.19	0	0.46493	8	8
39	2	22	18.10	0	0.60685	4	2
40	10	68	49.71	0	0.00216	8	8
41	2	31	47.14	0	0.43815	5	6

Table 28. Balance improvement in matching those to whom a child was born between 2003 and 2004/2005 (treatment group) with those to whom no children were born in this period (control group)

		Raw Data				Matched Data			
		Treatment		Control		Treatment		Control	
		Mea	SD	Mea	SD	Mea	SD	Mea	SD
Distance		0.21	0.14	0.04	0.08	0.21	0.14	0.20	0.14
Satisfaction with life		7.09	1.83	6.60	1.94	7.09	1.83	6.98	1.69
Recent perceived well-being		6.49	1.78	5.86	1.81	6.49	1.78	6.47	1.61
Sex	Male	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Female	0.53	0.50	0.59	0.49	0.53	0.50	0.50	0.50
Education	Primary or less	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Vocational secondary	0.32	0.47	0.29	0.45	0.32	0.47	0.33	0.47
	General secondary	0.34	0.47	0.31	0.46	0.34	0.47	0.32	0.47
	Tertiary	0.21	0.40	0.15	0.36	0.21	0.40	0.21	0.41
Satisfaction with housing		6.86	2.24	7.22	2.29	6.86	2.24	6.96	2.04
Age		27.5	5.01	45.9	14.2	27.5	5.01	27.9	5.94
Residence	Capital city	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Bigger city	0.24	0.43	0.22	0.41	0.24	0.43	0.21	0.40
	Smaller city	0.26	0.44	0.30	0.46	0.26	0.44	0.22	0.41
	Village	0.42	0.49	0.36	0.48	0.42	0.49	0.48	0.50
Subjective health status		8.41	1.55	7.00	2.33	8.41	1.55	8.57	1.39
Equivalent household		50.8	34.3	48.9	34.9	50.8	34.3	47.8	31.1
Labour market status	Employed	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Self-employed	0.07	0.26	0.06	0.23	0.07	0.26	0.05	0.22
	Unemployed	0.08	0.27	0.05	0.21	0.08	0.27	0.08	0.27
	Other non-working	0.18	0.39	0.38	0.49	0.18	0.39	0.18	0.38
Has ever experienced unemployment		0.49	0.50	0.50	0.47	0.49	0.49	0.50	0.50
Workplace	Owned by the state	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Private	0.50	0.06	0.33	0.47	0.50	0.06	0.52	0.50
	Non respond	0.27	0.50	0.44	0.50	0.27	0.50	0.25	0.43
Last (most important) work	Blue collar	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	White collar	0.38	0.49	0.37	0.48	0.38	0.49	0.39	0.49
Marital status	Single	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Married living together	0.57	0.50	0.66	0.47	0.57	0.50	0.61	0.49
	Married living apart	0.01	0.10	0.01	0.10	0.01	0.10	0.01	0.10
	Widow	0.00	0.00	0.08	0.28	0.00	0.00	0.00	0.00
	Divorced	0.07	0.25	0.09	0.29	0.07	0.25	0.03	0.17
Partner labour market status	Does not have partner	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Employed	0.46	0.50	0.35	0.48	0.46	0.50	0.49	0.50
	Self-employed	0.06	0.24	0.06	0.24	0.06	0.24	0.05	0.21
	Retired	0.01	0.1	0.21	0.41	0.01	0.1	0.00	0.00
	Unemployed	0.06	0.08	0.04	0.19	0.06	0.08	0.04	0.20
	Other non-working	0.15	0.24	0.05	0.23	0.15	0.24	0.14	0.35
	No answer	0.17	0.38	0.05	0.22	0.17	0.38	0.17	0.38
Satisfaction with partner	Does not have partner	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Dissatisfied	0.02	0.14	0.05	0.22	0.02	0.14	0.02	0.14
	Neutral	0.06	0.24	0.08	0.26	0.06	0.24	0.04	0.20
	Rather satisfied	0.28	0.45	0.26	0.44	0.28	0.45	0.24	0.43
	Very satisfied	0.43	0.50	0.34	0.47	0.43	0.50	0.48	0.50
	No answer	0.12	0.32	0.04	0.20	0.12	0.32	0.11	0.32
Number of children		0.82	1.00	1.62	1.12	0.82	1.00	0.81	0.95
Does not enjoy working	Completely disagree	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Disagree	0.25	0.43	0.21	0.41	0.25	0.43	0.19	0.39
	Rather agree	0.12	0.33	0.14	0.35	0.12	0.33	0.12	0.32
	Completely agree	0.04	0.19	0.05	0.22	0.04	0.19	0.03	0.16
Trust in the future	Completely disagree	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Disagree	0.07	0.25	0.10	0.29	0.07	0.25	0.06	0.24
	Rather agree	0.43	0.50	0.45	0.50	0.43	0.50	0.42	0.50
	Completely agree	0.50	0.50	0.41	0.49	0.50	0.50	0.51	0.50
Sample weights		0.93	0.20	0.99	0.20	0.93	0.20	0.94	0.18

Table 29. Regression models for matched data about the effect of having a child on change in life satisfaction (coefficients and significances[40])

		Short-term effect (change in life satisfaction between 2001/2002 and 2004/2005)		Long-term effect (change in life satisfaction between 2001/2002 and 2012/2013)	
(Intercept)		3.61	**	2.28	
Parenthood		**0.56**	*******	**0.39**	******
Satisfaction with life		-0.71	***	-0.83	***
Recent perceived well-being		-0.06		0.14	**
Sex	Male	(R)		(R)	
	Female	0.33	*	-0.22	
Education	Primary or less	(R)		(R)	
	Vocational secondary	0.28		0.19	
	General secondary	0.31		0.14	
	Tertiary	0.66	*	0.49	
Satisfaction with housing		0.06	*	0.05	
Age		-0.01		-0.02	
Residence	Capital city	(R)		(R)	
	Bigger city	-0.34		0.25	
	Smaller city	-0.27		0.25	
	Village	-0.32		0.02	
Subjective health status		0.09	*	0.06	
Equivalent household income		0.00		0.00	
Labour market status	Employed	(R)		(R)	
	Self-employed	0.15		0.37	
	Unemployed	0.53		0.48	
	Other non-working	0.79		0.20	
Has ever experienced unemployment		-0.13		-0.01	
Workplace	Owned by the state	(R)		(R)	
	Private	0.15		0.37	
	Non respond	-0.82		-0.17	
Last (most important) work	Blue collar	(R)		(R)	
	White collar	0.41	*	0.08	
Marital status	Single	(R)		(R)	
	Married living together	0.13		-0.27	
	Married living apart	-1.33	**	0.48	
	Divorced	-0.28		-0.16	
Partner labour market status	Does not have partner	(R)		(R)	
	Employed	0.17		0.77	
	Self-employed	0.62		0.46	
	Retired	-1.33		0.92	
	Unemployed	0.29		0.99	*
	Other non-working	0.13		0.15	
	No answer	0.15		0.19	
Satisfaction with partner	Does not have partner	(R)		(R)	
	Dissatisfied	-1.58	**	-1.02	
	Neutral	-0.12		-0.13	
	Rather satisfied	-0.20		-0.54	
	Very satisfied	0.11		-0.32	
Number of children		-0.10		0.23	*
Does not enjoy working	Completely disagree	(R)		(R)	
	Disagree	-0.21		0.11	
	Rather agree	0.10		0.20	
	Completely agree	0.30		-0.90	*
Trust in the future	Completely disagree	(R)		(R)	
	Disagree	0.92		2.23	*
	Rather agree	0.51		2.53	**
	Completely agree	0.54		2.78	**
Sample weights		0.21		-0.50	

[40] Levels of significance: ***<0.001, **<0.05, *<0.1

Table 30. Balance improvement in matching those to whom the first child was born between 2003 and 2004/2005 (treatment group) with those who remained childless (control group)

		Raw Data				Matched Data			
		Treatment		Control		Treatment		Control	
		Mea	SD	Mea	SD	Mea	SD	Mea	SD
Distance		0.45	0.20	0.10	0.15	0.45	0.20	0.42	0.23
Satisfaction with life		6.94	1.46	6.32	1.91	6.94	1.46	6.87	1.44
Recent perceived well-being		6.43	1.48	5.92	1.83	6.43	1.48	6.39	1.51
Sex	Male	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Female	0.54	0.50	0.43	0.50	0.54	0.50	0.55	0.50
Education	Primary or less	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Vocational secondary	0.38	0.48	0.31	0.46	0.38	0.48	0.37	0.48
	General secondary	0.42	0.49	0.37	0.48	0.42	0.49	0.49	0.50
	Tertiary	0.12	0.28	0.15	0.36	0.12	0.28	0.10	0.30
Satisfaction with housing		7.12	2.07	7.04	2.34	7.12	2.07	7.05	1.94
Age		25.8	4.23	36.7	15.1	25.8	4.23	26.0	5.61
Residence	Capital city	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Bigger city	0.22	0.41	0.21	0.41	0.22	0.41	0.19	0.40
	Smaller city	0.25	0.44	0.29	0.45	0.25	0.44	0.32	0.47
	Village	0.47	0.50	0.36	0.48	0.47	0.50	0.43	0.50
Subjective health status		8.47	1.61	7.55	2.28	8.47	1.61	8.49	1.49
Equivalent household		55.1	29.7	55.2	46.3	55.1	29.7	53.5	28.1
Labour market status	Employed	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Self-employed	0.10	0.21	0.06	0.23	0.10	0.21	0.08	0.28
	Unemployed	0.06	0.25	0.07	0.25	0.06	0.25	0.05	0.22
	Other non-working	0.04	0.38	0.24	0.42	0.04	0.38	0.02	0.15
Has ever experienced unemployment		0.50	0.50	0.50	0.49	0.50	0.50	0.50	0.50
Workplace	Owned by the state	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Private	0.68	0.50	0.43	0.50	0.68	0.50	0.68	0.47
	Non respond	0.10	0.43	0.31	0.46	0.10	0.43	0.07	0.26
Last (most important) work	Blue collar	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	White collar	0.29	0.44	0.38	0.49	0.29	0.44	0.34	0.47
Marital status	Single	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Married living together	0.34	0.40	0.19	0.39	0.34	0.40	0.35	0.48
	Married living apart	0.01	0.07	0.01	0.07	0.01	0.07	0.01	0.09
	Widow	0.00	0.00	0.05	0.22	0.00	0.00	0.00	0.00
	Divorced	0.02	0.18	0.05	0.22	0.02	0.18	0.00	0.00
Partner labour market status	Does not have partner	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Employed	0.43	0.43	0.14	0.35	0.43	0.43	0.43	0.50
	Self-employed	0.06	0.16	0.02	0.16	0.06	0.16	0.05	0.22
	Retired	0.00	0.00	0.09	0.29	0.00	0.00	0.00	0.00
	Unemployed	0.08	0.20	0.02	0.15	0.08	0.20	0.06	0.24
	Other non-working	0.04	0.06	0.01	0.12	0.04	0.06	0.04	0.19
	No answer	0.22	0.36	0.13	0.34	0.22	0.36	0.22	0.42
Satisfaction with partner	Does not have partner	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Dissatisfied	0.01	0.14	0.02	0.14	0.01	0.14	0.00	0.00
	Neutral	0.04	0.18	0.04	0.19	0.04	0.18	0.03	0.17
	Rather satisfied	0.23	0.29	0.10	0.30	0.23	0.29	0.24	0.43
	Very satisfied	0.33	0.38	0.14	0.35	0.33	0.38	0.32	0.47
	No answer	0.22	0.36	0.13	0.34	0.22	0.36	0.22	0.41
Does not enjoy working	Completely disagree	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Disagree	0.25	0.43	0.23	0.42	0.25	0.43	0.18	0.39
	Rather agree	0.16	0.37	0.17	0.37	0.16	0.37	0.19	0.39
	Completely agree	0.03	0.25	0.05	0.22	0.03	0.25	0.01	0.12
Trust in the future	Completely disagree	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Disagree	0.07	0.34	0.09	0.29	0.07	0.34	0.06	0.24
	Rather agree	0.50	0.50	0.44	0.50	0.50	0.50	0.49	0.50
	Completely agree	0.42	0.49	0.44	0.50	0.42	0.49	0.45	0.50
Sample weights		0.91	0.17	1.03	0.29	0.91	0.17	0.89	0.20

Table 31. Regression models for matched data about the effect of having a first child on change in life satisfaction (coefficients and significances[41])

		Short-term effect (change in life satisfaction between 2001/2002 and 2004/2005)		Long-term effect (change in life satisfaction between 2001/2002 and 2012/2013)	
(Intercept)		1.52		6.26	**
First child		**0.81**	*******	**0.54**	******
Satisfaction with life		-0.96	***	-0.71	***
Recent perceived well-being		0.18	**	-0.04	
Sex	Male	(R)		(R)	
	Female	0.28		-0.08	
Education	Primary or less	(R)		(R)	
	Vocational secondary	0.48		0.75	
	General secondary	0.61		0.43	
	Tertiary	1.08	*	1.13	
Satisfaction with housing		0.01		-0.03	
Age		0.02		-0.06	*
Residence	Capital city	(R)		(R)	
	Bigger city	0.43		1.04	*
	Smaller city	0.39		0.22	
	Village	0.93	*	0.53	
Subjective health status		0.09		0.01	
Equivalent household income		0.00		0.00	
Labour market status	Employed	(R)		(R)	
	Self-employed	-0.05		0.64	
	Unemployed	-1.21		-0.71	
	Other non-working	-3.02	*	-1.60	
Has ever experienced unemployment		0.15		-0.07	
Workplace	Owned by the state	(R)		(R)	
	Private	0.46		0.94	**
	Non respond	0.95		1.56	
Last (most important) work	Blue collar	(R)		(R)	
	White collar	0.09		0.27	
Marital status	Single	(R)		(R)	
	Married living together	-0.09		-0.39	
	Married living apart	-0.91		1.02	
	Divorced	-1.94	**	-0.70	
Partner labour market status	Does not have partner	(R)		(R)	
	Employed	0.11		0.79	
	Self-employed	0.47		0.88	
	Retired	-0.11		0.86	
	Unemployed	0.55		1.25	
	Other non-working	0.43		0.25	
Satisfaction with partner	Does not have partner	(R)		(R)	
	Dissatisfied	0.13		0.91	
	Neutral	0.56		0.28	
	Rather satisfied	-0.16		-0.62	
	Very satisfied	0.23		-0.95	
Does not enjoy working	Completely disagree	(R)		(R)	
	Disagree	-0.04		0.55	*
	Rather agree	0.05		0.10	
	Completely agree	1.08		-0.26	
Trust in the future	Completely disagree	(R)		(R)	
	Disagree	1.57		0.29	
	Rather agree	0.52		0.57	
	Completely agree	1.00		1.28	
Sample weights		0.42		-1.78	*

[41] Level of significance: ***<0.001, **<0.05, *<0.1

Table 32. Balance improvement in matching those to whom the second child was born between 2003 and 2004/2005 (treatment group) with those who had no additional child in this period but already had one from earlier (control group)

		Raw Data				Matched Data			
		Treatment		Control		Treatment		Control	
		Mea	SD	Mea	SD	Mea	SD	Mea	SD
Distance		0.47	0.27	0.04	0.11	0.47	0.27	0.39	0.23
Satisfaction with life		7.36	1.69	6.41	1.92	7.36	1.69	7.45	1.41
Recent perceived well-being		6.66	1.56	5.77	1.80	6.66	1.56	6.66	1.46
Sex	Male	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Female	0.60	0.49	0.64	0.48	0.60	0.49	0.61	0.49
Education	Primary or less	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Vocational secondary	0.07	0.26	0.24	0.43	0.07	0.26	0.04	0.19
	General secondary	0.31	0.46	0.28	0.45	0.31	0.46	0.26	0.44
	Tertiary	0.36	0.48	0.35	0.48	0.36	0.48	0.48	0.50
Satisfaction with housing		6.72	2.11	7.23	2.32	6.72	2.11	6.80	2.06
Age		28.1	3.51	49.6	13.7	28.1	3.51	29.2	5.38
Residence	Capital city	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Bigger city	0.38	0.49	0.25	0.43	0.38	0.49	0.34	0.48
	Smaller city	0.28	0.45	0.29	0.46	0.28	0.45	0.29	0.46
	Village	0.25	0.43	0.28	0.45	0.25	0.43	0.24	0.43
Subjective health status		8.41	1.27	6.68	2.33	8.41	1.27	8.64	1.51
Equivalent household		51.3	37.7	48.9	31.1	51.3	37.7	45.8	25.5
Labour market status	Employed	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Self-employed	0.06	0.24	0.06	0.24	0.06	0.24	0.04	0.19
	Unemployed	0.08	0.28	0.04	0.19	0.08	0.28	0.08	0.28
	Other non-working	0.32	0.47	0.44	0.50	0.32	0.47	0.32	0.47
Has ever experienced unemployment		0.51	0.50	0.31	0.46	0.51	0.50	0.53	0.50
Workplace	Owned by the state	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Private	0.39	0.49	0.30	0.46	0.39	0.49	0.32	0.47
	Non respond	0.42	0.50	0.48	0.50	0.42	0.50	0.40	0.49
Last (most important) work	Blue collar	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	White collar	0.53	0.50	0.42	0.49	0.53	0.50	0.55	0.50
Marital status	Single	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Married living together	0.82	0.38	0.68	0.47	0.82	0.38	0.78	0.42
	Married living apart	0.01	0.11	0.01	0.10	0.01	0.11	0.01	0.11
	Widow	0.00	0.00	0.10	0.30	0.00	0.00	0.00	0.00
	Divorced	0.04	0.19	0.15	0.36	0.04	0.19	0.01	0.11
Partner labour market status	Does not have partner	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Employed	0.53	0.50	0.36	0.48	0.53	0.50	0.62	0.49
	Self-employed	0.08	0.28	0.05	0.21	0.08	0.28	0.08	0.28
	Retired	0.01	0.11	0.27	0.44	0.01	0.11	0.00	0.00
	Unemployed	0.02	0.15	0.03	0.17	0.02	0.15	0.01	0.11
	Other non-working	0.18	0.38	0.04	0.20	0.18	0.38	0.13	0.34
	No answer	0.13	0.34	0.03	0.18	0.13	0.34	0.13	0.34
Satisfaction with partner	Does not have partner	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Dissatisfied	0.04	0.19	0.05	0.22	0.04	0.19	0.01	0.11
	Neutral	0.09	0.29	0.09	0.28	0.09	0.29	0.12	0.33
	Rather satisfied	0.36	0.48	0.26	0.44	0.36	0.48	0.32	0.47
	Very satisfied	0.46	0.50	0.36	0.48	0.46	0.50	0.47	0.50
	No answer	0.02	0.15	0.03	0.16	0.02	0.15	0.06	0.24
Does not enjoy working	Completely disagree	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Disagree	0.27	0.45	0.19	0.40	0.27	0.45	0.29	0.46
	Rather agree	0.06	0.24	0.14	0.34	0.06	0.24	0.06	0.24
	Completely agree	0.00	0.00	0.05	0.23	0.00	0.00	0.00	0.00
Trust in the future	Completely disagree	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Disagree	0.08	0.28	0.11	0.32	0.08	0.28	0.02	0.15
	Rather agree	0.35	0.48	0.45	0.50	0.35	0.48	0.35	0.48
	Completely agree	0.56	0.50	0.39	0.49	0.56	0.50	0.62	0.49
Sample weights		0.91	0.24	0.99	0.21	0.91	0.24	0.93	0.20

