Doctoral Thesis

Essays on the economic determinants of health and well-being

Chiara Costi

A thesis submitted in fulfilment of the requirements

for the degree of

Doctor of Philosophy

April, 2023

Declaration of Authorship

I, Chiara Costi, declare that this thesis titled, "Essays on the economic determinants of health and well-being" and the work presented in it are my own and has not been submitted in any other form for the award of a higher degree elsewhere.

Chapter 2 titled "The impact of parental education and prenatal smoking on infant health: an intergenerational approach" is a joint authored piece of work with myself, Giuseppe Migali and Eugenio Zucchelli. A signed declaration of joint authorship can be found in Appendix B.

Chapter 3 titled "Does caring for others affect our mental health? Evidence from the COVID-19 pandemic" is already published in Social Science and Medicine (SSM), Volume 321, March 2023, 115721. A declaration of joint authorship can be found in Appendix B and in the published version of the article at `https://www.sciencedirect.com/science/article/pii/S0277953623000771`.

Chiara Costi

April 2023

Acknowledgements

This PhD journey, and indeed writing this thesis, was possible thanks to the help of so many wonderful people.

Firstly, I would like to thank my supervisors Bruce Hollingsworth, Vincent O'Sullivan and Eugenio Zucchelli. Their friendly and positive guidance was fundamental for me. I have never felt on my own, so thank you for your constant support, useful feedback and our regular meetings. You shaped my academic path and helped me become the researcher I am today.

Starting a PhD with a pandemic was not easy and the first in-person events were like an epiphany for me. I want to thank everyone I have met in these occasions, being surrounded by so many amazing researchers is inspirational. The invaluable feedback I have received especially from health economists at HOPE (UoM) greatly improved this thesis, so thank you all. I also need to thank Themis Pavlidis, who never hesitated to give me the possibility to attend conferences and workshops.

I also need to thank all the people in the department of Economics at Lancaster University, I have learnt a lot from each one of you. I would like to particularly thank Jean-Francois Maystadt who was the one encouraging me to pursue the Master's degree in Economics and who first taught me how to conduct rigorous research.

A PhD in Economics could be cryptic and Machiavellian, but thanks to my PhD colleagues it was not the case. My PhD journey was a little easier and certainly more fun because of you. Special thanks go to Amanda De Pirro and Mario Martinez-Jimenez without whom this journey would not have been the same.

Last but not least, my greatest thank you goes to my family. Words cannot describe how grateful I am, I would not be who I am today without your limitless support and love. To mum, dad, Arianna, Simonetta, Mariuccia, Gino, Carmina and of course my grandma Adriana, my role model (I would not know what to do without our daily calls), this achievement is all yours.

Notes

This PhD has been funded by the Department of Economics at Lancaster University Management School (LUMS).

Each chapter of this thesis presents its own individual literature review. Additionally, this thesis uses a variety of econometric methods which are explained in each chapter, rather than in a dedicated methodology chapter. References for each chapter are combined and placed at the end of the dissertation. Additional figures and tables for each chapter are also combined under a common Appendix section, placed at the end of the thesis.

Chapters presented in conferences and seminars

I presented the chapters of this thesis in several conferences and seminar series.

An early version of chapter 2 titled "The impact of parental education and prenatal smoking on infant health: an intergenerational approach" has been presented in the seminar series of the Health Organisation, Policy and Economics (HOPE) at the University of Manchester. It has received helpful comments from the members of the department, especially from Prof Matthew Sutton. This chapter has also received suggestions from Prof Carol Propper and Prof Gabriella Conti.

Chapter 3 titled "Does caring for others affect our mental health? Evidence from the COVID-19 pandemic" has been presented at: the 16th Ruhr Graduate School in Economics (RGS Econ) at Bochum, DE (2023); the 100th Health Economics Study Group (HESG) meeting in Sheffield, UK (2022); the 9th European Health Economics Association (EuHEA) PhD Conference in Galway, IE (2022); the North West Social Science Doctoral Training PhD Conference (NWSSDTP) in the University of Manchester, UK (2022); the Health Studies User Conference, UK Data Service [virtual] (2022).

The published version of this chapter has benefited from helpful suggestions which have been made particularly by Dr Luke Munford, Dr Sara Rellstab, Prof Scott Cunningham and Prof Jeffrey Wooldridge. This chapter was also presented in the seminar series of the Academic Unit of Health Economics (AUHE) at the University of Leeds (2023) receiving further comments.

Chapter 4 titled "Health and quality of life in ageing populations: A structural equation modelling approach" has been presented at the 2022 French Stata Conference at Aix-Marseille School of Economics, Marseille, FR (2022); the 8th European Health Economics Association (EuHEA) PhD conference, Erasmus University [virtual] (2021); and at the Health Economics Study Group (HESG) meeting in Cambridge University [virtual] (2021).

Data availability

Chapter 2 uses data from the National Longitudinal Study of Adolescent to Adult Health (Add Health), a program directed by Kathleen Mullan Harris at the University of North Carolina. Information on how to obtain the Add Health data files is available on the Add Health website (`http://www.cpc.unc.edu/addhealth`).

Chapter 3 uses data from the UK Household Longitudinal Study (Understanding Society), which is an initiative funded by the Economic and Social Research Council and various Government Departments, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by NatCen Social Research and Kantar Public. Information on how to access the data files is available on the Understanding Society website (`https://www.understandingsociety.ac.uk/`).

Chapter 4 analyses data from the Survey of Health, Ageing, and Retirement in Europe (SHARE). This data can only be accessed upon request, but additional information to enable and facilitate the use of SHARE data is available on the SHARE website (`https://www.share-eric.eu/`).

These data distributors bear no responsibility for the analysis or its interpretation throughout this thesis.

Abstract

This thesis provides an in-depth analysis of socioeconomic determinants of individuals' health and well-being in three key moments of their lives. It contains three pieces of applied work, with each chapter investigating a different longitudinal data set (representative of the US, UK and European populations) while employing several econometric techniques (standard and multiple time periods difference-in-differences models, instrumental variable approaches, structural equation models, propensity score matching and factor analysis). Introduction and conclusions of the thesis are presented in chapter 1 and chapter 4, respectively.

Chapter 2 explores the effect of parental socioeconomic status and risky health behaviours on offspring's infant health. To identify such effects, an intergenerational instrumental variable approach is employed, using grandparents' education and smoking behaviour as instruments for parental characteristics. The National Longitudinal Study of Adolescent to Adult Health (Add Health) is analysed, which contains information on three generations followed over time.

Chapter 3 looks at a later life stage (adulthood) investigating the impact of providing informal care on caregivers' mental health during COVID-19. The UK Household Longitudinal Study (Understanding Society) is analysed, employing a mixture of traditional and novel difference-in-differences models combined with matching.

Chapter 4 is a methodological work assessing the performance of multiple-item scale scores formed with different weighting structures (i.e. composite indices and latent variables), while exploring correlations between socioeconomic factors and

quality of life in older adulthood. Structural equation modelling is employed on data drawn from the Survey of Health, Ageing and Retirement in Europe (SHARE).

The objective of this thesis is to explore socioeconomic factors which affect health and well-being during key periods of individual life-cycle, i.e. childhood, adulthood and older adulthood. This thesis aims at contributing to the literature focused on socioeconomic determinants of health and well-being, employing longitudinal data together with econometric and causal inference methods. The results found in this thesis will inform policymakers to promote timely investments targeting socioeconomic factors enhancing people's health and well-being in order to create (and maintain) a better functioning and healthier society.

Contents

List of Tables

List of Figures

To Gianni, Gigi, Carla and Ilio

Chapter 1

Introduction

For decades, economic research has been focused almost exclusively on material determinants of social welfare, trying to understand factors and mechanisms which could improve the economy of countries both at an aggregate level (e.g., aggregate demand and national economic outputs) and at an individual level (e.g., income, unemployment and education) (Keynes, 1936; Mas-Colell, 1985; Becker, 2017). Besides research around economic activity and financial gains, non-monetary aspects have gained popularity in social sciences playing an increasingly relevant role also within economic research (Easterlin, 1974; Hamermesh, 1977; Freeman, 1977; Grossman, 1972; Wagstaff, 1986). It can be safely said that much has changed in economics over the past few decades (Clark, 2018).

"The ultimate purpose of economics, of course, is to understand and promote the enhancement of well-being"

(Ben Bernanke, Chair of the US Federal Reserve in 2012).[1]

This statement emphasises the importance of understanding socioeconomic determinants of individuals' well-being, which is arguably the ultimate goal of people's lives (López Ulloa et al., 2013). But, what is well-being? And why is it important?

[1]Quote from Ben S. Bernanke's speech in 2012 at the 32nd General Conference of the International Association for Research in Income and Wealth

Well-being is defined as the combination of feeling good and functioning well, which is also a synonym of positive mental health (Ruggeri et al., 2020). The World Health Organization (2001) defines positive mental health as *"a state of well-being in which the individual realizes his or her own abilities, can cope with the normal stresses of life, can work productively and fruitfully, and is able to make a contribution to his or her community"*. Therefore, it is imperative to investigate factors contributing to individuals' health and well-being to create (and maintain) a better functioning and healthier society. Understanding the causes and contributors to such important aspects of people's lives is fundamental to promote timely investments which would produce a high rate of return for individuals' health and well-being, but also (financial) benefits to all of society (Layard et al., 2014).

This thesis investigates socioeconomic determinants of individuals' health and well-being in three key phases of people's lives. Specifically, each chapter of this thesis focuses on one life stage, analysing determinants of: health at birth; mental health in adulthood (specific to informal caregivers); and well-being of older adults. By analysing three different life stages separately, this thesis aims at contributing to the overall picture of different socioeconomic determinants of individuals' health and well-being.

Since causality is central to economic research (Hoover, 2006; Heckman, 2008), this thesis employs multiple econometric techniques attempting to support causal inference. For example, methodological tools such as standard and multiple time periods difference-in-differences models combined with matching techniques as well as instrumental variable approaches are implemented. This thesis also uses structural equation modelling, factor analysis and principal component analysis to assess the performance of indices with different weighting structures while investigating correlations.

More specifically, chapter 2 examines key factors affecting health outcomes in infancy and childhood. It exploits a rich US panel data set (Add Health) including information on multiple generations. The aim of this chapter is to investigate the effects of parental socioeconomic status and risky health behaviours on children's health. To identify such effects, an intergenerational instrumental variable (IV) approach is employed, using grandparents' education and smoking behaviour as instruments for parental education and prenatal smoking. Findings from this chapter show that when parental education and prenatal smoking are considered separately, higher levels of education reduce the probability of offspring's low birth weight while prenatal smoking (both maternal smoking during pregnancy and parental regular smoking) increases it. However, when the education and smoking channels are considered jointly, only the effect of parental education on low birth weight appears to persist. Heterogeneity analyses by gender, ethnicity and alternative early-life health outcomes appear to confirm the findings.

Chapter 3 moves on to a later life stage (adulthood) estimating the impact of providing informal care on caregivers' mental health during COVID-19. Here, longitudinal data are used from the UK Household Longitudinal Study (Understanding Society) employing a mixture of difference-in-differences models combined with matching. While matching accounts for selection on observables into caregiving, multiple period difference-in-differences specifications allow investigation of heterogeneous mental health effects of COVID-19 by timing and duration of informal care. The estimates suggest that while mental health fluctuated following the imposition of social restrictions, informal carers who started caregiving during the pandemic show a large mental health deterioration, especially during lockdowns.

Chapter 4 aims at assessing the performance of multiple-item scale scores formed with different weighting structures, while investigating correlations between socioeconomic factors and quality of life in older adulthood. To do so, the complex concepts of (self-assessed) health and quality of life are measured using latent variables, composite indices with equal weights and composite indices derived from principal component analysis. Structural equation modelling (SEM) is employed since it allows the imputation of latent variables and to separately include latent and composite indices. The Survey of Health, Ageing and Retirement in Europe is used to conduct this methodological exercise. Findings show that non-pecuniary factors such as physical health and participating in social activities seem to have a stronger association with higher quality of life as people age, compared to pecuniary factors such as income and financial assets. Importantly, the use of latent variables appears to be more appropriate when self-assessed variables are included, whereas latent variables and composite indices (constructed with principal component analysis) seem to perform equally well when objective variables are considered.

Overall, understanding which socioeconomic factors affect individuals' health and well-being in key stages of their lives is important to address risk factors through timely investments and effective policies.

Chapter 2

The impact of parental education and prenatal smoking on infant health: an intergenerational approach

2.1 Introduction

It is widely accepted that ill-health at birth is systematically associated with long-term poor health as well as adverse broader educational and labour-market outcomes (Almond and Currie, 2011; Conti and Heckman, 2012; Almond et al., 2018). Therefore, identifying the relative roles played by key factors affecting infant health is one of the most relevant policy issues in human development. In addition, estimating intergenerational effects, especially those of parental socioeconomic status and behaviours on children's health, has always been of great interest to researchers (Maystadt and Migali, 2021; Currie, 2009; Godah et al., 2021; Banderali et al., 2015; Pascal et al., 2018; Kane et al., 2018). Low birth weight, often defined as weighing less than 2,500g, is commonly used in the literature as a predictor of poor general health as it is strongly associated with several short- and long-term adverse health outcomes (Currie and Hyson, 1999; Marmot, 1997; Linsell et al., 2015; Mathewson et al., 2017; Figlio et al., 2014), including higher infant mortality and morbidity (Altman et al., 2012; Zhang and Kramer, 2009; Vilanova et al., 2019), and wider negative outcomes such as lower educational attainment and increased probabilities of unemployment and lower lifetime earnings (Case et al., 2005; Linnet et al., 2006; Behrman and Rosenzweig, 2004; Black et al., 2007; Royer, 2009; Trickett et al., 2020).

Although it is well-established that there is a strong correlation between parental socioeconomic characteristics and children's physical and mental health (Case et al., 2002; Currie and Stabile, 2006; Currie and Lin, 2007; Hedley et al., 2004; Guo and Harris, 2000; Warrington et al., 2019; Hines et al., 2021), it is often difficult to provide causal estimates of this important relationship. One way to identify a possible causal relationship between parents' characteristics and offspring's health would be to make use of instrumental variable approaches exploiting intergenerational data. This would help identify the causal effects of parental education and risky health behaviours on children's health outcomes, by for example exploiting information on grandparents' characteristics. Previous studies have rarely exploited intergenerational information to aid the identification of the causal impact between parental characteristics and offspring's health outcomes, most probably due to the lack of appropriate data. Moreover, no previous studies appear to have explored the effects of parental education and smoking on offspring's birth weight simultaneously.

The main objective of this chapter is to explore the causal impact of parental socioeconomic status (education) and risky health behaviours (smoking) on children's health outcomes by exploiting intergenerational information. To do so, an intergenerational IV approach is employed using rich data spanning three generations drawn from the US National Longitudinal Study of Adolescent to Adult Health (Add Health). This involves estimating the impact of parental education and smoking behaviour (generation II) on their children's birth weight (generation III) by employing grandparents' education and smoking behaviour (generation I) as instruments for the potentially endogenous parental variables. Initially, the effect of parental education and smoking behaviour (including maternal smoking during pregnancy and parental prenatal smoking) on birth weight is considered separately. Subsequently, a multiple IV estimation approach is employed using grandparents' education and smoking behaviour to jointly instrument parental education and prenatal smoking. While the first estimation should test the presence of education- and smoking-led causal pathways on birth weight independently, the second allows identifying the overall or net effect of these two potential channels on birth weight. Importantly, through the latter joint intergenerational IV approach this chapter enables exploring for the first time potential causal effects of both parental education and prenatal smoking simultaneously. This might be of relevance to policymakers devising policies targeted at improving children's health.

The idea behind this intergenerational IV approach is that while grandparents' education and smoking behaviour should influence parental education and smoking, they should not have a direct effect on grandchildren's birth weight (and overall health). This is explored through a series of robustness checks, including employing a sample where grandparents (generation I) died before the birth of their grandchildren (generation III), therefore avoiding any direct contact between individuals of generations I and III, as well as grandparents fixed effects models, testing the role of systematic differences between families of high vs low socioeconomic status.

Instrumental variable approaches have been previously used to overcome potential endogeneity concerns when attempting to identify a causal relationship between parental socioeconomic status, often measured via education, and low birth weight.

Currie and Moretti (2003) use the opening of colleges in the mothers' counties in the US as an instrument for mother's education, since openings might increase educational attainment without directly affecting children's health. Their results show that increased maternal education leads to lower probabilities of low birth weight and preterm birth. McCrary and Royer (2011) employ school entry policies in California and Texas as instruments for education. Their findings show that variations in maternal education due to entry policies do not appear to affect children's health, yet they also find heterogeneous effects by mothers' ethnicity. Chevalier and O'Sullivan (2007) and Lindeboom et al. (2009) exploit the 1947 UK Raising of the School Leaving Age (RoSLA), which increased the minimum school leaving age from 14 to 15, as an instrument for education. The former study finds that maternal education leads to higher mean birth weight (although it does not appear to have an effect on low birth weight) while the latter does not identify the presence of causal effects. Furthemore, Doyle et al. (2005) instrument parental education with the 1957 UK RoSLA, finding no significant effects of parental education on children's subjective general health. Only Kemptner and Marcus (2013) seem to use an instrumental variable approach exploiting information on multiple generations. Using SOEP, a German panel data set, they instrument maternal education via the number of her siblings. Findings show that higher maternal education does not affect offspring's low birth weight, but it improves adolescents' health outcomes.

In addition to the role of education, researchers have also placed great emphasis on understanding the link between parental risky health behaviours and children's health. Specifically, this strand of the literature has mainly focused on the role of parental smoking as a key factor influencing infant health and it is now broadly accepted that there is a strong association between maternal smoking during pregnancy and children's ill-health at birth. A large body of evidence finds that maternal smoking during pregnancy increases both the probability of preterm delivery and low birth weight (Hines et al., 2021; Banderali et al., 2015; Rebagliato et al., 1995; Ventura et al., 2003; Jaddoe et al., 2008; Osborne and Bailey, 2022). In this case, changes in tax rates and different types of smoking cessation policies are often used as instruments in the attempt to identify causal effects. For instance, Permutt and Hebel (1989) find a positive effect of smoking cessation on birth weight using data from a randomized controlled experiment of a smoking cessation intervention targeting pregnant women. Evans and Ringel (1999) and Lien and Evans (2005) employ changes in cigarette taxes in the US as instruments, confirming that smoking lowers birth weight. Wehby et al. (2011) instead make use of genetic markers correlated with smoking as instruments, also finding that prenatal smoking significantly reduces birth weight. Yet, no previous studies have employed an intergenerational IV approach to identify the causal impacts of education and risky health behaviours on children's health and birth weight simultaneously.

This chapter offers several contributions to the literature. First, instruments based on intergenerational information about grandparents' education and smoking behaviour are used to explore the causal impact of parental education and smoking on children's health including birth weight, a relevant early-life outcome with significant long-term effects. These intergenerational instruments have not been previously used in the literature. One of the main advantages of this intergenerational IV approach is that it would allow researchers to employ a causal framework on any panel data sets with detailed information on multiple generations without the need to necessarily exploit either natural experiments or IV based on policy changes.

Second, differently from previous works, this chapter merges and builds on two strands of literature based on the effects of education and risky health behaviours on early-life health. Until now, the literature has focused on either the effect of education or parental smoking on infant health separately. This multiple intergenerational IV approach may shed some further light on the relative roles played by parental education and risky health behaviours, allowing for a more comprehensive examination of different pathways affecting birth weight and other relevant early-life health outcomes.

Third, this chapter also contributes to the recent and growing literature on the effects across multiple generations of socioeconomic status and risky health behaviours. The results confirm that both channels passing through education and smoking behaviour (maternal smoking during pregnancy and parental regular smoking) have a significant effect on the main outcome variable (i.e. low birth weight) as well as on other early-life health outcomes. However, when the effects of parental education and smoking are considered jointly via a multiple IV estimation, only the effect of education appears to persist, decreasing the probability of low birth weight by 1pp. This might suggest that the effect of parental education could dominate the one of parental smoking, although it does not necessarily imply that higher levels of parental education may fully compensate the negative effect of smoking on low birth weight. Nevertheless, higher levels of parental education appear to improve offspring's health, perhaps via the potential mechanism of higher health literacy[1](Lee et al., 2020; Van Der Heide et al., 2013; Vamos et al., 2020), which could improve the health of current and future generations through better life-style choices and health decisions (e.g. Friedman and Hoffman-Goetz (2008); Lee et al. (2015)).

[1]Note that health literacy is defined as *"The degree to which individuals can obtain, process, understand, and communicate about health-related information needed to make informed health decisions"* (Berkman et al., 2010).

2.2 Data

The data used in this analysis are drawn from the National Longitudinal Study of Adolescent to Adult Health (Add Health). Add Health is a panel study representative of US students initially in high school and subsequently followed throughout adolescence and adulthood. The first four waves of Add Health are employed, that is when individuals in the sample are in grade 7-12 (wave I: 1994; wave II: 1995-1996) until they are aged 24-32 (wave III: 2001; wave IV: 2008)[2]. Overall, the sample is composed of 20,745 individuals (generation II in the chapter). This represents an augmented sample obtained from the core sample of the in-home questionnaire of wave I (including 12,105 students randomly selected from 132 schools) and booster samples based on ethnicity, adoption and disability status. Importantly, the additional sample of African-American students have highly educated parents. Information included in the parent questionnaire provides data on marriage, health, education and employment which is completed by parents (generation I) of the individuals responding to the wave I in-home questionnaire. The children-parenting questionnaire of waves III-IV includes instead detailed information about the main respondents' children (generation III). As such, Add Health is suited to this empirical approach as it includes a wide range of variables on three generations followed over time, allowing to explore the role of intergenerational transmission of education and health behaviours on offspring's health outcomes.

Out of the 20,745 main respondents (generation II), 8,234 individuals become parents[3] of 17,137 children. Importantly, Add Health allows to link each main respondent (generation II) with their parents (generation I). Furthermore, among main respondents (generation II) with children, there are 1,410 siblings who have 3,203 children (Table 2.1).

[2] See Harris (2013)

[3] On average, main respondents (generation II) become parents at the age of 24

Table 2.1: Add Health Longitudinal Design

	(Generation II)	
Wave I	Adolescents in grades 7-12	Parents (Generation I)
1994-1995	20,275	17,670
Wave II	Adolescents in grades 8-12	
1996	14,738	
Wave III	Young Adults aged 18-26	
2001-2002	15,197	Children (Generation III)*
Wave IV	Adults aged 24-32	17,137
2008	15,701	

*8,234 generation II individuals had 17,137 children. Of these children: 3,203 are cousins (children of 1,410 generation II siblings)

2.2.1 Key variables

The intergenerational instruments are built by drawing information on grandparents (generation I) from waves I and II in-home questionnaires answered by the 20,745 main respondents (generation II). The level of education of grandparents is defined using years of education corresponding to the highest education level of each grandparent. Following common practice in the education literature aimed at maximizing sample size (see Holmlund et al. (2011)), the empirical model includes a variable based on the sum of both grandparents' years of education. The other instrument used in the analysis is based on the grandparents' smoking status, i.e. a binary variable indicating whether any of the two grandparents were regular smokers. Another binary variable indicating whether grandparents were married is also included in the models as a control.

The four available waves of the in-home questionnaire are used to gather information about main respondents' (generation II) individual-level characteristics before their children were born. Parental education is also measured in years and based on the highest level achieved including dropout. This means that if a respondent is a dropout, an extra year is added to the years of education corresponding to the highest qualification achieved. Importantly, the occurrence of smoking during pregnancy is defined using information provided in waves III-IV and only refers to mothers' smoking behaviour[4], whereas the variable capturing regular smoking defines whether respondents (both mothers and fathers) smoked regularly throughout waves I-IV but before their children were born. The covariates considered in the analyses are: the Peabody vocabulary test score from wave I (a standard test capturing respondents' cognitive abilities, (Dunn and Markwardt, 1970)); wave I and II proxies for risk preferences/attitude and myopic behaviour[5]. These variables capture respondents' risk aversion, which can influence future choices of smoking behaviours and educational attainment. Ethnicity (defined as a binary variable indicating white or non-white) and sex are also included in the analysis (Table 2.2).

The children-parenting questionnaire part of waves III-IV where main respondents (generation II) answer questions related to their children is exploited to define relevant variables about children's health (generation III). The main outcome of interest is a binary variable defining low birth weight as weighting less than 2,500g (5lb 8oz). Alternative children's health outcomes have also been used, such as the continuous variable of birth weight and general health (self-assessed measure answered by the parents). The control variables related to children are: birth order; sex and prenatal care (a binary variable indicating if mothers attended pregnancy check-ups or doctor/nurse-midwife visits for prenatal care).

2.2.2 Descriptive Statistics

Table 2.2 shows descriptive statistics of key variables of the three generations included in the analysis. Grandfathers and grandmothers (generation I) present sim-

[4]When this variable is answered by male respondents (generation II), mothers are the daughters-in-law of grandparents (generation I)

[5]Risk attitude is a general risky attitude binary indicator capturing at least one of the following behaviours: no use of seat belts, no use of birth controls or not resisting to sex if no birth control is in use. It takes the value of 1 if at least one of these behaviours is in place, and 0 otherwise. Myopic behaviour is measured asking respondents if they live their life without much thought for the future, with answers on a 5-point scale. A binary variable is constructed, taking the value of 1 if respondents answered "strongly agree or agree", 0 otherwise

ilar levels of education (around 14 years). Grandfathers tend to smoke more than grandmothers, yet both grandparents report a high percentage of smokers.

Table 2.2: Descriptive statistics

	Mean	SD	Min	Max	N
Generation I					
Grandparents' education (sum of years)	24.58	7.49	0	40	7465
Grandmother education (years)	13.84	2.69	0	20	6876
Grandfather education (years)	13.81	2.78	0	20	5882
Grandparents' smoking	0.69	0.46	0	1	8065
Grandmother smoking	0.51	0.50	0	1	7816
Grandfather smoking	0.61	0.49	0	1	5898
Married	0.68	0.47	0	1	7011
Generation II					
Parental education (average years)	15.01	2.15	9	23	7857
Mother education (years)	15.21	2.10	9	23	4734
Father education (years)	14.70	2.17	9	23	3122
Maternal smoking during pregnancy	0.13	0.34	0	1	8219
Parents' regularly smoking	0.32	0.46	0	1	7911
Vocabulary test	63.28	10.25	0	87	7821
Myopic	0.14	0.35	0	1	8234
Risk attitude	0.40	0.49	0	1	8234
Ethnicity (White)	0.51	0.50	0	1	8234
Gender	0.60	0.49	0	1	8233
Generation III					
Low birth weight	0.10	0.30	0	1	16707
Birth weight (lb)	7.19	1.38	1	15	16004
General health	0.94	0.25	0	1	16565
Birth order	1.94	1.25	1	13	17137
Gender	0.50	0.50	0	1	16893
Prenatal care	0.90	0.29	0	1	17137

As expected, the average level of education of generation II individuals (of both men and women) as measured by the number of school years is higher if compared to the one of generation I. 13% of mothers smoke during pregnancy, whereas the percentage increases to 32% if looking at the variable capturing regular smokers among both mothers and fathers. 51% of individuals in generation II are white (vs non-white), although African-American adolescents with highly educated parents are oversampled (Harris, 2013). Females represent 60% of main respondents (generation II).

Children (generation III) present an even proportion of males and females and 90% of their mothers received prenatal care. Importantly, 10% of the children are born with low birth weight while the average birth weight is 7.19lb, corresponding to around 3,300g. This is in line with standard birth weight measures in the US at the time (Tilstra and Masters, 2020; Kennedy-Moulton et al., 2022; Currie and Moretti, 2007). A very high proportion of the children (94%) have excellent or very good health, while the remaining 6% have good, poor or very poor health. Here the self-assessed health variable is dichotomised with a positive focus on the healthiest group of children, as it is done also in other research (see. Currie and Lin (2007); Cullati et al. (2020); Doiron et al. (2015)).

2.3 Empirical Approach

To estimate the effects of parental education and smoking behaviour on children's early-life health outcomes, an intergenerational instrumental variable (IV) approach is used. Indeed, the simple correlation between parental socioeconomic characteristics and children's health outcomes is likely to be biased by unobserved factors which might affect both treatments (parental variables) and outcomes (children's health). In the presence of valid and relevant instruments, instrumental variables are often used to overcome endogeneity concerns. Validity essentially requires that the instrument should affect the outcome only through an exogenous variation of the treatment. Relevance instead indicates that the instrument has a sufficiently large explanatory power with respect to the treatment (or in other words that the instrument is highly correlated with the treatment). Using a 2 stage least squares (2SLS) approach, a valid and relevant IV produces a local average treatment effect (LATE) of the parameters of interest.