Table 33. Regression models for matched data about the effect of having a second child on change in life satisfaction (coefficients and significances[42])

		Short-term effect (change in life satisfaction between 2001/2002 and 2004/2005)		Long-term effect (change in life satisfaction between 2001/2002 and 2012/2013)	
(Intercept)		5.15	**	2.99	
Second child		**0.51**	**	**0.69**	**
Satisfaction with life		-0.74	***	-0.99	***
Recent perceived well-being		0.04		0.07	
Sex	Male	(R)		(R)	
	Female	0.23		0.91	**
Education	Primary or less	(R)		(R)	
	Vocational secondary	-0.21		-0.36	
	General secondary	-0.69		-0.32	
	Tertiary	-0.43		0.04	
Satisfaction with housing		0.01		0.07	
Age		0.01		-0.01	
Residence	Capital city	(R)		(R)	
	Bigger city	0.35		0.50	
	Smaller city	0.43		1.47	**
	Village	-0.44		-0.12	
Subjective health status		-0.14		0.04	
Equivalent household income		0.00		0.00	
Labour market status	Employed	(R)		(R)	
	Self-employed	0.49		-0.19	
	Unemployed	-1.53		-0.23	
	Other non-working	-0.86		0.44	
Has ever experienced unemployment		0.03		0.01	
Workplace	Owned by the state	(R)		(R)	
	Private	0.08		0.26	
	Non respond	0.92		-1.42	
Last (most important) work	Blue collar	(R)		(R)	
	White collar	0.57		0.02	
Marital status	Single	(R)		(R)	
	Married living together	0.99	**	0.48	
	Married living apart	-0.50		0.90	
	Divorced	1.09		1.68	*
Partner labour market status	Does not have partner	(R)		(R)	
	Employed	-1.80		0.96	
	Self-employed	-1.12		1.34	
	Retired	-0.73		0.20	
	Unemployed	-4.87	**	3.00	
	Other non-working	-2.44		1.77	
Satisfaction with partner	Does not have partner	(R)		(R)	
	Dissatisfied	-0.47		0.23	
	Neutral	1.31		0.33	
	Rather satisfied	2.11	**	0.84	
	Very satisfied	2.70	**	1.53	*
Does not enjoy working	Completely disagree	(R)		(R)	
	Disagree	0.49		0.67	**
	Rather agree	-0.06		0.65	
Trust in the future	Completely disagree	(R)		(R)	
	Rather agree	-0.76		0.63	
	Completely agree	-0.42		0.92	
Sample weights		0.57		0.15	

[42] Level of significance: ***<0.001, **<0.05, *<0.1

Table 34. Balance improvement in matching women to whom a child was born between 2003 and 2004/2005 (treatment group) with women to whom no children were born in this period (control group)

		Raw Data				Matched Data			
		Treatment		Control		Treatment		Control	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Distance		0.42	0.26	0.03	0.09	0.42	0.26	0.36	0.25
Satisfaction with life		7.37	1.78	6.59	1.95	7.37	1.78	7.19	1.82
Recent perceived well-being		6.71	1.69	5.86	1.82	6.71	1.69	6.56	1.53
Education	Primary or less	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Vocational secondary	0.28	0.45	0.20	0.40	0.28	0.45	0.28	0.45
	General secondary	0.37	0.49	0.35	0.48	0.37	0.49	0.38	0.49
	Tertiary	0.23	0.42	0.15	0.35	0.23	0.42	0.18	0.39
Satisfaction with housing		7.06	2.36	7.30	2.27	7.06	2.36	7.20	2.17
Age		26.6	3.89	48.4	12.9	26.6	3.89	27.7	6.44
Residence	Capital city	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Bigger city	0.24	0.43	0.23	0.42	0.24	0.43	0.20	0.40
	Smaller city	0.23	0.42	0.30	0.46	0.23	0.42	0.30	0.46
	Village	0.45	0.50	0.35	0.48	0.45	0.50	0.43	0.50
Subjective health status		8.31	1.52	6.73	2.35	8.31	1.52	8.37	1.58
Equivalent household		51.1	29.3	47.1	29.7	51.1	29.3	43.0	21.9
Labour market status	Employed	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Self-employed	0.04	0.20	0.03	0.18	0.04	0.20	0.02	0.14
	Unemployed	0.06	0.24	0.04	0.19	0.06	0.24	0.06	0.23
	Other non-working	0.30	0.46	0.46	0.50	0.30	0.46	0.33	0.47
Has ever experienced unemployment		0.45	0.50	0.50	0.46	0.45	0.45	0.50	0.50
Workplace	Owned by the state	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Private	0.43	0.50	0.25	0.43	0.43	0.50	0.41	0.49
	Non respond	0.37	0.48	0.51	0.50	0.37	0.48	0.38	0.49
Last (most important) work	Blue collar	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	White collar	0.47	0.50	0.44	0.50	0.47	0.50	0.44	0.50
Marital status	Single	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Married living together	0.56	0.50	0.66	0.47	0.56	0.50	0.60	0.49
	Married living apart	0.02	0.13	0.01	0.11	0.02	0.13	0.02	0.14
	Widow	0.00	0.00	0.13	0.34	0.00	0.00	0.00	0.00
	Divorced	0.06	0.24	0.12	0.32	0.06	0.24	0.06	0.23
Partner labour market status	Does not have partner	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Employed	0.59	0.50	0.34	0.47	0.59	0.50	0.61	0.49
	Self-employed	0.10	0.30	0.08	0.28	0.10	0.30	0.04	0.20
	Retired	0.01	0.08	0.24	0.43	0.01	0.08	0.01	0.08
	Unemployed	0.07	0.25	0.03	0.18	0.07	0.25	0.07	0.25
	Other non-working	0.02	0.13	0.01	0.10	0.02	0.13	0.02	0.14
	No answer	0.14	0.34	0.03	0.17	0.14	0.34	0.14	0.34
Satisfaction with partner	Does not have partner	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Dissatisfied	0.03	0.17	0.07	0.26	0.03	0.17	0.07	0.26
	Neutral	0.06	0.24	0.08	0.28	0.06	0.24	0.06	0.23
	Rather satisfied	0.32	0.47	0.26	0.44	0.32	0.47	0.26	0.44
	Very satisfied	0.37	0.48	0.30	0.46	0.37	0.48	0.37	0.49
	No answer	0.14	0.34	0.03	0.18	0.14	0.34	0.14	0.34
Number of children		0.75	0.90	1.83	1.03	0.75	0.90	0.92	0.93
Does not enjoy working	Completely disagree	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Disagree	0.25	0.44	0.20	0.40	0.25	0.44	0.09	0.29
	Rather agree	0.08	0.27	0.13	0.34	0.08	0.27	0.08	0.27
	Completely agree	0.02	0.13	0.05	0.22	0.02	0.13	0.01	0.08
Trust in the future	Completely disagree	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Disagree	0.07	0.25	0.10	0.30	0.07	0.25	0.06	0.24
	Rather agree	0.47	0.50	0.46	0.50	0.47	0.50	0.44	0.50
	Completely agree	0.45	0.50	0.40	0.49	0.45	0.50	0.48	0.50
Sample weights		0.86	0.17	0.97	0.19	0.86	0.17	0.87	0.17

Table 35. Regression models for matched data about the effect of having a child on change in life satisfaction among women (coefficients and significances[43])

		Short-term effect (change in life satisfaction between 2001/2002 and 2004/2005)		Long-term effect (change in life satisfaction between 2001/2002 and 2012/2013)	
(Intercept)		3.57	**	5.08	**
Motherhood		**0.64**	*******	**0.48**	******
Satisfaction with life		-0.67	***	-0.73	***
Recent perceived well-being		0.10		0.07	
Education	Primary or less	(R)		(R)	
	Vocational secondary	0.15		-0.38	
	General secondary	-0.74	*	-0.86	*
	Tertiary	-0.43		-0.82	
Satisfaction with housing		-0.05		-0.02	
Age		-0.01		-0.06	*
Residence	Capital city	(R)		(R)	
	Bigger city	-0.78	**	0.45	
	Smaller city	-0.85	**	-0.02	
	Village	-0.71	*	-0.29	
Subjective health status		0.08		0.02	
Equivalent household income		0.00		0.00	
Labour market status	Employed	(R)		(R)	
	Self-employed	-0.05		-0.72	
	Unemployed	0.57		-0.26	
	Other non-working	1.00		0.67	
Has ever experienced unemployment		-0.14		-0.14	
Workplace	Owned by the state	(R)		(R)	
	Private	0.12		0.48	
	Non respond	-1.03		-0.68	
Last (most important) work	Blue collar	(R)		(R)	
	White collar	0.66	**	0.79	**
Marital status	Single	(R)		(R)	
	Married living together	0.52	*	0.31	
	Married living apart	-1.44	**	0.73	
	Widow	-0.17		0.02	
	Divorced	0.52	*	0.31	
Partner labour market status	Does not have partner	(R)		(R)	
	Employed	-0.03		-0.36	
	Self-employed	0.30		-0.14	
	Retired	-0.05		-0.72	
	Unemployed	0.03		-0.81	
	Other non-working	-1.04		0.50	
	Does not answer	0.30		-0.50	
Satisfaction with partner	Does not have partner	(R)		(R)	
	Dissatisfied	-1.67	**	-0.89	
	Neutral	-0.67		0.34	
	Rather satisfied	-0.41	*	-0.04	
	Very satisfied	1.41		-1.11	
Number of children		-0.11		0.23	
Does not enjoy working	Completely disagree	(R)		(R)	
	Disagree	-0.49	*	-0.03	
	Rather agree	0.59	*	0.20	
	Completely agree	-0.98		-3.01	**
Trust in the future	Completely disagree	(R)		(R)	
	Disagree	3.21	***	3.21	**
	Rather agree	1.76	**	2.47	**
	Completely agree	1.98	**	3.04	**
Sample weights		-0.08		-0.84	

[43] Level of significance: ***<0.001, **<0.05, *<0.1

Table 36. Balance improvement in matching men to whom a child was born between 2003 and 2004/2005 (treatment group) with men to whom no children were born in this period (control group)

		Raw Data				Matched Data			
		Treatment		Control		Treatment		Control	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Distance		0.32	0.22	0.04	0.09	0.32	0.22	0.30	0.22
Satisfaction with life		6.81	1.84	6.55	1.89	6.81	1.84	6.72	1.54
Recent perceived well-being		6.23	1.83	5.80	1.79	6.23	1.83	6.29	1.71
Education	Primary or less	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Vocational secondary	0.37	0.48	0.41	0.49	0.37	0.48	0.39	0.49
	General secondary	0.32	0.47	0.25	0.43	0.32	0.47	0.26	0.44
	Tertiary	0.18	0.39	0.15	0.36	0.18	0.39	0.20	0.41
Satisfaction with housing		6.62	2.11	7.22	2.26	6.62	2.11	6.93	2.01
Age		28.8	5.90	46.8	13.7	28.8	5.90	29.9	9.08
Residence	Capital city	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Bigger city	0.28	0.45	0.21	0.41	0.28	0.45	0.24	0.43
	Smaller city	0.30	0.46	0.29	0.45	0.30	0.46	0.31	0.46
	Village	0.35	0.48	0.38	0.49	0.35	0.48	0.39	0.49
Subjective health status		8.50	1.59	7.13	2.27	8.50	1.59	8.50	1.63
Equivalent household		51.2	39.7	50.9	42.7	51.2	39.7	48.2	30.0
Labour market status	Employed	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Self-employed	0.12	0.32	0.10	0.30	0.12	0.32	0.07	0.26
	Unemployed	0.08	0.27	0.06	0.23	0.08	0.27	0.08	0.27
	Other non-working	0.05	0.21	0.29	0.46	0.05	0.21	0.05	0.21
Has ever experienced unemployment		0.53	0.50	0.50	0.48	0.53	0.50	0.50	0.50
Workplace	Owned by the state	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Private	0.61	0.49	0.42	0.49	0.61	0.49	0.59	0.49
	Non respond	0.14	0.35	0.36	0.48	0.14	0.35	0.13	0.33
Last (most important) work	Blue collar	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	White collar	0.29	0.46	0.26	0.44	0.29	0.46	0.32	0.47
Marital status	Single	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Married living together	0.61	0.49	0.72	0.45	0.61	0.49	0.66	0.48
	Married living apart	0.00	0.00	0.01	0.08	0.00	0.00	0.00	0.00
	Widow	0.00	0.00	0.03	0.16	0.00	0.00	0.00	0.00
	Divorced	0.06	0.24	0.07	0.25	0.06	0.24	0.04	0.20
Partner labour market status	Does not have partner	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Employed	0.32	0.47	0.38	0.48	0.32	0.47	0.36	0.48
	Self-employed	0.02	0.13	0.03	0.18	0.02	0.13	0.02	0.13
	Retired	0.01	0.09	0.21	0.41	0.01	0.09	0.01	0.09
	Unemployed	0.05	0.21	0.04	0.20	0.05	0.21	0.02	0.15
	Other non-working	0.28	0.45	0.11	0.31	0.28	0.45	0.31	0.46
	No answer	0.21	0.35	0.05	0.22	0.21	0.35	0.17	0.37
Satisfaction with partner	Does not have partner	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Dissatisfied	0.01	0.09	0.03	0.18	0.01	0.09	0.01	0.09
	Neutral	0.05	0.21	0.07	0.25	0.05	0.21	0.04	0.20
	Rather satisfied	0.24	0.43	0.26	0.44	0.24	0.43	0.22	0.42
	Very satisfied	0.50	0.50	0.42	0.49	0.50	0.50	0.51	0.50
	No answer	0.10	0.30	0.04	0.19	0.10	0.30	0.10	0.30
Number of children		0.94	1.15	1.61	1.07	0.94	1.15	0.98	0.92
Does not enjoy working	Completely disagree	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Disagree	0.26	0.44	0.22	0.42	0.26	0.44	0.22	0.42
	Rather agree	0.15	0.36	0.15	0.35	0.15	0.36	0.20	0.40
	Completely agree	0.06	0.24	0.05	0.22	0.06	0.24	0.05	0.21
Trust in the future	Completely disagree	(R)	(R)	(R)	(R)	(R)	(R)	(R)	(R)
	Disagree	0.06	0.24	0.10	0.30	0.06	0.24	0.03	0.18
	Rather agree	0.38	0.49	0.46	0.50	0.38	0.49	0.40	0.49
	Completely agree	0.55	0.50	0.41	0.49	0.55	0.50	0.57	0.50
Sample weights		1.00	0.21	1.04	0.20	1.00	0.21	0.99	0.20

Table 37. Regression models for matched data about the effect of having a child on change in life satisfaction among men (coefficients and significances[44])

		Short-term effect (change in life satisfaction between 2001/2002 and 2004/2005)	Long-term effect (change in life satisfaction between 2001/2002 and 2012/2013)
(Intercept)		4.08	-1.12
Fatherhood		**0.46 ****	**0.28**
Satisfaction with life		-0.71 ***	-0.78 ***
Recent perceived well-being		-0.12	0.14
Education	Primary or less	(R)	(R)
	Vocational secondary	0.50	-0.36
	General secondary	0.44	0.09
	Tertiary	1.07	0.89
Satisfaction with housing		0.10	0.01
Age		0.00	0.00
Residence	Capital city	(R)	(R)
	Bigger city	0.28	0.78
	Smaller city	0.33	0.65
	Village	-0.08	0.65
Subjective health status		0.07	0.07
Equivalent household income		0.00	0.00
Labour market status	Employed	(R)	(R)
	Self-employed	0.42	0.26
	Unemployed	-1.78	0.51
	Other non-working	-2.10	-0.35
Has ever experienced unemployment		0.19	0.06
Workplace	Owned by the state	(R)	(R)
	Private	0.29	0.09
	Non respond	1.11	-0.91
Last (most important) work	Blue collar	(R)	(R)
	White collar	0.19	-0.77
Marital status	Single	(R)	(R)
	Married living together	-0.22	-0.56
	Divorced	0.03	-0.64
Partner labour market status	Does not have partner	(R)	(R)
	Employed	0.68	1.19 *
	Self-employed	0.93	2.19 *
	Retired	1.02	5.35 **
	Unemployed	1.00	2.47 **
	Other non-working	1.14 .	0.91
	No answer	0.47	0.79
Satisfaction with partner	Does not have partner	(R)	(R)
	Dissatisfied	-1.16	-1.45
	Neutral	-0.20	-0.83
	Rather satisfied	-0.04	-1.00
	Very satisfied	0.30	-0.63
Number of children		-0.10	0.21
Does not enjoy working	Completely disagree	(R)	(R)
	Disagree	0.46	0.33
	Rather agree	0.23	0.04
	Completely agree	-0.04	-0.58
Trust in the future	Completely disagree	(R)	(R)
	Disagree	-0.50	4.09 **
	Rather agree	-0.92	3.93 **
	Completely agree	-0.51	4.28 **
Sample weights		-0.18	0.59