The main identification strategy initially considers education and smoking behaviour as two separate channels. Equation (2.1a) shows the first stage of the IV approach, which estimates the intergenerational correlation between grandparents' education ($YoEd_{genI}$) and parental education ($YoEd_{genII}$). Grandparents' education is assumed to generate an exogenous variation in parents' education, and this is exploited to instrument the effect of parental education ($Y\hat{o}Ed_{genII}$) on children's birth weight ($LowBW_{genIII}$), see equation (2.1b).

$$YoEd_{genII} = \gamma_0 + \gamma_1 YoEd_{genI} + \gamma_2 Ch_{genI} + \gamma_3 ChEd_{genII} + \epsilon \tag{2.1a}$$

$$LowBW_{genIII} = \beta_0 + \beta_1 Y\hat{o}Ed_{genII} + \beta_2 Ch_{genIII} + \xi \tag{2.1b}$$

The same approach is repeated with smoking behaviours, estimating in the first stage the intergenerational correlation between grandparents' smoking behaviour ($RegSmok_{genI}$) and parental smoking ($Smok_{genII}$), see equation (2.2a). This correlation is exploited to instrument parental smoking ($S\hat{m}ok_{genII}$) in the second stage, estimating its effect on children's birth weight ($LowBW_{genIII}$), see equation (2.2b).

$$Smok_{genII} = \gamma_0 + \gamma_1 RegSmok_{genI} + \gamma_2 Ch_{genI} + \gamma_3 ChSmk_{genII} + \epsilon \tag{2.2a}$$

$$LowBW_{genIII} = \beta_0 + \beta_1 S\hat{m}ok_{genII} + \beta_2 Ch_{genIII} + \xi \tag{2.2b}$$

The second identification strategy considers the joint effect of the two causal pathways of education and prenatal smoking on birth weight via a multiple IV approach. Two first stages are considered separately. Equation (2.3a) estimates the intergenerational correlation between grandparents' education and parental education, also including the effect of grandparents' smoking behaviour. Equation (2.3b) estimates the intergenerational correlation between grandparents smoking and parental smoking, including the effect of grandparents' education. In the second stage, the channels are first estimated separately (including one instrumented variable at a time[6]). Subsequently, the joint effect of the instrumented parental smoking and parental education is estimated on children's birth weight as equation (2.3c) shows.

$$YoEd_{genII} = \gamma_0 + \gamma_1 YoEd_{genI} + \gamma_2 RegSmok_{genI} + \gamma_3 Ch_{genI} + \gamma_4 ChEd_{genII} + \epsilon \quad (2.3a)$$

$$Smok_{genII} = \delta_0 + \delta_1 YoEd_{genI} + \delta_2 RegSmok_{genI} + \delta_3 Ch_{genI} + \delta_4 ChSmk_{genII} + \mu \quad (2.3b)$$

$$LowBW_{genIII} = \beta_0 + \beta_1 Yo\hat{E}d_{genII} + \beta_2 Sm\hat{o}k_{genII} + \beta_3 Ch_{genIII} + \xi \quad (2.3c)$$

The 2SLS models are initially estimated without covariates. Subsequently, further information is included in the first stage such as marital status of grandparents (generation I) (Ch_{genI}), and more specific covariates depending on the channel considered (education or smoking). Specifically, in the case of education, parents' (generation II) Peabody vocabulary test average scores, ethnicity and sex are included ($ChEd_{genII}$). As for the smoking pathway, the covariates considered are binary variables defining respondents' risk attitude and myopic behaviour ($ChSmk_{genII}$). In the second stage, the same set of controls in both channels is included, such as children's birth order, sex and mothers' prenatal care (generation III) (Ch_{genIII}).

2.3.1 Assessment of IV

The relevance of the instruments is evaluated relying on their statistical significance observed in the first stage, and on their F-statistics, which are well above the widely

[6]Note that differently from the first identification strategy, here both grandparents' education and smoking behaviour are used to instrument parents' education and smoking behaviour

accepted threshold (Keane and Neal, 2023; Lee et al., 2022).

In this intergenerational IV approach, the assumption is that grandparents' education generates a variation in parents' education, but has no direct effect on grandchildren's birth weight. It is also assumed that grandparents (regular) smoking, while affecting mothers' smoking during pregnancy and more broadly parents' (regular) smoking behaviour[7], should not directly affect grandchildren's birth weight. While this might be reasonable to assume for grandparents' education, it might not be categorically excluded that some of the grandchildren's health outcomes used in this study may be partly influenced by grandparents' smoking behaviour. For this reason, estimates of these relationships are also provided for a sub-sample of children whose grandparents died at the start of the study (see Tab 2.6, Section 2.5.1), that is before their grandchildren were born. This should avoid any direct influence of grandparents' smoking on their grandchildren's health.

It might also be argued that the genetic endowment between grandparents and grandchildren might violate the validity assumption. This would imply assuming that grandparents might have a direct effect on their grandchildren's birth weight via their partly shared genetic endowment. Although specific genetic variants are strongly correlated with some traits, they often explain a very small portion of the variation of offspring's health outcomes (Hirschhorn and Daly, 2005; Gibson, 2012; Price et al., 2015). In this case, grandparents' genetic traits related to education and smoking behaviour are not expected to have a strong residual direct effect on grandchildren's infant health status via their partly shared genetic endowment[8]. However, it might be possible that a combination of genetic endowment and environmental

[7]Note that individuals who start smoking after having children are excluded to ensure that smoking behaviour is defined before the birth of generation III

[8]In the case of grandparents' education, this would imply that the level of education of a grandparent might influence their grandchildren birth weight somehow directly and genetically and not just indirectly via the potentially improved socioeconomic status of their offspring (i.e. the grandchildren's parents). In the case of smoking, this would assume that the gene increasing the probability of smoking behaviour among the grandparents would have a residual direct effect on their grandchildren's infant health/birth weight not through the smoking behaviour of their offspring (the grandchildren's parents).

factors could have a direct effect (Jami et al., 2021; Bohacek and Mansuy, 2013). To further investigate this point and exclude the roles of unobserved gene-environment interactions via grandparents, grandparents fixed effects models are employed on families of higher vs lower socioeconomic status to explore the effect of a positive vs negative environment (and genetic endowment) on children's health.

In addition, alternative sample restrictions are also tested. Specifically, additional analyses further restrict the sample to the same set of children whose parents have information on both education and smoking behaviour; and first born children only. This allows considering whether the main results are confirmed throughout different (relevant) samples.

2.4 Results

2.4.1 Main results: effects of parental education and smoking

Table 2.3 shows the results of the main model considering parental education and smoking behaviour as separate channels potentially affecting children's birth weight. Estimates reported in Panel 1, columns 1 and 2, show a strong intergenerational correlation between grandparents' (generation I) and parents' (generation II) education. An additional year of grandparents' education increases by between 0.06 and 0.08 years the education of parents. The effect is statistically significant at 1% and confirms the relevance of the instrument in the first stage of the 2SLS estimation. In Panel 2, it can be observed that parental education (generation II) has a negative effect on the variable defining low birth weight. This implies that an additional year of parental education reduces the probability of low birth weight by 1pp (generation III). Importantly, the estimated effect remains highly statistically significant when further including covariates (col. 2).

As displayed in Panel 1 columns 3-6, grandparents regular smoking is associated

Table 2.3: Separate IG effects of parental education and smoking on low birth weight

	(1)	(2)	(3)	(4)	(5)	(6)
Panel 1 - IGT generation I and II						
	$YoEd_{genII}$		$Smokpreg_{genII}$		$RegSmok_{genII}$	
YoEd_{genI}	0.084***	0.063***				
	(0.004)	(0.004)				
RegSmok_{genI}			0.091***	0.089***	0.156***	0.151***
			(0.007)	(0.007)	(0.011)	(0.011)
Controls		✓		✓		✓
N	6222	6222	8050	8050	7756	7756
F	503.63	258.46	155.67	150.27	209.52	197.81
Panel 2 - Effects on Low birth weight						
YoEd_{genII}	-0.015***	-0.014***				
	(0.005)	(0.003)				
Smokpreg_{genII}			0.068	0.104**		
			(0.050)	(0.049)		
RegSmok_{genII}					0.048	0.085***
					(0.033)	(0.030)
$\text{Birthorder}_{genIII}$		0.003		0.006**		0.007***
		(0.003)		(0.002)		(0.002)
Female_{genIII}		0.019***		0.018***		0.018***
		(0.005)		(0.005)		(0.005)
$\text{Prenatalcare}_{genIII}$		-0.019*		-0.028***		-0.027***
		(0.010)		(0.010)		(0.009)
N	12686	12686	16182	16182	15586	15586

Clustered Standard errors at family level in the 1st stage
Robust Standard errors in the 2nd stage
Significance levels: *** 1% ** 5% * 10%

with a higher probability (around 9pp) that mothers[9] would smoke during pregnancy and that, more generally, parents would also regularly smoke (with a 15pp increase). These findings show a strong and highly statistically significant intergenerational effect, and confirm the relevance of this instrument. In Panel 2, an effect (of around 10pp) of maternal smoking during pregnancy on the probability of low birth weight is observed, though this effect becomes statistically significant only after adding children's controls (see col. 4). Indeed, children's birth order, sex and prenatal care are all statistically significant and with the expected signs, suggesting that those variables might also influence birth weight and should be considered when evaluating the effects of parents' smoking behaviours. Similarly, the effect of parental smoking also becomes highly statistically significant (increasing the probability of low birth weight by 8.5pp) when covariates are included (see col. 6). Birth order could influence birth weight through advanced maternal age and through differences in maternal weight gain, while attending prenatal care visits can improve in-utero conditions and hence affecting birth weight. Furthermore, male newborns weigh on average more than female newborns and so all these variables are important controls to consider in the analysis.

In Table 2.4 the results of the second identification strategy are presented, based on the joint use of the two intergenerational instrumental variables (when including covariates). In the first stage, the joint effect of grandparents' education and smoking behaviour (generation I) on corresponding parental variables (generation II) are considered simultaneously[10]. Looking at Panel 1, the positive intergenerational transmission of education (col. 1) does not appear to be affected by the inclusion of grandparents regularly smoking, which presents a large negative effect on parental education. The effect of parental education on smoking during pregnancy appears

[9] Note that grandparents' daughters-in-low are also included.

[10] Note that differently from the set up leading to results of the main identification strategy, here both grandparents' education and smoking behaviour are used to instrument parents' education and smoking behaviour.

to be small, while the effect on regular smoking is not statistically significant. Conversely, the intergenerational transmission of smoking does not appear to be affected by the inclusion of grandparents' education, showing highly statistically significant and positive estimates (cols. 2 and 3).

In Panel 2, the education and smoking channels are estimated separately. In this case, it can be noticed that the effect of parental education (generation II) on low birth weight appears to be very similar to the findings in Table 2.3, whereas the effect of maternal smoking during pregnancy (col. 2) and parental regular smoking (col. 3) seem slightly stronger (13.7pp and 9.5pp, respectively) and statistically significant at 1% level.

In Panel 3, the joint effect of parental education and smoking on low birth weight is explored. When the effects of parental education and smoking (either maternal smoking during pregnancy or parental regular smoking) are considered jointly, only the effect of education appears to persist with a magnitude and significance similar to what previously found. This may imply that the (positive) net effect on low birth is driven by education. However, it might be also that the sample used for the joint IV analysis is smaller and that the samples in columns 1-2 are different. To account for this, the multiple IV approach is also employed on a sample of children for whom there is information on both parental education and smoking behaviour (Tab A2 in the Appendix) confirming that education drives the effect on children's health and also showing a reduction in the probability of low birth weight by around 1pp.

Table 2.4: Combined IG effects of parental education and smoking on low birth weight

	(1)	(2)	(3)
Panel 1 - IGT generation I and II			
	$YoEd_{genII}$	$Smokpreg_{genII}$	$RegSmok_{genII}$
YoEd_{genI}	0.059***	-0.001**	0.001
	(0.004)	(0.001)	(0.001)
RegSmok_{genI}	-0.494***	0.089***	0.149***
	(0.055)	(0.008)	(0.011)
Controls	✓	✓	✓
N	6162	7293	7062
F	174.58	71.49	85.36
Panel 2 - Effects on Low birth weight			
Separate channels			
YoEd_{genII}	-0.014***		
	(0.003)		
Smokpreg_{genII}		0.137***	
		(0.051)	
RegSmok_{genII}			0.095***
			(0.031)
$\text{Birthorder}_{genIII}$	0.003	0.005*	0.006**
	(0.003)	(0.003)	(0.002)
Female_{genIII}	0.020***	0.020***	0.019***
	(0.005)	(0.005)	(0.005)
$\text{Prenatalcare}_{genIII}$	-0.019*	-0.028***	-0.025***
	(0.010)	(0.010)	(0.010)
N	12518	14640	14170
Panel 3 - Effects on Low birth weight			
Joint channels			
YoEd_{genII}	-0.014***	-0.012***	
	(0.003)	(0.003)	
Smokpreg_{genII}	-0.023		
	(0.060)		
RegSmok_{genII}		0.012	
		(0.037)	
$\text{Birthorder}_{genIII}$	0.003	0.002	
	(0.003)	(0.003)	
Female_{genIII}	0.020***	0.018***	
	(0.005)	(0.005)	
$\text{Prenatalcare}_{genIII}$	-0.017	-0.024**	
	(0.011)	(0.010)	
N	12508	12098	

Clustered Standard errors at family level in the 1st stage

Robust Standard errors in the 2nd stage

Significance levels: *** 1% ** 5% * 10%

2.4.2 Heterogeneity: the roles of sex and ethnicity

In Table 2.5 the model estimated in Table 2.3 is replicated focusing on sex. More specifically, Panel 1 explores the intergenerational correlation between individuals in generations I and II of the same sex, that is between grandmothers and mothers and grandfathers and fathers, respectively, for both education and smoking behaviour. Strong intergenerational correlations of education for both females and males are found, with an estimated coefficient of around 0.2 (cols. 1 and 4) which is in line with results from the specialised literature on the intergenerational transmission of education (Holmlund et al., 2011). There is also a strong and highly statistically significant intergenerational transmission of smoking for both sexes (cols. 2, 3 and 5).

To investigate the effects on low birth weight among grandchildren (generation III), Panel 2 shows estimates for models considering all grandchildren regardless of sex, while Panel 3 displays estimates for female and male grandchildren separately. Importantly, when looking at the triads of grandmothers-mothers-grandchildren and grandfathers-fathers-grandchildren (Panel 2, cols. 1 and 4), education has a small and weakly statistically significant effect for the former, but it has a highly statistically significant effect for the latter reducing birth weight by 3pp. These effects are confirmed also when the triads are further restricted to grandmothers-mothers-granddaughters and grandfathers-fathers-grandsons (Panel 3, cols. 1 and 4). Here, estimates for males are of (slightly) larger magnitude compared to those of Panel 2, while female education has no statistically significant effect. These results might suggest that paternal education is important in determining offspring health outcomes (Giuntella et al., 2022). A possible explanation underlying this finding could be that higher paternal education generates a rise in household income through higher returns to education and through assortative mating. This can contribute to the socioeconomic position of the household which is important in determining

Table 2.5: IG effects of parental education and smoking on low birth weight, by sex

	(1)	(2)	(3)	(4)	(5)
Panel 1 - IGT generation I and II					
		Female		***Male***	
	YoEd_{genII}	Smokpreg_{genII}	RegSmok_{genII}	YoEd_{genII}	RegSmok_{genII}
Mother's YoEd_{genI}	0.212***				
	(0.013)				
Mother's RegSmok_{genI}		0.126***	0.157***		
		(0.011)	(0.014)		
Father's YoEd_{genI}				0.199***	
				(0.018)	
Father's RegSmok_{genI}					0.121***
					(0.019)
Controls	✓	✓	✓	✓	✓
N	3636	4675	4537	1978	2351
F	247.30	130.65	131.82	117.85	39.99
Panel 2 - Effects on Low birth weight of female and male children					
YoEd_{genII}	-0.007*			-0.030***	
	(0.004)			(0.006)	
Smokpreg_{genII}		0.099**			
		(0.040)			
RegSmok_{genII}			0.094***		0.138***
			(0.035)		(0.051)
$\text{Birthorder}_{genIII}$	0.003	0.004	0.004	0.005	0.005
	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)
Female_{genIII}	0.023***	0.020***	0.017***	0.006	0.019**
	(0.007)	(0.006)	(0.006)	(0.010)	(0.009)
$\text{Prenatalcare}_{genIII}$	-0.024*	-0.019*	-0.018	-0.029	-0.031*
	(0.013)	(0.012)	(0.012)	(0.018)	(0.017)
N	7870	9966	9681	3594	4197
Panel 3 - Effects on Low birth weight of female or male children					
YoEd_{genII}	0.007			-0.037***	
	(0.006)			(0.008)	
Smokpreg_{genII}		0.137**			
		(0.055)			
RegSmok_{genII}			0.121***		0.142**
			(0.047)		(0.067)
$\text{Birthorder}_{genIII}$	0.011**	0.004	0.005	0.001	0.004
	(0.005)	(0.004)	(0.004)	(0.007)	(0.006)
$\text{Prenatalcare}_{genIII}$	-0.020	-0.026	-0.018	-0.025	-0.032
	(0.019)	(0.017)	(0.017)	(0.025)	(0.023)
N	3971	5034	4871	1867	2143

Clustered Standard errors at family level in the 1st stage

Robust Standard errors in the 2nd stage

Significance levels: *** 1% ** 5% * 10%

offspring health at birth.

Looking at smoking behaviour, the strong first stage intergenerational correlation, already observed in Panel 1 is confirmed. Considering the triads grandmothers-mothers-grandchildren (Panel 2, cols. 2 and 3) both smoking during pregnancy and smoking regularly increase the probability of low birth weight by around 9pp. Similar effects, though larger in magnitude, are obtained for the triads grandmothers-mothers-granddaughters (Panel 3, cols. 2 and 3). For males, there appears to be also an effect of smoking regularly when considering the triads grandfathers-fathers-grandchildren and grandfathers-fathers-grandsons, increasing the probability of low birth weight by around 14pp (Col.5 Panel 2 and 3, respectively).

The main model is further replicated focusing on the ethnicity of individuals in generations I and II, defined as white vs non-white. As shown in Table A1 in the Appendix, results for whites are similar to those reported in Table 2.3 for first and second stage, although estimates related to smoking are now weakly statistically significant. Non-whites show a significant effect of parental education on low birth weight as well as large and significant effects of maternal smoking during pregnancy and regular smoking. According to this heterogeneity analysis, it seems that smoking during pregnancy is particularly detrimental for non-white newborns. A possible explanation could be that, within the US context, racial segregation also plays a role in smoking behaviours, with higher probability of maternal smoking during pregnancy for non-whites when living in racially segregated areas (Yang et al., 2014). This could in turn contribute to poorer health at birth of their offspring.

2.5 Robustness checks

A series of findings is also presented using alternative sample restrictions and further early-life health outcomes to check the robustness of the main results.

The most restrictive analysis focuses on the same sample of children for whom there is information on both parental education and smoking behaviour (see Table A2 in the Appendix). Panel 1 of Table A2 shows quantitatively similar but less precise estimates of the education and smoking pathways if compared to the ones found in Panel 2 of Table 2.4. Panel 2 shows the joint effect of parental education and smoking on children's low birth weight. As previously found, only education appears to persist with a similar magnitude (around 1pp) and the same statistical significance. Therefore, this additional analysis appears to confirm that the education effect "dominates" the smoking effect.

So far, all available children (generation III) were included, controlling for birth order to consider the presence of siblings. When restricting the analysis to generation III first-born children only, as in Currie and Moretti (2003), results are also similar to those found in the main analysis (see Table A3 in the Appendix). The effect of the variable defining smoking during pregnancy does not appear to be statistically significant this time, though sign and magnitude are those expected.

Additional robustness checks are also performed using alternative health outcomes in the second stage. Keeping the same specification of the first stage as in the main model reported in Table 2.3, a continuous variable (standardised) measuring overall birth weight is now investigated (see Table A4). As expected, the sign and the effect of parental education is positive, implying that more educated parents have healthier children. Also, both maternal smoking during pregnancy and parental regular smoking reduces birth weight.

Finally, when using a measure of general health[11] for children (generation III),

[11]Parents (generation II) are asked to rate how good is their children's health on a 5-point scale,

findings show that higher levels of parental education increase the probability to have a child in excellent/very good health by around 1pp. Parental smoking has the opposite effect and reduces the same probability by 6.5pp when parents are regular smokers (see Table A5).

2.5.1 Robustness of Instruments

The relevance of the instruments used in the estimation approach can be evaluated looking, in the first stage, at the statistical significance of the correlation between the instrument and the endogenous variable and at their F-statistics. The F-statistics, reported in each table of the results, show values well above the widely accepted threshold (Keane and Neal, 2023; Lee et al., 2022), further confirming the relevance of the instruments used in this analysis.

However, there is no formal test to assess the validity of the instrument, testing for the absence of any possible correlation between the instrument and the outcome via unobservables. A potential threat to the validity of the instrument might be related to the fact that grandparents' smoking could affect grandchildren's low birth weight, if children were directly exposed to grandparents' smoking while in-utero. To overcome this concern, the models for parental smoking behaviours are re-estimated by restricting the sample to generation I grandparents who have died before their grandchildren were born. By doing so, grandparents' smoking behaviour could not directly affect the outcome of interest. Looking at Table 2.6, Panel 1 shows a strong intergenerational correlation between grandparents' (generation I) and parents' (generation II) smoking behaviour, especially related to regular smoking. In Panel 2, results suggest that parental regular smoking contribute to substantially higher probabilities of low birth weight of around 29pp (col. 2), while maternal smoking during pregnancy is statistically significant only at the 10% level.

ranging from excellent to poor. This self-assessed health variable is here dichotomised with a positive focus on the healthiest group of children (with excellent or very good health).

Table 2.6: IG effects on low birth weight, dead grandparents

	(1)	(2)	(3)	(4)
Panel 1 - IGT generation I and II				
	Wave (1-2-3)		***Wave (1-2)***	
	Smokpreg$_{genII}$	RegSmok$_{genII}$	Smokpreg$_{genII}$	RegSmok$_{genII}$
RegSmok$_{genI}$	0.062**	0.140***	0.049	0.122***
	(0.029)	(0.039)	(0.032)	(0.044)
N	615	590	481	466
F	4.66	12.72	2.33	7.84
Panel 2 - Effects on Low birth weight				
Smokpreg$_{genII}$	0.484*		0.606	
	(0.280)		(0.500)	
RegSmok$_{genII}$		0.298**		0.282
		(0.145)		(0.193)
Birthorder$_{genIII}$	0.006	0.007	-0.000	0.002
	(0.009)	(0.008)	(0.012)	(0.009)
Female$_{genIII}$	0.024	0.019	0.016	0.006
	(0.019)	(0.019)	(0.024)	(0.021)
Prenatalcare$_{genIII}$	-0.017	0.018	-0.079	0.014
	(0.033)	(0.031)	(0.064)	(0.040)
N	1341	1290	1024	999

Clustered Standard errors at family level in the 1st stage
Robust Standard errors in the 2nd stage
Significance levels: *** 1% ** 5% * 10%
Controls are not included in the first stage

Although the estimates are less precisely estimated, they still confirm the main results. Importantly, grandparents who died in waves I-II (cols. 3 and 4) passed away before the birth of the grandchildren, while this might not be the case for those grandparents reported to be deceased in wave III. While columns 1-2 of Table 2.6 refer to grandparents who were reported as deceased in all three waves, columns 3-4 instead includes observations for those grandparents who died in waves I-II (therefore excluding any possible direct interaction between grandparents and grandchildren). In this case, the estimates are still in the expected direction and magnitude (as in cols. 1 and 2) but not statistically significant due to the smaller sample size and the higher standard errors. Nevertheless, this test appears to provide evidence confirming the presence of the effect of interest, even in the absence of a direct effect of grandparents' smoking on grandchildren's birth weight (through a shared environment).

The effect of parental education on children's low birth weight is also estimated using grandparents fixed effects models. In this way, further unobserved grandparents' characteristics such as family environment and genetic endowment can be controlled for. Specifically, these fixed effects models allow exploiting the variation in parents' education (who are siblings) by keeping fixed the grandparents' environment and so also potentially accounting for the (partly) shared genetic endowment.

Specifically, grandparents with higher (lower) socioeconomic status, proxied by higher (lower) grandparents' educational attainment, could have a direct effect through a combination of environmental and genetic factors on children's health outcomes. Therefore, the fixed effects models are conducted on samples of grandparents (generation I) as divided into highly educated families (e.g. with at least one grandparent with a bachelor degree) vs low educated families (e.g. grandparents with secondary education or less). Using this approach, results from Table A6 in the Appendix do not show any statistical significant effect of parental education (generation II) on the probability of low birth weight (generation III) between families with a higher vs lower socioeconomic status. This might imply that there may not be an obvious genetic or environmental "premium" for children born in families with highly educated grandparents compared to those born into families with less educated grandparents. This also appears to support overall the intergenerational IV approach.

2.6 Discussion

This chapter explores the causal impact of parental socioeconomic status and risky health behaviours, via parental education and smoking, on children's early-life health outcomes. To identify these effects, an intergenerational IV approach is used on rich multi-generational information from Add Health. This involves using grandparents' education and smoking behaviour (generation I) as instruments for parental education and smoking behaviour (generation II). When the causal pathways of parental education and smoking behaviour are considered separately, higher levels of parental education have a positive effect on children's early life health (both in terms of birth weight and general health), while parental smoking (regular smoking of either parent before the child was born and maternal smoking during pregnancy) increases the likelihood of low birth weight and worse general health. Interestingly, when the parental education- and smoking-led pathways to children's health are considered jointly, the effect of education appears to dominate, suggesting a residual positive effect due to higher levels of parental education. A series of robustness checks support the main results and the validity of the instruments. These include an IV model based on a sample of grandparents who died before the birth of their grandchildren, to exclude any interactions between generations I and III, and grandparents fixed effects models, to account for (unobserved) shared genetic and environmental factors.

One of the main contributions of this intergenerational IV approach is that, in the presence of appropriate research designs, it could be applied to potentially all panel data sets including information on multiple generations, without the need of exploiting policy changes or natural experiments. This would therefore provide a more general, and potentially more universally applicable, set of instruments to identify the causal impact of either parental education or parental smoking on offspring's infant health.

The findings of this chapter seem to confirm established results previously found in the literature, with parental education contributing to better children's health (Godah et al., 2021; Kemptner and Marcus, 2013) and with parental prenatal smoking deteriorating offspring's health outcomes (Osborne and Bailey, 2022; Hines et al., 2021). Interestingly, this chapter presents estimates related to the education channel which are of the same magnitude as those found in Currie and Moretti (2003), with parental education decreasing the probability of low birth weight by 1pp. Also for parental smoking, the results found here are similar to the ones presented in Lien and Evans (2005), increasing the likelihood of low birth weight by 10pp.

Importantly, this chapter contributes directly to the current literature by providing a more comprehensive examination of different, yet potentially simultaneous, pathways affecting relevant early-life health outcomes by exploring the relative roles played by parental education and risky health behaviours. Although findings from the joint IV model appear to suggest that the (positive) effect of education might dominate over the (negative) effect of parental smoking on low birth weight, this does not necessarily imply that higher levels of parental education might fully compensate for the negative effect of smoking on low birth weight.

From a policy-making perspective, investing in educational programs and intervention strategies aimed at increasing education might have immediate and long-term health impacts (Lee et al., 2020). For example, interventions reducing schooling drop-out as well as health promotion initiatives targeting smoking prevention and cessation, especially during pregnancy, can play a significant role in increasing awareness of the negative impact that these choices have on one's own health as well as on their offspring health. Such initiatives could be particularly effective if coupled with health literacy programs carried out in schools. Developing health literacy early in life is important to adopt healthy behaviours, preventing health problems and better managing illnesses. Having a high health literacy could also impact offspring's health outcomes through an improved parents' ability to more correctly interpret or better adhere to medical advice, managing children's medications or adopting overall healthier behaviours during pregnancy (Harrington et al., 2015; Betz et al., 2008). For this reason, investing in educational programs and implementing intervention strategies aimed at promoting education and health literacy could be useful to improve the health of both current and future generations (Bayati et al., 2018; Solhi et al., 2019).