[44] Level of significance: ***<0.001, **<0.05, *<0.1

Table 38. Balance improvement in matching those who had retired between 2001/2002 and 2004/2005 (treatment group) with those who did not (control group)

		Raw Data				Matched data	
		Treatment		Control		Control	
		Mean	SD	Mean	SD	Mean	SD
Distance		0.39	0.24	0.05	0.10	0.31	0.21
Satisfaction with life		6.38	1.96	6.76	1.88	6.45	1.63
Recent perceived well-being		5.60	1.87	6.03	1.76	5.68	1.57
Sex	Male	(R)	(R)	(R)	(R)	(R)	(R)
	Female	0.50	0.50	0.56	0.50	0.50	0.50
Education	Primary or less	(R)	(R)	(R)	(R)	(R)	(R)
	Vocational secondary	0.30	0.46	0.32	0.47	0.30	0.46
	General secondary	0.30	0.46	0.36	0.48	0.30	0.46
	Tertiary	0.14	0.35	0.17	0.38	0.14	0.35
Satisfaction with housing		7.26	2.33	7.02	2.26	7.21	2.07
Age		52.70	6.56	35.95	10.16	50.70	6.00
Residence	Capital city	(R)	(R)	(R)	(R)	(R)	(R)
	Bigger city	0.20	0.40	0.22	0.41	0.20	0.40
	Smaller city	0.33	0.47	0.29	0.45	0.27	0.45
	Village	0.34	0.47	0.38	0.49	0.39	0.49
Subjective health status		6.34	2.26	7.84	1.90	6.43	1.84
Equivalent household income		54.47	48.87	49.91	37.79	53.38	36.89
Labour market status	Employed	(R)	(R)	(R)	(R)	(R)	(R)
	Self-employed	0.12	0.32	0.08	0.27	0.07	0.25
	Unemployed	0.07	0.26	0.07	0.25	0.05	0.22
	Other non-working	0.09	0.29	0.14	0.34	0.06	0.22
Has ever experienced unemployment		0.36	0.44	0.48	0.43	0.49	0.33
Workplace	Owned by the state	(R)	(R)	(R)	(R)	(R)	(R)
	Private	0.47	0.50	0.47	0.50	0.42	0.49
	Non respond	0.18	0.39	0.21	0.41	0.13	0.33
Last (most important) work	Blue collar	(R)	(R)	(R)	(R)	(R)	(R)
	White collar	0.31	0.46	0.39	0.49	0.33	0.47
	Never had. no respond	0.01	0.09	0.00	0.03	0.00	0.00
Marital status	Single	(R)	(R)	(R)	(R)	(R)	(R)
	Married living together	0.74	0.44	0.64	0.48	0.80	0.40
	Married living apart	0.02	0.13	0.01	0.09	0.01	0.12
	Widow	0.09	0.29	0.02	0.15	0.09	0.29
	Divorced	0.11	0.32	0.08	0.28	0.06	0.25
Partner labour market status	Does not have partner	(R)	(R)	(R)	(R)	(R)	(R)
	Employed	0.37	0.48	0.45	0.50	0.41	0.49
	Self-employed	0.05	0.22	0.08	0.27	0.05	0.22
	Retired	0.27	0.45	0.06	0.24	0.28	0.45
	Unemployed	0.04	0.19	0.05	0.21	0.03	0.17
	Other non-working	0.05	0.22	0.14	0.26	0.03	0.17
	Do not answer	0.03	0.17	0.08	0.27	0.02	0.13
Satisfaction with the partner	Does not have partner	(R)	(R)	(R)	(R)	(R)	(R)
	Dissatisfied	0.06	0.24	0.05	0.21	0.06	0.24
	Neutral	0.08	0.27	0.08	0.27	0.07	0.26
	Rather satisfied	0.28	0.45	0.28	0.45	0.30	0.46
	Very satisfied	0.36	0.48	0.32	0.47	0.38	0.49
	Does not answer	0.03	0.17	0.06	0.24	0.01	0.12
Number of female children		0.99	0.99	0.88	1.10	1.00	0.87
Number of male children		1.04	1.01	0.95	1.12	1.06	0.92
Number of grandchildren		0.06	0.24	0.02	0.15	0.05	0.22
Does not enjoy working	Completely disagree	(R)	(R)	(R)	(R)	(R)	(R)
	Disagree	0.19	0.39	0.22	0.42	0.17	0.38
	Rather agree	0.16	0.36	0.15	0.35	0.16	0.36
	Completely agree	0.06	0.25	0.04	0.20	0.06	0.24
Trust in the future	Completely disagree	(R)	(R)	(R)	(R)	(R)	(R)
	Disagree	0.10	0.30	0.08	0.27	0.09	0.28
	Rather agree	0.52	0.50	0.46	0.50	0.60	0.49
	Completely agree	0.32	0.47	0.44	0.50	0.30	0.46
Sample weights		1.02	0.20	1.00	0.21	1.00	0.16

Table 39. Balance improvement in matching those who retired voluntary between 2001/2002 and 2004/2005 (treatment group) with those who retired involuntary (control group)

		Raw Data				Matched data	
		Treatment		Control		Control	
		Mean	SD	Mean	SD	Mean	SD
Distance		0.68	0.22	0.35	0.25	0.63	0.23
Satisfaction with life		6.80	1.86	6.07	1.99	6.59	1.84
Recent perceived well-being		5.96	1.69	5.32	1.89	6.03	1.48
Sex	Male	(R)	(R)	(R)	(R)	(R)	(R)
	Female	0.47	0.50	0.53	0.50	0.55	0.50
Education	Primary or less	(R)	0.37	(R)	(R)	(R)	(R)
	Vocational secondary	0.24	0.43	0.35	0.48	0.30	0.46
	General secondary	0.34	0.47	0.28	0.45	0.32	0.47
	Tertiary	0.24	0.43	0.07	0.25	0.22	0.41
Satisfaction with housing		7.54	2.17	7.07	2.35	7.46	2.23
Age		55.42	4.01	51.20	6.86	54.40	4.20
Residence	Capital city	(R)	(R)	(R)	(R)	(R)	(R)
	Bigger city	0.25	0.43	0.16	0.36	0.22	0.42
	Smaller city	0.30	0.46	0.35	0.48	0.23	0.42
	Village	0.24	0.43	0.41	0.49	0.29	0.46
Subjective health status		7.06	2.03	5.72	2.29	7.22	1.74
Equivalent household income		68.02	64.09	43.11	21.04	58.25	25.01
Labour market status	Employed	(R)	(R)	(R)	(R)	(R)	(R)
	Self-employed	0.13	0.34	0.09	0.28	0.07	0.26
	Unemployed	0.01	0.11	0.12	0.33	0.01	0.11
	Other non-working	0.06	0.24	0.11	0.30	0.07	0.26
Has ever experienced unemployment		0.25	0.43	0.46	0.50	0.30	0.46
Workplace	Owned by the state	(R)	(R)	(R)	(R)	(R)	(R)
	Private	0.53	0.50	0.42	0.50	0.57	0.50
	Non respond	0.09	0.29	0.24	0.43	0.08	0.28
Last (most important) work	Blue collar	(R)	(R)	(R)	(R)	(R)	(R)
	White collar	0.44	0.50	0.23	0.42	0.39	0.49
	Never had. no respond	0.01	0.08	0.01	0.12	0.00	0.00
Marital status	Single	(R)	(R)	(R)	(R)	(R)	(R)
	Married living together	0.77	0.42	0.70	0.46	0.80	0.40
	Married living apart	0.01	0.11	0.02	0.14	0.01	0.11
	Widow	0.10	0.30	0.09	0.28	0.09	0.29
	Divorced	0.09	0.29	0.14	0.35	0.09	0.30
Partner labour market status	Does not have partner	(R)	(R)	(R)	(R)	(R)	(R)
	Employed	0.35	0.48	0.36	0.48	0.38	0.49
	Self-employed	0.07	0.26	0.05	0.21	0.06	0.23
	Retired	0.33	0.47	0.22	0.41	0.34	0.48
	Unemployed	0.02	0.14	0.05	0.23	0.01	0.08
	Other non-working	0.03	0.22	0.08	0.26	0.02	0.14
	Do not answer	0.02	0.17	0.01	0.08	0.00	0.00
Satisfaction with the partner	Does not have partner	(R)	(R)	(R)	(R)	(R)	(R)
	Dissatisfied	0.06	0.23	0.05	0.23	0.03	0.18
	Neutral	0.07	0.26	0.07	0.26	0.04	0.19
	Rather satisfied	0.30	0.46	0.25	0.43	0.40	0.49
	Very satisfied	0.37	0.49	0.37	0.48	0.34	0.48
	Does not answer	0.02	0.49	0.03	0.16	0.00	0.00
Number of female children		0.97	0.96	0.99	0.95	0.87	0.78
Number of male children		1.04	0.99	1.06	1.04	0.92	0.74
Number of grandchildren		0.13	0.34	0.09	0.29	0.13	0.34
Does not enjoy working	Completely disagree	(R)	(R)	(R)	(R)	(R)	(R)
	Disagree	0.14	0.31	0.24	0.43	0.11	0.32
	Rather agree	0.11	0.31	0.20	0.40	0.07	0.26
	Completely agree	0.08	0.28	0.05	0.21	0.06	0.25
Trust in the future	Completely disagree	(R)	(R)	(R)	(R)	(R)	(R)
	Disagree	0.10	0.30	0.11	0.31	0.06	0.23
	Rather agree	0.48	0.50	0.54	0.50	0.46	0.50
	Completely agree	0.36	0.48	0.28	0.45	0.44	0.50
Sample weights		0.98	0.17	1.04	0.22	0.95	0.18

Table 40. Regression models about the effect of retirement on change in life satisfaction between 2001/2002 and 2004/2005 (coefficients and significances[45])

		Raw data correlation	Raw data regression adjustment	Matched data regression adjustment
(Intercept)		0.04	3.79 ***	2.46 **
Retirement		**0.13**	**0.07**	**-0.07**
Satisfaction with life			-0.74 ***	-0.85 ***
Recent perceived well-being			0.08 ***	0.12 **
Sex	Male		(R)	(R)
	Female		0.15 **	-0.02
Education	Primary or less		(R)	(R)
	Vocational secondary		0.22 **	0.11
	General secondary		0.22 **	0.31
	Tertiary		0.50 ***	0.58
Satisfaction with housing			0.06 ***	0.11 **
Age			-0.01 **	0.01
Residence	Capital city		(R)	(R)
	Bigger city		0.00	0.19
	Smaller city		-0.08	0.12
	Village		-0.06	0.11
Subjective health status			0.04 **	0.04
Equivalent household income			0.01 ***	0.01 *
Labour market status	Employed		(R)	(R)
	Self-employed		0.00	-0.49
	Unemployed		-0.05	0.30
	Other non-working		0.13	0.48
Has ever experienced unemployment			-0.15 **	-0.48 **
Workplace	Owned by the state		(R)	(R)
	Private		0.01	0.02
	Non respond		-0.13	-0.27
Last (most important) work	Blue collar		(R)	(R)
	White collar		0.06	-0.24
	Never had. no respond		0.97 *	0.49 **
Marital status	Single		(R)	(R)
	Married living together		0.17 *	-0.32
	Married living apart		-0.06	0.42
	Widow		-0.20	0.14
	Divorced		-0.18 *	-0.21
Partner labour market status	Does not have partner		(R)	(R)
	Employed		0.11	1.63 **
	Self-employed		0.09	1.17
	Retired		0.02	1.37 *
	Unemployed		-0.05	1.58 *
	Other non-working		0.12	1.70 **
	Does not answer		0.13	0.42
Satisfaction with the partner	Does not have partner		(R)	(R)
	Dissatisfied		-0.30	-0.70
	Neutral		-0.03	-0.74
	Rather satisfied		0.10	-0.85
	Very satisfied		0.32	-0.43
Number of female children			0.01	0.23 **
Number of male children			0.01	0.12
Number of grandchildren			0.32 **	0.27
Does not enjoy working	Completely disagree		(R)	(R)
	Disagree		-0.02	0.20
	Rather agree		-0.05	0.07
	Completely agree		0.02	0.05
Trust in the future	Completely disagree		(R)	(R)
	Disagree		-0.03	-0.40
	Rather agree		0.10	-0.33
	Completely agree		0.22	-0.16
Sample weights			0.04	0.34

[45] The level of significance: ***<0.001, **<0.05, *<0.1

Table 41. Regression models about the effect of retirement on change in life satisfaction between 2001/2002 and 2012/2013 (coefficients and significances[46])

		Raw data correlation		Raw data regression adjustment		Matched data regression adjustment	
(Intercept)		0.34	***	5.27	***	4.72	***
Retirement		**0.14**		**-0.01**		**-0.16**	
Satisfaction with life				-0.81	***	-0.83	***
Recent perceived well-being				0.08	***	-0.02	
Sex	Male			(R)		(R)	
	Female			-0.04		0.08	
Education	Primary or less			(R)		(R)	
	Vocational secondary			0.05		0.19	
	General secondary			0.06		0.66	**
	Tertiary			0.23	*	0.54	
Satisfaction with housing				0.08	***	0.10	**
Age				0.00		0.03	**
Residence	Capital city			(R)		(R)	
	Bigger city			-0.09		0.03	
	Smaller city			-0.11		0.07	
	Village			-0.08		0.03	
Subjective health status				0.05	**	0.07	
Equivalent household income				0.01	***	0.00	
Labour market status	Employed			(R)		(R)	
	Self-employed			-0.04		0.06	
	Unemployed			-0.33		0.21	
	Other non-working			-0.35	*	0.09	**
Has ever experienced unemployment				-0.15	**	-0.63	
Workplace	Owned by the state			(R)		(R)	
	Private			0.10		0.39	
	Non respond			0.13		0.05	**
Last (most important) work	Blue collar			(R)		(R)	
	White collar			0.09		0.22	
	Never had. no respond			-1.11		-0.54	***
Marital status	Single			(R)		(R)	
	Married living together			-0.11		0.19	
	Married living apart			-0.13		-0.72	
	Widow			-0.52	**	-0.43	
	Divorced			-0.45	***	-0.36	
Partner labour market status	Does not have partner			(R)		(R)	
	Employed			0.15		-0.07	
	Self-employed			0.15		-0.17	
	Retired			-0.14		-0.34	
	Unemployed			0.19		-0.11	
	Other non-working			-0.20		0.54	
	Does not answer			0.12		0.20	
Satisfaction with the partner	Does not have partner			(R)		(R)	
	Dissatisfied			-0.35		0.78	
	Neutral			-0.11		0.09	
	Rather satisfied			0.15		0.05	
	Very satisfied			0.35		0.71	
Number of female children				0.05	*	0.17	*
Number of male children				0.06	*	0.02	
Number of grandchildren				0.22		-0.09	
Does not enjoy working	Completely disagree			(R)		(R)	
	Disagree			-0.06		-0.20	
	Rather agree			-0.12		0.04	
	Completely agree			0.06		0.20	
Trust in the future	Completely disagree			(R)		(R)	
	Disagree			-0.39	**	-0.99	**
	Rather agree			-0.03		-0.22	
	Completely agree			0.11		-0.47	
Sample weights				-0.70	***	-1.37	**

[46] The level of significance: ***<0.001, **<0.05, *<0.1

Table 42. Regression models about the effect of voluntarism retirement on change in life satisfaction between 2001/2002 and 2004/2005 (coefficients and significances[47])

		Raw data correlation	Raw data regression adjustment	Matched data regression adjustment
(Intercept)		0.35 **	3.97 *	1.41
Voluntarism of retirement		**-0.28**	**0.31**	**0.82 *****
Satisfaction with life			-0.79 ***	-0.83 ***
Recent perceived well-being			0.14	0.20 **
Sex	Male		(R)	(R)
	Female		-0.22	-0.08
Education	Primary or less		(R)	(R)
	Vocational secondary		-0.23	-0.76 **
	General secondary		0.11	-0.16
	Tertiary		0.26	-0.13
Satisfaction with housing			0.09 *	0.15 **
Age			-0.02	0.02
Residence	Capital city		(R)	(R)
	Bigger city		0.56	0.88 **
	Smaller city		0.20	0.29
	Village		0.05	0.01
Subjective health status			-0.05	-0.13 *
Equivalent household income			0.00	0.00
Labour market status	Employed		(R)	(R)
	Self-employed		-0.65 *	-0.79 *
	Unemployed		0.69	-0.12
	Other non-working		0.89	0.13
Has ever experienced unemployment			-0.71 **	-0.57 **
Workplace	Owned by the state		(R)	(R)
	Private		-0.28	-0.12
	Non respond		-0.64	-0.09
Last (most important) work	Blue collar		(R)	(R)
	White collar		-0.24	-0.55
	Never had. no respond		0.20	0.78
Marital status	Single		(R)	(R)
	Married living together		0.52	2.00 *
	Married living apart		0.63	2.08
	Widow		0.62	1.60
	Divorced		0.50	1.80 **
Partner labour market status	Does not have partner		(R)	(R)
	Employed		-0.08	2.56 *
	Self-employed		-0.07	2.84 **
	Retired		0.13	2.86 **
	Unemployed		-0.26	3.41 **
	Other non-working		0.29	2.91 *
	Does not answer		0.19	1.47
Satisfaction with the partner	Does not have partner		(R)	(R)
	Dissatisfied		0.12	-3.26 **
	Neutral		-0.25	-3.09 **
	Rather satisfied		0.29	-2.45 *
	Very satisfied		0.54	-2.33 *
Number of female children			0.27 **	0.07
Number of male children			0.19 *	0.13
Number of grandchildren			0.44	0.41
Does not enjoy working	Completely disagree		(R)	(R)
	Disagree		0.56 *	0.18
	Rather agree		0.35	-0.36
	Completely agree		0.05	-0.11
Trust in the future	Completely disagree		(R)	(R)
	Disagree		-0.58	-0.48
	Rather agree		-0.83 *	-0.48
	Completely agree		-0.49	-0.37
Sample weights			0.95	0.42