Chapter 3

Does caring for others affect our mental health? Evidence from the COVID-19 pandemic

3.1 Introduction

There is increasing evidence that the COVID-19 pandemic and related social restrictions are having an effect on mental health (Banks and Xu, 2020; Le and Nguyen, 2021; Lorenz-Dant and Comas-Herrera, 2021; Muldrew et al., 2022; Zhou and Kan, 2021). Recent studies show increases in loneliness, worry and boredom (Brodeur et al., 2021) together with overall worse mental health following lockdowns (Serrano-Alarcón et al., 2022). This fast-growing literature has mainly focused either on the general population or specific sub-groups, including young adults, women, ethnic minorities, households with children, and the least educated (Anaya et al., 2021; Daly et al., 2022; Etheridge and Spantig, 2020; Niedzwiedz et al., 2021; Proto and Quintana-Domeque, 2021; Zhou and Kan, 2021). However, less is known about the effect of COVID-19 on the psychological well-being of informal carers, one of the most affected yet potentially vulnerable groups of individuals (Rodrigues et al., 2021).

Before the COVID-19 outbreak, informal care was already considered essential for the sustainability of publicly funded healthcare systems (Lorenz-Dant and Comas-Herrera, 2021) because informal care is seen as a 'cost saving' alternative to formal care, for instance saving the UK Government £132 billion annually (Carers UK, 2015). In addition, it is often preferred by care recipients when provided by relatives or friends (Carers UK, 2020). When the pandemic began, the formal healthcare sector was overwhelmed by COVID-19 and long-term care systems were heavily disrupted (Giebel et al., 2021). Since then, informal carers are viewed more as frontline healthcare workers, caring for an ever-increasing number of vulnerable individuals (Kent et al., 2020). More generally, given the fast-growing older population and the increasing prevalence of age-related illnesses (Onwumere, 2020), governments are increasingly relying on this form of assistance (Lacey et al., 2019). Thus, it would be relevant to assess whether the pandemic causally affected the mental health of such an important category of healthcare workers.

While during the pandemic governments recommended avoiding close contact with the elderly and frail, the sudden disruption of most formal care services led existing carers to provide additional care, as well as many individuals starting to provide care (Carers UK, 2020). It is estimated that around 26% of the individuals in the UK population are currently providing some form of informal care (Onwumere et al., 2021) and 4.5 million people became informal carers after the COVID-19 outbreak (Carers UK, 2020). Since providing informal care is often associated with an increase in psychological distress (Adelman et al., 2014), it is plausible that the pandemic harmed the mental health of informal carers, especially among those who started providing care during the pandemic. New carers might have been concerned about infecting care recipients or have experienced greater psychological distress by suddenly having greater health-related responsibilities by providing care (Irani et al., 2021; Kent et al., 2020; Lorenz-Dant and Comas-Herrera, 2021; Muldrew et al., 2022).

The objective of this chapter is to identify potential causal effects of COVID-19 and related social restrictions on the mental health of informal carers. Detailed data on mental health and informal care are examined, drawn from the UK Household Longitudinal Study (Understanding Society), collected between 2016 and March 2021. The empirical approach relies on two difference-in-differences (DD) specifications, a standard two-way fixed effects model as well as a multi-period fixed effects DD model. The former DD model explores the causal effect of the COVID-19 pandemic on the mental health of informal carers while the latter is based on recent advancements in the quasi-experimental literature (Callaway and Sant'Anna, 2021) investigating heterogeneous effects driven by timing and duration of caregiving. Propensity score matching to pre-process the data is also employed to make treated and control groups (i.e., informal carers vs non-carers) more comparable by accounting for selection into informal caregiving through observable characteristics. The 12-item General Health Questionnaire (Goldberg et al., 1997), a psychometrically validated and widely used index of psychological distress, is used as measure of mental health.

The results show that new carers - those who started providing informal care after the COVID-19 outbreak - experienced a significant mental health deterioration especially when lockdown and social restrictions were in place. Specifically, their mental health deteriorated by around 0.37 points on the GHQ scale at the start of the pandemic (April 2020), 0.27 points during the second national lockdown (November 2020) and 0.25 points during the third national lockdown (January 2021), while it seems unaffected when social restrictions were lifted. These estimates are comparable to the mental health deterioration associated with major life events such as divorce and unemployment (Clark and Georgellis, 2013). Existing carers appear to have coped relatively well during the pandemic, mostly showing changes in mental health that are not statistically significant.

This chapter offers several contributions to the growing literature on the mental health effects of the COVID-19 pandemic as well as to the broader literature on the determinants of the mental health of (informal) carers. First, a quasi-experimental approach is employed to provide causal evidence on the mental health burden suffered by informal carers during the COVID-19 pandemic. Despite a rapidly increasing number of studies of mental health effects among healthcare workers, whether the pandemic had a causal impact on the mental health of informal carers remains an open empirical question. Second, this empirical approach is used to estimate standard as well as multiple period difference-in-differences models combined with matching. This simultaneously accounts for several important issues in identifying causal effects, including observed self-selection into treatment (via matching) and individual-level unobserved heterogeneity (via fixed effects). In addition, the difference-in-differences approach with multiple time periods recently proposed by Callaway and Sant'Anna (2021) allows estimation of the mental health effect of COVID-19 on multiple groups of informal carers during different periods and by duration of care. This has not been explored so far in the current literature. Third, the analysis clearly distinguishes between existing, and therefore more experienced, informal carers and new carers, that is those who started providing informal care during the pandemic. Though potentially relevant, the mental health effects on these different groups of informal carers during the pandemic have not been examined yet.

3.2 Literature

It is well-established that informal caregiving is associated with increased physical strain (Pinquart and Sörensen, 2003) and psychological distress (Lacey et al., 2019) among carers. However, establishing whether informal caregiving had a causal effect on mental health is challenging given potential endogeneity issues, including self-selection into caregiving (Bom et al., 2019). To account for selection into caregiving, recent papers used different matching techniques. Bom and Stöckel (2021) and Stöckel and Bom (2022) employed propensity score matching to explore the health effects of informal caregiving using data from the UK and the Netherlands. They found negative mental health effects of caregiving, especially among caregivers who provided care for longer, and with higher intensity as measured by additional hours of caregiving. De Zwart et al. (2017) also employed matching to deal with selection issues when studying health outcomes of spousal informal caregivers. Analysing data from the Survey of Health Ageing and Retirement in Europe (SHARE), they found that caregiving has a negative short-term effect on informal carers' overall health. Previous evidence also appears to show that the negative effect of informal caregiving seems to be larger among specific groups of caregivers, such as older caregivers (Bom et al., 2019); women (De Zwart et al., 2017; Lacey et al., 2019; Brenna, 2021); and people with more intense caregiving duties (Bom and Stöckel, 2021; Stöckel and Bom, 2022). However, an important limitation of these studies is that they appear to infer causality using matching as their main identification strategy and therefore rely on the conditional independence assumption (effectively assuming that there are no unobservable confounders).

Although there are many studies on the mental health effects of COVID-19, empirical research identifying causal effects of COVID-19 on informal carers' mental health is still limited. While existing evidence suggests an increase in psychological distress during the pandemic, previous studies mainly focus on convenience samples of informal caregivers using information from ad-hoc interviews conducted during the pandemic, thus having limited external and general validity (Azevedo et al., 2021; Lightfoot et al., 2021; Ng et al., 2020; Greaney et al., 2022; Irani et al., 2021). Other studies are based on cross-sectional data and do not include pre-pandemic information (Borelli et al., 2021; Beach et al., 2021; Leggett et al., 2021; Todorovic et al., 2020; Li et al., 2021); hence, they lack natural control groups.

A few studies used longitudinal data to explore mental health among informal carers. However, such studies do not appear to account for self-selection into informal caregiving nor provide causal estimates of mental health effects while accounting for timing and duration of informal care. Gallagher and Wetherell (2020) used two waves of the UK Household Longitudinal Study (Understanding Society) to explore whether depression increased among caregivers during the first month of the pandemic (April 2020). They found that caregivers were more likely to experience depressive symptoms compared to non-caregivers. Whitley et al. (2021) used the same dataset to investigate informal caregivers' mental health, as measured by the GHQ-12. Using OLS, they estimated differences in changes in mental health between home-carers and non-carers, and found that GHQ-12 scores of informal carers were already higher pre-pandemic compared to non-carers and that mental health further deteriorated after the COVID-19 outbreak. Wister et al. (2022) used the Canadian Longitudinal Study on Aging to explore depression and anxiety among informal caregivers during the first nine months of the pandemic, finding, using linear mixed models, that informal caregivers had worse mental health compared to non-carers. Park (2021) estimated logistic and negative binomial models to examine differences in mental distress among groups of non-caregivers; short-term ($\leq$ 1 year); and long-term ($>$ 1 year) informal caregivers, analysing data from the Understanding America Study. Findings showed that long-term caregivers were the most affected by the pandemic, possibly due to caring for longer. Truskinovsky et al.

(2022) estimated linear probability models to examine associations between care disruptions and mental health in US. They found that caregiving arrangements were disrupted by COVID-19, and that these disruptions were associated with increased depression, anxiety, and loneliness among caregivers.

Only two studies try to account for self-selection into caregiving while focusing on informal caregivers' mental health during COVID-19. These studies control for selection into caregiving, through matching on observable factors rather than considering both observable and unobservable factors. Mak et al. (2021) used propensity score matching to compare informal caregivers and non-caregivers' mental health using primary data collected during the pandemic in the UK. They found that carers experienced higher depressive symptoms and anxiety. Their empirical approach ignored potential unobservable differences between the two groups of carers and relied on a convenience sample collected during the COVID-19 pandemic, meaning that their results cannot be necessarily generalised. Furthermore, they cannot observe pre-pandemic levels of outcomes. Bergmann and Wagner (2021) investigated the effect of COVID-19 on informal caregivers' health using two waves of SHARE. They found that the mental and physical health of informal caregivers deteriorated during the first months of the pandemic. However, their statistical approach is similar to that of propensity score matching, relying on the conditional independence assumption and selection on observables (i.e., all observed health differences are attributable to the caregiving role). In addition, the observed characteristics they used do not appear to account for how people decide to become informal caregivers, and therefore self-selection into caregiving does not seem to be appropriately accounted for.

3.3 Data

The UK Household Longitudinal Study (Understanding Society) is here analysed, which contains detailed individual-level information of a representative sample of the UK's population. Specifically, this chapter analysed three mainstage questionnaires collected before COVID-19 (i.e., Wave 8, 9 and 10, covering 2016-18; 2017-2019; and 2018-2020) together with eight COVID-19 surveys specifically designed to collect relevant information on a monthly or bi-monthly basis (from April 2020 until March 2021). The sample analysed included 4698 respondents who were interviewed in all eleven waves considered in the analysis. This sample was obtained after omitting those interviewed in the mainstage questionnaire (pre-pandemic waves) but during or after March 2020 (when the first national lockdown was imposed in the UK). In this way, all the interviews of the mainstage questionnaire were conducted before COVID-19. Only those who always answered questions related to mental health (the outcome of interest) and informal care are then retained. Furthermore, those who were already informal carers during the first wave analysed (Wave 8) are dropped, so that nobody was providing informal care during the first period considered, which is essential for matching requiring pre-treatment variables for both treated and control groups. Importantly, only caregivers who continuously provided care are here considered, i.e. those who started caregiving and continued caregiving throughout the period analysed. Specifically, the question about caregiving is asked in the mainstage questionnaires and in the first COVID survey (April 2020). Given the social restrictions imposed, if someone provides care during the first lockdown, it is reasonable to assume they will also provide care later during the pandemic (at least during the period observed in this chapter).

Importantly, given that in the first COVID survey (April 2020) the question about informal care is related to people providing care to someone living outside their household, the same question is used in the mainstage questionnaires to identify informal carers. Moreover, only "external" informal carers are considered because caregiving tasks such as grocery shopping and helping around the house might be misreported if caregivers and care recipients live together. It is also likely that home-carers may have been already providing some form of care pre-pandemic which might have changed only in caregiving intensity, given formal care services disruptions. Compared to more general samples (i.e., the mainstage questionnaire sample before COVID and the COVID questionnaire sample with missing data), this estimating sample includes individuals who are older; more likely to be white; better educated; better off economically (in terms of employment and income); more likely to be married; physically healthier, and in better mental health. Given that individuals included in the estimating sample have greater social and financial resources, it is plausible that their mental health might have been less affected when providing informal care. As such, the estimated effects on mental health may provide a lower bound of the mental health effects of interest.

3.3.1 Key variables

To estimate the effect of caregiving on the mental health of informal carers with different caregiving experience and duration, this chapter exploited information included in both the mainstage UKHLS questionnaire and the COVID-19 surveys. Accordingly, two treatment groups are defined, existing carers (349 individuals) and new carers (1655 individuals), and a control group of never-carers (2694 individuals). More specifically, new carers were defined as those who started caregiving after the COVID-19 outbreak, as recorded by the related question in Wave 1 of the COVID survey (April 2020). These definitions were based on the variable *"Do you provide some regular service or help for any sick, disabled or elderly person not living with you?"*, identifying the provision of any type of informal care outside an individual's own household.

The outcome is mental health as measured by the 12-item General Health Questionnaire (GHQ-12) (Goldberg et al., 1997). GHQ is a self-completion psychometrically validated questionnaire, which has been extensively used in different research fields, including economics, to measure mental ill-health. GHQ-12 contains twelve questions asking whether respondents experienced specific symptoms or feelings, with each item rated on a four-point scale ranging from "Not at all" to "Much more than usual". Following previous literature, the main analysis used GHQ-caseness as the outcome (Banks and Xu, 2020; Chandola et al., 2022; Serrano-Alarcón et al., 2022). In GHQ-caseness, each of the 12 questions is rated with a binary variable: 1 for each answer with a score of 3 or 4 (corresponding to the answers "More than usual" and "Much more than usual"), and 0 otherwise. Scores range from 0 to 12 with higher values indicating worse mental health (i.e., this measure indicates an increase in mental ill-health). As a robustness check and following earlier studies (Banks and Xu, 2020; Serrano-Alarcón et al., 2022), a binary indicator of a GHQ score of 4 or more is also used, which indicates likely mental ill-health (NHS Digital, 2017). The results of the robustness checks were similar to those of the main analysis (Table B6, Appendix).

3.3.2 Other variables

Additionally, the following control variables (measured pre-treatment) were also included: age and age squared (to account for the U-shaped well-being curve, Easterlin (2003)); marital status and living alone (binary variables indicating if respondents were in single households or in a partnership); number of dependent children (to proxy the degree of loneliness of respondents); the demographic variables of sex and ethnicity; educational attainment (with value of 1 if respondents achieved an educational level higher than A-level, and 0 otherwise); being employed; and household income measured using quintiles. This chapter also included the governmental regions of the UK in which respondents live (regions of England, Wales, Scotland and Northern Ireland) and the specific COVID-19 status in each region using the daily cumulative number of deaths at a regional level and whether respondents reported being tested for COVID-19. The number of COVID-19 deaths were drawn from UK governmental websites (UK Government, 2022), while information about testing was found in the COVID surveys.

3.4 Empirical approach

This chapter employed a combination of propensity score matching (PSM) to pre-process the data and difference-in-differences (DD) models to identify the causal effect of COVID-19 on the mental health of informal carers. Propensity score matching is used to pre-process the data and hence account for observed characteristics potentially driving selection into informal caregiving. Matching combined with fixed effects in DD models helps produce more comparable treatment and control groups while also accounting for individual-level unobservable characteristics (via fixed effects in the DD models). Note that models employing the longitudinal weights provided in Understanding Society are also estimated as an alternative to those produced by matching (Table B5 in the Appendix, showing similar results)[1].

In addition, this chapter also considered using a Regression Discontinuity Design (RDD) to identify the mental health effects of COVID-19 among informal carers. However, there would not be enough observations around the cut-off created by the first COVID-19 wave (and corresponding national lockdown), where treatment and comparison units/individuals are most similar, and the treatment would be considered as good as random. Thus, an RDD study would have to rely on too few observations around the threshold, especially after the imposition of COVID

[1]The longitudinal weights available in UKHLS are design weights, which account for sample selection and reflect differences in data collection. These weights are also adjusted to compensate for attrition and non-response and they are calibrated to represent the whole population

restrictions, on which to draw any conclusions.

3.4.1 Propensity score matching

PSM is increasingly used when the treatment of interest is providing informal care (Bom et al., 2019; Stöckel and Bom, 2022; De Zwart et al., 2017). This is because people may self-select into informal caregiving due to specific circumstances, and this could result in systematic differences between informal carers versus non-carers. This may ultimately lead to biased estimates of the treatment effect (Cunningham, 2021). Essentially, matching assumes that after conditioning on observables, the difference in the outcome of interest can be solely attributed to the treatment (Rubin, 1974), which in this case is the provision of informal care.

PSM (Leuven and Sianesi, 2003) is here used to make the groups of informal carers and never-carers balanced on observed variables related to the decision to provide care (Berg, 2011). As for the variables used in matching, this chapter followed Schmitz and Westphal (2017) who argue that the decision to provide care depends on characteristics concerning three main areas: (i) the need to, (ii) the willingness to and (iii) the ability to provide care. Indeed, deciding to become an informal carer is driven by corresponding conditions: (i) someone close needs assistance; (ii) the future caregiver is prone to provide care, given their household situation and personality; and (iii) they are in good health and can provide assistance to someone. Here, PSM is employed using pre-treatment variables for both groups of informal carers (existing and new carers). Note that these matching models also included baseline (pre-treatment) mental health. As a result, after matching, both caregivers and never-carers would also present similar baseline mental health.

One-to-one matching was estimated, where each respondent in the treated group was matched to a respondent in the control group. As an alternative, Kernel PSM is also performed, whereby each respondent is matched to a weighted average of all control group members (Heckman and Vytlacil, 2007). While both methods produced very similar results, since there is a large control group, this chapter can rely on one-to-one matching. In this way, a unique control group member can be matched to a corresponding treated individual. Results obtained using one-to-one PSM are displayed in Table B1 (Appendix). The pairwise t-tests show, after matching, no statistically significant differences in the observed variables (at the 5% level for all variables, with one variable presenting a weakly significant difference at the 10% level). It should be noted that even if the bias attributable to observed confounding factors has been minimized, matching does not eliminate the potential bias driven by unobservables. To ease concerns around the role of unobservables and to account for individual-level unobserved heterogeneity, fixed effects are included in the DD models. In addition, and to further account for the potential role of unobserved heterogeneity, an Oster test (Oster, 2019) is employed to explore whether results would change in the presence of selection on unobservables. Results indicate that it would require a proportional level of selection on unobservables well beyond the conventional value of 1 (usually defined as an upper bound value for the proportional level of selection on unobservables) to fully confound the mental health effects in these difference-in-differences models (Table B4, Appendix).

3.4.2 Difference-in-differences models

This empirical approach relies on estimating a series of difference-in-differences (DD) models on the sample obtained using matching. In these DD models the treatment is represented by providing informal care during COVID while different treatment groups are defined according to when respondents started caregiving for the first time, either during the pandemic (new carers) or before the pandemic (existing carers). More specifically, generalised DD via two-way fixed effects (TWFE) models are first estimated[2]. The first approach allows us to identify the mental health effects of caregiving during the COVID-19 pandemic alternatively on existing and new carers using never-carers as a control group. Subsequently, a difference-in-differences model with multiple periods is estimated as proposed by Callaway and Sant'Anna (2021). This second approach enables one to estimate the mental health effects among the two groups of informal carers at different points in time, therefore accounting for timing and duration of care (i.e., when carers started to provide care and for how long) as well as the effects of the different COVID waves. However, even in this case the corresponding average treatment effects are always separately estimated for new and existing carers using never-carers as a control group.

Generalised difference-in-differences

A generalised difference-in-differences approach is used, estimating standard two-way fixed effects (TWFE) regression models (Wooldridge, 2010):

$$\mathrm{Y}_{it} = \alpha_i + \lambda_t + \delta^{DD} D_{it} + \gamma X_{it} + \varepsilon_{it} \tag{3.1}$$

Equation (3.1) specifies the effect of providing informal care during the COVID-19 pandemic (δ^{DD}) on the outcome of interest (Y_{it}), i.e., caregivers' mental health as measured by the GHQ-12. Individual and time fixed effects (α_i and λ_t) are included, along with a wide set of observed covariates (X_{it}). Standard errors (ε_{it}) are clustered at the primary sampling unit level (Abadie et al., 2017). Importantly, separate TWFE regressions are estimated for different groups of informal caregivers. This follows recent papers (Baker et al., 2022; Callaway and Sant'Anna, 2021; De Chaisemartin and d'Haultfoeuille, 2020; Goodman-Bacon, 2021; Sun and Abraham, 2021) which suggest that standard TWFE estimators are not robust to treatment effect

[2]Time fixed effects are expressed as dummy variables for each time period considered in the analysis. The change in the outcome of interest can be assessed by interacting the dummy variables of the time FEs with the treatment status (i.e. caregiver status), while accounting for unobservable characteristics

heterogeneity, i.e. when groups are treated at different times. Therefore, existing carers vs never-carers are first considered, and subsequently new carers vs never-carers are considered. D_{it} is a binary variable defining treated units during treatment, so that it will equal one for groups of informal caregivers (existing carers and new carers in separate TWFE analyses) and it will be zero for never-carers. Specifically, the effect of interest is the interaction between the time of the pandemic and the variable defining the different informal caregivers' groups.

3.4.3 Difference-in-differences with multiple time periods

A difference-in-differences model with multiple time periods is estimated, as proposed by Callaway and Sant'Anna (2021). Following this approach, it is possible to generalise the average treatment effect on the treated (ATT) to multiple periods. That is, the ATT can be directly estimated at any time for group g at time t, where a group is defined based on when it is first treated. Under the parallel trends assumption on never-treated units, the ATT is estimated as Equation (3.2) shows.

$$ATT_{(g,t)} = E[Y_t - Y_{g-1}|G = g] - E[Y_t - Y_{g-1}|C = 1] \tag{3.2}$$

Where: Y_t is the observed outcome at time t; Y_{g-1} is the observed outcome just before the unit becomes treated; G indicates each treated group as defined by when units are first treated (g); and C is an indicator variable for individuals who are part of the never-treated group (the control group of never-carers). Therefore, this difference-in-differences model with multiple time periods simultaneously estimates the ATT for each treated group at any point in time included in the analysis. Another advantage of this framework is that researchers can test for parallel trends based on never-treated units for all treated groups, while conditioning on observed covariates.

This additional DD specification accounts for potential differences in the mental health effects driven by the timing and duration of informal care. In this case, the analysis considered a treated group including both existing carers and new carers, and this group of treated individuals was divided into different sub-groups according to the timing (i.e., wave) in which they started to provide care, as displayed in Table 3.1. Another feature of this framework is that nobody is treated in the first period, as shown in Table 3.1 (Wave 8). Moreover, once units are treated, they are assumed to be treated during all subsequent periods. Because of this assumption, this chapter also made sure that informal caregivers provided continuous care throughout the period analysed. Finally, two groups of existing carers are considered: those who started caregiving in Wave 9 and those who started in Wave 10.

Table 3.1: Informal caregiving patterns

Period	Years	Treated			Control
		Existing carers		New-carers	Never-carers
		Wave 9	Wave 10		
	WAVE 8 (2016-2018)	0	0	0	0
Pre-COVID-19	WAVE 9 (2017-2019)	1	0	0	0
	WAVE 10 (2018-2020)	1	1	0	0
After-COVID-19	April 2020	1	1	1	0

Note: April 2020 is the first wave of the COVID questionnaire
Existing carers are grouped according to when they start caregiving.
The Wave 9 group comprises caregivers mainly starting in 2017-18.
The Wave 10 group comprises caregivers mainly starting in 2018-19.

3.5 Descriptive statistics

Figures 3.1 and 3.2 present the weighted mean of mental health over time by groups of existing caregivers and new carers versus never-carers obtained using the weights provided by the PSM. It is clear from these figures that informal carers always have worse mental health compared to never-carers. During the first year of COVID-19, mental health fluctuated according to social restrictions, improving when restrictions were lifted (May-September 2020) and worsening during national lockdowns (March-April 2020; September 2020-March 2021).

Figure 3.1: GHQ weighted mean over time by existing carers and never-carers

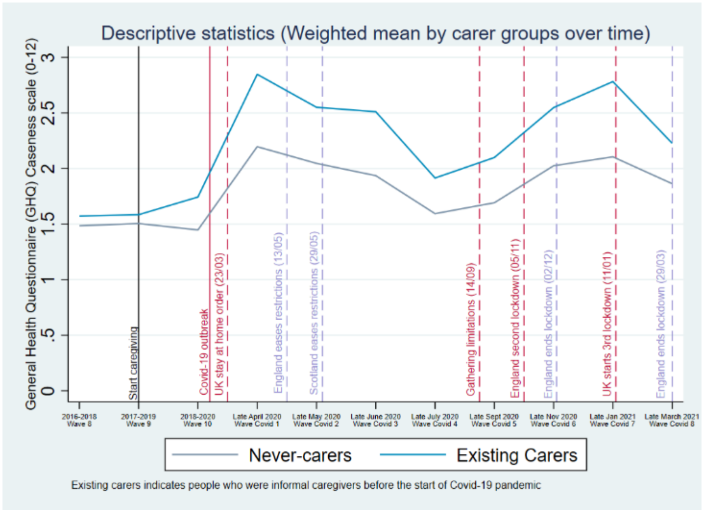
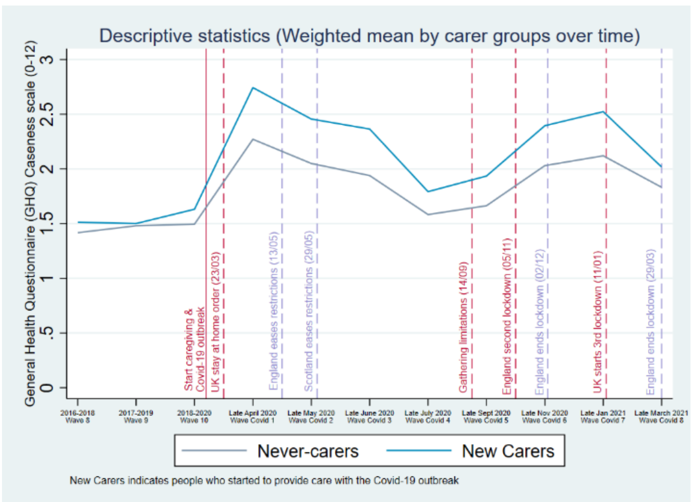


Figure 3.2: GHQ weighted mean over time by new carers and never-carers



Figures 3.1 and 3.2 show plots of the weighted average of the outcome over time, displaying pre-treatment trends. Although not a formal test (Cunningham, 2021), these figures suggest the plausibility of the assumption, by showing parallel trends before the treatment (i.e., providing informal care). This implies that, in the absence of treatment, the difference in levels of mental health between treated and control groups would have been constant over time and therefore researchers could attribute the difference in mental health post-treatment to the treatment itself. To further check the validity of this assumption, the treatment variable is interacted with pre-treatment periods in the TWFE models and multiple time periods DD models, finding coefficients that are not statistically significant. Overall, this supports the assumption of parallel trends. Note that while in Figure 3.2 the pre-COVID trends of mental health of the two groups seem to intersect, their corresponding 95% confidence intervals overlap (see Figure B1, Appendix). This implies that the null hypothesis, that the levels of GHQ scores before COVID were similar across never carers and new carers, cannot be rejected.