[47] The level of significance: ***<0.001, **<0.05, *<0.1

Table 43. Regression models about the effect of voluntarism retirement on change in life satisfaction between 2001/2002 and 2012/2013 (coefficients and significances[48])

		Raw data correlation	Raw data regression adjustment	Matched data regression adjustment
(Intercept)		0.57 ***	5.17 **	-0.10
Voluntarism of retirement		**-0.09**	**-0.12**	**0.58 ****
Satisfaction with life			-0.90 ***	-0.82 ***
Recent perceived well-being			-0.15 *	-0.07
Sex	Male		(R)	(R)
	Female		0.07	0.08
Education	Primary or less		(R)	(R)
	Vocational secondary		0.02	-0.11
	General secondary		0.80 *	1.12 **
	Tertiary		0.78	0.63
Satisfaction with housing			0.11 *	0.09
Age			0.01	0.01
Residence	Capital city		(R)	(R)
	Bigger city		0.44	0.78 *
	Smaller city		0.39	0.10
	Village		0.43	0.08
Subjective health status			0.16 **	0.12
Equivalent household income			0.00	0.00
Labour market status	Employed		(R)	(R)
	Self-employed		0.49	0.40
	Unemployed		-0.15	-1.47
	Other non-working		-0.95	0.26
Has ever experienced unemployment			-0.85 **	-0.99
Workplace	Owned by the state		(R)	(R)
	Private		0.35	0.51
	Non respond		0.98	0.16
Last (most important) work	Blue collar		(R)	(R)
	White collar		-0.02	-0.10
	Never had. no respond		-1.04	-1.66 **
Marital status	Single		(R)	(R)
	Married living together		1.44 *	3.45 **
	Married living apart		1.39	3.42 *
	Widow		0.48	1.43
	Divorced		0.63	1.86 *
Partner labour market status	Does not have partner		(R)	(R)
	Employed		0.16	-1.24
	Self-employed		0.02	-1.20
	Retired		-0.26	-1.60
	Unemployed		-0.22	-3.07
	Other non-working		1.00	0.47
	Does not answer		-0.45	-0.88
Satisfaction with the partner	Does not have partner		(R)	(R)
	Dissatisfied		1.04	1.17
	Neutral		-0.30	0.02
	Rather satisfied		0.11	0.35
	Very satisfied		0.19	0.16
Number of female children			-0.04	-0.06
Number of male children			-0.15	-0.07
Number of grandchildren			-0.23	-0.14
Does not enjoy working	Completely disagree		(R)	(R)
	Disagree		-0.11	-0.13
	Rather agree		0.26	-0.46
	Completely agree		0.40	0.22
Trust in the future	Completely disagree		(R)	(R)
	Disagree		-1.00 *	-0.51
	Rather agree		-0.29	0.45
	Completely agree		-0.16	0.39
Sample weights			-1.09	1.65

48 The level of significance: ***<0.001, **<0.05, *<0.1

Figure 16. The difference in propensity score between the treatment group (those to whom a child was born in the observation period) and control groups for estimating the effect of parenthood

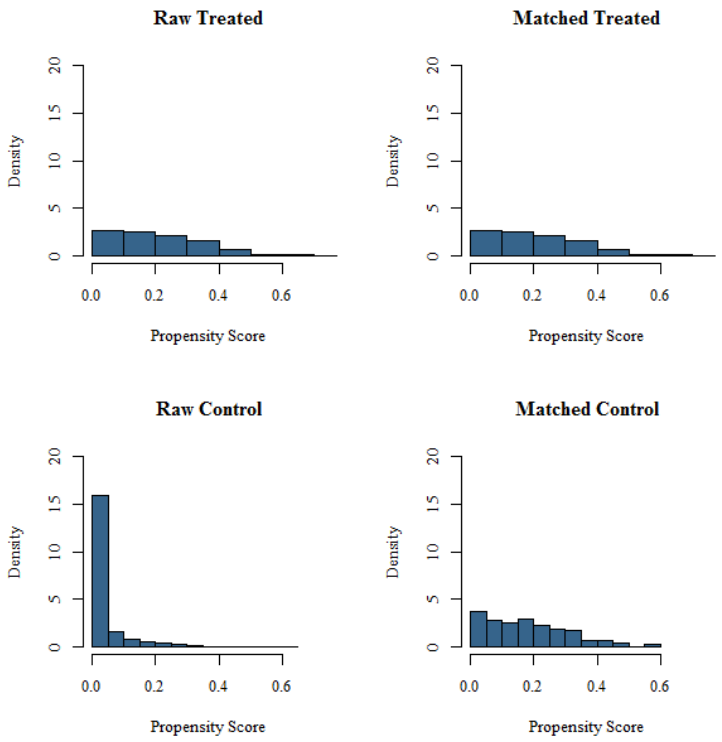


Figure 17. The difference in propensity score between the treatment group (those to whom the first child was born in the observation period) and control groups for estimating the first child effect

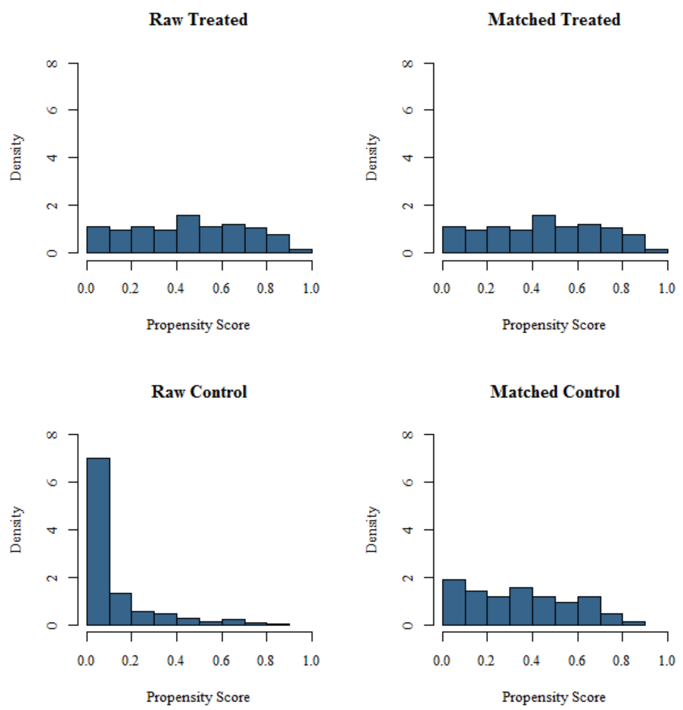


Figure 18. The difference in propensity score between the treatment group (those to whom the second child was born in the observation period) and controls group for estimating the second child effect

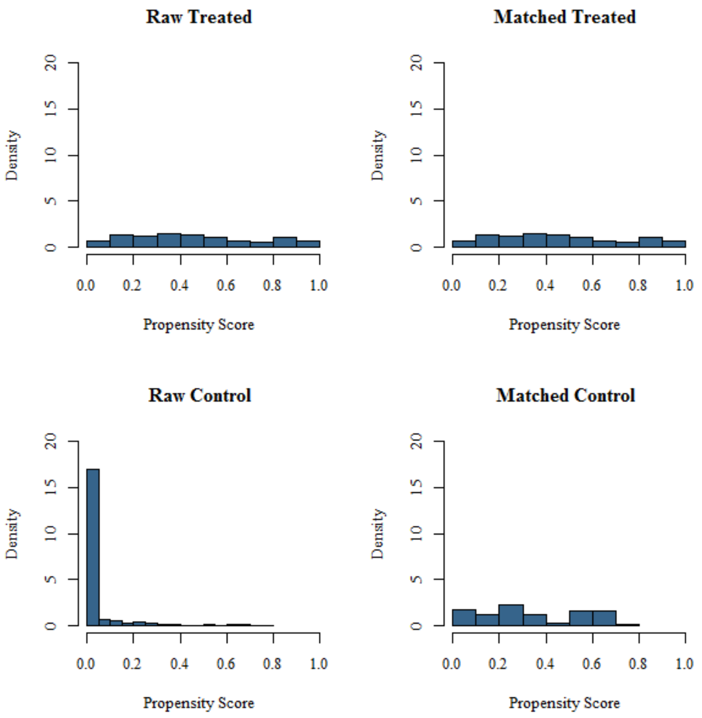


Figure 19. The difference in propensity score between the treatment group (women to whom a child was born in the observation period) and control groups for estimating the motherhood effect

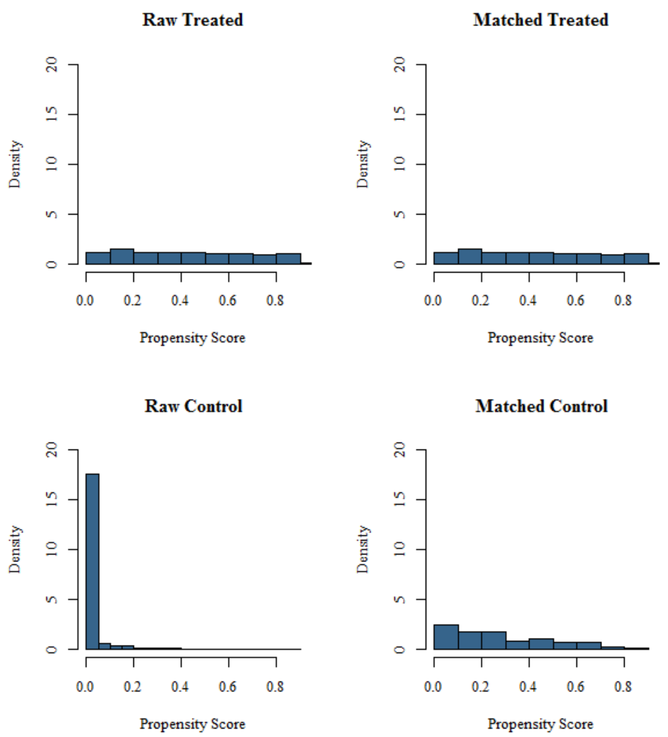


Figure 20. The difference in propensity score between the treatment (men to whom a child was born in the observation period) and control groups for estimating the fatherhood effect

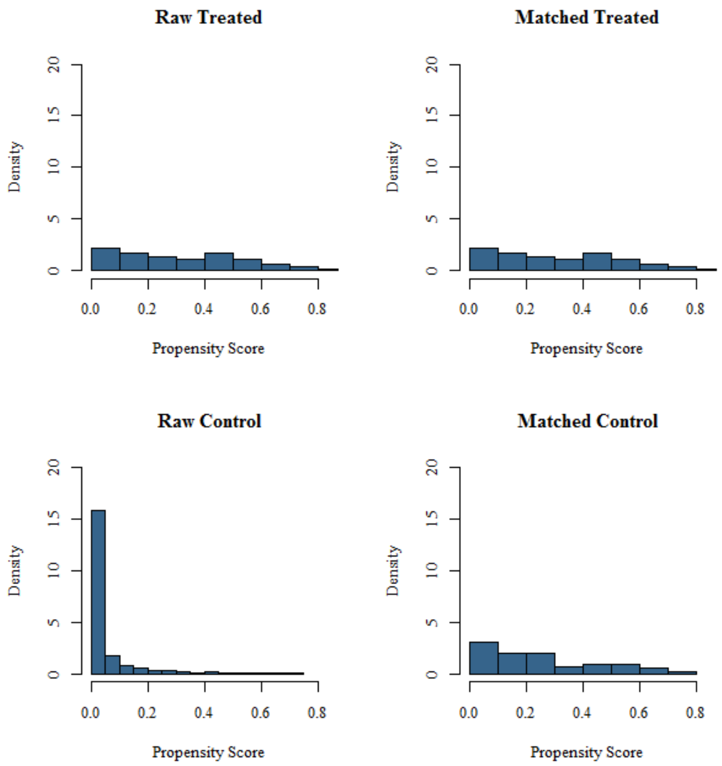


Figure 21. The difference in propensity score between the treatment (retirees) and control groups (non-retirees) for estimating the retirement effect

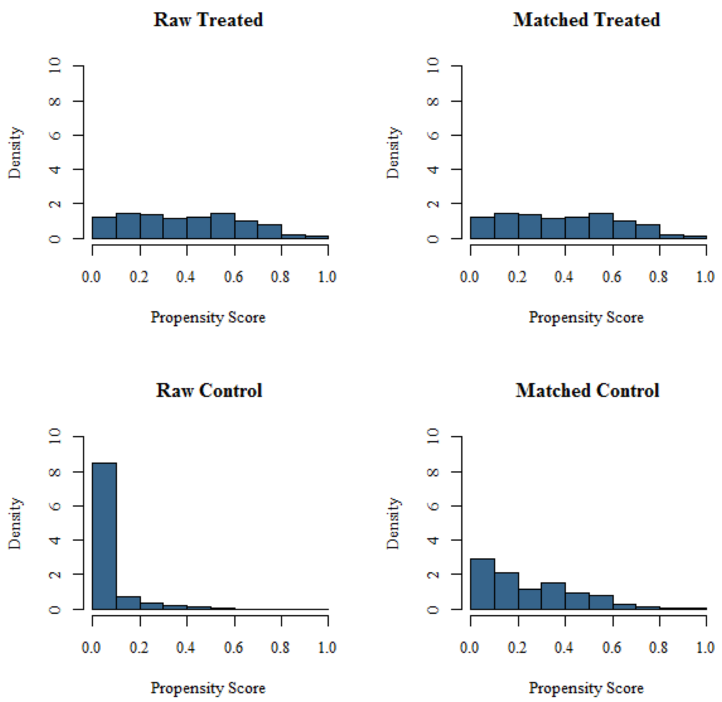


Figure 22. The difference in propensity score between the treatment (voluntary retires) and control groups (involuntary retires) for estimating the voluntary retirement effect

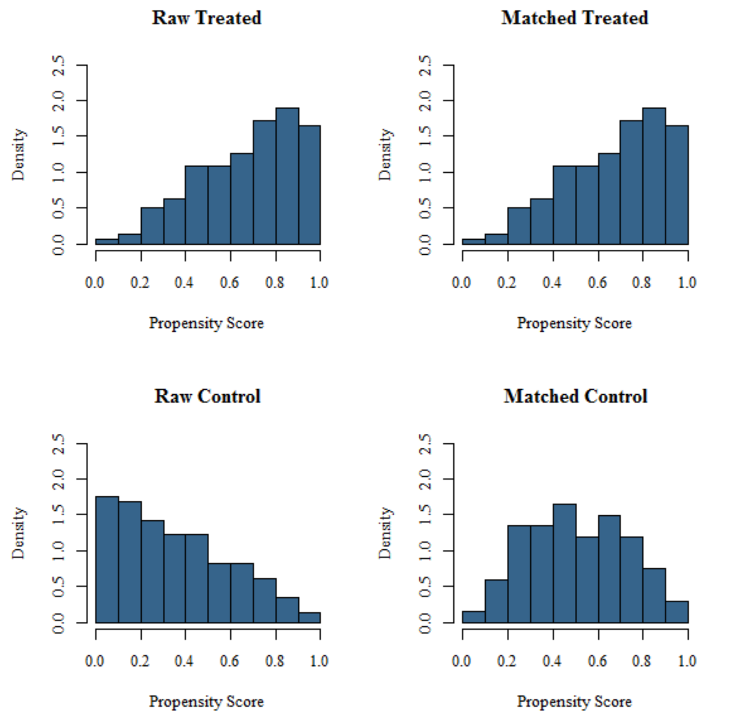
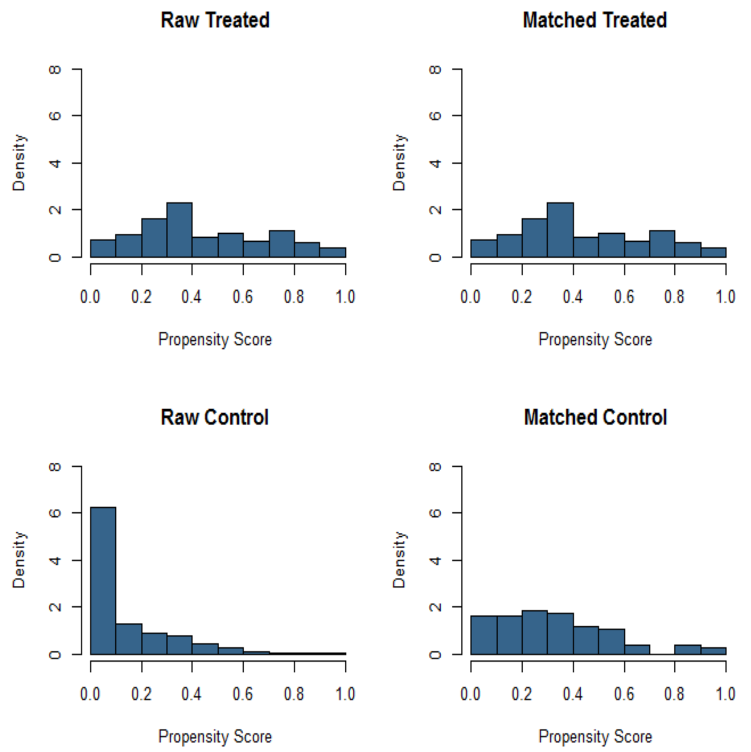


Figure 23. The difference in propensity score between the treatment (full nest 2) and control groups (the rest of the population) for estimating the household life-cycle effect



References

Aassve, A., Arpino, B., & Balbo, N. (2016). It Takes Two to Tango: Couples' Happiness and Childbearing. *European Journal of Population*, 32(3), 339-354. doi: 10.1007/s10680-016-9385-1

Aassve, A., Goisis, A., & Sironi, M. (2012). Happiness and childbearing across Europe. *Social Indicators Research*, 108(1), 65-86. doi: 10.1007/s11205-011-9866-x

Abadie, A., & Imbens, G. W. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica*, 74(1), 235-267. doi: 10.1111/j.1468-0262.2006.00655.x

Abolhassani, M., & Alessie, R. (2013). Subjective well-being around retirement. *De Economist*, 161(3), 349-366. doi: 10.1007/s10645-013-9209-1

Adams, R. E. (1992). Is happiness a home in the suburbs?: The influence of urban versus suburban neighborhoods on psychological health. *Journal of Community Psychology*, 20(4), 353-372. doi: 10.1002/1520-6629(199210)20:4<353::AID-JCOP2290200409>3.0.CO;2-Z

Ahuvia, A. C., & Friedman, D. C. (1998). Income, consumption, and subjective well-being: Toward a composite macromarketing model. *Journal of Macromarketing*, 18(2), 153-168. doi: 10.1177/027614679801800207

Allison, P. D. (1990). Change scores as dependent variables in regression analysis. *Sociological Methodology*, 93-114. doi: 10.2307/271083.

Amato, P. R. (2004). Tension between institutional and individual views of marriage. *Journal of Marriage and Family*, 66(4), 959-965. doi: 10.1111/j.0022-2445.2004.00065.x

Amato, P. R., & Hohmann-Marriott, B. (2007). A comparison of high-and low-distress marriages that end in divorce. *Journal of Marriage and Family*, 69(3), 621-638. doi: 10.1111/j.1741-3737.2007.00396.x

Andrews, D., & Sánchez, A. C. (2011). The evolution of homeownership rates in selected OECD countries: Demographic and public policy influences. *OECD Journal: Economic Studies*, 2011(1), 1-37.

Angeles, L. (2010). Children and life satisfaction. *Journal of Happiness Studies*, 11(4), 523-538. doi: 10.1007/s10680-016-9385-1

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434), 444-455. doi: 10.3386/t0136

Arpino, B., & Aassve, A. (2013). Estimating the causal effect of fertility on economic wellbeing: data requirements, identifying assumptions and estimation methods. *Empirical Economics*, 44(1), 355-385. doi: 10.1007/s00181-010-0356-9.

Arpino, B., & Mealli, F. (2011). The specification of the propensity score in multilevel observational studies. *Computational Statistics & Data Analysis*, 55(4), 1770-1780. doi: 10.1016/j.csda.2010.11.008

Ateca-Amestoy, V.A., Aguilar, A. C., & Moro-Egido, A. I., (2014). Social Interactions and Life Satisfaction: Evidencefrom Latin America, *Journal of Happiness Studies*. 15, 527–554. doi: 10.1007/s10902-013-9434-y

Atchley, R. C. (1971). Retirement and leisure participation: Continuity or crisis? *The Gerontologist*, 11, 13–17. doi: 10.1093/geront/11.1_part_1.13

Atchley, R. C. (1989). A continuity theory of normative aging. *Gerontologist*, 29(2): 183–190. doi: 10.1093/geront/29.2.183

Athey, S., & Imbens, G. W (2006). Identification and inference in nonlinear difference-in-differences models. *Econometrica* 74, 2: 431-497.doi: 10.1111/j.1468-0262.2006.00668.x

Baetschmann, G., Staub, K. E., & Studer, R. (2016). Does the stork deliver happiness? Parenthood and life satisfaction. *Journal of Economic Behavior & Organization*, 130, 242-260. doi: 10.2139/ssrn.2167277

Baker, M., Gruber, J., & Milligan, K. S. (2009). Retirement income security and well-being in Canada. *NBER Working Paper Series* National Bureau of Economic Research.

Balbo, N., & Arpino, B. (2016). The role of family orientations in shaping the effect of fertility on subjective well-being: A propensity score matching approach. *Demography*, 53(4), 955-978. doi: 10.1007/s13524-016-0480-z

Ballweg, J. A. (1967). Resolution of conjugal role adjustment after retirement. *Journal of Marriage and Family*, 29, 277–281. doi: 10.2307/349688

Baltagi, B. H. (2014). *The Oxford handbook of panel data*: Oxford Handbooks.

Baranowska, A., & Matysiak, A. (2011). Does parenthood increase happiness? Evidence for Poland. *Vienna Yearbook of Population Research*, 9, 307-325. doi: 10.1553/populationyearbook2011s307

Bartley, M., Power, C., Blane, D., Smith, G. D., & Shipley, M. (1994). Birth weight and later socioeconomic disadvantage: evidence from the 1958 British cohort study. *British Medical Journa*l, 309(6967), 1475-1478.

Bardo, A. R. (2017). A life course model for a domains-of-life approach to happiness: Evidence from the United States. *Advances in Life Course Research*, 33, 11-22. doi: 10.1016/j.alcr.2017.06.002

Bauer, G., & Kneip, T. (2012). Fertility from a couple perspective: A test of competing decision rules on proceptive behaviour. *European Sociological Review*, 29(3), 535-548. doi: 10.1093/esr/jcr095

Barrett, G. F., & Kecmanovic, M. (2013). Changes in subjective well-being with retirement: assessing savings adequacy. *Applied Economics*, 45(35), 4883-4893. doi: 10.1080/00036846.2013.806786

Bartus, T. (2015). Lemorzsolódás és súlyozás az Életünk Fordulópontjai panelfelvételben [Drop-out and weighting in Turning Points of Life Course survey]. *Demográfia*, 58(4), 287-308. doi: 10.21543/Dem.58.4.3

Bartus, T., Murinkó, L., Szalma, I., & Szél, B. (2013). The effect of education on second births in Hungary: A test of the time-squeeze, self-selection, and partner-effect hypotheses. *Demographic Research*, 28, 0_1. doi: 10.4054/DemRes.2013.28.1.

Baxter, J., Hewitt, B., & Haynes, M. (2008). Life course transitions and housework: Marriage, parenthood, and time on housework. *Journal of Marriage and Family*, 70(2), 259-272. doi: 10.1111/j.1741-3737.2008.00479.x

Becker, G. (1981). *A treatise on the family* (Vol. 30): Harvard University Press.

Becker, G. S., & Barro, R. J. (1988). A reformulation of the economic theory of fertility. *The Quarterly Journal of Economics*, 103(1), 1-25. doi:

10.2307/1882640

Becker, G. S., Landes, E. M., & Michael, R. T. (1977). An economic analysis of marital instability. *Journal of Political Economy*, 85(6), 1141-1187. doi: 10.1086/260631

Bender, K. A. (2012). An analysis of well-being in retirement: The role of pensions, health, and ‘voluntariness’ of retirement. *The Journal of Socio-Economics*, 41(4), 424-433. doi: 10.1016/j.socec.2011.05.010

Bernardi, L., Bollmann, G., Potarca, G., & Rossier, J. (2017). Multidimensionality of well-being and spillover effects across life domains: How do parenthood and personality affect changes in domain-specific satisfaction? *Research in Human Development*, 14(1), 26-51. doi: 10.1080/15427609.2016.1268893

Bernardi, L., & Klärner, A. (2014). Social networks and fertility. *Demographic Research*, 30(22), 641-670. doi: 10.4054/DemRes.2014.30.22.

Besen, E., Matz-Costa, C., Brown, M., Smyer, M. A., & Pitt-Catsouphes, M. (2013). Job characteristics, core self-evaluations, and job satisfaction: what's age got to do with it? *The International Journal of Aging and Human Development*, 76(4), 269-295. doi: 10.2190/AG.76.4.a?journalCode=ahdb

Bianchi, S. M. (2000). Maternal employment and time with children: Dramatic change or surprising continuity? *Demography*, 37(4), 401-414. doi: 10.1353/dem.2000.0001

Bianchi, S. M., Robinson, J. P., & Milke, M. A. (2006). *The changing rhythms of American family life*: Russell Sage Foundation.

Bierman, A., Fazio, E. M., & Milkie, M. A. (2006). A multifaceted approach to the mental health advantage of the married: Assessing how explanations vary by outcome measure and unmarried group. *Journal of Family Issues*, 27(4), 554-582. doi: 10.1177/0192513X05284111

Billari, F. C. (2009). The happiness commonality: Fertility decisions in low-fertility settings. How generations and gender shape demographic change, Keynote address at *Conference on How Generations and Gender Shape Demographic Change*: Toward Policies Based on Better Knowledge, UNECE, Geneva, Switzerland. Retrieved from www.unece.org/fileadmin/DAM/pau/_docs/ggp/2008/GGP_2008_GGConf_Publ_1_Chapter-1.pdf

Bird, C. E. (1999). Gender, household labor, and psychological distress: The impact of the amount and division of housework. *Journal of Health and Social Behavior*, 32-45. doi: 10.2307/2676377

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2007). From the cradle to the labor market? The effect of birth weight on adult outcomes. *The Quarterly Journal of Economics*, 122(1), 409-439. doi: 10.1162/qjec.122.1.409

Blackwell, M., Iacus, S., King, G., & Porro, G. (2009). CEM: Coarsened Exact Matching in Stata. *The Stata Journal*, 9, Pp. 524–546. Copy at http://j.mp/2q0JxRP

Blanchflower, D. G., & Oswald, A. J. (2008). Is well-being U-shaped over the life cycle? *Social Science & Medicine*, 66(8), 1733-1749. doi: 10.1016/j.socscimed.2008.01.030

Blood Jr, R. O., & Wolfe, D. M. (1960). *Husbands and wives: The dynamics of family* *living*. Oxford, England: Free Press Glencoe.

Boldrin, M., De Nardi, M., & Jones, L. E. (2015). Fertility and social security. *Journal of Demographic Economics*, 81(3), 261-299. doi: 10.3386/w11146

Boldrin, M., & Jones, L. E. (2002). Mortality, fertility, and saving in a Malthusian economy. *Review of Economic Dynamics*, 5(4), 775-814. doi: 10.1006/redy.2002.0186

Bonsang, E., & Klein, T. J. (2012). Retirement and subjective well-being. *Journal of Economic Behavior & Organization*, 83(3), 311-329. doi: 10.2139/ssrn.1625757

Bossé, R., Aldwin, C. M., Levenson, M. R., & Ekerdt, D. J. (1987). Mental health differences among retirees and workers: findings from the Normative Aging Study. *Psychology and Aging*, 2(4), 383. doi: 10.1037//0882-7974.2.4.383

Bowen, C. F., Radhakrishna, R., & Keyser, R. (1994). Job satisfaction and commitment of 4-H agents. *Journal of Extension*, 32(1), 1-22.

Börsch-Supan, A., & Jürges, H. (2006). Early retirement, social security and well-being in Germany. In W. David (Ed.), *Developments in Economics of Aging*: University of Chicago Press.

Brickman, P., Coates, D., & Janoff-Bulman, R. (1978). Lottery winners and accident victims: Is happiness relative? *Journal of Personality and Social Psychology*, 36(8), 917. doi: 10.1037/0022-3514.36.8.917

Brükner, H., & Mayer, K. U. (2005). The Stucture of the Life Course, Standardized? Individualized? Differentiated? *Advanced in Life Course Research*, 9: 27-53. doi: 10.1016/S1040-2608(04)09002-1

Bryson, A., Forth, J., & Stokes, L. (2015). Does worker wellbeing affect workplace performance? *IZA Discussion Paper*.

Burden, D. S. (1986). Single parents and the work setting: The impact of multiple job and homelife responsibilities. *Family Relations*, 37-43. doi: 10.2307/584280

Caldwell, J. C. (1978). A theory of fertility: from high plateau to destabilization. *Population and Development Review*, 553-577. doi: 10.2307/1971727

Caldwell, J. C. (1982). *Theory of fertility decline*. New York: Academic Press.

Caldwell, J. C., & Schindlmayr, T. (2003). Explanations of the fertility crisis in modern societies: A search for commonalities. *Population Studies*, 57(3), 241-263. doi: 10.1080/0032472032000137790

Carter, M. A. T., & Cook, K. (1995). Adaptation to retirement: Role changes and psychological resources. *The Career Development Quarterly*, 44(1), 67-82. doi: 10.1002/j.2161-0045.1995.tb00530.x

Chabé-Ferret, S. (2015). Analysis of the bias of Matching and Difference-in-Difference under alternative earnings and selection processes. *Journal of Econometrics*, 185(1), 110-123. doi: 10.1016/j.jeconom.2014.09.013

Charles, K. K. (2002). Is retirement depressing?: Labor force inactivity and psychological well-being in later life NBER Working Paper (Vol. 9033): *National Bureau of Economic Research*.

Clark, A. E., Diener, E., Georgellis, Y., & Lucas, R. E. (2008). Lags and leads in life satisfaction: A test of the baseline hypothesis. *The Economic Journal*, 118(529). doi: 10.1111/j.1468-0297.2008.02150.x

Clark, A. E., & Fawaz, Y. (2009). Valuing jobs via retirement: European evidence. *National Institute Economic Review*, 209(1), 88-103. doi: 131.172.36.29

Clark, A. E., & Georgellis, Y. (2013). Back to baseline in Britain: adaptation in the British household panel survey. *Economica*, 80(319), 496-512. doi: 10.1111/ecca.12007

Clark, A. E., & Oswald, A. J. (1994). Unhappiness and unemployment. *The Economic Journal*, 104(424), 648-659. doi: 10.2307/2234639

Clark, A., Oswald, A., & Warr, P. (1996). Is job satisfaction U-shaped in age? *Journal of Occupational and Organizational Psychology*, 69(1), 57-81. doi: 10.1111/j.2044-8325.1996.tb00600.x

Clark, A. E., & Oswald, A. J. (2006). The curved relationship between subjective well-being and age. *Paris-Jourdan Sciences Economiques Working Paper*(29).

Claxton, A., & Perry-Jenkins, M. (2008). No fun anymore: Leisure and marital quality across the transition to parenthood. *Journal of Marriage and Family*, 70(1), 28-43. doi: 10.1111/j.1741-3737.2007.00459.x

Cochran, W. G. (1968). The effectiveness of adjustment by subclassification in removing bias in observational studies. *Biometrics*, 24(2), 295-313. doi: 10.2307/2528036

Cochran, W. G., & Rubin, D. B. (1973). Controlling bias in observational studies: A review. Sankhyā: *The Indian Journal of Statistics, Series A*, 417-446. doi: 10.1017/CBO9780511810725.005

Costa, P. T., Zonderman, A. B., McCrae, R. R., Huntley, J. C., Locke, B. Z., & Barbano, H. E. (1987). Longitudinal analyses of psychological well-being in a national sample: Stability of mean levels. *Journal of Gerontology*, 42(1), 50-55. doi: 10.1093/geronj/42.1.50

Coursolle, K. M., Sweeney, M. M., Raymo, J. M., & Ho, J.-H. (2010). The association between retirement and emotional well-being: Does prior work–family conflict matter? The *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 65(5), 609–620. doi: 10.1093/geronb/gbp116

Craig, L. (2016). *Contemporary motherhood: The impact of children on adult time*: Routledge.

Craig, L., & Mullan, K. (2013). Parental leisure time: A gender comparison in five countries. *Social Politics*, 20(3), 329-357. doi: 10.1093/sp/jxt002

Crump, R. K., Hotz, V. J., Imbens, G. W., & Mitnik, O. A. (2009). Dealing with limited overlap in estimation of average treatment effects. *Biometrika*, 96(1), 187-199. doi: 10.1093/biomet/asn055

Cummins, R. A. (2000). Objective and subjective auality of life: An interactive model. *Social Indicators Research*, 52(1), 55-72. doi: 10.1023/A:100702782

Cummins, R. A., Eckersley, R., Pallant, J., Van Vugt, J., & Misajon, R. (2003). Developing a national index of subjective wellbeing: The Australian Unity Wellbeing Index. *Social Indicators Research*, 64(2), 159-190. doi: 10.1023/A:102470432

Czajka, J. L., Hirabayashi, S. M., Little, R. J., & Rubin, D. B. (1992). Projecting from advance data using propensity modeling: An application to income and tax statistics. *Journal of Business & Economic Statistics*, 10(2), 117-131. doi: 10.2307/1391671

Davis, G. D. (2012). Looking toward the future: predicting retirement satisfaction. *Journal of Social and Psychological Sciences*, 5(2), 1-22. doi: 10.3389/fpsyg.2017.00891

Deaton, A. (2008). Income, health, and well-being around the world: Evidence from the Gallup World Poll. *The Journal of Economic Perspectives*, 22(2), 53-72. doi: 10.1257/jep.22.2.53

Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and Statistics*, 84(1), 151-161. doi: 10.3386/w6829

DeMaris, A. (2017). The marriage advantage in subjective well-being: Causal effect or unmeasured heterogeneity? *Marriage & Family Review*, 1-16. doi: 10.1080/01494929.2017.1359812

Derrick, F. W., & Lehfeld, A. K. (1980). The family life cycle: An alternative approach. *Journal of Consumer Research*, 7(2), 214-217. doi: 10.1086/208809

Desmet, R., Jousten, A., & Perelman, S. (2005). The benefits of separating early retirees from the unemployed: simulation results for Belgian wage earners. *IZA Discussion Papers* 1571

Deza, M. M., & Deza, E. (2009). *Encyclopedia of distances* (Vol. 94): Springer.