Table 3.2 shows the descriptive statistics of variables in the mainstage Wave 8 (when individuals were not providing informal care) for each group of informal caregivers (existing and new carers) versus never-carers. Table 3.2 shows that, on average, existing carers are older than never-carers, while new carers are the youngest group. Compared to existing carers, new carers and never-carers are more likely to live alone, be single and less likely to be widowed or divorced. New carers tend to have more dependent children and are more likely to be employed. Existing caregivers have poorer physical health with greater functional limitations and long-standing illnesses. Importantly, all groups have similar levels of mental health before providing informal care. This suggests that becoming an informal carer may affect mental health. Interestingly, while all groups have similar demographic characteristics in terms of sex, ethnicity, education, and nation of residence, existing carers

Table 3.2: Descriptive statistics of groups in Wave 8

Variables	Never-Carers (2,694)		Existing Carers (349)		New Carers (1,655)	
	Mean	SD	Mean	SD	Mean	SD
Age	52.509	16.364	55.476	10.994	48.593	13.326
Female	0.499	0.500	0.625	0.485	0.628	0.483
White	0.933	0.250	0.966	0.182	0.928	0.259
Education:						
Less than O-level or eq.	0.574	0.495	0.605	0.490	0.593	0.491
O-level, A-level or eq.	0.134	0.341	0.123	0.330	0.134	0.341
Higher than A-level	0.291	0.455	0.272	0.446	0.273	0.445
Live alone	0.844	0.363	0.771	0.421	0.866	0.340
Marital status:						
Single	0.235	0.424	0.147	0.354	0.224	0.417
Married/Civil partnership	0.612	0.487	0.638	0.481	0.636	0.481
Divorced/Widowed	0.153	0.360	0.216	0.412	0.140	0.348
Number of children under16	0.317	0.722	0.223	0.564	0.505	0.866
Job status:						
In paid employment	0.584	0.493	0.630	0.483	0.732	0.443
Unemployed	0.019	0.135	0.029	0.167	0.013	0.115
Retired	0.308	0.462	0.281	0.450	0.177	0.382
Other non-paid activities	0.089	0.285	0.060	0.238	0.077	0.267
Income (quintiles)	2.286	1.396	2.313	1.340	2.406	1.363
SF-12 physical score	51.217	9.716	49.518	10.208	52.284	8.897
SF-12 mental score	50.179	9.846	50.003	8.875	49.717	9.236
Number functional limits	0.417	1.116	0.549	1.06	0.309	0.887
Long-standing illness	0.323	0.468	0.372	0.484	0.283	0.451
Self-rated health	2.540	0.990	2.670	0.978	2.422	0.941
Nation:						
England	0.824	0.381	0.794	0.405	0.828	0.378
Wales	0.057	0.231	0.066	0.248	0.057	0.232
Scotland	0.092	0.289	0.095	0.293	0.082	0.274
Northern Ireland	0.027	0.162	0.046	0.209	0.034	0.181
Life satisfaction	5.395	1.386	5.398	1.317	5.404	1.289
GHQ-Caseness	1.389	2.819	1.479	2.708	1.514	2.812
GHQ-Likert	10.496	5.289	10.762	4.941	10.718	4.986

Note: Descriptive statistics are calculated in Wave 8, when nobody was caregiving are on average older compared to new carers. As such, it is plausible to assume that existing carers might have started to provide care earlier to support older parents or relatives.

Since it might be relevant to provide information about the type of care provided as well as about the person cared for, the Appendix reports a series of additional tables (see Tables B2-B3). These include more specific information on the tasks performed by informal carers and the type of person they cared for during the first COVID wave (broken down by existing and new carers). Table B2 suggests that new and existing carers tended to mostly perform similar tasks (e.g., gardening and shopping); however, existing carers seemed to more frequently "look after personal affairs" and this might be a reflection of longer standing arrangements. Table B3 also shows that both existing and new carers appeared to have provided care mostly to neighbour/friends and older individuals. While this information might be useful, this question was answered by a relatively small proportion of individuals in the sample and only during COVID Waves 1, 6 and 8.

3.6 Results

3.6.1 Two-way fixed effects models

Estimates of the two-way fixed effects (TWFE) models are presented in Table 3.3. These were separately estimated using alternative treatment groups of existing (Columns 1-3) and new carers (Columns 4-6). Note that in all models, the control group included never-carers; the outcome of interest is GHQ-12, which indicates an increase in mental ill-health; and Wave 10 is the baseline wave, i.e. the last wave before the COVID-19 outbreak. TWFE regressions were estimated with an incremental number of controls (columns 1-3) and the estimates in grey correspond to when lockdowns or strict social restrictions were in place.

Table 3.3: DD results. GHQ-caseness

DD interactions	GHQ Existing carers			GHQ New carers		
(Wave 10 as baseline)	(1)	(2)	(3)	(1)	(2)	(3)
Carer x Wave 8 (2016-2018)	-0.312	-0.342	-0.343	-0.040	-0.024	-0.024
	(0.358)	(0.251)	(0.251)	(0.104)	(0.121)	(0.121)
Carer x Wave 9 (2017-2019)	-0.428	-0.199	-0.199	-0.116	-0.071	-0.071
	(0.347)	(0.223)	(0.223)	(0.099)	(0.115)	(0.115)
Carer x Wave COVID 1 (April 2020)	0.480	0.296	0.296	0.336***	0.369***	0.369***
	(0.414)	(0.283)	(0.283)	(0.119)	(0.137)	(0.137)
Carer x Wave COVID 2 (May 2020)	0.512	0.125	0.126	0.270***	0.319***	0.320***
	(0.356)	(0.243)	(0.243)	(0.117)	(0.136)	(0.136)
Carer x Wave COVID 3 (June 2020)	0.370	0.359	0.361	0.289**	0.387***	0.386***
	(0.378)	(0.256)	(0.257)	(0.119)	(0.138)	(0.138)
Carer x Wave COVID 4 (July 2020)	0.107	0.054	0.053	0.073	0.180	0.181
	(0.342)	(0.244)	(0.244)	(0.113)	(0.131)	(0.132)
Carer x Wave COVID 5 (Sept 2020)	0.225	0.064	0.066	0.135	0.193	0.193
	(0.331)	(0.241)	(0.242)	(0.114)	(0.130)	(0.130)
Carer x Wave COVID 6 (Nov 2020)	-0.007	0.317	0.314	0.228*	0.265*	0.266*
	(0.275)	(0.246)	(0.246)	(0.118)	(0.137)	(0.137)
Carer x Wave COVID 7 (Jan 2021)	0.323	0.476*	0.476*	0.267**	0.246*	0.246*
	(0.309)	(0.244)	(0.244)	(0.123)	(0.140)	(0.140)
Carer x Wave COVID 8 (Mar 2021)	-0.034	0.090	0.091	0.049	0.092	0.092
	(0.354)	(0.246)	(0.245)	(0.121)	(0.138)	(0.138)
Age		-0.098	-0.097		0.064	0.068
		(0.080)	(0.080)		(0.061)	(0.061)
Age squared		0.001	0.001		0.000	0.000
		(0.001)	(0.001)		(0.000)	(0.000)
Living alone		0.104	0.102		-0.019	-0.021
		(0.116)	(0.116)		(0.113)	(0.113)
Children(<16) in household		0.429*	0.428*		0.377***	0.376***
		(0.226)	(0.227)		(0.144)	(0.144)
Female		-0.597***	-0.709***		-0.005	-0.021
		(0.105)	(0.131)		(0.386)	(0.398)
White		-1.199***	-1.194***		-0.829***	-0.810***
		(0.112)	(0.113)		(0.110)	(0.112)
Education		-0.005	-0.004		0.004	0.004
		(0.394)	(0.397)		(0.300)	(0.301)
Paid employment		-0.061	-0.062		-0.361***	-0.360***
		(0.112)	(0.112)		(0.101)	(0.101)
Income (quint) ++		0.038	0.038		0.018	0.018
		(0.051)	(0.051)		(0.040)	(0.040)
Cum daily deaths by region			0.001*			0.001**
			(0.001)			(0.001)
Being tested			0.097			0.028
			(0.086)			(0.062)
Observations	30360	23176	23176	47322	34472	34472
N of respondents	2760	2112	2112	4302	3143	3143

Note Carer equals one if respondents are existing caregivers, or new caregivers and zero if they are never-carers.
*** p<0.01, ** p<0.05, * p<0.1
++ indicates variables are projected from the mainstage waves
Robust standard errors in parenthesis, clustered at primary sampling unit.
Grey estimates indicate when the UK was under strict social restrictions while white rows indicate more relaxed periods.
All specifications included time and individual fixed effects. Models (2) and (3) also included regional fixed effects.

Table 3.3 shows that pre-treatment periods (first two rows) are not statistically significant, and this further supports the parallel trends assumption. Table 3.3 also appears to show that existing carers present generally worse mental health compared to never-carers during the COVID-19 pandemic. However, the only statistically significant estimate (at a 10% level) corresponds to COVID Wave 7, where being an existing caregiver contributes to an increase in psychological distress by 0.48 points on the GHQ scale. Since COVID Wave 7 was collected during January 2021 (when the UK was in its third national lockdown), this implies that while existing carers were coping relatively well with the pandemic, possibly due to their previous experience, imposing another lockdown almost one year after the start of the pandemic might have negatively affected their mental health. Although this coefficient is only weakly significant, to have a sense of its magnitude, we can compare it with the effect of other major life events on the GHQ-12 score in the UK as shown by Clark and Georgellis (2013). For example, experiencing unemployment increases mental ill-health by 0.41 points on the GHQ scale for men and 0.60 points for women, while the death of a partner leads to a worse mental health by 0.51 points for women and 0.53 points for men. This suggests that a deterioration of 0.47 GHQ points is a sizeable effect.

Concerning new caregivers, it can be noticed that their mental well-being more clearly deteriorated during the pandemic. Specifically, mental health is significantly worse compared to never-carers during the first national lockdown (COVID Waves 1-2: April-May 2020), persisting also throughout June 2020 (COVID Wave 3), and during the second and third national lockdowns (COVID Waves 6-7: November 2020-January 2021). These results suggest that COVID-19 and its related social restrictions led to a statistically significant mental health decline among informal caregivers who started to provide care after the COVID-19 outbreak. In terms of the size of the corresponding coefficients, during the first national lockdown, being a new caregiver contributed to a highly statistically significant decline in mental health by around 0.37 points on the GHQ scale in April 2020; 0.32 points in May 2020; and 0.39 points in June 2020. In November 2020 and January 2021, when further lockdowns and stay-at-home orders were reintroduced, new caregivers' mental health also weakened compared to never-carers by about 0.27 and 0.25 points on the GHQ scale, respectively, although these estimates are only weakly significant. A possible explanation of these results might be linked to the lack of experience of new caregivers combined with the strict lockdown rules limiting social interactions as well as external support. Overall, it should be noted that the mental health effects of new carers appear to be more precisely estimated compared to the effect on existing carers. This might imply that we cannot categorically exclude the presence of some mental health effects among existing carers as well and that these could be potentially more precisely identified using a larger sample.

All models controlled for covariates that might influence mental health. In line with the previous literature, white people have better mental health (1.19 GHQ points for existing carers and 0.81 GHQ points for new carers). Women present better mental health (a change of 0.71 GHQ points among existing carers). Finally, being in paid employment is correlated with better mental health (0.36 GHQ points for new carers), while having dependent children appears to decrease mental health (0.43 and 0.38 GHQ points for existing and new carers). As expected, a higher cumulative number of daily COVID-19 deaths measured at a regional level is associated with worse mental health, although the effect is of a small magnitude, while being tested for COVID is positively but insignificantly associated with a mental health deterioration.

3.6.2 Multiple time periods difference-in-differences

Difference-in-differences models with multiple periods were also estimated (Callaway and Sant'Anna, 2021). Using this approach, this chapter simultaneously estimated the group-time average treatment effect (ATT) on each treated group of carers. This ultimately accounts for timing and duration of care when estimating the mental health effects of COVID-19 among informal carers. Accordingly, existing caregivers are divided into two sub-groups based on when they started caregiving, either in Wave 9 or Wave 10. Figures 3.3-3.5 show the group-time average treatment effects over time for each treated group when the entire set of covariates were included in the analysis. Figures 3.3 and 3.4, which are related to existing caregivers, show that for the group of existing carers starting in Wave 9 the ATT is always positive. This suggests this group of carers experienced worse mental health compared to never-carers; however, coefficients are not statistically significant. Existing carers starting in Wave 10 show a statistically significant ATT, although only at the 10% level, of about 0.56 points on the GHQ scale in COVID Wave 7 (time 7 in Figure 4). This is in line with our previous finding, suggesting that existing caregivers starting in Wave 10 (2018-2020) might explain the results of the TWFE models relative to existing carers.

Figure 3.5 shows that new carers present significantly worse mental health, especially during the first national lockdown of about 0.29 GHQ points in both COVID Wave 1 (time 0 in Figure 3.5) and COVID Wave 2 (time 1 in Figure 3.5), and 0.35 GHQ points in COVID Wave 3 (time 2 in Figure 3.5). However, the statistically significant effect previously found during the second and third national lockdowns does not persist in the most comprehensive specification of this DD analysis.

Figure 3.3: ATT over time by existing carers (Wave 9) and never-carers

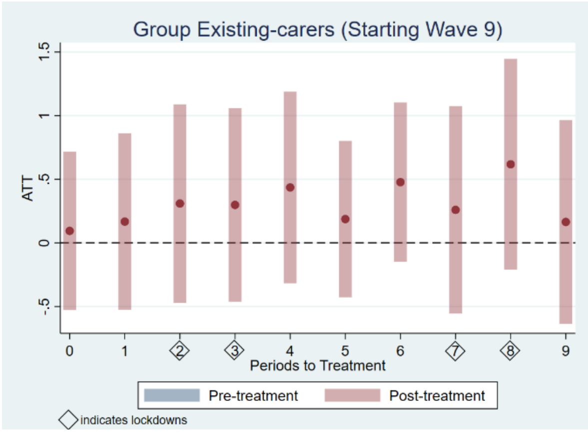


Figure 3.4: ATT over time by existing carers (Wave 10) and never-carers

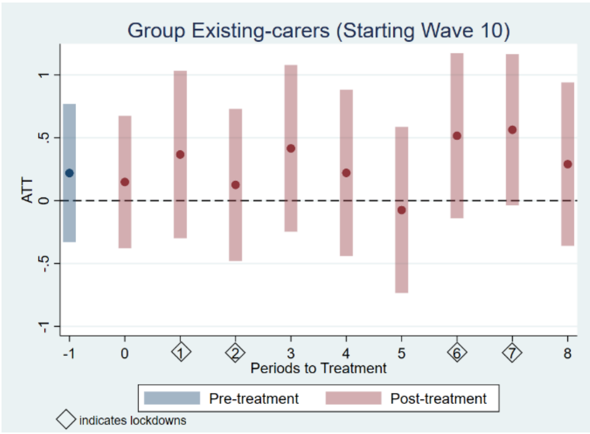
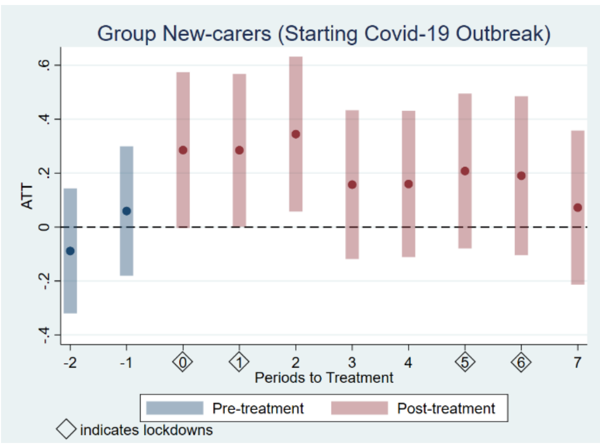


Figure 3.5: ATT over time by new carers and never-carers



3.7 Conclusions

This chapter investigated whether the current COVID-19 pandemic had a causal effect on mental health among informal carers. Unlike previous studies, a quasi-experimental approach is here employed combining difference-in-differences with matching to account for the potential roles played by timing and duration of caregiving. This is the first analysis employing multiple time periods difference-in-differences specifications while also looking at different types of informal carers: existing and new informal carers. Results about the harm to the mental health of new carers, i.e. those who started providing informal care only during the pandemic, might be of particular interest to policymakers as they refer to a sizeable group of individuals that has been overlooked by most previous studies.

These findings suggest that mental health fluctuated according to social restrictions, but informal carers had consistently worse outcomes during the pandemic compared to never-carers. Specifically, the estimates show that new carers were the most affected, with a statistically significant and sizeable deterioration in their mental health when lockdowns and social restrictions were in place. Therefore, even if social restrictions were essential to curb infection rates by limiting COVID-19 transmission, these results imply that imposing stay-at-home orders harmed mental health, especially for those who became informal caregivers only after the COVID-19 outbreak. New informal caregivers were lacking the social support needed due to the imposed social isolation during lockdowns. Hence, the lack of social support coupled with the anxiety of having new caregiving responsibilities, especially at the start of the caregiving role and during a public health crisis, were the possible mechanisms underlying these findings.

This chapter has limitations. First, only informal caregivers providing care to someone living outside their household were investigated. However, considering the stay-at-home orders and travel bans, informal caregivers providing care to someone living outside their household might have been less affected by COVID-19 compared to home-carers. External caregivers might provide less intense care compared to informal carers living with care recipients, especially when social restrictions were in place. Thus, these results might represent the lower bound of the true mental health effects on informal carers. Second, the intensity of care is not examined. The duration of care is examined, i.e. when someone became a caregiver for the first time, but care intensity (such as the number of hours of care provided) is not considered. This is because such information is not available in each wave of the survey. Finally, there is not information about use of therapy or medications that might have influenced mental health.

Overall, the results of this chapter are of interest to those considering the policy relevant need to better understand how informal caregivers' mental health was affected during COVID-19. Estimates suggest that mental health was mostly affected when social restrictions were imposed, especially among those people who started caregiving after the COVID-19 outbreak. Of course, that is not to say that social restrictions were necessarily damaging in aggregate. These findings add to recent evidence which suggests that caregivers in general had significantly worse mental health than non-carers during the pandemic, including long-term mental health issues (Park, 2021; Dhiman et al., 2020). Moreover, these results reinforce the need to implement policies to provide psychological support for new informal caregivers, potentially online (Bertuzzi et al., 2021). This might be particularly effective at the start of informal care provision, especially for those who become informal carers for the first time during a public health crisis and for those who might be socially isolated, thus lacking the support of a social network. The type of support offered during a pandemic might have to be very different to other forms of support offered to new informal carers, where more traditionally this would have taken the form of financial support or respite care (Courtin et al., 2014). This may further highlight the need for research into the availability and cost effectiveness of digital tools, the reopening of face-to-face services post pandemic (Giebel et al., 2021), and the relevance of communication (Bailey et al., 2022) to support informal carers, especially in emergency situations and their longer-term aftermath.

Chapter 4

Health and quality of life in ageing populations: A structural equation modelling approach

4.1 Introduction

The global population is undergoing an unprecedented ageing process, representing a major challenge for governments in developed and developing countries (United Nations, 2020). The joint effect of higher life expectancy and lower fertility rates is changing the European population structure (European Commission, 2018), and in response to this, the European Union has increased investment in projects focused on healthy ageing, supporting better health and quality of life of older adults. Among rapidly ageing societies, the improvement of quality of life in advanced age is not only a key public policy issue for policymakers (OECD, 2013; Steptoe et al., 2015), but it is also relevant to all of society, being an ultimate goal of people's lives (Frey and Stutzer, 2002; López Ulloa et al., 2013). Therefore, understanding which factors are related to quality of life in advanced age is important in shaping government policy and advising individuals on how to allocate resources. Although this is a policy-relevant research question, there is discussion among researchers concerning how to define and assess unobserved concepts, i.e. quality of life and self-assessed health, which cannot be directly measured through observed data (Bollen and Bauldry, 2011; McNeish and Wolf, 2020).

A number of studies employ self-assessed single-item questionnaires as proxies of unobservable constructs. For instance, well-being is often proxied with the single question *"How satisfied are you with your life?"* (Collins et al., 2008; Deaton, 2008; Graham et al., 2011). However, given the complex nature of such unobservable concepts, it might be reasonable to think that a single answer may not capture their multidimensionality. As a result, composite indices formed by sets of self-assessed items are often employed to investigate unobserved constructs. For example, the CASP-19 index is often used to measure quality of life specific to older people (Hyde et al., 2003). However, some authors have also questioned the validity of such composite indices, debating the reliability of self-assessed variables to measure unobserved concepts (Althubaiti, 2016; Bollen and Lennox, 1991).

The use of latent variables has been proposed as a solution to overcome measurement issues related to self-assessed items, since the unique measurement error of self-assessed observed items is excluded while constructing latent variables (Bollen, 2002). For this reason these variables are considered to be less biased and more reliable than composite indices (McNeish and Wolf, 2020) because by removing the measurement error of each self-assessed item, the estimate of the true effect is improved (Kline, 2015). Composite indices are instead subject to measurement unreliability assuming that self-assessed items are measured without measurement error (Hoyle, 2012; Bollen, 1989). Specifically, if not accounting for measurement error, the results are biased downward in both coefficients and explained variance (Hoyle, 2012; Kline, 2015).[1]

[1]Table C1 in the Appendix presents a graphical representation of both latent variables and composite indices.

The aim of this chapter is to assess the performance of multiple-item scale scores formed with different weighting structures, while investigating the role of socioeconomic factors towards quality of life in older adulthood. Specifically, the complex concepts of (self-assessed) health and quality of life are measured using composite indices with equal weights, composite indices with weights drawn from principal component analysis and latent variables. Structural equation modelling (SEM) is employed since it allows to include (and construct) latent variables and composite indices. In other words, linear regressions are estimated through SEM including in separate analyses either latent variables or composite indices (with equal weights or with principal component weights). Since the only difference among these regressions is the inclusion of latent variables or composite indices in separate analyses, it is therefore assessed whether different weighting structures can lead to different results.

This chapter contributes to the literature in a number of ways. Conducting this empirical exercise (i.e. employing structural equation modelling with latent variables or composite indices) contributes to the on-going debate related to measurement issues of unobservable concepts. As previously mentioned, unobservable factors, i.e. self-assessed health and quality of life, are often measured using multiple-item scales formed by adding a set of observed indicators or questions. This chapter aims at investigating the performance of multiple-item scale scores formed with different weighting structures while investigating a wide set of factors associated with quality of life among older populations, which is a policy-relevant research question.

In line with theory, this chapter confirms that when SEM with latent variables is employed, the estimates of latent variables are bigger in magnitude compared to composite indices, especially if these concepts are indirectly measured through self-assessed observed items (e.g. physical health in this chapter). Although the estimates are different in magnitude, they all seem to be highly statistically significant.

This suggest that there is not a universal answer to when to use latent variables or composite indices, but it depends on the specific research question. If a particular study aims to accurately measure unobserved concepts proxied with self-assessed items, then the use of latent variables might be more appropriate. This could be relevant when the size of a variable is important (e.g. to detect cases of probable mental ill health or clinical depression according to thresholds in the relative self-assessed scales). Otherwise, if researchers are more interested in the effect of factors on an outcome variable, the use of summed scores might be more appropriate due to their simplicity and common applicability in all statistical models. Therefore, this chapter highlights the importance of accurately choosing how to form scale scores, encouraging researchers to more critically evaluate their choice in the use of multiple items indicators considering their research question since different weighting structures can alter conclusions.

Finally, the overall results of this chapter suggest that non-pecuniary factors such as physical health and participating in social activities play the largest role for a better quality of life as people age compared to pecuniary factors such as income and financial assets. Therefore, policies focused on non-pecuniary aspects, such as improving physical health or widening social networks, could be helpful to increase quality of life in older adulthood.

4.2 Background

Although there has been some significant work on quality of life in old age (Boggatz, 2016; Raggi et al., 2016), there still remain significant gaps around its conceptualisation and measurement (Van Hecke et al., 2018; Vanleerberghe et al., 2017; del Rocío Santana-Berlanga et al., 2020). Some studies have focused on the specific role of some non-pecuniary factors, such as age (Blanchflower and Oswald, 2008; Cheng et al., 2017; Van Landeghem, 2012), sex (Green et al., 2018; Kahneman and Deaton, 2010; Stevenson and Wolfers, 2009) and education (Clark, 2018). Health has gained popularity as a contributing factor to better quality of life in advanced age, with studies focusing on specific health aspects. For instance, Weber et al. (2015) and Freedman et al. (2017) considered physical health and found that it had a strong and positive association with quality of life in advanced age, which was also confirmed by a systematic review by Fortin et al. (2004). Other authors have focused only on cognitive aspects of health, showing that cognitive deterioration was associated with lower quality of life (Allerhand et al., 2014; Comijs et al., 2005; Jetten et al., 2010; Pan et al., 2015). Graham et al. (2011) studied the determinants of quality of life considering the overall health of the respondents while also accounting for some socioeconomic variables. However, these authors employed standard linear regressions, measuring health and quality of life with self-assessed single-items or additive indices, perhaps not ideal as these additive scale scores assume that the unobserved concepts are measured without measurement errors (Hoyle, 2012). Structural equation modelling has been suggested as an alternative approach to assess the effects of demographic and socioeconomic variables on specific quality of life dimensions (for example, Mataria et al. (2009)).

Structural equation modelling frameworks are becoming increasingly popular within social science disciplines such as gerontology and psychology, where several studies implemented this statistical method to understand the role of health towards quality of life using latent variables (Cho et al., 2011; Hirve et al., 2014; León et al., 2020). However, in these studies only the relationship between health and quality of life has been analysed, omitting other socioeconomic factors that might be related to quality of life. The use of structural equation modelling is not widely adopted in economics to date. Previous research within health economics used multiple causes multiple indicators (MIMIC) models to measure unobservable constructs (Wagstaff, 1986; Van de Ven and Van Der Gaag, 1982). These authors have employed MIMIC models to uniquely model health as a latent variable, without estimating the effect of one latent variable on another. The recent availability of high performing computers and dedicated packages in popular statistical software, such as Stata, has contributed to a resurgence of interest in the use of SEM related techniques (Tarka, 2018).[2]

This chapter makes use of SEM to investigate how socioeconomic factors relate to quality of life in older adulthood, which is a policy-relevant topic. SEM is a highly flexible framework, allowing researchers to measure complex unobserved concepts (i.e. self-assessed health and quality of life) constructing both latent and composite variables, while investigating their relationships. The implementation of this framework allows assessing the performance of multiple-item scale scores formed with different weighting structures by including latent and composite variables of the same underlying concept. The SEM approach adds a new dimension to the understanding of using scores with different weighting structures, showing that choosing latent variables or composite indices to investigate the same concept can alter conclusions.

4.3 Data

The data used in this chapter are drawn from the Survey of Health, Ageing and Retirement in Europe (SHARE), which is a longitudinal survey currently conducted in 28 European countries, plus Israel. It includes rich individual-level information about health, employment, housing and socioeconomic status (Börsch-Supan et al., 2005). Each wave of SHARE is collected every two years, with eight waves available at present from Wave 1 (2004) until Wave 8 (2020). Only people older than 50 years old are eligible for the survey.[3]

[2]The first SEM feature of Stata was introduced with Stata 12, with user-friendly improvements in Stata 16. Freeware packages are also available in R, such as LAVAAN.

[3]Spouses or partners regardless of their age are also interviewed. Hospitalized patients and those who are unable to speak the local language are excluded from the survey.

4.3.1 Key variables

Cognitive ability is thought of as an underlying concept since it cannot be directly observed. In fact, it is often inferred from other variables which can be directly measured, such as cognitive tests performed on survey participants. Hence, the latent variable of cognition is constructed using answers related to cognitive tests administered by SHARE interviewers. The choice of variables to construct the unobserved concept of cognition has been driven by the exploratory factor analysis, where all variables related to cognition have been included. According to the EFA, the variables in Table 4.1 were the ones that most represented the underlying concept (having the highest shared variance), and therefore these variables are used to construct the latent variable of cognition.

Table 4.1: SHARE selected variables for cognition

Latent variable	Observed indicators	Description	Value
Cognition	Immediate recall	10 words list to recall immediately	0-10
	Delayed recall	10 words list to recall after a period of time	0-10
	Verbal fluency	Name as many animals as possible	0-100
	Sub numeracy	Subtract a series of numbers	0-5

Note: Higher scores mean better cognitive ability

SHARE includes several self-assessed items asking respondents about their physical health. Here, the latent variable of physical health is measured with items related to mobility limitations (Table 4.2 shows the detailed set of binary questions about difficulties in mobility asked in SHARE). This choice is supported by the fact that in advanced age, physical status and mobility can be considered closely related (Rosso et al., 2013; Webber et al., 2010). Since all mobility items are self-assessed, a latent variable could be considered conceptually more appropriate to measure such constructs accounting for measurement error. Similarly to the construction of the latent variable of cognition, all items related to physical health are included in the exploratory factor analysis. In this case, all the items load on the underlying factor of physical health and so they are all used to construct this latent variable.

Table 4.2: SHARE selected variables for physical status

Latent variable	Observed indicators	Description (Difficulties in:)
	Item 1	Walking 100 meters
	Item 2	Sitting two hours
	Item 3	Getting up from chair
	Item 4	Climbing several flights of stairs
Physical status	Item 5	Climbing one flight of stairs
	Item 6	Stooping, kneeling, crouching
	Item 7	Reaching/extending arms above shoulder
	Item 8	Pulling or pushing large objects
	Item 9	Lifting or carrying weights over 5 kilos
	Item 10	Picking up a small coin from a table

Note: Variables recoded (0-1) so that higher scores mean better physical ability

SHARE contains CASP-12 which is a self-completion questionnaire designed to measure quality of life of older individuals across the four domains of Control; Autonomy; Self-realisation; and Pleasure (Hyde et al., 2003). Within each CASP-12 domain, respondents are asked to answer questions rating how often they experience specific feelings on a 4-point scale (ranging from often to never). The overall score, which is the sum of all the items, form the additive index which is usually used to assess the respondents' quality of life (Gale et al., 2014; Okely et al., 2017; Pascual-Sáez et al., 2019). However, CASP-12 items are here used to form the latent variable of quality of life, given the measurement issues related to self-assessed questionnaires and that well-being might be better thought as a latent factor. Specifically, the latent construct of quality of life is constructed considering the items related to the dimensions of self-realisation and pleasure (Table 4.3), which load on a single latent factor.[4] As suggested by Sexton et al. (2013), pleasure and self-realisation involve the pursuit of happiness and personal fulfilment, capturing the hedonic aspect of well-being.