Diamond, A., & Sekhon, J. S. (2013). Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies. *Review of Economics and Statistics*, 95(3), 932-945. doi: 10.1162/REST_a_00318

Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, 95, 542-575. doi: 10.1037/0033-2909.95.3.542

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, 49(1), 71-75. doi: 10.1207/s15327752jpa4901_13

Diener, E., & Eunkook Suh, M. (1997). Subjective well-being and age: An international analysis. *Annual Review of Gerontology and Geriatrics*, 17, 304-324.

Diener, E., Heintzelman, S. J., Kushlev, K., Tay, L., Wirtz, D., Lutes, L. D., & Oishi, S. (2016). Findings all psychologists should know from the new science on subjective well-being. *Canadian Psychology*, 58, 87–104. doi: 10.1037/cap0000063

Diener, E., Lucas, R. E., & Scollon, C. N. (2006). Beyond the hedonic treadmill: revising the adaptation theory of well-being. *American Psychologist*, 61, 305. doi: 10.1007/978-90-481-2350-6 5

Diener, E., Lucas, R. E., & Scollon, C. N. (2006). Beyond the hedonic treadmill: revising the adaptation theory of well-being. *American Psychologist*, 61, 305. doi: 10.1007/978-90-481-2350-6 5

Diener, E., Sandvik, E., Seidlitz, L., & Diener, M. (1993). The relationship between income and subjective well-being: Relative or absolute? *Social Indicators Research*, 28(3), 195-223. doi: 10.1007/BF01079018

Diener, E., & Seligman, M. E. (2002). Very happy people. *Psychological Science*, 13(1), 81-84. doi: 10.1111/1467-9280.00415

Diener, E., Suh, E., & Oishi, S. (1997). Recent Finding on Subjective Well-being, *Indian Journal of Clinical Psychology*, 24(1): 25-41.

DiPrete, T. A., & Gangl, M. (2004). Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments. *Sociological Methodology*, 34(1), 271-310. doi: 10.2307/3649376

Dorn, D., & Sousa-Poza, A. (2010). 'Voluntary'and 'involuntary'early retirement: an international analysis. *Applied Economics*, 42(4), 427-438. doi: 10.1080/00036840701663277

Draxler, J., & Van Vliet, O. (2010). European social model: No convergence from the East. *Journal of European Integration*, 32(1), 115-135. doi: 10.1080/07036330903375230

DuGoff, E. H., Schuler, M., & Stuart, E. A. (2014). Generalizing observational study results: applying propensity score methods to complex surveys. *Health Services Research*, 49(1), 284-303. doi: 10.1111/1475-6773.12090

Dwyer, J. W., & Coward, R. T. (1991). A multivariate comparison of the involvement of adult sons versus daughters in the care of impaired parents. *Journal of Gerontology*, 46(5), S259-S269. doi: 10.1093/geronj/46.5.S259

Dykstra, P. A., & Fokkema, T. (2011). Relationships between parents and their adult children: A West European typology of late-life families. *Ageing & Society*, 31(4), 545-569. doi: 10.1017/ S0144686X10001108.

Easterlin, R. A. (1974). Does economic growth improve the human lot? Some empirical evidence. *Nations and Households in Economic Growth*, 89, 89-125. doi: 10.1016/b978-0-12-205050-3.50008-7

Easterlin, R. A. (2001). Income and happiness: Towards a unified theory. *The Economic Journal*, 111(473), 465-484. doi: 10.1111/1468-0297.00646

Easterlin, R. A. (2006). Life cycle happiness and its sources: Intersections of psychology, economics, and demography. *Journal of Economic Psychology*, 27(4), 463-482. doi: 10.1016/j.joep.2006.05.002

Easterlin, R. A., & Sawangfa, O. (2007). Happiness and domain satisfaction: Theory and evidence. *IZA Discussion Paper*, No. 2584.

Ebbinghaus, B., & Hofäcker, D. (2013). Reversing early retirement in advanced welfare economies a paradigm shift to overcome push and pull factors. *Comparative Population Studies*, 38(4). doi: 10.12765/CPoS-2013-24en

Ekman, P., Davidson, R. J., & Friesen, W. V. (1990). The Duchenne smile: Emotional expression and brain physiology. *Journal of Personality and Social Psychology*, 58(2), 342. doi: 10.1037/0022-3514.58.2.342

Elder Jr, G. H., Johnson, M. K., & Crosnoe, R. (2003). The emergence and development of life course theory. In J. T. Mortimer & M. J. Shanahan (Eds.), *Handbook of the Life Course* (pp. 3-19): Springer.

Ellison, D. L. (1968). Work, retirement and the sick role. *The Gerontologist*, 8, 189–192. doi 10.1093/geront/8.3_part_1.189

Elsenburg, L. K., Smidt, N., & Liefbroer, A. C. (2017). The longitudinal relation between accumulation of adverse life events and body mass index from early adolescence to young adulthood. *Psychosomatic Medicine*, 79(3), 365-373. doi: 10.1097/PSY.0000000000000401.

Elwert, F., & Winship, C. (2014). Endogenous selection bias: The problem of conditioning on a collider variable. *Annual Review of Sociology*, 40, 31-53.

doi: 10.1146/annurev-soc-071913-043455

Erdogan, B., Bauer, T. N., Truxillo, D. M., & Mansfield, L. R. (2012). Whistle while you work: A review of the life satisfaction literature. *Journal of Management*, 38(4), 1038-1083. doi: 10.1177/0149206311429379

Erikson, R. (1974). Welfare as a planning goal. *Acta Sociologica*, 17(3), 273-288. doi: 10.1177/000169937401700305

Evertsson, M. (2016). Parental leave and careers: Women's and men's wages after parental leave in Sweden. *Advances in Life Course Research*, 29, 26-40. doi: 10.1016/j.alcr.2016.02.002

Fernández-Carro, C., Módenes, J. A., & Spijker, J. (2015). Living conditions as predictor of elderly residential satisfaction. A cross-European view by poverty status. *European Journal of Ageing*, 12(3), 187-202. doi: 10.1007/s10433-015-0338-z

Ferraro, K. F., & Shippee, T. P. (2009). Aging and cumulative inequality: How does inequality get under the skin?. *The Gerontologist*, 49(3), 333-343.

Ferrer-i-Carbonell, A., & Frijters, P. (2004). How important is methodology for the estimates of the determinants of happiness? *The Economic Journal*, 114, 497: 641-659. doi: 10.1111/j.1468-0297.2004.00235.x

Fisher, R. A. (1925). *Statistical Methods for Research Workers*. 1st Edition. Oliver and Boyd, Edinburgh.

Fonseca, R., Kapteyn, A., Lee, J., & Zamarro, G. (2017). Does Retirement Make you Happy? A Simultaneous Equations Approach. In D. A. Wise (Ed.), *Insights in the Economics of Aging*. Chigago, USA: University of Chicago Press.

Foster, C. (2000). The limits to low fertility: A biosocial approach. *Population and Development Review*, 26(2), 209-234. doi: 10.1111/j.1728-4457.2000.00209.x

Frijters, P., & Beatton, T. (2012). The mystery of the U-shaped relationship between happiness and age. *Journal of Economic Behavior & Organization*, 82(2), 525-542. doi: 10.1016/j.jebo.2012.03.008

Gall, T. L., Evans, D. R., & Howard, J. (1997). The retirement adjustment process: Changes in the well-being of male retirees across time. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 52(3), 110-117. doi: 10.1093/geronb/52B.3.P110

Ganglmair-Wooliscroft, A., & Lawson, R. (2011). Subjective well-being of different consumer lifestyle segments. *Journal of Macromarketing*, 31(2), 172-183. doi: 10.1177/0276146710393251

Gazioglu, S., & Tansel, A. (2006). Job satisfaction in Britain: individual and job related factors. *Applied Economics*, 38(10), 1163-1171. doi: 10.1080/00036840500392987

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*: Cambridge University Press.

George, L. K. (1992). Economic status and subjective well-being: A review of the literature and an agenda for future research. In N. E. Cutler, D. W. Gregg & P. M. Lawton (Eds.), *Aging, money, and life satisfaction*. Philadelphia: Boettner Institute of Financial Gerontology.

George, L. K., & Maddox, G. L. (1977). Subjective adaptation to the loss of the workrole: A longitudinal study. *Journal of Gerontology*, 32, 456–462. doi: 10.1093/geronj/32.4.456

Giczi, J. (2008). Szubjektív jólét időskorban (Subjective well-ebing at old age). In Á. Grábics (Ed.), *Aktív időskor* [Active old age]. Budapest: KSH, Szociális és Munkaügyi Minisztérium.

Gilly, M. C., & Enis, B. M. (1982). *Recycling the family life cycle: A proposal for redefinition.* ACR North American Advances.

Glenn, N. D., Taylor, P. A., & Weaver, C. N. (1977). Age and job satisfaction among males and females: A multivariate, multisurvey study. *Journal of Applied Psychology*, 62(2), 189.

Goldsteen, K., & Ross, C. E. (1989). The perceived burden of children. *Journal of Family Issues*, 10(4), 504-526. doi: 10.1177/019251389010004005

Graham, C., & Nikolova, M. (2015). Bentham or Aristotle in the development process? An empirical investigation of capabilities and subjective well-being. *World Development*, 68, 163-179. doi: 10.1016/j.worlddev.2014.11.018

Graham, C., & Pozuelo, J. R. (2017). Happiness, stress, and age: How the U curve varies across people and places. *Journal of Population Economics*, 30(1), 225-264. doi: 10.1007/s00148-016-0611-2

Greenwood, E. (1945). *Experimental sociology: A study in method*. New York: King's crown Press.

Gu, X. S., & Rosenbaum, P. R. (1993). Comparison of multivariate matching methods: Structures, distances, and algorithms. *Journal of Computational and Graphical Statistics*, 2(4), 405-420. doi: 10.1080/10618600.1993.10474623

Guriev, S., & Zhuravskaya, E. (2009). (Un)happiness in transition. *The Journal of Economic Perspectives*, 23(2), 143-168. doi: 10.1257/jep.23.2.143

Gwozdz, W., & Sousa-Poza, A. (2010). Ageing, health and life satisfaction of the oldest old: An analysis for Germany. *Social Indicators Research*, 97(3), 397-417. doi: 10.1007/s11205-009-9508-8

Hajdu, T., Hajdu, G. (2013). Are more equal societies happier? Subjective well-being, income inequality, and redistribution. MT-DP – 2013/20, *MTA Discussion Papers*

Hannan, M. T., Tuma, N. B., & Groeneveld, L. P. (1978). Income and independence effects on marital dissolution: Results from the Seattle and Denver income-maintenance experiments. *American Journal of Sociology*, 84(3), 611-633. doi: 10.1086/226829

Hansen, B. B. (2004). Full matching in an observational study of coaching for the SAT. *Journal of the American Statistical Association*, 99(467), 609-618. doi: 10.1198/016214504000000647

Hansen, T. (2012). Parenthood and happiness: A review of folk theories versus empirical evidence. *Social Indicators Research*, 108(1), 29-64. doi: 10.1007/s11205-011-9865-y

Hausman, J. A., & Wise, D. A. (1979). Attrition bias in experimental and panel data: the Gary income maintenance experiment. *Econometrica: Journal of the Econometric Society*, 455-473. doi: 10.2307/1914193

Headey, B. (2006). Subjective well-being: Revisions to dynamic equilibrium theory using national panel data and panel regression methods. *Social Indicators Research*, 79(3), 369-403. doi: 10.1007/s11205-005-5381-2

Headey, B., & Wearing, A. (1989). Personality, life events, and subjective well-being: Toward a dynamic equilibrium model. *Journal of Personality and Social Psychology*, 57(4), 731. doi: 10.1037/0022-3514.57.4.731

Heckhausen, J., Dixon, R. A., & Baltes, P. B. (1989). Gains and losses in development throughout adulthood as perceived by different adult age groups. *Developmental Psychology,* 25(1), 109. doi: 10.1037/0012-1649.25.1.109

Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The Review of Economic Studies*, 64(4), 605-654. doi: 10.2307/2971733

Hegedűs, R. (2001). Szubjektív társadalmi indikátorok–szelektív áttekintés a téma irodalmából [Subjective well-being indicators- selective review of the literature]. *Szociológia Szemle*, 2, 58-71.

Hegedűs, R., & Lengyel, G. (2002). A szubjektív jólét objektív tényezői nemzetközi összehasonlításban [The objective indicators of subjective well-bein in international comparision]. In G. Lengyel (Ed.), *Indikátorok és elemzések. Műhelytanulmányok a társadalmi jelzőszámok témaköréből* (pp. 87-104). Budapest: BKÁE.

Heller, D., Watson, D., & Ilies, R. (2004). The role of person versus situation in life satisfaction: a critical examination. *Psychological Bulletin*, 130(4), 574. doi: 0.1037/0033-2909.130.4.574

Henning, G., Lindwall, M., & Johansson, B. (2016). Continuity in well-being in the transition to retirement. *GeroPsych : The Journal of Gerontopsychology and Geriatric Psychiatry*, 29(4):225-237. doi: 10.1024/1662-9647/a000155

Herbst, C. M. (2012). Footloose and fancy free? Two decades of single mothers' subjective well-being. *Social Service Review*, 86(2), 189-222. doi: 10.1086/666390

Heybroek, L., Haynes, M., & Baxter, J. (2015). Life satisfaction and retirement in Australia: A longitudinal approach. *Work, Aging and Retirement*, 1(2), 166-180. doi: 10.1093/workar/wav006

Hitlin, S., & Kirkpatrick, J. M. (2015). Reconceptualizing agency within the life course: The power of looking ahead. *American Journal of Sociology*, 120(5), 1429-1472. doi: 10.1086/681216

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, 42(8), 1-28. doi: 10.18637/jss.v042.i08

Hobcraft, J. (2006). The ABC of demographic behaviour: How the interplays of alleles, brains, and contexts over the life course should shape research aimed at understanding population processes. *Population Studies*, 60(2), 153-187. doi: 10.1080/00324720600646410

Hochwarter, W. A., Ferris, G. R., Perrewe, P. L., Witt, L. A., & Kiewitz, C. (2001). A Note on the Nonlinearity of the Age-Job-Satisfaction Relationship. *Journal of Applied Social Psychology*, 31(6), 1223-1237. doi: 10.1111/j.1559-1816.2001.tb02671.x

Hoffman, L. W., & Hoffman, M. L. (1973). The value of children to parents. In J. T.

Fawcett (Ed.), *Psychological Perspectives on Population*. New York: Basic Books.

Hoffman, L. W., Thornton, A., & Manis, J. D. (1978). The value of children to parents in the United States. *Population & Environment*, 1(2), 91-131. doi: 10.1007/BF01277597

Hohmann-Marriott, B. E., & Amato, P. (2008). Relationship quality in interethnic marriages and cohabitations. *Social Forces*, 87(2), 825-855. doi: 10.1353/sof.0.0151

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American statistical Association*, 81(396), 945-960. doi: 10.1080/01621459.1986.10478354

Holland, J. H. (1992). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*: MIT Press.

Holtzman, M., & Glass, J. (1999). Explaining changes in mothers' job satisfaction following childbirth. *Work and Occupations*, 26(3), 365-404. doi: 10.1177/0730888499026003005

Horvitz, D. G., & Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47(260), 663-685. doi: 10.1080/01621459.1952.10483446

Horowitz, A. (1985). Sons and daughters as caregivers to older parents: Differences in role performance and consequences. *The Gerontologist*, 25(6), 612-617. doi:10.1093/geront/25.6.612.

Hughes, M. E., & Waite, L. J. (2009). Marital Biography and Health at Mid-Life. *Journal of Health and Social Behavior*, 50(3), 344-358. doi: 10.1177/002214650905000307

Idler, E. L., & Benyamini, Y. (1997). Self-rated health and mortality: a review of twenty-seven community studies. *Journal of Health and Social Behavior*, 21-37. doi: 10.2307/2955359

Imai, K., King, G., & Stuart, E. A. (2007). Misunderstandings among Experimentalists and Observationalists about Causal Inference1. *Journal of the Royal Statistical Society*, 171(2), 481-502. doi: 10.1111/j.1467-985X.2007.00527.x

Imai, K., & Van Dyk, D. A. (2004). Causal inference with general treatment regimes: Generalizing the propensity score. *Journal of the American Statistical Association*, 99(467), 854-866. doi: 10.1198/016214504000001187

Imbens, G., & Angrist, J. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2), 467-475. doi: 10.2307/2951620

Imbens, G., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*. 142: 615–635. doi: 10.3386/w13039

Imbens, G., & Rubin, D. B. (2015). *Causal inference in statistics, social, and biomedical sciences*: Cambridge University Press.

Imbens, G., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1), 5-86. doi: 10.3386/w14251 doi: 10.3386/w14251

Ito, Y., Sagara, J., Ikeda, M., & Kawaura, Y. (2003). Reliability and validity of subjective well-being scale. *The Japanese Journal of Psychology*, 74(3), 276-281. doi: 10.4992/jjpsy.74.276

Johnson, D. R., & Wu, J. (2002). An empirical test of crisis, social selection, and role explanations of the relationship between marital disruption and psychological distress: A pooled time-series analysis of four-wave panel data. *Journal of Marriage and Family*, 64(1), 211-224. doi: 10.1111/j.1741-3737.2002.00211.x

Johnston, D. W., & Lee, W.-S. (2009). Retiring to the good life? The short-term effects of retirement on health. *Economics Letters*, 103(1), 8-11. doi: 10.1016/j.econlet.2009.01.015

Kahneman, D., Diener, E., & Schwarz, N. (1999). *Well-being: Foundations of hedonic psychology*: Russell Sage Foundation.

Kahneman, D., & Krueger, A. B. (2006). Developments in the measurement of subjective well-being. *Journal of Economic Perspectives*, 20(1), 3-24.doi: 10.1257/089533006776526030

Kammann, R. (1983). Across Time and Place1. *Journal of Psychology*, 12, 14-22.

Kapitány, B., & Spéder, Z. (2015). Fertility. *Demographic Portrait of Hungary*, 41-55.