[4]As shown in Tables C4-C5, in the Appendix

Table 4.3: SHARE items of Quality of life (CASP-12)

Latent variable	Dimensions	Observed indicators	Description
Quality of life	Control	Age prevents doing	Age prevents you from doing
		Out of control	Feel what happens is out of your control
		Left out	Feel left out of things
	Autonomy	Do things	You can do things you want
		Family responsibility	Family responsibilities prevent you from doing
		Money shortage	Shortage of money prevent you from doing
	Self-realisation	Energy	Feel full of energy
		Future opportunities	Feel life is full of opportunities
		Good future	Feel future looks good for you
	Pleasure	Look forward	Look forward to each day
		Life meaning	Feel life has meaning
		Happiness	Look back on life with happiness

Note: Variables recoded (0-3) so that higher scores mean higher quality of life

Apart from the individual characteristics of age, sex, education (measured in years of education) and the economic variables of individual income and monetary assets[5], other observed variables are included since they represent factors which are usually thought to be associated with quality of life, especially in the ageing process. These include marital status (a binary variable indicating if respondents are in a partnership or not); household size (Kotwal et al., 2016; Rosso et al., 2013; Warner and Adams, 2016; Warner and Kelley-Moore, 2012); area of living (with the value of 1 if respondents live in urban areas and 0 if they live in rural areas) and participating in social activities (indicating if respondents have participated in at least one activity among charity, sport, religion and political activities) (Berkman et al., 2000; Rowe and Kahn, 1997).

4.4 Descriptive statistics

The sample is composed of older Europeans, including only individuals who are not clinically depressed and who are not affected by severe cognitive or physical disor-

[5]Income and monetary assets are variables asked at a household level. The individual value is firstly obtained by dividing the household income and monetary assets by the number of household members. Subsequently this value is adjusted at the purchasing power parity index (as displayed by the World Bank International Comparison Program, https://www.worldbank.org/en/programs/icp). Individual income and monetary assets are included in the analysis as quintiles

ders.[6] The reason for this exclusion is that clinical depression could directly impact quality of life (Wilson et al., 2013); cognition would be influenced by chronic cognitive illnesses such as Alzheimer's disease, Parkinson disease and severe dementia; and physical ability would be impacted by serious physical illnesses such as cancer, osteoporosis and hip, femoral or other fractures (Okely et al., 2017; Perrino et al., 2010; León et al., 2020). Therefore, if diagnosed patients are included in the sample, quality of life could be directly driven by these conditions, making it difficult to estimate the associations with the other factors. Conventionally, older individuals have been defined as 65 years old or older (Orimo et al., 2006) and therefore only participants of such age are retained in the final sample. Moreover, in order to avoid complications of endogenous labour supply, only retired individuals are analysed. In fact, the working environment might be a source of work-related stress and rewards which might influence quality of life (Babu et al., 2016; Tzeng et al., 2012).

The aim of this chapter is to assess how latent and manifest variables perform while investigating quality of life among older adults. Therefore, the emphasis is not placed on a longitudinal analysis, but rather on the measurement of unobserved concepts and on the fit of structural models to the analysed data. For these reasons, a cross-sectional analysis is implemented using SHARE Wave 4, Wave 5 and Wave 6.[7] The structural equation model is firstly employed using Wave 4 since it is the first wave of SHARE containing cognitive tests. Wave 5 and Wave 6 are also analysed to assess if the proposed SEM model performs equally well when it is fitted in different waves.Table 4.4 shows the descriptive statistics of the sample divided by the three waves analysed. All the variables included in the analysis have very similar averages across all waves. Residents of North European countries always represent the greatest proportion of respondents, followed by Eastern European countries and

[6]A further analysis conducted on all retired respondents, including also diagnosed patients, shows similar results to the main estimates.

[7]Wave 7 and Wave 8 are not considered in the analysis since Wave 7 presents many missing data of quality of life and Wave 8 became available after this estimation.

Mediterranean countries.

Table 4.4: Descriptive statistics for Wave 4, 5 and 6

	Wave 4		Wave 5		Wave 6	
	Mean	SD	Mean	SD	Mean	SD
Quality of Life	4.81	1.049	4.89	0.996	4.83	1.001
Cognition	1.88	0.538	1.93	0.526	1.94	0.539
Physical status	9.00	1.643	9.05	1.700	9.05	1.622
Gender	0.42	0.494	0.42	0.494	0.41	0.493
Years of education	10.48	4.227	10.87	4.276	10.81	4.399
Age	73.18	6.358	73.24	6.401	73.40	6.418
Marital status	0.71	0.453	0.72	0.447	0.72	0.450
Household size	1.93	0.798	1.93	0.740	1.97	0.824
Urban	0.67	0.470	0.70	0.458	0.68	0.465
Social activities	0.46	0.498	0.45	0.497	0.41	0.491
Country (%):						
North Europe	46.70%		53.22%		42.48%	
Mediterranean	21.26%		23.99%		29.97%	
East Europe	32.04%		22.79%		27.55%	
Income (quintile)	3.00	1.404	3.01	1.406	3.02	1.395
Assets (quintile)	2.07	1.696	2.17	1.717	2.12	1.711
Tot obs.	10,990		14,305		15,520	

4.5 Empirical Analysis

When employing structural equation models with latent variables, the first step is selecting the observed items to form each latent construct through exploratory factor analysis (EFA). Subsequently, the full structural equation model is performed. SEM consists of two parts: i) the measurement model, which is also known as confirmatory factor analysis (CFA), assessing how well the proposed model fits the analysed data; and ii) the structural model, which adds the hypothesised relationships among latent and observed variables. Stata version 17 is used for the analysis.

4.5.1 Exploratory factor analysis

Exploratory factor analysis is used to determine which indicators form each latent variable. The observed items retained for each latent construct are selected from the sets of variables related to each unobserved concept (presented in the Data section 4.3). Specifically, EFA is employed to select the observed items by looking at the factors' composition (Eigenvalues) and at their reliability indices (Cronbach's alpha α and the composite reliability index ρ). Conventionally, a factor is formed if its eigenvalue is greater than 1, meaning that the greatest proportion of overall variance of the unobserved concept is explained by that factor. Moreover, the items are also selected based on thresholds to be met for the reliability indices. These thresholds are 0.8 (or greater) for Cronbach's α and 0.6 (or greater) for composite reliability index ρ (Acock et al., 2013; Kline, 2015).

In this chapter, each latent variable has a simple structure, meaning that conceptually related items load on a single latent variable ensuring in this way a clear conceptual interpretation of the latent construct. Firstly, exploratory factor analysis is conducted on the totality of older adults present in Wave 4 (Tables C2-C7 in the Appendix). However, given that the selection of observed items is data-driven, the same EFA is also conducted on the other two waves (Wave 5 and Wave 6) to assess if the same observed items are still obtained (Tables C8-C11 in the Appendix). The EFA conducted on the three analysed waves confirms that the same observed items are obtained to form the latent variables of interest. This indicates that the factor structure of the latent variables is acceptable and valid for a wide sample of elderly populations and not only for the specific sample analysed in one wave of SHARE.

4.5.2 Confirmatory factor analysis

After the selection of observed items through EFA, confirmatory factor analysis is implemented. The aim of this measurement part of SEM is assessing how well the proposed measurement model fits the data. If the fit of the measurement model is acceptable, the analysis can proceed with the structural part of SEM. In order to assess the fit of the measurement model, goodness-of-fit statistics are estimated to measure how closely the model-implied covariance matrix matches the observed covariance matrix. Table 4.5 shows the most commonly used statistics, with their cut-off thresholds, when the full measurement model is analysed. As it can be noticed, all the goodness-of-fit statistics are acceptable, apart from the Chi-square index which is rejected. However, the Chi-square index is overly sensitive in model testing for large samples (Fan et al., 1999); hence, common practice in SEM to justify retaining models with large sample size is to ignore a failed Chi-square test as long as the other local fit tests are acceptable (Kline, 2015).

Table 4.5: Fit statistics full measurement model on Wave 4, 5 and 6

	Cut-off value	Wave 4	Wave 5	Wave 6
Chi-square	P>0.001	P<0.001	P<0.001	P<0.001
RMSEA	$\leq$ 0.08	0.045	0.048	0.047
CFI	$\geq$ 0.9	0.919	0.915	0.918
SRMR	$\leq$ 0.1	0.039	0.042	0.039

Note: Root Mean Square Error of Approximation (RMSEA);
Comparative Fit Index (CFI);
Standardized Root Mean Squared Residual (SRMR)

4.5.3 Structural equation model

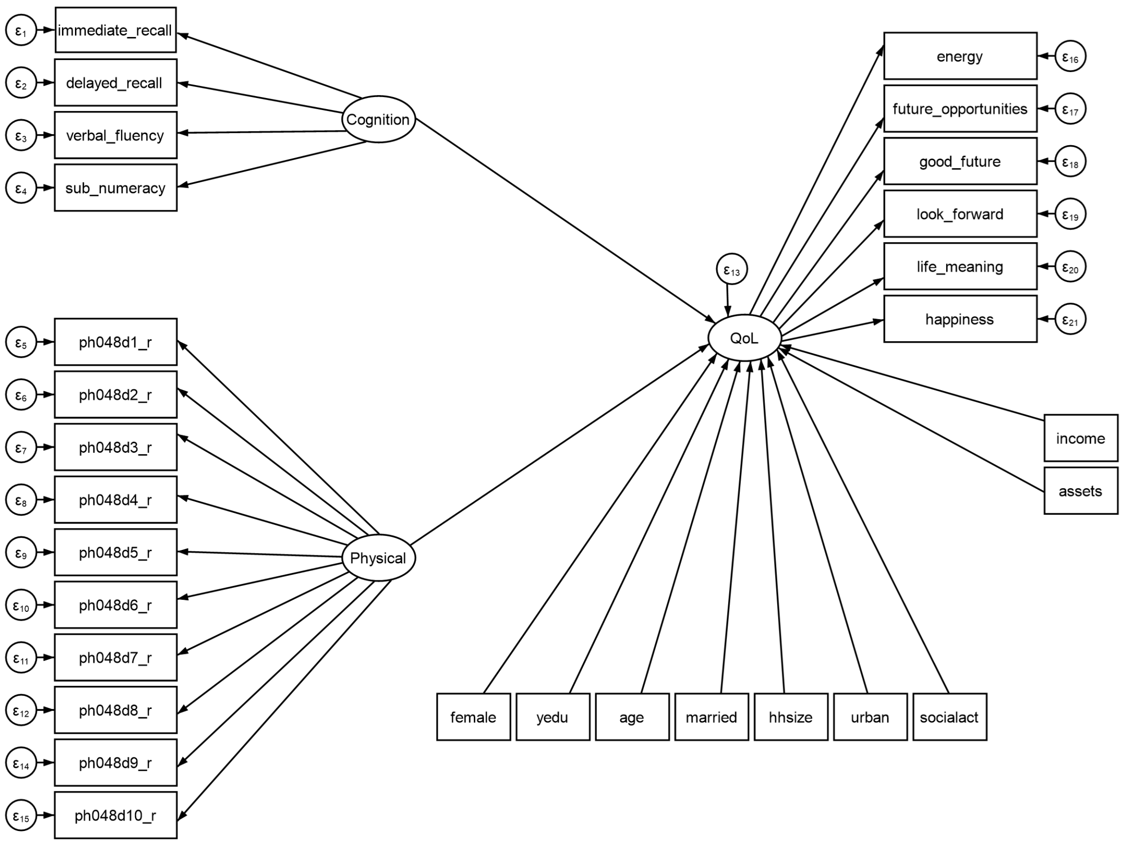
Following the measurement part of SEM, the structural model of Equation 4.1 is employed. The aim is assessing the role of health (as divided into cognition (η_{1i}) and physical health (η_{2i})) along with the socioeconomic factors of personal characteristics (X_i) and economic variables, (Z_i) towards quality of life (Υ_i) of older adults.

$$\Upsilon_i = \alpha_i + \gamma_1\eta_{1i} + \gamma_2\eta_{2i} + \gamma_3X_i + \gamma_4Z_i + \theta_i + \zeta_i \tag{4.1}$$

Where α_i is the intercept and ζ_i is the error term. To fully understand the effect of each factor, Equation 1 is implemented in four specifications. Firstly, Model 1 only estimates the associations of the latent variables of cognition (η_{1i}) and physical health status (η_{2i}) towards the latent variable of quality of life (Υ_i). Model 2 adds a set of individual demographic characteristics (X_i) including sex, years of education, age, marital status, household size, area of living and social activities. Model 3 further adds a set of PPP adjusted economic variables (Z_i) including individual income and individual monetary assets. In the most stringent specification, country-fixed effects (θ_i) are further included to control for different socioeconomic environments. Figure 4.1 is a graphical representation of the full structural model.

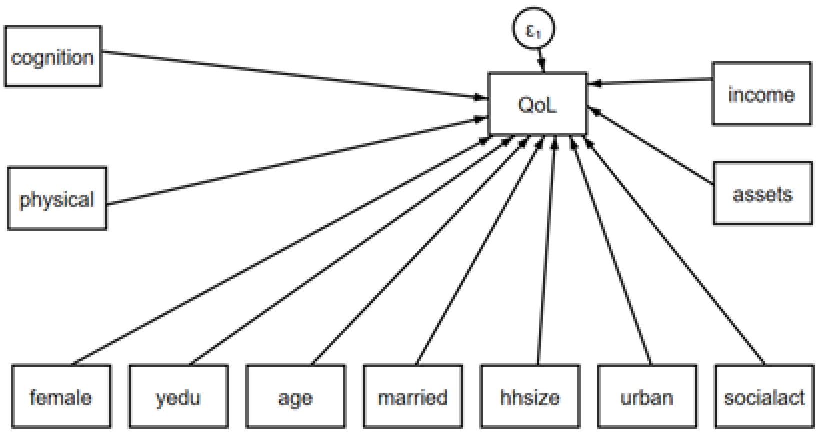
Along with the main SEM estimation with latent variables, composite indices are included in the model instead of latent variables. Specifically, the analysis is firstly conducted with sum scores (also named additive indices) which are obtained by adding the selected observed items using equal weights. Subsequently, weighted scores are used in the analysis. These composite indices are formed through principal component analysis (PCA), with different weights assigned to each item related to each unobserved concept (Table C12-C14 in the Appendix). Figure 4.2 shows a graphical representation of the standard approach with composite indices (either with equal or PCA weights). The use of SEM with latent variables as well as the standard approach with composite indices is useful to assess which statistical tool is the most appropriate to measure such unobserved constructs while investigating quality of life among older adults.

Figure 4.1: Graphical representation of the proposed structural model (SEM)



Note: Unobserved concepts are treated as observed using composite indices. Squared boxes represent observed variables. ph048d are the 10 items of physical status as displayed in Table 4.2

Figure 4.2: Graphical representation of the standard approach with composite indices (linear regression)



4.6 Results

Table 4.6 shows the results of SEM with latent variables alongside the estimates found with standard linear regressions using composite indices with equal weights or PCA indices. Here, only the most stringent specification of the model is displayed, but the results hold for all other model specifications as well (Table C15, Appendix). Table 4.6 shows that both cognition and physical health are important factors for quality of life in older age, with positive and highly statistically significant estimates. SEM is useful to compare the magnitude of the estimates for different indices used to measure the same unobserved concept. Here, it can be noticed that the estimates found when using latent variables are larger in magnitude compared to the ones found when using composite indices. Specifically, the estimates found with additive indices with equal weights are downward biased compared to latent variables as the theory suggests (Kline, 2015). Indices formed with principal component analysis do better than additive indices with equal weights. However, the estimates found are still smaller in magnitude compared to latent variable estimates especially related to physical health which is measured with self-assessed items. On the other hand, the results found for cognition are not as straightforward, where it seems that latent variables and PCA indices perform equally well. This might be explained by the fact that cognition can be thought as an underlying concept, but it is still proxied by objective cognitive tests performed on the respondents.

Table 4.6: Results of Structural Equation Modelling and linear regressions

	Wave 4 QoL			Wave 5 QoL			Wave 6 QoL		
Using:	LV	AInd	PCA	LV	AInd	PCA	LV	AInd	PCA
Cognition	0.114***	0.087***	0.102***	0.088***	0.081***	0.091***	0.111***	0.081***	0.097***
	(0.020)	(0.015)	(0.015)	(0.018)	(0.014)	(0.014)	(0.017)	(0.013)	(0.013)
Physical Health	0.253***	0.185***	0.184***	0.264***	0.190***	0.191***	0.243***	0.186***	0.181***
	(0.017)	(0.013)	(0.013)	(0.015)	(0.012)	(0.012)	(0.015)	(0.011)	(0.011)
Error term	0.699	0.768	0.775	0.684	0.757	0.763	0.663	0.752	0.752

Note that LV (Latent Variables); AInd (Additive indices; PCA (PCA indices)

Hence, these results might indicate that if researchers are interested in accurately measuring unobserved concepts proxied by self-assessed items, the use of latent variables might be more appropriate. By removing the measurement error of each self-assessed item, the estimate of the true effect is improved. This might be useful in research areas focused on the measurement of unobserved concepts. For instance, national healthcare systems often use self-report additive indices to detect cases of probable mental ill health (e.g. GHQ-12) or clinical depression (e.g. CES-D). Therefore being able to accurately measure such unobserved concepts is important to correctly detect cases and to target factors with the larger role towards the outcome of interest.

If instead unobservable concepts are measured through objective tests, as it is the case for cognition, or if the research question does not focus on the size and measurement of an unobserved factor, the use of weighted sum scores might be more appropriate.[8] This is because composite indices are commonly used in all statistical models given their simplicity and widely accepted interpretability.

In terms of the overall findings, Table 4.7 shows the results of all model specifications when using SEM with latent variables. Similar results are found when employing standard regressions, apart from the estimates related to composite indices (cognition and physical health) as previously discussed. According to these estimates, health is the factor with the highest association to quality of life with physical status contributing the most. This might be explained by the significant role that autonomy plays in older adults, allowing them to retain their independence and social contacts (Rosso et al., 2013). Cognition is also positively associated with quality of life but to a lesser extent than physical ability. In terms of the numerical results, looking at the most stringent specification (Model 4), higher cognition of one standard deviation is associated, on average, with a 0.1 standard deviation increase in quality of life, while a higher physical status is associated with a 0.25 increase.

[8] As shown in Table 4.6, the magnitude of the estimates found with PCA indices is really similar to the one found with latent variables

Table 4.7: Standardized results of SEM analysis

	WAVE 4				WAVE 5				WAVE 6			
	Structural equation models				Structural equation models				Structural equation models			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Cognition	0.209***	0.170***	0.147***	0.114***	0.190***	0.126***	0.099***	0.088***	0.216***	0.150***	0.108***	0.111***
	(0.011)	(0.017)	(0.020)	(0.020)	(0.010)	(0.015)	(0.018)	(0.018)	(0.010)	(0.014)	(0.016)	(0.017)
Physical	0.302***	0.274***	0.264***	0.253***	0.320***	0.290***	0.268***	0.264***	0.306***	0.279***	0.258***	0.243***
	(0.012)	(0.015)	(0.018)	(0.017)	(0.010)	(0.013)	(0.016)	(0.015)	(0.010)	(0.013)	(0.015)	(0.015)
Female		0.006	-0.006	0.035**		0.032***	0.042***	0.044***		0.041***	0.066***	0.061***
		(0.014)	(0.016)	(0.016)		(0.012)	(0.014)	(0.014)		(0.011)	(0.013)	(0.013)
Education		0.019	0.009	0.036**		0.028**	0.013	0.017		0.033***	0.007	0.014
		(0.014)	(0.016)	(0.017)		(0.012)	(0.014)	(0.015)		(0.011)	(0.013)	(0.014)
Age		0.003	0.001	-0.016		-0.009	0.001	-0.013		-0.019	-0.014	-0.020
		(0.015)	(0.017)	(0.017)		(0.013)	(0.015)	(0.015)		(0.012)	(0.014)	(0.014)
Married		0.051***	0.037**	0.029		0.082***	0.076***	0.057***		0.068***	0.053***	0.058***
		(0.016)	(0.019)	(0.018)		(0.014)	(0.017)	(0.016)		(0.013)	(0.015)	(0.015)
Household size		-0.034**	-0.025	0.023		-0.030**	0.011	0.020		-0.027**	0.033**	0.003
		(0.015)	(0.018)	(0.018)		(0.014)	(0.017)	(0.017)		(0.012)	(0.015)	(0.015)
Urban		-0.050***	-0.065***	-0.027*		-0.008	-0.018	0.006		-0.019*	-0.029**	-0.003
		(0.013)	(0.015)	(0.015)		(0.011)	(0.013)	(0.013)		(0.011)	(0.012)	(0.012)
Social activities		0.163***	0.155***	0.089***		0.187***	0.166***	0.127***		0.194***	0.169***	0.134***
		(0.013)	(0.015)	(0.015)		(0.012)	(0.014)	(0.014)		(0.011)	(0.013)	(0.013)
Income			0.022	0.049**			0.201	0.082***			0.191***	0.057***
			(0.016)	(0.020)			(0.015)	(0.017)			(0.014)	(0.017)
Assets			0.082***	0.031*			0.052	0.024*			0.082***	0.046***
			(0.016)	(0.016)			(0.014)	(0.015)			(0.014)	(0.014)
Country FE				✓				✓				✓
N(obs)	N(10,615)	N(6,833)	N(4,836)	N(4,836)	N(13,650)	N(8,552)	N(5,902)	N(5,902)	N(15,101)	N(9,480)	N(6,707)	N(6,707)
Chi-squared	P<.001	P<.001	P<.001	P<.001	P<.001	P<.001	P<.001	P<.001	P<.001	P<.001	P<.001	P<.001
RMSEA	0.045	0.040	0.039	0.042	0.048	0.042	0.041	0.044	0.047	0.040	0.040	0.037
CFI	0.919	0.902	0.901	0.812	0.915	0.901	0.897	0.818	0.918	0.907	0.902	0.848
SRMR	0.039	0.035	0.033	0.031	0.042	0.036	0.036	0.033	0.039	0.034	0.033	0.027

Note: Cut-off values for fit indices: RMSEA$\leq .08; CFI \geq .9; SRMR \leq 0.1$

Standard error included in parenthesis

*** p$\leq 0.01, **p \leq 0.05, *p \leq 0.1$

It can be noticed that, compared to the other factors included in the analysis, the overall health status plays the greatest role in enhancing older adults' well-being. Importantly, the results found within a SEM framework are standardized. Therefore, it is possible to compare the extent to which different factors affect the outcome variable, identifying which aspect contributes the most to increasing the outcome of interest.

As expected, participating in social activities is positively associated with quality of life, contributing to the same extent as cognition. In fact, participating in social activities is associated with an overall increase of 0.12 standard deviation in quality of life. However, it can be seen that its magnitude is reduced when country fixed effects are included in the analysis. Age does not seem to be associated with quality of life. This result does not support the theory of age-related decline leading to lower quality of life, nor the widely documented age-related paradox of a U-shaped relationship (Blanchflower and Oswald, 2008; Easterlin, 2006). If the paradox was holding here, a positive association between age and quality of life should have been found since older people are located in the upward sloping end of the U-shaped relationship. As expected from past research on loneliness among older individuals, being married is associated with a higher quality of life (Warner and Adams, 2016; Warner and Kelley-Moore, 2012), even if here it seems that living in a household with many members is not significantly associated with quality of life. Finally, the economic variables of income and monetary assets, are positively and significantly associated with quality of life. However, economic status does not seem to play a major role towards higher quality of life. Both economic variables are positively and significantly associated with quality of life but to a lesser extent than other non-economic factors, quantified as an increase of 0.9 standard deviation if combining both economic variables.

These results are consistent for all waves considered in the analysis. Accordingly, it can be noticed from Table 4.7 that the increase of quality of life is mainly associated with an increase in the overall health status and participation in social activities while economic assets do not seem to play a major role. Since the results are consistent throughout all waves, this suggests that the identified factors are important for the entire period analysed and not only for a specific point in time.

Hence, individuals should allocate their resources on non-pecuniary factors such as maintaining good health and having a wide social network. These aspects seem to contribute the most to have a higher quality of life, while an emphasis on accumulating economic assets might not give the expected payoffs in increasing quality of life in later ages. At a governmental level, social activities involving older adults as well as health policies focused on enhancing their physical health might have a positive effect for increasing quality of life in advanced age.

4.6.1 Robustness Checks

The default estimation for both SEM and linear regressions is based on complete case analysis. Full information maximum likelihood (FIML) can be performed within the SEM framework, where the estimates are imputed using all the available information (Allison, 2003; Hoyle, 2012). Here, despite the percentage of missing data, the remaining sample of respondents with complete information is still large enough to conduct a complete case analysis without risking losing statistical power. However, SEM with full information maximum likelihood estimation (FIML) is also performed as sensitivity analysis, confirming the results previously presented (see Table C21 in the Appendix).

4.6.2 Limitations

There are some limitations to consider. Firstly, the direction of the relationship between health and quality of life is widely debated. For instance, Guven and

Saloumidis (2009), Chei et al. (2018) and Lawrence et al. (2019) show that happiness contributes to longevity and others show that greater well-being contributes to better cognitive functions (Allerhand et al., 2014; Boyle et al., 2012; Llewellyn et al., 2008). However, when we consider these multi-dimensional concepts in their entirety, it is reasonable to assume that poor overall health status affects individuals' quality of life rather than vice versa (Easterlin, 2003; Hill et al., 2017; Foroughan et al., 2018). Given that SEM is a confirmatory tool, this is the tested hypothesis when fitting the proposed model to the analysed data.

Another limitation concerns the fact that structural equation modelling, as well as factor analysis, is heavily data based. Therefore, some might argue that the latent variables are only specific to the analysed sample. One way to overcome this concern is to do a cross-sectional validation as is common practice in machine learning techniques. Following this, the sample is randomly divided into two parts. The analysis is then conducted on these two random subsamples finding similar results (Tables C16-C20, Appendix). A further sensitivity analysis is conducted considering participants present only in one wave, either Wave 5 or Wave 6, to have a completely different set of respondents compared to the ones present in Wave 4 (Tables C8-C11, Appendix). All these sensitivity analyses confirm that the latent variables are acceptable. However there might still be concerns about the external validity of the results found. According to this research, the proposed SEM model is a good fit for the three waves of SHARE analysed, but it might be possible that this SEM model and the results found in this chapter might not hold when considering different data sets. Since SHARE has comparable data sets from other countries, such as TILDA (The Irish Longitudinal Study on Ageing) and HRS (Health and Retirement Study), one way to overcome the external validity concern might be to conduct the same SEM analysis on these other data sets to see if the model is still acceptable.

4.7 Discussion

The aim of this chapter is to contribute to the general debate concerning the theory and applicability of manifest and latent variables, while investigating factors associated with quality of life among older populations. It provides a contribution to this controversial and on-going debate around the measurement of unobservable concepts. Specifically, latent and manifest variables (sum and weighted sum score via principal component analysis) are here used to measure unobserved concepts i.e. health and quality of life within a SEM framework. This empirical exercise is applied on a timely research question investigating quality of life among older populations.

Overall, the results suggest that the estimates of additive indices with equal weights are biased downward compared to latent variables. This is also the case for scale scores with principal component weights if the unobserved concepts are measured by self-assessed items (i.e. physical health). When unobserved concepts are instead inferred from objective items (i.e. cognition), it seems that latent variables and PCA indices perform equally well. Importantly, the magnitude of the estimates differs if using multiple item indicators with different weighting structures, but the significance of the results does not seem to be affected. This may suggest that the choice of using latent variables or composite indices depends on the specific research question. For example, if a study aims to accurately value and measure unobserved concepts, then the use of latent variables might be more appropriate. If instead the interest lies on the relationship among variables, composite indices might be more useful since they are widely used and implemented in all statistical models. Therefore, this chapter contributes to the argument that there is not a universal answer to the question of when to use composite scores and when to use latent variables, but it depends on the research question at hand (Jacobs and Goddard, 2007; Goddard and Jacobs, 2009), and this is an important finding for future research.

Finally, this empirical exercise shows that health, divided into the latent variables of cognition and physical status, is the factor associated the most with quality of life. Participating in social activities has also a great effect. Interestingly, income and monetary assets do not contribute as much to individuals' quality of life in old age. Importantly, the estimates represent changes in standard deviations, therefore it is possible to compare different factors looking at their associations with the outcome variable directly assessing which factor is playing the bigger role for quality of life among older adults. According to the results found, older individuals and related policies should focus more on non-pecuniary aspects of their lives, such as improving their physical health or widening their social network, rather than on accumulating economic resources which does not seem to contribute as much to quality of life in advanced age.

For example, local authorities could organise social activities at the municipality level. Encouraging older individuals to attend such activities could widen their social network and improve their social participation, contributing to a higher quality of life. Moreover, activities to improve fine motor skills under the supervision of health professionals such as geriatricians or physiotherapists could be particularly helpful to improve older people's physical health and hence increasing their well-being.

Chapter 5

Conclusions

This thesis examines socioeconomic factors influencing health and well-being at specific phases of people's lives. In other words, each chapter investigates socioeconomic factors specific to three key life stages of individuals' life-cycle (i.e. early childhood, adulthood and older adulthood). Findings of this thesis suggest that there is an intergenerational transmission of parental life-style choices and behaviours (e.g. education and smoking) on offspring's infant health. Specifically, higher parental education contributes to better children's health, while parental smoking behaviour leads to worse offspring's health outcomes. Additionally, providing informal care is detrimental for informal caregivers' mental health especially at the start of their caregiving provision period and when caregivers do not have a wide social network to support them (e.g. during COVID-19 lockdown periods). Finally, having better health and participating in social activities is positively associated to a higher quality of life in older adulthood, while accumulating financial and monetary assets does not seem to give the expected payoffs in terms of higher quality of life.

Overall, these results fall within the "cross-sectional" aspect of a life course approach, where social elements interact with each other and the (dis)advantages tend to cluster cross-sectionally (Marmot and Wilkinson, 2005). In other words, (dis)advantages in one sphere (e.g., low-paying job) are likely to be accompanied by (dis)advantages in another sphere (e.g., worse housing conditions) within the same life stage. Therefore, from a policy-making perspective, understanding which socioeconomic determinants are specific to key stages across people's lives is important to address such risk factors and enhance individuals' health and well-being through timely investments (Ashton et al., 2020).

In other words, policies targeting the factors which contribute the most to individuals' health and well-being in each life stage could be particularly effective. For example, implementing intervention strategies in schools promoting the importance of education and the risks associated to smoking behaviours can improve individuals' health as well as the one of their offspring. If policymakers are interested in improving informal caregivers' mental health, providing psychological support and a social network around informal caregivers could be particularly effective for those at the start of their caregiving provision and for those who might be socially isolated. Finally, if funding is allocated to improve older people's quality of life, organising social activities at the municipality level also involving fine motor activities supervised by health professionals could be important. This would increase individual social participation while improving their physical health, which are both key factors for higher well-being as people age.

One avenue for future research could be to consider the "longitudinal" aspect of a life course approach, where one phase of life influences the next and the (dis)advantages accumulate over time (Marmot and Wilkinson, 2005). Here, a "longitudinal" life course model could be implemented to investigate determinants of health and well-being considering how different life phases are interconnected and influence each other. Research of this kind is important to shed further light on mechanisms such as social accumulation, social mobility and social protection and to tackle issues such as inequality and equity among societies.

Another distinctive feature of this thesis is that it explores three different longitudinal data sets: chapter 2 uses the National Longitudinal Study of Adolescent to Adult Health (Add Health) in the US; chapter 3 analyses the UK Household Longitudinal Study (Understanding Society); while chapter 4 employs the Survey of Health, Ageing and Retirement in Europe (SHARE). Exploiting longitudinal data together with causal inference methods is essential to properly address questions concerning the impact of (risk) factors and relevant (health) outcomes of interest. Despite the benefits of longitudinal designs for science and policy, data collection is resource-intensive given the long-term nature of these studies as well as the amount of funding and time investment required. For these reasons, longitudinal research has mostly been conducted in developed countries, while research efforts in developing countries have mainly focused on cross-sectional studies being less expensive and easier to implement avoiding the process of tracking respondents over time (National Academies of Sciences and Medicine, 2002).

This thesis presents analyses of longitudinal data sets collected in developed countries given their ready availability and (relatively) easy accessibility. Indeed, an avenue of future research in this area would be to contribute to the scientific knowledge around individuals' health and well-being also considering developing countries. A thorough understanding of such topics would be possible if greater funding is allocated to longitudinal research in developing countries. This would allow researchers to analyse such data, hence being able to advice policymakers on effective policies targeted at improving health and well-being of their populations.

Although longitudinal studies are important to address research questions concerning causal relationships, a limitation often incurred by researchers is the inconsistency of the questions asked in different waves of the same data sets. For example, in chapter 3 it would have been interesting to investigate the intensity of care provided during COVID-19. The main stage questionnaire waves include a specific question asking how many hours of care respondents provide on a weekly basis. However, this question was not asked when the COVID survey waves started, making it impossible to estimate the mental health effect of caregivers with a higher care burden. Also questions regarding caregiving tasks or the care recipients' identity were not consistently asked throughout the survey waves. Indeed, this set of information would have allowed one to draw a more complete picture of informal caregivers' mental health.

Another issue that researchers need to deal with is missing data. Responding to survey questions is voluntary, and as such, individual level information of some respondents' characteristics is often missing. For example, the data set used in chapter 2 also contains information on non-cognitive traits of main respondents including the so called Big Five personality traits (i.e. neuroticism, extroversion, consciousness, openness and agreeableness). However, there is a high amount of missing information about these variables which considerably reduces the sample size and therefore lowers the precision of the estimates found. Since it is costly to conduct longitudinal studies, as previously mentioned, it could be useful to implement strategies during the data collection phase in order to avoid missing answers especially related to non-sensitive information. Although methods for dealing with missing data (including listwise deletion and imputation methods such as multiple imputation, maximum likelihood and Bayesian simulation) could be employed, minimising missing data during data collection would help researchers in drawing a wider picture of the mechanisms of interest while increasing the statistical power of their work.

The timing of the collection of survey data is also important for researchers. For example, it would have been interesting to employ a regression discontinuity design (RDD) approach in chapter 3. This method would be useful to estimate causal effects if data are available in the time frame just around the time point of an intervention or event. Specific to chapter 3, it would have been interesting to isolate the effect of the start of the COVID-19 pandemic if mental health information of informal caregivers and never-carers was available in the days just before and after the implementation of the first national lockdown. However, the timing of the collection of the questionnaires of the COVID surveys made it impossible to conduct such analysis. More generally, a higher coordination between researchers and teams in charge of data collection could be really useful to collect data in ways that policy experiments or quasi-experimental designes could be more easily implemented.

A limitation specific to chapter 4 is that the three analysed survey waves, which are part of a longitudinal data set, are considered separately. A longitudinal approach is not needed in this chapter since its aim is to assess the performance of indicators with different weighting structures. However, future research could explore these associations across time as well. The implementation of standard econometric techniques using longitudinal strategies is widely known and implemented. Using structural models with latent variables is instead more complicated since a latent growth model should be employed. Latent growth models are broadly implemented when all the variables of the model are manifest and only the intercept and the slope are considered as latent variables. However, the computational complexity of the model increases when the aim is to estimate associations among factors measured with latent variables over time. The implementation of such an approach has not been considered in this thesis, but it could be an interesting future research step.

5.1 Final remarks

This thesis aims at contributing to the economics literature by shedding further light on socioeconomic determinants of health and well-being. Each chapter draws conclusions from research questions investigated in the key life stages of childhood, adulthood and older adulthood.

Chapter 2 suggests that higher education and avoidance of risky health behaviours, such as smoking, of one generation have impacts on better health of the next generation. This is extremely important from a public policy perspective, where educational investments and prevention policies are fundamental not only for the health of the targeted individuals, but also for their offspring. This goes hand in hand with tackling intergenerational mobility, which is fundamental to reduce inequalities among populations.

Chapter 3 provides evidence that becoming an informal caregiver during COVID-19 had detrimental effects on caregivers' mental health. This could be due to the new responsibility of providing care to vulnerable (older) people, coupled with enforced social isolation and lockdown policies. Results of this chapter suggest that psychological support for informal caregivers is needed, especially at the start of their caregiving provision period and for those who are socially isolated. Therefore, policies aimed at creating a wide social network around informal caregivers could be effective to improve their mental health and well-being. Understanding how to best support unpaid caregivers is of primary importance for governments since the informal care sector plays an essential role for the well-being of our societies.

Chapter 4 suggests that the use of indicators constructed with different weighting structures can lead to different results. The idea behind this chapter is that unobservable factors, such as (self-assessed) health and quality of life, are often measured using multiple-item scales formed by adding a set of observed indicators or questions. However, this chapter along with other research (Jacobs and Goddard, 2007; Bollen and Bauldry, 2011; McNeish and Wolf, 2020) shows that the weighting structure of such indicators can determine the estimates of the outcome of interest. This chapter draws attention on the importance to critically evaluate the indices used while investigating correlations among socioeconomic factors and quality of life of older adults. Although there is not a formal and definite solution for this methodological question, it is indeed something that should be considered by policymakers and researchers.

Overall, the evidence produced by this thesis will be useful for policymakers to implement timely investments targeting socioeconomic factors which improve health and well-being during key life stages of the population.

Appendix

Chapter 2, Appendix

Table A1: IG effects of parental education and smoking on low birth weight, by ethnicity

	(1)	(2)	(3)	(4)	(5)	(6)
Panel 1 - IGT generation I and II						
		White			***Non-White***	
	YoEd	Smokpreg	RegSmok	YoEd	Smokpreg	RegSmok
YoEd_{genI}	0.071***			0.056***		
	(0.006)			(0.005)		
RegSmok_{genI}		0.110***	0.176***		0.045***	0.069***
		(0.012)	(0.017)		(0.008)	(0.013)
Controls	✓	✓	✓	✓	✓	✓
N	3362	4121	3959	2860	3929	3797
F	168.27	78.87	102.50	110.90	30.08	28.16
Panel 2 - Effects on Low birth weight						
YoEd_{genII}	-0.015***			-0.013**		
	(0.004)			(0.005)		
Smokpreg_{genII}		0.076			0.360***	
		(0.059)			(0.123)	
RegSmok_{genII}			0.054*			0.249***
			(0.032)			(0.080)
$\text{Birthorder}_{genIII}$	0.003	0.005	0.006*	0.001	-0.000	0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Female_{genIII}	0.019***	0.018***	0.018***	0.020**	0.021***	0.021***
	(0.007)	(0.006)	(0.006)	(0.008)	(0.007)	(0.008)
$\text{Prenatalcare}_{genIII}$	-0.006	-0.015	-0.013	-0.028*	-0.037***	-0.022
	(0.013)	(0.014)	(0.012)	(0.015)	(0.014)	(0.015)
N	6635	7983	7650	6051	8199	7936

Clustered Standard errors at family level in the 1st stage
Robust Standard errors in the 2nd stage
Significance levels: *** 1% ** 5% * 10%

Table A2: IG effects on same sample of children

	(1)	(2)	(3)
Panel 1 - Effects on Low birth weight			
Separate channels			
YoEd_{genII}	-0.013***		
	(0.003)		
Smokpreg_{genII}		0.090*	
		(0.054)	
RegSmok_{genII}			0.083**
			(0.033)
$\text{Birthorder}_{genIII}$	0.003	0.004	0.005*
	(0.003)	(0.003)	(0.003)
Female_{genIII}	0.018***	0.018***	0.018***
	(0.005)	(0.005)	(0.006)
$\text{Prenatalcare}_{genIII}$	-0.024**	-0.030***	-0.025**
	(0.010)	(0.011)	(0.010)
N	12088	12088	12088
Panel 2 - Effects on Low birth weight			
Joint channels			
YoEd_{genII}	-0.013***	-0.012***	
	(0.003)	(0.003)	
Smokpreg_{genII}	-0.009		
	(0.060)		
RegSmok_{genII}		0.011	
		(0.037)	
$\text{Birthorder}_{genIII}$	0.003	0.002	
	(0.003)	(0.003)	
Female_{genIII}	0.018***	0.018***	
	(0.005)	(0.005)	
$\text{Prenatalcare}_{genIII}$	-0.023**	-0.024**	
	(0.012)	(0.010)	
N	12088	12088	

Clustered Standard errors at family level in the 1st stage
Robust Standard errors in the 2nd stage
Significance levels: *** 1% ** 5% * 10%
Controls are included in the first stage

Table A3: IG effects considering first born children only

	(1)	(2)	(3)
Effects on Low birth weight			
$YoEd_{genII}$	-0.014***		
	(0.004)		
$Smokpreg_{genII}$		0.104	
		(0.078)	
$RegSmok_{genII}$			0.083**
			(0.042)
$Female_{genIII}$	0.015*	0.017**	0.017**
	(0.008)	(0.007)	(0.007)
$Prenatalcare_{genIII}$	-0.010	-0.016	-0.014
	(0.012)	(0.012)	(0.011)
N	6157	7949	7654

Significance levels: *** 1% ** 5% * 10%

Controls are included in the first stage

Table A4: IG effects using alternative outcome: continuous birth weight

	(1)	(2)	(3)
Effects on birth weight (standardized)			
$YoEd_{genII}$	0.047***		
	(0.010)		
$Smokpreg_{genII}$		-0.478***	
		(0.162)	
$RegSmok_{genII}$			-0.372***
			(0.096)
$Birthorder_{genIII}$	-0.021**	-0.022***	-0.024***
	(0.009)	(0.008)	(0.008)
$Female_{genIII}$	-0.232***	-0.214***	-0.214***
	(0.018)	(0.016)	(0.016)
$Prenatalcare_{genIII}$	0.083**	0.097***	0.077***
	(0.033)	(0.032)	(0.030)
N	12156	15504	14926

Significance levels: *** 1% ** 5% * 10%

Controls are included in the first stage

Table A5: IG effects alternative outcome: general health

	(1)	(2)	(3)
Effects on (good) health			
YoEd_{genII}	0.012***		
	(0.002)		
Smokpreg_{genII}		-0.050	
		(0.040)	
RegSmok_{genII}			-0.065***
			(0.023)
$\text{Birthorder}_{genIII}$	-0.008***	-0.009***	-0.008***
	(0.002)	(0.002)	(0.002)
Female_{genIII}	0.009**	0.009**	0.006
	(0.004)	(0.004)	(0.004)
$\text{Prenatalcare}_{genIII}$	0.012	0.012	0.007
	(0.008)	(0.008)	(0.007)
N	12560	16025	15437

Significance levels: *** 1% ** 5% * 10%

Controls are included in the first stage

Table A6: Grandparents fixed effect models, by family education

	(1)	(2)
Effects on Low birth weight		
	$\text{Fam}_{HighEdu}$	Fam_{LowEdu}
YoEd_{genII}	0.007	-0.006
	(0.007)	(0.006)
$\text{Birthorder}_{genIII}$	0.004	0.000
	(0.003)	(0.003)
Female_{genIII}	0.020**	0.023***
	(0.008)	(0.009)
$\text{Prenatalcare}_{genIII}$	-0.118***	-0.094**
	(0.041)	(0.040)
N	5710	6245

Significance levels: *** 1% ** 5% * 10%

Chapter 3, Appendix

Table B1: Propensity Score Matching on pre-treatment variables

	Existing carers VS never-carers		New carers VS never-carers	
	Before PSM: p-value	After PSM: p-value	Before PSM: p-value	After PSM: p-value
Need to provide care:				
Married/civil partner	0.3452	0.753	**0.0336**	0.511
Living alone	**0.0005**	0.648	**0.0050**	0.127
Number of children under 16	**0.0204**	0.602	**0.0000**	0.507
Willingness to provide care:				
Paid employment	0.3130	0.814	**0.0000**	0.908
Job type	0.4261	0.625	**0.0389**	0.847
Income (quintiles)	0.7268	0.267	**0.0042**	1.000
Ability to provide care:				
Age	**0.0010**	0.495	**0.0000**	0.403
Female	**0.0000**	0.815	**0.0000**	0.885
White	**0.0180**	0.690	0.5379	0.333
Long standing illness or disability	**0.0664**	0.875	**0.0071**	0.165
Self-assessed health	**0.0202**	0.593	**0.0001**	0.370
SF-12 physical health	**0.0023**	0.507	**0.0006**	0.614
SF-12 mental health	0.7500	0.778	**0.0912**	0.353
Number of functional limitations	**0.0369**	0.823	**0.0012**	0.291
Satisfaction with health	**0.0457**	0.677	**0.0114**	0.587
Satisfaction with income	**0.3640**	0.315	0.7616	0.445
Satisfaction with life overall	0.9661	0.779	0.2632	0.772
GHQ likert scale (baseline)	0.3734	0.988	**0.0054**	0.263
GHQ caseness scale (baseline)	0.5735	0.795	**0.0662**	**0.094**

Note: p-value of difference

Table B2: Caregiving tasks performed in Covid wave 1, by caregiver status

Caregiving task:	Existing carers	New carers
Giving lifts	0.059	0.027
Shopping	0.826	0.908
Cooking	0.233	0.138
Helping with personal needs	0.042	0.006
Washing/cleaning	0.127	0.028
Dealing personal affairs	0.258	0.078
Assisting online or internet access	0.157	0.118
Gardening/house repairs	0.097	0.076
Looking after children	0.03	0.023
Something else	0.131	0.125
Sum of tasks	1.962	1.527
N. observations	236	1403

Table B3: Identity of person being cared for in Covid wave 1, by caregiver status

Care-recipients:	Existing carers	New carers
Adult children (also in-law)	0.148	0.134
Parents or grandparents (also in-law)	0.619	0.504
Siblings	0.047	0.088
Spouse or partner	0.038	0.023
Former spouse or partner	0.004	0.010
Friends	0.195	0.247
Neighbours	0.415	0.432
Someone else	0.076	0.080
Sum of number of care recipients	1.542	1.518
N. observations	236	1404

Table B4: Oster Test Results

	Mental health, GHQ-12			
	(1)	(2)	(3)	(4)
Treatment variable:	Baseline Effect	Controlled Effect	Fixed Effects	δ for β=0
Carer after Covid	(Std. Error), [R^2]	(Std. Error), [R^2]	(Std. Error), [R^2]	Given R_max]
	0.220*** (.080) [.5102]	0.522*** (.063) [.4991]	0.369** (.144) [.5038]	17.13639

Note: (1) regression without controls ; (2) with observed controls;
(3) including all time and individual FE; (4) (δ for β=0.
The bounding set using R_max and (δ=1 is [0.209, 0.830]

Table B5: DD results. GHQ-caseness, using Longitudinal Weights

DD interactions	GHQ Existing carers			GHQ New carers		
(Wave 10 as baseline)	(1)	(2)	(3)	(1)	(2)	(3)
Carer x Wave 8 (2016-2018)	-0.475	-0.264	-0.263	0.121	0.033	0.033
	0.162	0.364	0.364	0.486	0.858	0.858
Carer x Wave 9 (2017-2019)	-0.383	0.277	0.277	0.015	0.000	0.000
	0.217	0.499	0.499	0.921	0.998	1.000
Carer x Wave Covid 1 (April 2020)	0.590	0.820*	0.822*	0.410**	0.390*	0.387*
	0.181	0.076	0.076	0.025	0.060	0.063
Carer x Wave Covid 2 (May 2020)	0.425	0.199	0.200	0.280	0.316	0.312
	0.270	0.564	0.561	0.168	0.151	0.157
Carer x Wave Covid 3 (June 2020)	0.427	0.070	0.072	0.688***	0.755***	0.752***
	0.302	0.826	0.820	0.000	0.000	0.000
Carer x Wave Covid 4 (July 2020)	0.331	-0.054	-0.057	0.228	0.247	0.241
	0.400	0.857	0.850	0.281	0.245	0.255
Carer x Wave Covid 5 (Sept 2020)	0.507	0.068	0.066	0.387*	0.412*	0.409*
	0.225	0.810	0.815	0.088	0.075	0.075
Carer x Wave Covid 6 (Nov 2020)	0.089	0.485	0.484	0.381*	0.516**	0.513**
	0.801	0.213	0.214	0.058	0.013	0.013
Carer x Wave Covid 7 (Jan 2021)	0.409	0.935*	0.933*	0.364*	0.432*	0.429*
	0.301	0.057	0.056	0.077	0.059	0.058
Carer x Wave Covid 8 (Mar 2021)	0.236	0.200	0.198	0.336	0.442*	0.440*
	0.635	0.526	0.530	0.151	0.066	0.065
Individual Characteristics		✓	✓		✓	✓
Covid variables			✓			✓
Observations	27764	20951	20951	43065	31131	31131
N of respondents	2524	1909	1909	3915	2838	2838

Note: All models include individual and time fixed effects.

Grey estimates indicate when the UK was under strict social restrictions while white rows indicate more relaxed periods.

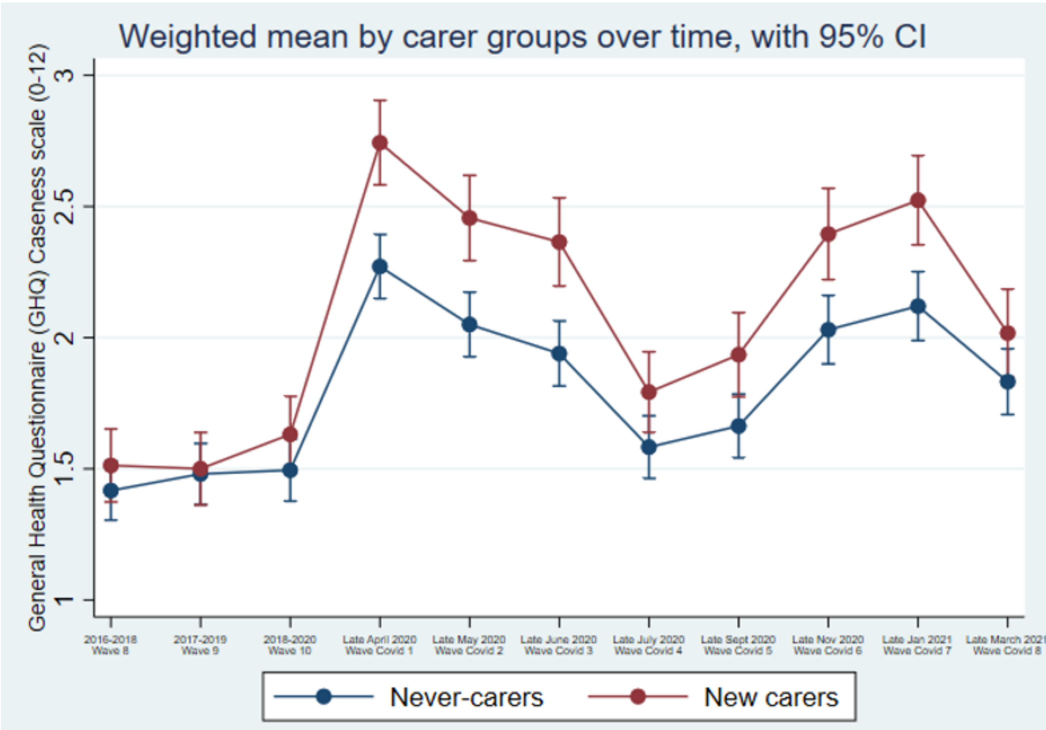
Table B6: DD results. GHQ-caseness binary

DD interactions	GHQ Existing carers			GHQ New carers		
(Wave 10 as baseline)	(1)	(2)	(3)	(1)	(2)	(3)
Carer x Wave 8	-0.049	-0.027	-0.027	-0.004	0.004	0.004
	(0.046)	(0.037)	(0.037)	(0.014)	(0.017)	(0.017)
Carer x Wave 9	-0.069	-0.021	-0.021	-0.012	-0.001	-0.001
	(0.046)	(0.033)	(0.033)	(0.014)	(0.016)	(0.016)
Carer x Wave COVID 1 (April 2020)	0.048	0.002	0.002	0.043**	0.052**	0.052**
	(0.062)	(0.041)	(0.041)	(0.017)	(0.020)	(0.020)
Carer x Wave COVID 2 (May 2020)	0.043	0.011	0.012	0.026	0.037*	0.037*
	(0.048)	(0.036)	(0.036)	(0.016)	(0.019)	(0.019)
Carer x Wave COVID 3 (June 2020)	0.024	0.042	0.042	0.033**	0.046**	0.046**
	(0.051)	(0.038)	(0.038)	(0.016)	(0.019)	(0.019)
Carer x Wave COVID 4 (July 2020)	0.014	-0.004	-0.004	0.002	0.013	0.012
	(0.051)	(0.036)	(0.036)	(0.015)	(0.018)	(0.018)
Carer x Wave COVID 5 (Sept 2020)	0.017	-0.006	-0.006	0.002	0.016	0.015
	(0.053)	(0.037)	(0.037)	(0.015)	(0.018)	(0.018)
Carer x Wave COVID 6 (Nov 2020)	0.027	0.018	0.017	0.015	0.023	0.022
	(0.043)	(0.035)	(0.035)	(0.016)	(0.019)	(0.019)
Carer x Wave COVID 7 (Jan 2021)	0.062	0.081**	0.080**	0.029*	0.028	0.027
	(0.054)	(0.038)	(0.038)	(0.017)	(0.019)	(0.019)
Carer x Wave COVID 8 (Mar 2021)	-0.030	-0.013	-0.013	-0.005	0.005	0.004
	(0.049)	(0.035)	(0.035)	(0.016)	(0.018)	(0.018)
Individual Characteristics		✓	✓		✓	✓
Covid variables			✓			✓
Observations	30360	23176	23176	47322	34472	34472
N of respondents	2760	2112	2112	4302	3143	3143

Note: All models include individual and time fixed effects.

Grey estimates indicate when the UK was under strict social restrictions while white rows indicate more relaxed periods.

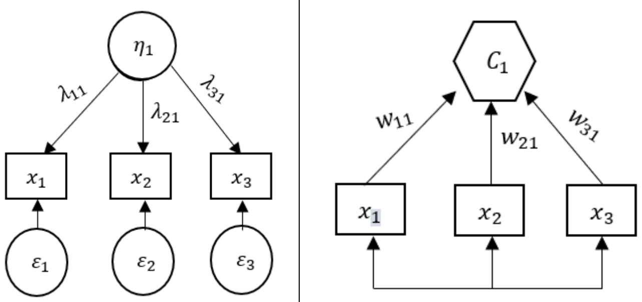
Figure B1: GHQ weighted mean over time by new carers and never-carers with confidence intervals



Chapter 4, Appendix

Table C1: Example of latent variable and composite variable with three observed items

Latent variable | Composite variable



Eq. Latent variable
$X_n = \alpha + \Lambda\eta_1 + \varepsilon_n$

Eq. Composite variable
$C_1 = \alpha + w_{11}x_1 + w_{n1}x_n$

Note: Rectangles for observed variables, circles for latent variables and error terms and hexagons for composite indices.

Table C2: Factor loadings and unique variances of cognition (Wave 4)

	Factor 1	Factor 2	Uniqueness
Reading self-assessed		0.92	0.13
Writing self-assessed		0.91	0.13
Memory test		0.53	0.64
Orientation	0.39		0.79
Numeracy skills	0.59	0.31	0.55
Immediate recall	0.85		0.25
Delayed recall	0.83		0.28
Verbal fluency	0.67		0.47

Objective measurements of cognition reflect a single Factor (1).

Note: The reliability indices for cognition indicate to drop the orientation variable (C7)

Table C3: Factor loadings and unique variances of physical health (Wave 4)

Difficulties in:		Factor 1	Uniqueness
Walking 100 mt	Item1	0.70	0.51
Sitting 2 hours	Item2	0.53	0.72
Getting up from chair	Item3	0.68	0.54
Climbing several flights of stairs	Item4	0.70	0.51
Climbing one flight of stairs	Item5	0.70	0.51
Kneeling, crouching	Item6	0.68	0.54
Extend arms above shoulder	Item7	0.60	0.64
Pulling or pushing large objects	Item8	0.73	0.47
Lifting or carrying weights over 5 kilos	Item9	0.72	0.48
Picking up a coin from a table	Item10	0.46	0.79

Note: All items load on a single Factor

Table C4: Factor loadings and unique variances of quality of life (Wave 4)

		Factor 1	Factor 2	Factor 3	Uniqueness
Control	Age prevents doing		0.77		0.39
	Out of control		0.76		0.36
	Left out		0.64		0.46
Autonomy	Do things	0.46	0.34		0.65
	Family responsibility			0.81	0.32
	Money shortage			0.54	0.61
Self-realisation	Energy	0.63	0.48		0.35
	Future opportunities	0.72	0.33		0.37
	Good future	0.73	0.33		0.36
Pleasure	Look forward	0.67			0.53
	Life meaning	0.76			0.40
	Happiness	0.67			0.49

Note: We include only items related to self-realisation and pleasure, Factor 1 (Sexton et al., 2013).