Kassenboehmer, S. C., & Haisken-DeNew, J. P. (2012). Heresy or enlightenment? The well-being age U-shape effect is flat. *Economics Letters*, 117(1), 235-238. doi: 10.1016/j.econlet.2012.05.013

Keele, L. (2010). An overview of rbounds: An R package for Rosenbaum bounds sensitivity analysis with matched data. White Paper. Columbus, OH, 1-15.

Keizer, R., Dykstra, P. A., & Poortman, A.-R. (2010). The transition to parenthood and well-being: the impact of partner status and work hour transitions. *Journal of Family Psychology*, 24(4), 429. doi: 10.1037/a0020414.

Kesavayuth, D., Rosenman, R. E., & Zikos, V. (2016). Retirement, Personality, and Well-Being. *Economic Inquiry*, 54(2), 733-750. doi: 10.1111/ecin.12307

Kézdi, G. (2004). *Az aktív foglalkoztatáspolitikai programok hatásvizsgálatának módszertani kérdései*: Közgazdaságtudományi Kutatóközpont.

Khaleque, A., & Rahman, M. A. (1987). Perceived importance of job facets and overall job satisfaction of industrial workers. Human Relations, 40(7), 401-415. doi: 10.1177/001872678704000701

Kim, J. E., & Moen, P. (2002). Retirement transitions, gender, and psychological well-being a life-course, ecological model. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 57(3), 212-222. doi: 10.1093/geronb/57.3.P212

Kimmel, J., & Connelly, R. (2007). Mothers' time choices caregiving, leisure, home production, and paid work. *Journal of Human Resources*, 42(3), 643-681. doi: 10.2307/40057322

King, G., & Nielsen, R. (2016). Why propensity scores should not be used for matching. j.mp/PScore, 378.

King, G., & Zeng, L. (2006). The dangers of extreme counterfactuals. *Political Analysis*, 131-159. doi: 10.1093/pan/mpj004

Kohler, H.-P., Behrman, J. R., & Skytthe, A. (2005). Partner+ children= happiness? An assessment of the effect of fertility and partnerships on subjective well-being in Danish twins. *Population and Development Review*, 31(3), 407-445. doi:

10.1111/j.1728-4457.2005.00078

Kohler, H. P., Billari, F. C., & Ortega, J. A. (2002). The emergence of lowest-low fertility in Europe during the 1990s. *Population and Development Review*, 28(4), 641-680. doi: 10.1111/j.1728-4457.2002.00641.x

Kohli, M. (1993). Public solidarity between generations: Historical and Comparative Elements. *Forschungsgruppe Altern und Lebenslauf* (FALL), Forschungsbericht, 39.

Kohli, M. (2007). The institutionalization of the life course: Looking back to look ahead. *Research in Human Development*, 4(3-4), 253-271. doi: 10.1080/15427600701663122

Kohli, M. (2014). Later retirement? Patterns. preferences and policies. *Studia Humanistyczne AGH. Contribution to Humanties*, 13(4), 19-34. doi: http://dx.doi.org/10.7494/human.2014.13.4.19-32

Koivumaa-Honkanen, H., Honkanen, R., Viinamaeki, H., Heikkilae, K., Kaprio, J., & Koskenvuo, M. (2001). Life satisfaction and suicide: a 20-year follow-up study. *American Journal of Psychiatry*, 158(3), 433-439. doi: 10.1176/appi.ajp.158.3.433

Kok, J. (2007). Principles and prospects of the life course paradigm. Paper presented at the *Annales de Démographie Historique*.

Kravdal, Ø. (2014). The estimation of fertility effects on happiness: Even more difficult than usually acknowledged. *European Journal of Population*, 30(3), 263-290.

Krueger, A. B., & Schkade, D. A. (2008). The reliability of subjective well-being measures. *Journal of Public Economics*, 92(8), 1833-1845. doi: 10.1016/j.jpubeco.2007.12.015

Krüger, H., & Baldus, B. (1999). Work, gender and the life course: Social construction and individual experience. *Canadian Journal of Sociology/Cahiers Canadiens de Sociologie*, 355-379. doi: 10.2307/3341394

Kuo, Y.-H. (2001). Extrapolation of Association between Two Variables in Four General Medical Journals. Paper presented at the *Fourth International Congress on Peer Review in Biomedical Publication*, Barcelona, Spain.

Lachman, M. E., Röcke, C., Rosnick, C., & Ryff, C. D. (2008). Realism and illusion in Americans' temporal views of their life satisfaction: Age differences in reconstructing the past and anticipating the future. *Psychological Science*, 19(9), 889-897. doi: 10.1111/j.1467-9280.2008.02173.x

LaLonde, R. J. (1986). Evaluating the econometric evaluations of training programs with experimental data. *American Economic Review*, 604-620.

Lang, I., Llewellyn, D., Hubbard, R., Langa, K., & Melzer, D. (2011). Income and the midlife peak in common mental disorder prevalence. *Psychological Medicine*, 41(7), 1365-1372. doi: 10.1017/S0033291710002060.

Lapa, T. Y. (2013). Life satisfaction, leisure satisfaction and perceived freedom of park recreation participants. *Procedia-Social and Behavioral Sciences*, 93, 1985-1993. doi: 10.1016/j.sbspro.2013.10.153

Larsen, R. J., Diener, E., & Emmons, R. A. (1985). An evaluation of subjective well-being measures. *Social Indicators Research*, 17(1), 1-17. doi: 10.1007/BF00354108

Latif, E. (2011). The impact of retirement on psychological well-being in Canada. *The Journal of Socio-Economics*, 40(4), 373-380. doi: 10.1016/j.socec.2010.12.011

Lavee, Y., Sharlin, S., & Katz, R. (1996). The effect of parenting stress on marital quality: An integrated mother-father model. *Journal of Family Issues*, 17(1), 114-135. doi: 10.1177/019251396017001007

Lawrence, E., Rothman, A. D., Cobb, R. J., Rothman, M. T., & Bradbury, T. N. (2008). Marital satisfaction across the transition to parenthood. *Journal of Family Psychology*, 22(1), 41. doi: 10.1037/0893-3200.22.1.41.

Lechner, M. (2011). The estimation of causal effects by difference-in-difference methods. *Econometrics*, 4(3), 165-224. doi: 10.1561/0800000014

Lee, B. K., Lessler, J., & Stuart, E. A. (2010). Improving propensity score weighting using machine learning. *Statistics in Medicine*, 29(3), 337-346. doi: 10.1002/sim.3782.

Leibenstein, H. (1957). *Economic Backwardness and Economic Growth*, New York: John Wiley

Leisering, L. (2003). Government and the life course. *Handbook of the Life Course*, 205-225.

Le Moglie, M., Mencarini, L., & Rapallini, C. (2015). Is it just a matter of personality? On the role of subjective well-being in childbearing behavior. *Journal of Economic Behavior and Organization, 117*, 453-475. doi: 10.1016/j.jebo.2015.07.006

Lengyel, G., & Janky, B. (2002). A szubjektív jólét társadalmi feltételei [The social background of subjective well-being]. *Esély*, 14, 3-26.

Lepper, H. S. (1998). Use of other-reports to validate subjective well-being measures. *Social Indicators Research*, 44(3), 367-379. doi: 10.1023/A:1006872027638

Lindström, K. (1988). Age-related differences in job characteristics and in their relation to job satisfaction. *Scandinavian Journal of Work, Environment & Health*, 14(1), 24-26. doi: 10.1037/0033-2909.93.2.328

Lorenz, F. O., Wickrama, K., Conger, R. D., & Elder Jr, G. H. (2006). The Short-Term and Decade-Long Effects of Divorce on Women's Midlife Health. *Journal of Health and Social Behavior*, 47(2), 111-125. doi: 10.1177/002214650604700202

Lucas, R. E. (2007). Adaptation and the set-point model of subjective well-being: Does happiness change after major life events? *Current Directions in Psychological Science*, 16(2), 75-79. doi: 10.1111/j.1467-8721.2007.00479.x

Lucas, R. E., Diener, E., & Suh, E. (1996). Discriminant validity of well-being measures. *Journal of Personality and Social Psychology*, 71(3), 616. doi: 10.1037/0022-3514.71.3.616

Lucas, R.E., Clark, A.E., Georgellis, Y., & Diener, E. (2004). Unemployment alters the set-point for life satisfaction, *Psychological Science*, 15, 8–13. doi: 10.1111/j.0963-7214.2004.01501002.x

Luhmann, M., Hofmann, W., Eid, M., & Lucas, R. E. (2012). Subjective well-being and adaptation to life events: a meta-analysis. *Journal of Personality and Social Psychology*, 102(3), 592. doi: 10.1037/a0025948

Lunceford, J. K., & Davidian, M. (2004). Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. *Statistics* *in Medicine*, 23(19), 2937-2960. doi: 10.1002/sim.1903

Luppi, F. (2016). When is the second one coming? The effect of couple's subjective well-being following the onset of parenthood. *European Journal of Population*, 32, 421-444. doi: 10.1007/s10680-016-9388-y

Lykken, D., & Tellegen, A. (1996). Happiness is a stochastic phenomenon. *Psychological Science*, 7(3), 186-189. doi: 10.1111/j.1467-9280.1996.tb00355.x

Macmillan, R. (2005). *The structure of the life course: Standardized? Individualized? Differentiated?* Elsevier. ISBN: 0-7623- 1 193-2

Makay, Z. (2015). Family support system-Childrising-Employment. *Demographic Portrait of Hungary*.

Manning, N. (2004). Diversity and change in pre-accession Central and Eastern Europe since 1989. *Journal of European Social Policy*, 14(3), 211-232. doi: 10.1177/0958928704044620

Margolis, R., & Myrskylä, M. (2011). A global perspective on happiness and fertility. *Population and Development Review*, 37(1), 29-56. doi: 10.1111/j.1728-4457.2011.00389.x

Margolis, R., & Myrskylä, M. (2015). Parental well-being surrounding first birth as a determinant of further parity progression. *Demography*, 52(4), 1147-1166. doi: 10.1007/s13524-015-0413-2

Margolis, R., & Myrskyla, M. (2016). Children's Sex and the Happiness of Parents. *European Journal of Population*, 32(3), 403-420. doi: 10.1007/s10680-016-9387-z

Mason, A., & Lee, R. (2006). Reform and support systems for the elderly in developing countries: capturing the second demographic dividend. *Genus*, 11-35. doi: 10.1.1.466.9987

Matjeke, K. T. (2016). *Factors influencing work satisfaction of single parents in the South African National Defence Force: an exploratory study* (Doctoral dissertation, Stellenbosch: Stellenbosch University).

Mattingly, M. J., & Bianchi, S. M. (2003). Gender differences in the quantity and quality of free time: The US experience. *Social Forces*, 81(3), 999-1030. doi: 10.1353/sof.2003.0036

Matysiak, A., Mencarini, L., & Vignoli, D. (2016). Work–family conflict moderates the relationship between childbearing and subjective well-being. *European Journal of Population*, 32, 355-379. doi: 10.1007/s10680-016-9390-4

Mayer, K. U., & Schoepflin, U. (1989). The state and the life course. *Annual Review of Sociology*, 15(1), 187-209. doi: 10.1146/annurev.so.15.080189.001155

McAdams, K. K., Lucas, R. E., & Donnellan, M. B. (2012). The role of domain satisfaction in explaining the paradoxical association between life satisfaction and age. *Social Indicators Research*, 109(2), 295-303. doi: 10.1007/s11205-011-9903-9

McLanahan, S., & Adams, J. (1987). Parenthood and psychological well-being. *Annual Review of Sociology*, 13(1), 237-257. doi: 10.1146/annurev.so.13.080187.001321

Medgyesi, M., & Róbert, P. (2000). A munkával való elégedettség nemzetközi összehasonlításban [Satisfaction with work in international comparision]. *Társadalmi Riport* [Social Report], 591-616.

Medgyesi, M., & Zólyomi, E. (2016). Job satisfaction and satisfaction in financial situation and their impact on life satisfaction. *Social Situation Monitor Research* Note 6/2016. Directorate-General for Employment, Social Affairs and Inclusion 2016

Michalos, A. C. (1985). Multiple Discrepancies Theory (MDT), *Social Indicators Research*, 16: 347–413. doi: 10.1007/BF00333288

Micheel, F., Roloff, J., & Wickenheiser, I. (2011). The impact of socioeconomic characteristics on older employees' willingness to continue working in retirement age. *Comparative Population Studies*, 35(4). doi: 10.4232/10.CPoS-2010-19en

Miettinen, A., & Szalma, I. (2014). Childlessness intentions and ideals in Europe. *Finnish Yearbook of Population Research*, 49, 31-55.

Mikucka, M. (2016). How does parenthood affect life satisfaction in Russia? *Advances in Life Course Research*, 30, 16-29. doi: 10.13140/RG.2.1.2636.5285

Miller, M. K., & Crader, K. W. (1979). Rural-urban differences in two dimensions of community satisfaction. *Rural Sociology*, 44(3), 489.

Molnár, E. S. (2004). Életmód és közérzet az idősödés korában [Lifestyle and General Conditions in the Stage of Ageing]. *Társadalmi riport* [Social Report], 152-164.

Molnár, E. S. (2009). A gyermekszám-preferenciák alakulása Magyarországon az elmúlt évtizedekben [Preferences in number of children in the last decades in Hungary]. *Demográfia*, 52(4), 283-312.

Molnár, G., & Kapitány, Z. (2006). Mobilitás, bizonytalanság és szubjektív jóllét Magyarországon [Mobility, uncertainty and subjective well-being in Hungary]. *Közgazdasági Szemle*, 53, 845-872.

Molnár, G., & Kapitány, Z. (2013). Miért elégedetlenek annyira a magyarok az életükkel? A szubjektív jóllétet befolyásoló tényezők mikroszintű összehasonlító elemzése magyar és osztrák adatokon: *IEHAS Discussion Papers.*

Monostori, J. (2008). *Korai nyugdíjba vonulás. Okok és következmények [Early retirement. reasons and consequences]. Műhelytanulmányok* 07.

Monostori, J. (2015). Aging and retirement. In M. Judit, Ő. Péter & S. Zsolt (Eds.), Aging and retirement: *Demographic Portrait of Hungary*

Montizaan, R. M., & Vendrik, M. C. (2014). Misery loves company: Exogenous shocks in retirement expectations and social comparison effects on subjective well-being. *Journal of Economic Behavior & Organization*, 97, 1-26. doi: 10.1016/j.jebo.2013.10.009

Mortimer, J. T., & Shanahan, M. J. (2007). *Handbook of the life course*: Springer Science & Business Media.

Mroczek, D. K., & Spiro, A. (2005). Change in life satisfaction during adulthood: findings from the veterans affairs normative aging study. *Journal of Personality and Social Psychology*, 88(1), 189-2002. doi: 10.1037/0022-3514.88.1.189.

Munch, A., McPherson, J. M., & Smith-Lovin, L. (1997). Gender, children, and social contact: The effects of childrearing for men and women. *American Sociological Review*, 62(4), 509-520. doi: 10.2307/2657423

Murinkó, L. (2007). Életkor és szubjektív életminőség (Age and subjective well-being). In Á. Utasi (Ed.), Az életminőség feltételei Budapest, Hungary: MTA Politikai Tudományok Intézete.

Musick, K., Meier, A., & Flood, S. (2016). How parents fare: Mothers' and fathers' subjective well-being in time with children. *American Sociological Review*, 81(5), 1069-1095. doi: 10.1177/0003122416663917

Myers, D. (1999). *Well-Being: the foundations of hedonic psychology*. New York: Russell sage foundation.

Myrskylä, M., & Margolis, R. (2014). Happiness: Before and after the kids. *Demography,* 51(5), 1843-1866. doi: 10.1007/s13524-014-0321-x.

Nauck, B. (2007). Value of children and the framing of fertility: Results from a cross-cultural comparative survey in 10 societies. *European Sociological Review,* 23(5), 615-629. doi: 10.1093/esr/jcm028

Neher, P.A. (1971). Peasants, Procreation, and Pensions, *American Economic Review*, 61: 380-389.

Nelson, S. K., Kushlev, K., English, T., Dunn, E. W., & Lyubomirsky, S. (2013). In defense of parenthood: Children are associated with more joy than misery. *Psychological Science*, 24(1), 3-10. doi: 10.1177/0956797612447798

Nelson, S. K., Kushlev, K., & Lyubomirsky, S. (2014). The pains and pleasures of parenting: When, why, and how is parenthood associated with more or less well-being? *Psychological Bulletin*, 140(3), 846. doi: 10.1037/a0035444

Neulinger, A., & Simon, J. (2011). Food consumption patterns and healthy eating across the household life cycle in Hungary. *International Journal of Consumer Studies,* 35(5), 538-544. doi: 10.1111/j.1470-6431.2011.01015.x

Newman, L. (2008). How parenthood experiences influence desire for more children in Australia: A qualitative study. *Journal of Population Research*, 25(1), 1-27. doi: 10.1007/BF03031938

Neyman, J. (1923). On the application of probability theory to agricultural experiments: essay on principles, section 9. Translated in Statistical Science, 5(4), 465–480, (1990).

Ng, T. W., & Feldman, D. C. (2010). The relationships of age with job attitudes: a meta-analysis. *Personnel Psychology*, 63(3), 677-718. doi: 10.1111/j.1744-6570.2010.01184.x

Nikolova, M., & Graham, C. (2014). Employment, late-life work, retirement, and well-being in Europe and the United States. *IZA Journal of European Labor Studies*, 3(1), 1.

Nomaguchi, K. M. (2009). Change in Work-Family Conflict Among Employed Parents Between 1977 and 1997. *Journal of Marriage and Family,* 71(1), 15-32. doi: 10.1111/j.1741-3737.2008.00577.x

Nomaguchi, K. M. (2012). Parenthood and psychological well-being: Clarifying the role of child age and parent–child relationship quality. *Social Science Research*, 41(2), 489-498. doi: 10.1016/j.ssresearch.2011.08.001

Nomaguchi, K. M., & Milkie, M. A. (2003). Costs and rewards of children: The effects of becoming a parent on adults' lives. *Journal of Marriage and Family*, 65(2), 356-374. doi: 10.1111/j.1741-3737.2003.00356.x

O'brien, G. E., & Dowling, P. (1981). Age and job satisfaction. *Australian Psychologist*, 16(1), 49-61. doi: 10.1080/00050068108254415

Oreopoulos, P., Stabile, M., Walld, R., & Roos, L. L. (2008). Short-, medium-, and long-term consequences of poor infant health an analysis using siblings and twins. *Journal of Human Resources*, 43(1), 88-138. doi: 10.3368/jhr.43.1.88

Örkény, A., Koltai, J., & Székelyi, M. (2011). A köznapi igazságossági ítéletek és a generációk közötti igazságossági elvek Közép-Európában= Everyday justice judgements and generational justice in Central Europe. *OTKA Kutatási Jelentések| OTKA Research Reports*. Budapest

Osborne, J. W. (2008). *Best practices in quantitative methods*: Sage.

Palmore, E. (1969). Predicting longevity: a follow-up controlling for age. *The Gerontologist*. 9(4), 247–250. doi: 10.1093/geront/22.6.513

Parr, N. (2010). Satisfaction with life as an antecedent of fertility: Partner+ Happiness= Children? *Demographic Research*, 22, 635-662. doi: 10.4054/DemRes.2010.22.21

Pavot, W., & Diener, E. (1993). Review of the satisfaction with life scale. *Psychological Assessment*, 5(2), 164. doi: 10.1007/978-90-481-2354-4_5

Pedersen, P., & Schmidt, T. (2014). Life events and subjective well-being: The case of having children. *IZA* DP No. 8207

Plagnol, A. C. (2011). Financial satisfaction over the life course: The influence of assets and liabilities. *Journal of Economic Psychology*, 32(1), 45-64. doi: 10.1016/j.joep.2010.10.006

Plagnol, A. C., & Easterlin, R. A. (2008). Aspirations, attainments, and satisfaction: Life cycle differences between American women and men. *Journal of Happiness Studies*, 9(4), 601-619. doi: 10.1007/s10902-008-9106-5

Polese, A., Morris, J., Kovács, B., & Harboe, I. (2014). 'Welfare states' and social policies in Eastern Europe and the former USSR: where informality fits in? *Journal of Contemporary European Stu*dies, 22(2), 184-198. doi: 10.1080/14782804.2014.902368

Pollmann-Schult, M. (2013). Parenthood and life satisfaction in Germany. *Comparative Population Studies*, 38(1). doi: 10.4232/10.CPoS-2013-05en

Pollmann-Schult, M. (2014). Parenthood and Life Satisfaction: Why Don't Children Make People Happy? *Journal of Marriage and Family*, 76(2), 319-336. doi: 10.1111/jomf.12095

Radó, M. (2012). A nyugdíjba vonulási döntések önkéntessége [Voluntary retirement descions]. *Munkaügyi Szemle*, 56(2), 52-63.

Radó, M., Nagy, B., & Király, G. (2016). Work-to-family spillover: Gender differences in Hungary. *Demográfia English Edition*, 58(5). doi: 10.21543/DEE.2015.2

Raley, S., Bianchi, S. M., & Wang, W. (2012). When do fathers care? Mothers' economic contribution and fathers' involvement in child care. *American Journal of Sociology*, 117(5), 1422-1459. doi: 10.1086/663354

Reizer, B. (2011). A gyermekvállalás hatása a család jövedelmére Magyarországon

[Parenthood effect on household income in Hungary]. *Demográfia*, 54(2-3), 160-175.

Rollins, B. C., & Cannon, K. L. (1974). Marital satisfaction over the family life cycle: A reevaluation. *Journal of Marriage and the Family*, 271-282. doi: 10.2307/351153

Rosenbaum, P. R. (1989). Optimal matching for observational studies. *Journal of the American Statistical Association*, 84(408), 1024-1032. doi: 10.1080/01621459.1989.10478868

Rosenbaum, P. R. (2002). *Observational Studies*, Springer: New York.

Rosenbaum, P. R. (1984). The consequences of adjustment for a concomitant variable that has been affected by the treatment. *Journal of the Royal Statistical Society. Series A (General)*, 656-666. doi: 10.2307/2981697

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55. doi: 10.1093/biomet/70.1.41

Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association*, 79(387), 516-524. doi: 10.1080/01621459.1984.10478078

Rosenbaum, P. R., & Rubin, D. B. (1985a). The bias due to incomplete matching. *Biometrics*, 103-116. doi: 10.2307/2530647

Rosenbaum, P. R., & Rubin, D. B. (1985b). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, 39(1), 33-38. doi: 10.2307/2683903

Rothstein, B. (2010). Happiness and the welfare state. *Social Research: An International Quarterly*, 77(2), 441-468. doi: 10.1093/eurpub/ckx186.346

Röcke, C., & Lachman, M. E. (2008). Perceived trajectories of life satisfaction across past, present, and future: profiles and correlates of subjective change in young, middle-aged, and older adults. *Psychology and Aging*, 23(4), 833. doi: 10.1037/a0013680

Rubin, D. B. (1973a). Matching to remove bias in observational studies. *Biometrics*, 159-183. doi: 10.2307/2529684

Rubin, D. B. (1973b). The use of matched sampling and regression adjustment to remove bias in observational studies. *Biometrics*, 185-203. doi: 10.2307/2529685

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5), 688. doi: 10.1037/h0037350

Rubin, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *The Annals of Statistics*, 34-58. doi: 10.1214/aos/1176344064

Russell, J. N. (2008). Residential satisfaction of elderly tenants in apartment housing. *Social Indicators Research*, 89(3), 421-437. doi: 10.1007/s11205-008-9241-8

Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, 57(6), 1069. doi: 10.1037%2F0022-3514.57.6.1069

Sacks, D. W., Stevenson, B., & Wolfers, J. (2010). Subjective well-being, income, economic development and growth. *NBER Working Paper* No. 16441, doi: 10.3386/w16441

Sales, S. M., & House, J. (1971). Job dissatisfaction as a possible risk factor in coronary heart disease. *Journal of Chronic Diseases*, 23(12), 861-873. doi: 10.1016/0021-9681(71)90015-4

Sanchez, L., & Thomson, E. (1997). Becoming mothers and fathers: Parenthood, gender, and the division of labor. *Gender & Society*, 11(6), 747-772. doi: 10.1177/089124397011006003

Sayer, L. C., Bianchi, S. M., & Robinson, J. P. (2004). Are parents investing less in children? Trends in mothers' and fathers' time with children. *American Journal of Sociology*, 110(1), 1-43. doi: 10.1086/386270

Schafer, M. H., Mustillo, S. A., & Ferraro, K. F. (2013). Age and the tenses of life satisfaction. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 68(4), 571-579. doi: 10.1093/geronb/gbt038.

Scharfstein, D. O., Rotnitzky, A., & Robins, J. M. (1999). Adjusting for nonignorable drop-out using semiparametric nonresponse models. *Journal of the American Statistical Association*, 94(448), 1096-1120.

Scharle, Á. (2012). Az aktív korúak ellátásainak jóléti és munkapiaci hatása elméletben. In K. Fazekas, Scarle, Ágota (Ed.), *A magyar foglalkoztatáspolitika két évtizede*, 1990–2010. Budapest: Országos Foglalkoztatási Közalapítvány

Schimmack, U., Oishi, S., Furr, R. M., & Funder, D. C. (2004). Personality and life satisfaction: A facet-level analysis. *Personality and Social Psychology Bulletin*, 30(8), 1062-1075. doi: 10.1177/0146167204264292

Schulz, W. (2000). Explaining Quality of Life -The Controversy between Objective and Subjective Variables. *EuReporting Working Paper*, 10.

Schwarz, N. (2013). *Stimmung als Information: Untersuchungen zum Einfluß von Stimmungen auf die Bewertung des eigenen Lebens* (Vol. 24): Springer-Verlag.

Scott, M., Swortzel, K. A., & Taylor, W. N. (2005). The relationships between selected demographic factors and the level of job satisfaction of extension agents. *Journal of Southern Agricultural Education Research*, 55(1), 102-115. doi: 10.21061/jcte.v27i2.715

Sekhon, J. S. (2008). Multivariate and propensity score matching software with automated balance optimization: the matching package for R. *Journal of Statistical Software*, 42(7), 1–52. doi: 10.18637/jss.v042.i07

Sekhon, J. S. (2009). Opiates for the matches: Matching methods for causal inference. *Annual Review of Political Science*, 12, 487-508. doi: 10.1146/annurev.polisci.11.060606.135444

Sekhon, J. S., & Grieve, R. (2008). A new non-parametric matching method for bias adjustment with applications to economic evaluations. *SSRN Electronic Journal*, doi: 10.2139/ssrn.1138926

Selman, B., & Gomes, C. P. (2006). Hill-climbing Search. Encyclopedia of Cognitive Science. *Encyclopedia of Cognitive Science*, doi: 10.1002/0470018860.s00015

Shultz, K. S., Morton, K. R., & Weckerle, J. R. (1998). The influence of push and pull factors on voluntary and involuntary early retirees' retirement decision and adjustment. *Journal of Vocational Behavior*, 53(1), 45-57. doi: 10.1006/jvbe.1997.1610

Simonovits, A. (2009). Népességöregedés, tb-nyugdíj és megtakarítás-parametrikus nyugdíjreformok [Population aging, the public pension system, and savings: parametric pension reforms]. *Közgazdasagi Szemle [Economic Review]*, 56.

Sirgy, M. J. (2012). *The psychology of quality of life: Hedonic well-being, life satisfaction, and eudaimonia* (Vol. 50): Springer Science & Business Media.

Silverstein, M., & Bengtson, V. L. (1997). Intergenerational solidarity and the structure of adult child-parent relationships in American families. *American Journal of Sociology*, 103(2), 429-60. doi: 10.1017/S0144686X10001108

Sironi, E., & Billari, F. C. (2013). Do union formation and childbearing improve subjective well-being? An application of propensity score matching to a Bulgarian panel Advances in Theoretical and Applied Statistics, *Studies in Theoretical and Applied Statistics*, 351-360: Springer. doi: 10.1007%2F978-3-642-35588-2_32

Spéder, Z., & Kapitány, B. (2002). A magyar lakosság elégedettségének meghatározó tényezői nemzetközi összehasonlításban [Indicators of the Hungarians' life satisfaction in internationakl comparision]. *Társadalmi Riport*, 162-172.

Spéder, Z., & Kapitány, B. (2014). Failure to realize fertility intentions: A key aspect of the post-communist fertility transition. *Population Research and Policy Review*, 33(3), 393-418. doi: 10.1007/s11113-013-9313-6

Spitze, G., & South, S. J. (1985). Women's employment, time expenditure, and divorce. *Journal of Family Issues*, 6(3), 307-329. doi: 10.1177/019251385006003004

Suitor, J. J., & Pillemer, K. (2006). Choosing daughters: Exploring why mothers favor adult daughters over sons. *Sociological Perspectives*, 49(2), 139-161. doi: 10.1525/sop.2006.49.2.139.

Stanca, L. (2012). Suffer the little children: Measuring the effects of parenthood on well-being worldwide. *Journal of Economic Behavior & Organization*, 81(3), 742-750. doi: 10.1016/j.jebo.2010.12.019

Steel, P., Schmidt, J., & Shultz, J. (2008). Refining the relationship between personality and subjective well-being. *Psychological Bulletin*, 134(1), 138. doi: 10.1037/0033-2909.134.1.138.

Steptoe, A., Deaton, A., & Stone, A. A. (2015). Subjective wellbeing, health, and ageing. *The Lancet*, 385(9968), 640-648. doi: 10.1016/S0140-6736(13)61489-0

Stevenson, B., & Wolfers, J. (2009). The paradox of declining female happiness. American Economic Journal: *Economic Policy*, 1(2), 190-225. doi: 10.3386/w14969

Stone, R., Cafferata, G. L., & Sangl, J. (1987). Caregivers of the frail elderly: A national profile. *The Gerontologist*, 27(5), 616-626. doi: 10.1093/geront/27.5.616

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. Statistical science, *Review Journal of the Institute of Mathematical Statistics*, 25(1), 1. doi: 10.1214/09-STS313

Szalai, J. (1991). Hungary: Exit from the state economy. In M. e. a. Kohli (Ed.), *Time for retirement* (pp. 324-362). Cambridge: Cambridge University Press.

Szalma, I., & Takács, J. (2015). Who Remains Childless? Unrealised Fertility Plans in Hungary. *Sociologicky Casopis*, 51(6), 1047. doi: 10.13060/00380288.2015.51.6.228

Szinovacz, M. E., & Davey, A. (2004). Honeymoons and joint lunches: Effects of retirement and spouse's employment on depressive symptoms. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 59(5), 233-245. doi: 10.1093/geronb/59.5.P233

Thévenon, O. (2010). *Fertility in OECD countries: An assessment of macro-level trends and policy responses*. REPRO Reproductive Decision-Making in a Macro-Micro Perspective. Vienna: Vienna Institute of Demography.

Thomas, W. I., & Thomas, D. S. (1928). *The child in America; behavior problems and programs*. New York: Knopf.

Troske, K. R., & Voicu, A. (2009). The effect of children on the level of labor market involvement of married women: what is the role of education? *IZA Discussion Papers* 4074

Tversky, A., & Kahneman, D. (1974). *Judgment under uncertainty: Heuristics and biases Utility, probability, and human decision making* (pp. 141-162): Springer.

Twenge, J. M., Campbell, W. K., & Foster, C. A. (2003). Parenthood and marital satisfaction: a meta-analytic review. *Journal of Marriage and Family*, 65(3), 574-583. doi: 10.1111/j.1741-3737.2003.00574.x

Ubesekera, D., & Luo, J. (2008). Marriage and family life satisfaction: A literature review. *Sabaramuwa University Journal*, 8(1), 1-17. doi: 10.4038/suslj.v8i1.1847

Uhlenberg, P., & Mueller, M. (2003). *Family Context and Individual Weil-Being Handbook of the life course* (pp. 123-148): Springer.

Umberson, D., Pudrovska, T., & Reczek, C. (2010). Parenthood, childlessness, and well-being: A life course perspective. *Journal of Marriage and Family,* 72(3), 612-629. doi: 10.1111/j.1741-3737.2010.00721.x

Vanassche, S., Swicegood, G., & Matthijs, K. (2013). Marriage and children as a key to happiness? Cross-national differences in the effects of marital status and children on well-being. *Journal of Happiness Studies*, 14(2), 501-524. doi: 10.1007/s10902-012-9340-8

Van Landeghem, B. (2008). Human Well-Being over the Life Cycle: Longitudinal Evidence from a 20-Year Panel. *LICOS Discussion Paper* No. 213 doi: 10.2139/ssrn.1360731

Van Landeghem, B. (2012). A test for the convexity of human well-being over the life cycle: Longitudinal evidence from a 20-year panel. *Journal of Economic Behavior & Organization*, 81(2), 571-582. doi: 10.1016/j.jebo.2011.08.001

Van Laningham, J., Johnson, D. R., & Amato, P. (2001). Marital happiness, marital duration, and the U-shaped curve: Evidence from a five-wave panel study. *Social Forces*, 79(4), 1313-1341. doi: 10.1353/sof.2001.0055

Veenhoven, R. (2015). Social conditions for human happiness: A review of research. *International Journal of Psychology*, 50(5), 379-391. doi: 10.1002/ijop.12161

Wang, M., & Hesketh, B. (2012). Achieving well-being in retirement: Recommendations from 20 years of research. *SIOP White Paper Series*, 1(1), 11-22.

Wang, M., Henkens, K., & van Solinge, H. (2011). Retirement adjustment: A review of theoretical and empirical advancements. *American Psychologist*, 66, 991–1009.

doi: 10.1037/a0022414

Weaver, C. N. (1980). Job satisfaction in the United States in the 1970s. *Journal of Applied Psychology*, 65(3), 364. doi: 10.1037/0021-9010.65.3.364

Wells, W. D., & Gubar, G. (1966). Life cycle concept in marketing research. *Journal of Marketing Research*, 355-363. doi: 10.2307/3149851

White, A. T., & Spector, P. E. (1987). An investigation of age-related factors in the age-job-satisfaction relationship. *Psychology and Aging*, 2(3), 261. doi: 10.1037/0882-7974.2.3.261

Winkelmann, R. (2005). Subjective well-being and the family: Results from an ordered probit model with multiple random effects. *Empirical Economics*, 30(3), 749-761. doi: 10.1007/s00181-005-0255-7

Wooldridge, J. M. (2015). *Introductory econometrics*: A modern approach: Nelson Education.

World Health Organization (2006). Reproductive health indicators: guidelines for their generation, interpretation and analysis for global monitoring. ISBN: 92 4 156315 X

Yang, S., Imbens, G. W., Cui, Z., Faries, D. E., & Kadziola, Z. (2016). Propensity score matching and subclassification in observational studies with multi-level treatments. *Biometrics*, 72(4), 1055-1065. doi: 10.1016/j.jval.2014.08.2002

Zacher, H., Jimmieson, N. L., & Bordia, P. (2014). Time pressure and coworker support mediate the curvilinear relationship between age and occupational well-being. *Journal of Occupational Health Psychology*, 19(4), 462. doi: 10.1037/a0036995

Zaidi, A. (2009). Poverty and income of older people in OECD countries.

Zaidi, A. (2011). *Exclusion from material resources: poverty and capability deprivation among older people in EU countries*. European Centre for Social Welfare Policy and Research, Vienna.

Zimmermann, A. C., & Easterlin, R. A. (2006). Happily ever after? Cohabitation, marriage, divorce, and happiness in Germany. *Population and Development Review*, 32(3), 511-528. doi: 10.1111/j.1728-4457.2006.00135.x