Table C5: Factor loadings and unique variances of "hedonic" QoL (Self-realisation and Pleasure items in Casp-12) (Wave 4)

		Factor 1	Uniqueness
Self-realisation	Energy	0.77	0.41
	Future opportunities	0.81	0.35
	Good future	0.81	0.34
Pleasure	Look forward	0.61	0.63
	Life meaning	0.76	0.43
	Happiness	0.63	0.60

Note: These items load on one factor

Table C6: Reliability indices for latent variables in Wave 4

Indices	Cut-off value	Cognition	Physical Status	Quality of life
Cronbach's alpha	$\alpha \geq .8$	0.80	0.85	0.82
Composite reliability	$\rho \geq .6$	0.65	0.86	0.83

Table C7: Cronbach's alpha for Cognition

	Obs	Alpha
Orientation	29172	0.800
Sub_numeracy	29527	0.735
Immediate_recall	28674	0.678
Delayed_recall	28661	0.696
Verbal_fluency	28501	0.729
Test scale		0.772

Dropping orientation, Cronbach's alpha is 0.8

Table C8: Factor loadings and unique variances of cognition (Wave 5 & Wave 6)

	Wave 5			Wave 6		
	Factor 1	Factor 2	Uniqueness	Factor 1	Factor 2	Uniqueness
Reading self-assessed		0.91	0.14		0.90	0.15
Writing self-assessed		0.91	0.13		0.91	0.14
Memory test		0.53	0.64		0.62	0.60
Orientation	0.40		0.78	0.49		0.73
Numeracy skills	0.56	0.35	0.56	0.55	0.33	0.58
Immediate recall	0.85		0.24	0.84		0.25
Delayed recall	0.84		0.28	0.82		0.29
Verbal fluency	0.66	0.31	0.47	0.64		0.52

Note: Only respondents present for the first time in Wave 5 or Wave 6 are here considered.

Table C9: Factor loadings and unique variances of physical health (Wave 5 & Wave 6)

		Wave 5		Wave 6	
Difficulties in:		Factor 1	Uniqueness	Factor 1	Uniqueness
Walking 100 mt	Item1	0.71	0.49	0.71	0.49
Sitting 2 hours	Item2	0.54	0.71	0.57	0.67
Getting up from chair	Item3	0.69	0.52	0.68	0.54
Climbing several flights of stairs	Item4	0.72	0.48	0.68	0.54
Climbing one flight of stairs	Item5	0.75	0.44	0.70	0.52
Kneeling, crouching	Item6	0.68	0.54	0.69	0.52
Extend arms above shoulder	Item7	0.62	0.61	0.60	0.64
Pulling or pushing large objects	Item8	0.73	0.46	0.71	0.50
Lifting or carrying weights over 5 kilos	Item9	0.73	0.47	0.70	0.50
Picking up a coin from a table	Item10	0.44	0.80	0.44	0.80

Note: Only respondents present for the first time in Wave 5 or Wave 6 are here considered.

Table C10: Factor loadings and unique variances of "hedonic" QoL (Self-realisation and Pleasure items in Casp-12) (Wave 5 & Wave 6)

		Wave 5		Wave 6	
		Factor 1	Uniqueness	Factor 1	Uniqueness
Self-realisation	Energy	0.76	0.42	0.77	0.41
	Future opportunities	0.80	0.36	0.81	0.35
	Good future	0.81	0.34	0.82	0.33
Pleasure	Look forward	0.59	0.65	0.70	0.51
	Life meaning	0.75	0.43	0.79	0.37
	Happiness	0.62	0.61	0.60	0.64

Note: Here, respondents present for the first time in Wave 5 or Wave 6 are considered for the respective analyses. In this way, the sample is completely different from Wave 4.

Table C11: Reliability indices for latent variables in Wave 5 & Wave 6

		Wave 5			Wave 6		
Indices	Cut-off value	Cog	Phy Status	QoL	Cog	Phy Status	QoL
Cronbach's alpha	$\alpha \geq .8$	0.80	0.86	0.82	0.80	0.85	0.85
Composite reliability	$\rho \geq .6$	0.64	0.87	0.83	0.61	0.86	0.85

Note: Only respondents present for the first time in Wave 5 or Wave 6 are considered for the respective analyses. In this way, the sample is completely different from Wave 4.

Table C12: Weighted score of cognition via principal component analysis (PCA)

	Comp 1
Numeracy skills	0.40
Immediate recall	0.56
Delayed recall	0.55
Verbal fluency	0.47

Note: Eigenvalue is 2.28

Table C13: Weighted score of physical health via principal component analysis (PCA)

Difficulties in:		Comp 1
Walking 100 mt	Item1	0.34
Sitting 2 hours	Item2	0.23
Getting up from chair	Item3	0.32
Climbing several flights of stairs	Item4	0.36
Climbing one flight of stairs	Item5	0.34
Kneeling, crouching	Item6	0.34
Extend arms above shoulder	Item7	0.27
Pulling or pushing large objects	Item8	0.36
Lifting or carrying weights over 5 kilos	Item9	0.36
Picking up a coin from a table	Item10	0.20

Note: Eigenvalue is 3.25

Table C14: Weighted score of Quality of Life (hedonic) via principal component analysis (PCA)

		Factor 1
Self-realisation	Energy	0.42
	Future opportunities	0.46
	Good future	0.46
Pleasure	Look forward	0.31
	Life meaning	0.41
	Happiness	0.37

Note: Eigenvalue is 2.84

Table C15: Results of Structural Equation Modelling (latent) and linear regressions (manifest variables) of all model specifications

	Wave 4				Wave 5				Wave 6			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Cognition (Latent)	0.209***	0.170***	0.147***	0.114***	0.190***	0.126***	0.099***	0.088***	0.216***	0.150***	0.108***	0.111***
	(0.011)	(0.017)	(0.020)	(0.020)	(0.010)	(0.015)	(0.018)	(0.018)	(0.010)	(0.014)	(0.016)	(0.017)
Cognition (Add Index)	0.199***	0.152***	0.125***	0.087***	0.181***	0.120***	0.098***	0.081***	0.191***	0.122***	0.082***	0.081***
	(0.009)	(0.013)	(0.015)	(0.015)	(0.008)	(0.011)	(0.014)	(0.014)	(0.008)	(0.011)	(0.013)	(0.013)
Cognition (PCA Index)	0.204***	0.160***	0.136***	0.102***	0.192***	0.131***	0.110***	0.091***	0.202***	0.135***	0.098	0.097***
	(0.009)	(0.013)	(0.016)	(0.015)	(0.008)	(0.011)	(0.014)	(0.014)	(0.008)	(0.011)	(0.013)	(0.013)
Physical (Latent)	0.302***	0.274***	0.264***	0.253***	0.320***	0.290***	0.268***	0.264***	0.306***	0.279***	0.258***	0.243***
	(0.012)	(0.015)	(0.018)	(0.017)	(0.010)	(0.013)	(0.016)	(0.015)	(0.010)	(0.013)	(0.015)	(0.015)
Physical (Add Index)	0.213***	0.190***	0.191***	0.185***	0.231***	0.205***	0.189***	0.190***	0.235***	0.209***	0.197***	0.186***
	(0.009)	(0.012)	(0.014)	(0.013)	(0.008)	(0.010)	(0.012)	(0.012)	(0.008)	(0.010)	(0.011)	(0.011)
Physical (PCA Index)	0.215***	0.191***	0.190***	0.184***	0.233***	0.205***	0.190***	0.191***	0.229***	0.202***	0.190	0.181***
	(0.009)	(0.012)	(0.014)	(0.013)	(0.008)	(0.010)	(0.012)	(0.012)	(0.008)	(0.010)	(0.011)	(0.011)
Error term (SEM)	0.838	0.813	0.810	0.699	0.832	0.789	0.743	0.684	0.828	0.780	0.731	0.663
Error term (Add Index)	0.901	0.876	0.871	0.768	0.897	0.861	0.824	0.757	0.890	0.848	0.805	0.752
Error term (PCA Index)	0.896	0.873	0.868	0.775	0.889	0.856	0.817	0.763	0.888	0.848	0.806	0.752
Controls:												
Demographic Ch.		✓	✓	✓		✓	✓	✓		✓	✓	✓
Economic Variables			✓	✓			✓	✓			✓	✓
Country fixed effect				✓				✓				✓

Note: Add Index denotes composite indices with equal weights, while PCA Index denotes weighted scores via principal component analysis

Table C16: Factor loadings and unique variances of cognition (random subgroups of Wave 4)

	Subgroup A			Subgroup B		
	Factor 1	Factor 2	Uniqueness	Factor 1	Factor 2	Uniqueness
Reading self-assessed		0.92	0.13		0.92	0.13
Writing self-assessed		0.91	0.13		0.92	0.13
Memory test		0.54	0.64		0.53	0.64
Orientation	0.40		0.78	0.38		0.80
Numeracy skills	0.60	0.32	0.54	0.59		0.56
Immediate recall	0.85		0.25	0.84		0.25
Delayed recall	0.83		0.28	0.83		0.28
Verbal fluency	0.67		0.46	0.67		0.48

Table C17: Factor loadings and unique variances of physical health (random subgroups of Wave 4)

		Subgroup A		Subgroup B	
Difficulties in:		Factor 1	Uniqueness	Factor 1	Uniqueness
Walking 100 mt	Item1	0.70	0.51	0.70	0.51
Sitting 2 hours	Item2	0.52	0.73	0.53	0.72
Getting up from chair	Item3	0.67	0.55	0.68	0.54
Climbing several flights of stairs	Item4	0.70	0.51	0.70	0.52
Climbing one flight of stairs	Item5	0.70	0.51	0.70	0.52
Kneeling, crouching	Item6	0.68	0.54	0.68	0.53
Extend arms above shoulder	Item7	0.61	0.63	0.59	0.65
Pulling or pushing large objects	Item8	0.73	0.47	0.72	0.48
Lifting or carrying weights over 5 kilos	Item9	0.72	0.48	0.72	0.48
Picking up a coin from a table	Item10	0.46	0.79	0.47	0.78

Table C18: Factor loadings and unique variances of "hedonic" QoL (Self-realisation and Pleasure items in Casp-12) (random subgroups of Wave 4)

		Subgroup A		Subgroup B	
		Factor 1	Uniqueness	Factor 1	Uniqueness
Self-realisation	Energy	0.77	0.41	0.76	0.42
	Future opportunities	0.81	0.35	0.80	0.36
	Good future	0.82	0.33	0.81	0.34
Pleasure	Look forward	0.61	0.62	0.61	0.63
	Life meaning	0.76	0.43	0.75	0.43
	Happiness	0.64	0.60	0.62	0.61

Table C19: Reliability indices for latent variables in random subgroups of Wave 4

		Wave 5			Wave 6		
Indices	Cut-off value	Cog	Phy Status	QoL	Cog	Phy Status	QoL
Cronbach's alpha	$\alpha \geq .8$	0.80	0.85	0.83	0.80	0.85	0.82
Composite reliability	$\rho \geq .6$	0.66	0.86	0.84	0.64	0.86	0.83

Table C20: Fit statistics full measurement model on random subgroups (Wave 4)

	Cut-off value	Subgroup A	Subgroup B
Chi-square	P>0.001	P<0.001	P<0.001
RMSEA	≤ 0.08	0.051	0.049
CFI	≥ 0.9	0.933	0.936
SRMR	≤ 0.1	0.043	0.041

Note: Root Mean Square Error of Approximation (RMSEA);
Comparative Fit Index (CFI);
Standardized Root Mean Squared Residual (SRMR)

Table C21: Standardized results of SEM analysis using FIML estimation

	WAVE 4				WAVE 5				WAVE 6			
	Structural equation models				Structural equation models				Structural equation models			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Cognition	0.210***	0.165***	0.146***	0.117***	0.196***	0.140***	0.104***	0.086***	0.217***	0.149***	0.107***	0.107***
	(0.012)	(0.014)	(0.014)	(0.014)	(0.010)	(0.011)	(0.011)	(0.012)	(0.010)	(0.011)	(0.011)	(0.012)
Physical	0.304***	0.282***	0.275***	0.257***	0.330***	0.298***	0.280***	0.283***	0.309***	0.280***	0.263***	0.249***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Female		0.024**	0.026**	0.036***		0.049***	0.062***	0.061***		0.047***	0.056***	0.048***
		(0.011)	(0.011)	(0.010)		(0.008)	(0.008)	(0.008)		(0.009)	(0.009)	(0.008)
Education		0.020*	0.012	0.037***		0.026***	0.002	0.021**		0.028***	0.001	0.011
		(0.011)	(0.011)	(0.011)		(0.009)	(0.009)	(0.009)		(0.009)	(0.009)	(0.010)
Age		-0.011	-0.010	-0.041***		-0.011	-0.005	-0.034***		-0.031***	-0.031***	-0.046***
		(0.012)	(0.012)	(0.012)		(0.010)	(0.010)	(0.009)		(0.010)	(0.010)	(0.010)
Married		0.060***	0.051***	0.038***		0.078***	0.064***	0.056***		0.072***	0.055***	0.058***
		(0.012)	(0.012)	(0.012)		(0.010)	(0.010)	(0.010)		(0.010)	(0.010)	(0.010)
Household size		-0.040***	-0.030**	0.013		-0.028***	0.028***	0.032***		-0.029***	0.037***	0.015
		(0.012)	(0.012)	(0.012)		(0.010)	(0.010)	(0.010)		(0.011)	(0.011)	(0.011)
Urban		-0.046***	-0.046***	-0.008		-0.009	-0.014	0.013		-0.023**	-0.029***	0.003
		(0.012)	(0.012)	(0.012)		(0.010)	(0.010)	(0.010)		(0.010)	(0.010)	(0.010)
Social activities		0.159***	0.145***	0.078***		0.186***	0.146***	0.106***		0.185***	0.136***	0.106***
		(0.010)	(0.010)	(0.010)		(0.008)	(0.008)	(0.008)		(0.008)	(0.009)	(0.008)
Income			0.026**	0.037***			0.183***	0.080***			0.199***	0.073***
			(0.012)	(0.014)			(0.010)	(0.012)			(0.011)	(0.012)
Assets			0.091***	0.036***			0.061***	0.036***			0.067***	0.031***
			(0.010)	(0.011)			(0.009)	(0.009)			(0.009)	(0.010)
Country FE				✓				✓				✓
N(obs)	10,990	10,990	10,990	10,990	16,571	16,571	16,571	16,571	15,520	15,520	15,520	15,520
CD	0.962	0.963	0.963	0.968	0.970	0.971	0.972	0.973	0.966	0.967	0.969	0.971

Note: Robust standard error included in parenthesis
*** p$\leq 0.01, **p \leq 0.05, *p \leq 0.1$

Appendix B

Joint Authorship

We certify that Chiara Costi was involved in the conception and design of the work in Chapter 2 titled: "The impact of parental education and prenatal smoking on infant health: an intergenerational approach". She has contributed significantly to the data analysis and its interpretation. She is the main contributor to the drafting the chapter, the critical revision of the chapter, and the final approval of the article version set to be sent for publication.

Giuseppe Migali

Eugenio Zucchelli

Joint Authorship

We certify that Chiara Costi was involved in the conception and design of the work in Chapter 3 titled: "Does caring for others affect our mental health? Evidence from the COVID-19 pandemic". She has contributed significantly to the data analysis and its interpretation. She is the main contributor to the drafting the chapter, the critical revision of the chapter, and the final approval of the article version published in Social Science and Medicine, Vol. 321, March 2023, 115721.

Bruce Hollingsworth

Vincent O'Sullivan

Eugenio Zucchelli

Bibliography

Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. (2017). When should you adjust standard errors for clustering? Technical report, National Bureau of Economic Research.

Acock, A. C. et al. (2013). Discovering structural equation modeling using stata. *Stata Press Books*.

Adelman, R. D., Tmanova, L. L., Delgado, D., Dion, S., and Lachs, M. S. (2014). Caregiver burden: a clinical review. *Jama*, 311(10):1052–1060.

Allerhand, M., Gale, C. R., and Deary, I. J. (2014). The dynamic relationship between cognitive function and positive well-being in older people: a prospective study using the english longitudinal study of aging. *Psychology and aging*, 29(2):306.

Allison, P. D. (2003). Missing data techniques for structural equation modeling. *Journal of abnormal psychology*, 112(4):545.

Almond, D. and Currie, J. (2011). Killing me softly: The fetal origins hypothesis. *Journal of economic perspectives*, 25(3):153–172.

Almond, D., Currie, J., and Duque, V. (2018). Childhood circumstances and adult outcomes: Act ii. *Journal of Economic Literature*, 56(4):1360–1446.

Althubaiti, A. (2016). Information bias in health research: definition, pitfalls, and adjustment methods. *Journal of multidisciplinary healthcare*, pages 211–217.

Altman, M., Bonamy, A.-K. E., Wikström, A.-K., and Cnattingius, S. (2012). Cause-specific infant mortality in a population-based swedish study of term and post-term births: the contribution of gestational age and birth weight. *BMJ open*, 2(4):e001152.

Anaya, L., Howley, P., Waqas, M., and Yalonetzky, G. (2021). Locked down in distress: a causal estimation of the mental-health fallout from the covid-19 pandemic. *Available at SSRN 3873973*.

Ashton, K., Schröder-Bäck, P., Clemens, T., Dyakova, M., Stielke, A., and Bellis, M. A. (2020). The social value of investing in public health across the life course: a systematic scoping review. *BMC public health*, 20(1):1–18.

Azevedo, L. V. d. S., Calandri, I. L., Slachevsky, A., Graviotto, H. G., Vieira, M. C. S., Andrade, C. B. d., Rossetti, A. P., Generoso, A. B., Carmona, K. C., Pinto, L. A. C., et al. (2021). Impact of social isolation on people with dementia and their family caregivers. *Journal of Alzheimer's Disease*, 81(2):607–617.

Babu, G. R., Sudhir, P. M., Mahapatra, T., Das, A., Rathnaiah, M., Anand, I., and Detels, R. (2016). Association of quality of life and job stress in occupational workforce of india: Findings from a cross-sectional study on software professionals. *Indian Journal of Occupational and Environmental Medicine*, 20(2):109.

Bailey, C., Guo, P., MacArtney, J., Finucane, A., Swan, S., Meade, R., and Wagstaff, E. (2022). The experiences of informal carers during the covid-19 pandemic: A qualitative systematic review. *International Journal of Environmental Research and Public Health*, 19(20):13455.

Baker, A. C., Larcker, D. F., and Wang, C. C. (2022). How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics*, 144(2):370–395.

Banderali, G., Martelli, A., Landi, M., Moretti, F., Betti, F., Radaelli, G., Lassandro, C., and Verduci, E. (2015). Short and long term health effects of parental tobacco smoking during pregnancy and lactation: a descriptive review. *Journal of translational medicine*, 13(1):1–7.

Banks, J. and Xu, X. (2020). The mental health effects of the first two months of lockdown during the covid-19 pandemic in the uk. *Fiscal Studies*, 41(3):685–708.

Bayati, T., Dehghan, A., Bonyadi, F., and Bazrafkan, L. (2018). Investigating the effect of education on health literacy and its relation to health-promoting behaviors in health center. *Journal of education and health promotion*, 7.

Beach, S. R., Schulz, R., Donovan, H., and Rosland, A.-M. (2021). Family caregiving during the covid-19 pandemic. *The gerontologist*, 61(5):650–660.

Becker, G. (2017). *Economic theory*. Routledge.

Behrman, J. R. and Rosenzweig, M. R. (2004). Returns to birthweight. *Review of Economics and statistics*, 86(2):586–601.

Berg, G. D. (2011). An application of kernel-based versus one-to-one propensity score matching for a nonexperimental causal study: example from a disease management program evaluation. *Applied Economics Letters*, 18(5):439–447.

Bergmann, M. and Wagner, M. (2021). The impact of covid-19 on informal caregiving and care receiving across europe during the first phase of the pandemic. *Frontiers in Public Health*, 9:673874.

Berkman, L. F., Glass, T., Brissette, I., and Seeman, T. E. (2000). From social integration to health: Durkheim in the new millennium. *Social science & medicine*, 51(6):843–857.

Berkman, N. D., Davis, T. C., and McCormack, L. (2010). Health literacy: what is it? *Journal of health communication*, 15(S2):9–19.

Bertuzzi, V., Semonella, M., Bruno, D., Manna, C., Edbrook-Childs, J., Giusti, E. M., Castelnuovo, G., and Pietrabissa, G. (2021). Psychological support interventions for healthcare providers and informal caregivers during the covid-19 pandemic: a systematic review of the literature. *International Journal of Environmental Research and Public Health*, 18(13):6939.

Betz, C. L., Ruccione, K., Meeske, K., Smith, K., and Chang, N. (2008). Health literacy: a pediatric nursing concern. *Pediatric nursing*, 34(3).

Black, S. E., Devereux, P. J., and 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.

Blanchflower, D. G. and Oswald, A. J. (2008). Is well-being u-shaped over the life cycle? *Social science & medicine*, 66(8):1733–1749.

Boggatz, T. (2016). Quality of life in old age–a concept analysis. *International journal of older people nursing*, 11(1):55–69.

Bohacek, J. and Mansuy, I. M. (2013). Epigenetic inheritance of disease and disease risk. *Neuropsychopharmacology*, 38(1):220–236.

Bollen, K. and Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological bulletin*, 110(2):305.

Bollen, K. A. (1989). *Structural equations with latent variables*, volume 210. John Wiley & Sons.

Bollen, K. A. (2002). Latent variables in psychology and the social sciences. *Annual review of psychology*, 53(1):605–634.

Bollen, K. A. and Bauldry, S. (2011). Three cs in measurement models: causal indicators, composite indicators, and covariates. *Psychological methods*, 16(3):265.

Bom, J., Bakx, P., Schut, F., and Van Doorslaer, E. (2019). Health effects of caring for and about parents and spouses. *The Journal of the Economics of Ageing*, 14:100196.

Bom, J. and Stöckel, J. (2021). Is the grass greener on the other side? the health impact of providing informal care in the uk and the netherlands. *Social science & medicine*, 269:113562.

Borelli, W. V., Augustin, M. C., De Oliveira, P. B. F., Reggiani, L. C., Bandeira-de Mello, R. G., Schumacher-Schuh, A. F., Chaves, M. L. F., and Castilhos, R. M. (2021). Neuropsychiatric symptoms in patients with dementia associated with increased psychological distress in caregivers during the covid-19 pandemic. *Journal of Alzheimer's Disease*, 80(4):1705–1712.

Börsch-Supan, A., Hank, K., and Jürges, H. (2005). A new comprehensive and international view on ageing: introducing the âsurvey of health, ageing and retirement in europeâ. *European Journal of Ageing*, 2(4):245–253.

Boyle, P. A., Buchman, A. S., Wilson, R. S., Yu, L., Schneider, J. A., and Bennett, D. A. (2012). Effect of purpose in life on the relation between alzheimer disease pathologic changes on cognitive function in advanced age. *Archives of general psychiatry*, 69(5):499–504.

Brenna, E. (2021). Should i care for my mum or for my kid? sandwich generation and depression burden in italy. *Health Policy*, 125(3):415–423.

Brodeur, A., Clark, A. E., Fleche, S., and Powdthavee, N. (2021). Covid-19, lockdowns and well-being: Evidence from google trends. *Journal of public economics*, 193:104346.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

Carers UK (2015). Valuing carers 2015. *Carers UK, London.*

Carers UK (2020). Carers week research report: The rise in the number of unpaid carers during the coronavirus (covid-19) outbreak. *Carers UK, London.*

Case, A., Fertig, A., and Paxson, C. (2005). The lasting impact of childhood health and circumstance. *Journal of health economics*, 24(2):365–389.

Case, A., Lubotsky, D., and Paxson, C. (2002). Economic status and health in childhood: The origins of the gradient. *American Economic Review*, 92(5):1308–1334.

Chandola, T., Kumari, M., Booker, C. L., and Benzeval, M. (2022). The mental health impact of covid-19 and lockdown-related stressors among adults in the uk. *Psychological medicine*, 52(14):2997–3006.

Chei, C.-L., Lee, J. M.-L., Ma, S., and Malhotra, R. (2018). Happy older people live longer. *Age and Ageing*, 47(6):860–866.

Cheng, T. C., Powdthavee, N., and Oswald, A. J. (2017). Longitudinal evidence for a midlife nadir in human well-being: Results from four data sets. *The Economic Journal*, 127(599):126–142.

Chevalier, A. and O'Sullivan, V. (2007). Mother's education and birth weight.

Cho, J., Martin, P., Margrett, J., MacDonald, M., and Poon, L. W. (2011). Research article the relationship between physical health and psychological well-being among oldest-old adults.

Clark, A. E. (2018). Four decades of the economics of happiness: Where next? *Review of Income and Wealth*, 64(2):245–269.

Clark, A. E. and Georgellis, Y. (2013). Back to baseline in britain: adaptation in the british household panel survey. *Economica*, 80(319):496–512.

Collins, A. L., Goldman, N., and Rodríguez, G. (2008). Is positive well-being protective of mobility limitations among older adults? *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 63(6):P321–P327.

Comijs, H. C., Dik, M. G., Aartsen, M. J., Deeg, D. J., and Jonker, C. (2005). The impact of change in cognitive functioning and cognitive decline on disability, well-being, and the use of healthcare services in older persons. *Dementia and geriatric cognitive disorders*, 19(5-6):316–323.

Conti, G. and Heckman, J. J. (2012). The economics of child well-being. Technical report, National Bureau of Economic Research.

Courtin, E., Jemiai, N., and Mossialos, E. (2014). Mapping support policies for informal carers across the european union. *Health policy*, 118(1):84–94.

Cullati, S., Bochatay, N., Rossier, C., Guessous, I., Burton-Jeangros, C., and Courvoisier, D. S. (2020). Does the single-item self-rated health measure the same thing across different wordings? construct validity study. *Quality of Life Research*, 29:2593–2604.

Cunningham, S. (2021). *Causal inference: The mixtape*. Yale university press.

Currie, J. (2009). Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development. *JoUrnal of economIc lIteratUre*, 47(1):87–122.

Currie, J. and Hyson, R. (1999). Is the impact of health shocks cushioned by socioeconomic status? the case of low birthweight. *American economic review*, 89(2):245–250.

Currie, J. and Lin, W. (2007). Chipping away at health: more on the relationship between income and child health. *Health Affairs*, 26(2):331–344.

Currie, J. and Moretti, E. (2003). Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *The Quarterly journal of economics*, 118(4):1495–1532.

Currie, J. and Moretti, E. (2007). Biology as destiny? short-and long-run determinants of intergenerational transmission of birth weight. *Journal of Labor economics*, 25(2):231–264.

Currie, J. and Stabile, M. (2006). Child mental health and human capital accumulation: the case of adhd. *Journal of health economics*, 25(6):1094–1118.

Daly, M., Sutin, A. R., and Robinson, E. (2022). Longitudinal changes in mental health and the covid-19 pandemic: Evidence from the uk household longitudinal study. *Psychological medicine*, 52(13):2549–2558.

De Chaisemartin, C. and d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–2996.

De Zwart, P., Bakx, P., and Van Doorslaer, E. (2017). Will you still need me, will you still feed me when i'm 64? the health impact of caregiving to one's spouse. *Health economics*, 26:127–138.

Deaton, A. (2008). Income, health, and well-being around the world: Evidence from the gallup world poll. *Journal of Economic perspectives*, 22(2):53–72.

del Rocío Santana-Berlanga, N., Porcel-Gálvez, A. M., Botello-Hermosa, A., and Barrientos-Trigo, S. (2020). Instruments to measure quality of life in institutionalised older adults: systematic review. *Geriatric Nursing*, 41(4):445–462.

Dhiman, S., Sahu, P. K., Reed, W. R., Ganesh, G. S., Goyal, R. K., and Jain, S. (2020). Impact of covid-19 outbreak on mental health and perceived strain among caregivers tending children with special needs. *Research in Developmental Disabilities*, 107:103790.

Doiron, D., Fiebig, D. G., Johar, M., and Suziedelyte, A. (2015). Does self-assessed health measure health? *Applied Economics*, 47(2):180–194.

Doyle, O., Harmon, C. P., and Walker, I. (2005). The impact of parental income and education on the health of their children.

Dunn, L. M. and Markwardt, F. C. (1970). *Peabody individual achievement test.* American Guidance Service, Incorporated.

Easterlin, R. A. (1974). Does economic growth improve the human lot? some empirical evidence. In *Nations and households in economic growth*, pages 89–125. Elsevier.

Easterlin, R. A. (2003). Explaining happiness. *Proceedings of the National Academy of Sciences*, 100(19):11176–11183.

Easterlin, R. A. (2006). Life cycle happiness and its sources: Intersections of psychology, economics, and demography. *Journal of economic psychology*, 27(4):463–482.

Etheridge, B. and Spantig, L. (2020). The gender gap in mental well-being during the covid-19 outbreak: Evidence from the uk. Technical report, ISER Working paper series.

European Commission (2018). *The 2018 ageing report : economic budgetary projections for the 28 EU Member States (2016-2070)*. Publications Office, European Commission and Directorate-General for Economic and Financial Affairs.

Evans, W. N. and Ringel, J. S. (1999). Can higher cigarette taxes improve birth outcomes? *Journal of public Economics*, 72(1):135–154.

Fan, X., Thompson, B., and Wang, L. (1999). Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. *Structural equation modeling: a multidisciplinary journal*, 6(1):56–83.

Figlio, D., Guryan, J., Karbownik, K., and Roth, J. (2014). The effects of poor neonatal health on children's cognitive development. *American Economic Review*, 104(12):3921–3955.

Foroughan, M., Nejati, V., Shati, M., et al. (2018). Association between age associated cognitive decline and health related quality of life among iranian older individuals. *Electronic Physician*, 10(4):6663.

Fortin, M., Lapointe, L., Hudon, C., Vanasse, A., Ntetu, A. L., and Maltais, D. (2004). Multimorbidity and quality of life in primary care: a systematic review. *Health and Quality of life Outcomes*, 2(1):1–12.

Freedman, V. A., Carr, D., Cornman, J. C., and Lucas, R. E. (2017). Aging, mobility impairments and subjective wellbeing. *Disability and Health Journal*, 10(4):525–531.

Freeman, R. B. (1977). Job satisfaction as an economic variable. Technical report, National Bureau of Economic Research.

Frey, B. S. and Stutzer, A. (2002). What can economists learn from happiness research? *Journal of Economic literature*, 40(2):402–435.

Friedman, D. B. and Hoffman-Goetz, L. (2008). Literacy and health literacy as defined in cancer education research: a systematic review. *Health Education Journal*, 67(4):285–304.

Gale, C. R., Allerhand, M., Sayer, A. A., Cooper, C., and Deary, I. J. (2014). The dynamic relationship between cognitive function and walking speed: the english longitudinal study of ageing. *Age*, 36:1–11.

Gallagher, S. and Wetherell, M. A. (2020). Risk of depression in family caregivers: unintended consequence of covid-19. *BJPsych open*, 6(6):e119.

Gibson, G. (2012). Rare and common variants: twenty arguments. *Nature Reviews Genetics*, 13(2):135–145.

Giebel, C., Cannon, J., Hanna, K., Butchard, S., Eley, R., Gaughan, A., Komuravelli, A., Shenton, J., Callaghan, S., Tetlow, H., et al. (2021). Impact of covid-19 related social support service closures on people with dementia and unpaid carers: a qualitative study. *Aging & mental health*, 25(7):1281–1288.

Giuntella, O., La Mattina, G., and Quintana-Domeque, C. (2022). Intergenerational transmission of health at birth: Fathers matter too! Technical report, National Bureau of Economic Research.

Godah, M. W., Beydoun, Z., Abdul-Khalek, R. A., Safieddine, B., Khamis, A. M., and Abdulrahim, S. (2021). Maternal education and low birth weight in low-and middle-income countries: systematic review and meta-analysis. *Maternal and Child Health Journal*, 25(8):1305–1315.

Goddard, M. and Jacobs, R. (2009). 3.4 using composite indicators to measure performance in health care. *Performance measurement for health system improvement*, page 339.

Goldberg, D. P., Gater, R., Sartorius, N., Ustun, T. B., Piccinelli, M., Gureje, O., and Rutter, C. (1997). The validity of two versions of the ghq in the who study of mental illness in general health care. *Psychological medicine*, 27(1):191–197.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277.

Graham, C., Higuera, L., and Lora, E. (2011). Which health conditions cause the most unhappiness? *Health economics*, 20(12):1431–1447.

Greaney, M. L., Kunicki, Z. J., Drohan, M. M., Nash, C. C., and Cohen, S. A. (2022). Sleep quality among informal caregivers during the covid-19 pandemic: a cross-sectional study. *Gerontology and Geriatric Medicine*, 8:23337214211057387.

Green, C. P., Heywood, J. S., Kler, P., and Leeves, G. (2018). Paradox lost: The disappearing female job satisfaction premium. *British Journal of Industrial Relations*, 56(3):484–502.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political economy*, 80(2):223–255.

Guo, G. and Harris, K. M. (2000). The mechanisms mediating the effects of poverty on childrenâs intellectual development. *Demography*, 37(4):431–447.

Guven, C. and Saloumidis, R. (2009). Why is the world getting older? the influence of happiness on mortality. *The Influence of Happiness on Mortality (June 2009). SOEPpaper*, (198).

Hamermesh, D. S. (1977). Economic aspects of job satisfaction. *Essays in labor market analysis*, pages 53–72.

Harrington, K. F., Zhang, B., Magruder, T., Bailey, W. C., and Gerald, L. B. (2015). The impact of parent's health literacy on pediatric asthma outcomes. *Pediatric allergy, immunology, and pulmonology*, 28(1):20–26.

Harris, K. M. (2013). The add health study: Design and accomplishments.

Heckman, J. J. (2008). Econometric causality. *International statistical review*, 76(1):1–27.

Heckman, J. J. and Vytlacil, E. J. (2007). Econometric evaluation of social programs, part ii: Using the marginal treatment effect to organize alternative econometric estimators to evaluate social programs, and to forecast their effects in new environments. *Handbook of econometrics*, 6:4875–5143.

Hedley, A. A., Ogden, C. L., Johnson, C. L., Carroll, M. D., Curtin, L. R., and Flegal, K. M. (2004). Prevalence of overweight and obesity among us children, adolescents, and adults, 1999-2002. *Jama*, 291(23):2847–2850.

Hill, N. L., McDermott, C., Mogle, J., Munoz, E., DePasquale, N., Wion, R., and Whitaker, E. (2017). Subjective cognitive impairment and quality of life: a systematic review. *International Psychogeriatrics*, 29(12):1965–1977.

Hines, L. A., Spry, E. A., Moreno-Betancur, M., Mohamad Husin, H., Becker, D., Middleton, M., Craig, J. M., Doyle, L. W., Olsson, C. A., and Patton, G. (2021). Cannabis and tobacco use prior to pregnancy and subsequent offspring birth outcomes: a 20-year intergenerational prospective cohort study. *Scientific Reports*, 11(1):16826.

Hirschhorn, J. N. and Daly, M. J. (2005). Genome-wide association studies for common diseases and complex traits. *Nature reviews genetics*, 6(2):95–108.

Hirve, S., Oud, J. H., Sambhudas, S., Juvekar, S., Blomstedt, Y., Tollman, S., Wall, S., and Ng, N. (2014). Unpacking self-rated health and quality of life in older adults and elderly in india: a structural equation modelling approach. *Social indicators research*, 117:105–119.

Holmlund, H., Lindahl, M., and Plug, E. (2011). The causal effect of parents' schooling on children's schooling: A comparison of estimation methods. *Journal of economic literature*, 49(3):615–651.

Hoover, K. D. (2006). Causality in economics and econometrics.

Hoyle, R. H. (2012). *Handbook of structural equation modeling*. Guilford press.

Hyde, M., Wiggins, R. D., Higgs, P., and Blane, D. B. (2003). A measure of quality of life in early old age: the theory, development and properties of a needs satisfaction model (casp-19). *Aging & mental health*, 7(3):186–194.

Irani, E., Niyomyart, A., and Hickman Jr, R. L. (2021). Family caregiversâ experiences and changes in caregiving tasks during the covid-19 pandemic. *Clinical nursing research*, 30(7):1088–1097.

Jacobs, R. and Goddard, M. (2007). How do performance indicators add up? an examination of composite indicators in public services. *Public Money and Management*, 27(2):103–110.

Jaddoe, V. W., Troe, E.-J. W., Hofman, A., Mackenbach, J. P., Moll, H. A., Steegers, E. A., and Witteman, J. C. (2008). Active and passive maternal smoking during pregnancy and the risks of low birthweight and preterm birth: the generation r study. *Paediatric and perinatal epidemiology*, 22(2):162–171.

Jami, E. S., Hammerschlag, A. R., Bartels, M., and Middeldorp, C. M. (2021). Parental characteristics and offspring mental health and related outcomes: A systematic review of genetically informative literature. *Translational Psychiatry*, 11(1):197.

Jetten, J., Haslam, C., Pugliese, C., Tonks, J., and Haslam, S. A. (2010). Declining autobiographical memory and the loss of identity: Effects on well-being. *Journal of clinical and experimental neuropsychology*, 32(4):408–416.

Kahneman, D. and Deaton, A. (2010). High income improves evaluation of life but not emotional well-being. *Proceedings of the national academy of sciences*, 107(38):16489–16493.

Kane, J. B., Harris, K. M., and Siega-Riz, A. M. (2018). Intergenerational pathways linking maternal early life adversity to offspring birthweight. *Social Science & Medicine*, 207:89–96.

Keane, M. and Neal, T. (2023). Instrument strength in iv estimation and inference: A guide to theory and practice. *Journal of Econometrics*.

Kemptner, D. and Marcus, J. (2013). Spillover effects of maternal education on childâs health and health behavior. *Review of Economics of the Household*, 11:29–52.

Kennedy-Moulton, K., Miller, S., Persson, P., Rossin-Slater, M., Wherry, L., and Aldana, G. (2022). Maternal and infant health inequality: New evidence from linked administrative data. Technical report, National Bureau of Economic Research.

Kent, E. E., Ornstein, K. A., and Dionne-Odom, J. N. (2020). The family caregiving crisis meets an actual pandemic. *Journal of pain and symptom management*, 60(1):e66–e69.

Keynes, J. M. (1936). *The general theory of employment, interest and money.* Macmillan and Co., Ltd., London.

Kline, R. B. (2015). *Principles and practice of structural equation modeling.* Guilford publications.

Kotwal, A. A., Kim, J., Waite, L., and Dale, W. (2016). Social function and cognitive status: Results from a us nationally representative survey of older adults. *Journal of general internal medicine*, 31:854–862.

Lacey, R. E., McMunn, A., and Webb, E. (2019). Informal caregiving patterns and trajectories of psychological distress in the uk household longitudinal study. *Psychological medicine*, 49(10):1652–1660.

Lawrence, E. M., Rogers, R. G., Zajacova, A., and Wadsworth, T. (2019). Marital happiness, marital status, health, and longevity. *Journal of Happiness Studies*, 20(5):1539–1561.

Layard, R., Clark, A. E., Cornaglia, F., Powdthavee, N., and Vernoit, J. (2014). What predicts a successful life? a life-course model of well-being. *The Economic Journal*, 124(580):F720–F738.

Le, K. and Nguyen, M. (2021). The psychological burden of the covid-19 pandemic severity. *Economics & Human Biology*, 41:100979.

Lee, D. S., McCrary, J., Moreira, M. J., and Porter, J. (2022). Valid t-ratio inference for iv. *American Economic Review*, 112(10):3260–90.

Lee, H. Y., Rhee, T. G., Kim, N. K., and Ahluwalia, J. S. (2015). Health literacy as a social determinant of health in asian american immigrants: findings from a population-based survey in california. *Journal of general internal medicine*, 30:1118–1124.

Lee, H. Y., Zhou, A. Q., Lee, R. M., and Dillon, A. L. (2020). Parentsâ functional health literacy is associated with childrenâs health outcomes: Implications for health practice, policy, and research. *Children and youth services review*, 110:104801.

Leggett, A. N., Carmichael, A., Leonard, N., Jackson, J., Kirch, M., Solway, E., Kullgren, J. T., Singer, D., Malani, P. N., and Gonzalez, R. (2021). Care challenges due to covid-19 and mental health among caregivers of us adults with a chronic or disabling condition. *Innovation in Aging*, 5(3):igab031.

León, L. P. d., Mangin, J. P. L., and Ballesteros, S. (2020). Psychosocial determinants of quality of life and active aging. a structural equation model. *International journal of environmental research and public health*, 17(17):6023.

Leuven, E. and Sianesi, B. (2003). Psmatch2: Stata module to perform full mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing, version 3.0. 0. *URL http://ideas. repec. org/c/boc/bocode/s432001. html*.

Li, Q., Zhang, H., Zhang, M., Li, T., Ma, W., An, C., Chen, Y., Liu, S., Kuang, W., Yu, X., et al. (2021). Mental health multimorbidity among caregivers of older adults during the covid-19 epidemic. *The American Journal of Geriatric Psychiatry*, 29(7):687–697.

Lien, D. S. and Evans, W. N. (2005). Estimating the impact of large cigarette tax hikes the case of maternal smoking and infant birth weight. *Journal of Human resources*, 40(2):373–392.

Lightfoot, E., Moone, R., Suleiman, K., Otis, J., Yun, H., Kutzler, C., and Turck, K. (2021). Concerns of family caregivers during covid-19: The concerns of caregivers and the surprising silver linings. *Journal of Gerontological Social Work*, 64(6):656–675.

Lindeboom, M., Llena-Nozal, A., and van Der Klaauw, B. (2009). Parental education and child health: Evidence from a schooling reform. *Journal of health Economics*, 28(1):109–131.

Linnet, K. M., Wisborg, K., Agerbo, E., Secher, N.-J., Thomsen, P. H., and Henriksen, T. B. (2006). Gestational age, birth weight, and the risk of hyperkinetic disorder. *Archives of disease in childhood*, 91(8):655–660.

Linsell, L., Malouf, R., Morris, J., Kurinczuk, J. J., and Marlow, N. (2015). Prognostic factors for poor cognitive development in children born very preterm or with very low birth weight: a systematic review. *JAMA pediatrics*, 169(12):1162–1172.

Llewellyn, D. J., Lang, I. A., Langa, K. M., and Huppert, F. A. (2008). Cognitive function and psychological well-being: findings from a population-based cohort. *Age and ageing*, 37(6):685–689.

López Ulloa, B. F., Møller, V., and Sousa-Poza, A. (2013). How does subjective well-being evolve with age? a literature review. *Journal of Population Ageing*, 6:227–246.

Lorenz-Dant, K. and Comas-Herrera, A. (2021). The impacts of covid-19 on unpaid carers of adults with long-term care needs and measures to address these impacts: A rapid review of evidence up to november 2020. *Journal of Long-Term Care*, 2021:124–153.

Mak, H. W., Bu, F., and Fancourt, D. (2021). Mental health and wellbeing among people with informal caring responsibilities across different time points during the covid-19 pandemic: a population-based propensity score matching analysis. *Perspectives in Public Health*, page 17579139221104973.

Marmot, M. and Wilkinson, R. (2005). *Social determinants of health*. Oup Oxford.

Marmot, M. G. (1997). Early life and adult disorder: research themes. *British medical bulletin*, 53(1):3–9.

Mas-Colell, A. (1985). *The theory of general economic equilibrium: A differentiable approach*. Number 9. Cambridge University Press.

Mataria, A., Giacaman, R., Stefanini, A., Naidoo, N., Kowal, P., and Chatterji, S. (2009). The quality of life of palestinians living in chronic conflict: assessment and determinants. *The European Journal of Health Economics*, 10:93–101.

Mathewson, K. J., Chow, C. H., Dobson, K. G., Pope, E. I., Schmidt, L. A., and Van Lieshout, R. J. (2017). Mental health of extremely low birth weight survivors: A systematic review and meta-analysis. *Psychological bulletin*, 143(4):347.

Maystadt, J.-F. and Migali, G. (2021). The transmission of health across 7 generations in china, 1789–1906. *Journal of Health Economics*, 79:102493.

McCrary, J. and Royer, H. (2011). The effect of female education on fertility and infant health: evidence from school entry policies using exact date of birth. *American economic review*, 101(1):158–195.

McNeish, D. and Wolf, M. G. (2020). Thinking twice about sum scores. *Behavior research methods*, 52:2287–2305.

Muldrew, D. H., Fee, A., and Coates, V. (2022). Impact of the covid-19 pandemic on family carers in the community: a scoping review. *Health & Social Care in the Community*, 30(4):1275–1285.

National Academies of Sciences, E. and Medicine (2002). Leveraging longitudinal data in developing countries: Report of a workshop.

Ng, K. Y. Y., Zhou, S., Tan, S. H., Ishak, N. D. B., Goh, Z. Z. S., Chua, Z. Y., Chia, J. M. X., Chew, E. L., Shwe, T., Mok, J. K. Y., et al. (2020). Understanding the psychological impact of covid-19 pandemic on patients with cancer, their caregivers, and health care workers in singapore. *JCO global oncology*, 6:1494–1509.

NHS Digital (2017). Health survey for england 2016 well-being and mental health. *Health and Social Care Information Centre*.

Niedzwiedz, C. L., Green, M. J., Benzeval, M., Campbell, D., Craig, P., Demou, E., Leyland, A., Pearce, A., Thomson, R., Whitley, E., et al. (2021). Mental health and health behaviours before and during the initial phase of the covid-19 lockdown: longitudinal analyses of the uk household longitudinal study. *J Epidemiol Community Health*, 75(3):224–231.

OECD (2013). *A Good Life in Old Age?*

Okely, J. A., Shaheen, S. O., Weiss, A., and Gale, C. R. (2017). Wellbeing and chronic lung disease incidence: The survey of health, ageing and retirement in europe. *PLoS One*, 12(7):e0181320.

Onwumere, J. (2020). Informal carers in severe mental health conditions: Issues raised by the united kingdom sars-cov-2 (covid-19) pandemic.

Onwumere, J., Creswell, C., Livingston, G., Shiers, D., Tchanturia, K., Charman, T., Russell, A., Treasure, J., Di Forti, M., Wildman, E., et al. (2021). Covid-19 and uk family carers: policy implications. *The Lancet Psychiatry*, 8(10):929–936.

Orimo, H., Ito, H., Suzuki, T., Araki, A., Hosoi, T., and Sawabe, M. (2006). Reviewing the definition of âelderlyâ. *Geriatrics & gerontology international*, 6(3):149–158.

Osborne, J. B. and Bailey, B. A. (2022). Does it matter when i quit? could i just cut down some? links between trimester-specific smoking amount, preterm birth, and low birth weight. *Birth Defects Research*, 114(1):5–12.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2):187–204.

Pan, C.-W., Wang, X., Ma, Q., Sun, H.-P., Xu, Y., and Wang, P. (2015). Cognitive dysfunction and health-related quality of life among older chinese. *Scientific reports*, 5(1):17301.

Park, S. S. (2021). Caregiversâ mental health and somatic symptoms during covid-19. *The Journals of Gerontology: Series B*, 76(4):e235–e240.

Pascal, A., Govaert, P., Oostra, A., Naulaers, G., Ortibus, E., and Van den Broeck, C. (2018). Neurodevelopmental outcome in very preterm and very-low-birthweight infants born over the past decade: a meta-analytic review. *Developmental Medicine & Child Neurology*, 60(4):342–355.

Pascual-Sáez, M., Cantarero-Prieto, D., and Blázquez-Fernández, C. (2019). Partnerâs depression and quality of life among older europeans. *The European Journal of Health Economics*, 20:1093–1101.

Permutt, T. and Hebel, J. R. (1989). Simultaneous-equation estimation in a clinical trial of the effect of smoking on birth weight. *Biometrics*, pages 619–622.

Perrino, T., Mason, C. A., Brown, S. C., and Szapocznik, J. (2010). The relationship between depressive symptoms and walking among hispanic older adults: A longitudinal, cross-lagged panel analysis. *Aging & mental health*, 14(2):211–219.

Pinquart, M. and Sörensen, S. (2003). Differences between caregivers and noncaregivers in psychological health and physical health: a meta-analysis. *Psychology and aging*, 18(2):250.

Price, A. L., Spencer, C. C., and Donnelly, P. (2015). Progress and promise in understanding the genetic basis of common diseases. *Proceedings of the Royal Society B: Biological Sciences*, 282(1821):20151684.

Proto, E. and Quintana-Domeque, C. (2021). Covid-19 and mental health deterioration by ethnicity and gender in the uk. *PloS one*, 16(1):e0244419.

Raggi, A., Corso, B., Minicuci, N., Quintas, R., Sattin, D., De Torres, L., Chatterji, S., Frisoni, G. B., Haro, J. M., Koskinen, S., et al. (2016). Determinants of quality of life in ageing populations: Results from a cross-sectional study in finland, poland and spain. *PLoS One*, 11(7):e0159293.

Rebagliato, M., Florey, C. d. V., and Bolumar, F. (1995). Exposure to environmental tobacco smoke in nonsmoking pregnant women in relation to birth weight. *American Journal of Epidemiology*, 142(5):531–537.

Rodrigues, R., Simmons, C., Schmidt, A. E., and Steiber, N. (2021). Care in times of covid-19: The impact of the pandemic on informal caregiving in austria. *European journal of ageing*, 18(2):195–205.

Rosso, A. L., Taylor, J. A., Tabb, L. P., and Michael, Y. L. (2013). Mobility, disability, and social engagement in older adults. *Journal of aging and health*, 25(4):617–637.

Rowe, J. W. and Kahn, R. L. (1997). Successful aging. *The gerontologist*, 37(4):433–440.

Royer, H. (2009). Separated at girth: Us twin estimates of the effects of birth weight. *American Economic Journal: Applied Economics*, 1(1):49–85.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, 66(5):688.

Ruggeri, K., Garcia-Garzon, E., Maguire, Á., Matz, S., and Huppert, F. A. (2020). Well-being is more than happiness and life satisfaction: a multidimensional analysis of 21 countries. *Health and quality of life outcomes*, 18(1):1–16.

Schmitz, H. and Westphal, M. (2017). Informal care and long-term labor market outcomes. *Journal of health economics*, 56:1–18.

Serrano-Alarcón, M., Kentikelenis, A., Mckee, M., and Stuckler, D. (2022). Impact of covid-19 lockdowns on mental health: Evidence from a quasi-natural experiment in england and scotland. *Health economics*, 31(2):284–296.

Sexton, E., King-Kallimanis, B. L., Conroy, R. M., and Hickey, A. (2013). Psychometric evaluation of the casp-19 quality of life scale in an older irish cohort. *Quality of Life Research*, 22:2549–2559.

Solhi, M., Abbasi, K., Azar, F. E. F., and Hosseini, A. (2019). Effect of health literacy education on self-care in pregnant women: a randomized controlled clinical trial. *International journal of community based nursing and midwifery*, 7(1):2.

Steptoe, A., Deaton, A., and Stone, A. A. (2015). Subjective wellbeing, health, and ageing. *The Lancet*, 385(9968):640–648.

Stevenson, B. and Wolfers, J. (2009). The paradox of declining female happiness. *American Economic Journal: Economic Policy*, 1(2):190–225.

Stöckel, J. and Bom, J. (2022). Revisiting longer-term health effects of informal caregiving: Evidence from the uk. *The Journal of the Economics of Ageing*, 21:100343.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.

Tarka, P. (2018). An overview of structural equation modeling: its beginnings, historical development, usefulness and controversies in the social sciences. *Quality & quantity*, 52:313–354.

Tilstra, A. M. and Masters, R. K. (2020). Worth the weight? recent trends in obstetric practices, gestational age, and birth weight in the united states. *Demography*, 57(1):99–121.

Todorovic, N., Vracevic, M., Rajovic, N., Pavlovic, V., Madzarevic, P., Cumic, J., Mostic, T., Milic, N., Rajovic, T., Sapic, R., et al. (2020). Quality of life of informal caregivers behind the scene of the covid-19 epidemic in serbia. *Medicina*, 56(12):647.

Trickett, J., Johnson, S., and Wolke, D. (2020). Behavioural and educational outcomes following extremely preterm birth: current controversies and future directions. *Emerging Topics and Controversies in Neonatology*, pages 367–385.

Truskinovsky, Y., Finlay, J. M., and Kobayashi, L. C. (2022). Caregiving in a pandemic: Covid-19 and the well-being of family caregivers 55+ in the united states. *Medical Care Research and Review*, 79(5):663–675.

Tzeng, D.-S., Chung, W.-C., Lin, C.-H., and Yang, C.-Y. (2012). Effort-reward imbalance and quality of life of healthcare workers in military hospitals: a cross-sectional study. *BMC Health Services Research*, 12:1–9.

UK Government (2022). Coronovirus in the uk dashboard. `https://coronavirus.data.gov.uk/details/download` [Accessed: November 2022].

United Nations (2020). *World Population Ageing 2019: Highlights*. United Nations.

Vamos, S., Okan, O., Sentell, T., and Rootman, I. (2020). Making a case for âeducation for health literacyâ: An international perspective. *International journal of environmental research and public health*, 17(4):1436.

Van de Ven, W. P. and Van Der Gaag, J. (1982). Health as an unobservable: a mimic-model of demand for health care. *Journal of Health Economics*, 1(2):157–183.

Van Der Heide, I., Wang, J., Droomers, M., Spreeuwenberg, P., Rademakers, J., and Uiters, E. (2013). The relationship between health, education, and health literacy: results from the dutch adult literacy and life skills survey. *Journal of health communication*, 18(sup1):172–184.

Van Hecke, N., Claes, C., Vanderplasschen, W., De Maeyer, J., De Witte, N., and Vandevelde, S. (2018). Conceptualisation and measurement of quality of life based on schalock and verdugoâs model: A cross-disciplinary review of the literature. *Social Indicators Research*, 137:335–351.

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.

Vanleerberghe, P., De Witte, N., Claes, C., Schalock, R. L., and Verté, D. (2017). The quality of life of older people aging in place: a literature review. *Quality of Life Research*, 26:2899–2907.

Ventura, S. J., Hamilton, B. E., Mathews, T., and Chandra, A. (2003). Trends and variations in smoking during pregnancy and low birth weight: evidence from the birth certificate, 1990–2000. *Pediatrics*, 111(Supplement_1):1176–1180.

Vilanova, C. S., Hirakata, V. N., de Souza Buriol, V. C., Nunes, M., Goldani, M. Z., and da Silva, C. H. (2019). The relationship between the different low birth weight strata of newborns with infant mortality and the influence of the main health determinants in the extreme south of brazil. *Population health metrics*, 17(1):1–12.

Wagstaff, A. (1986). The demand for health: some new empirical evidence. *Journal of Health economics*, 5(3):195–233.

Warner, D. F. and Adams, S. A. (2016). Physical disability and increased loneliness among married older adults: The role of changing social relations. *Society and Mental Health*, 6(2):106–128.

Warner, D. F. and Kelley-Moore, J. (2012). The social context of disablement among older adults: Does marital quality matter for loneliness? *Journal of health and social behavior*, 53(1):50–66.

Warrington, N. M., Beaumont, R. N., Horikoshi, M., Day, F. R., Helgeland, Ø., Laurin, C., Bacelis, J., Peng, S., Hao, K., Feenstra, B., et al. (2019). Maternal and fetal genetic effects on birth weight and their relevance to cardio-metabolic risk factors. *Nature genetics*, 51(5):804–814.

Webber, S. C., Porter, M. M., and Menec, V. H. (2010). Mobility in older adults: a comprehensive framework. *The gerontologist*, 50(4):443–450.

Weber, K., Canuto, A., Giannakopoulos, P., Mouchian, A., Meiler-Mititelu, C., Meiler, A., Herrmann, F. R., Delaloye, C., Ghisletta, P., Lecerf, T., et al. (2015). Personality, psychosocial and health-related predictors of quality of life in old age. *Aging & Mental Health*, 19(2):151–158.

Wehby, G. L., Fletcher, J. M., Lehrer, S. F., Moreno, L. M., Murray, J. C., Wilcox, A., and Lie, R. T. (2011). A genetic instrumental variables analysis of the effects of prenatal smoking on birth weight: evidence from two samples. *Biodemography and social biology*, 57(1):3–32.

Whitley, E., Reeve, K., and Benzeval, M. (2021). Tracking the mental health of home-carers during the first covid-19 national lockdown: evidence from a nationally representative uk survey. *Psychological Medicine*, pages 1–10.

Wilson, R. S., Boyle, P. A., Segawa, E., Yu, L., Begeny, C. T., Anagnos, S. E., and Bennett, D. A. (2013). The influence of cognitive decline on well-being in old age. *Psychology and aging*, 28(2):304.

Wister, A., Li, L., Mitchell, B., Wolfson, C., McMillan, J., Griffith, L. E., Kirkland, S., and Raina, P. (2022). Levels of depression and anxiety among informal caregivers during the covid-19 pandemic: A study based on the canadian longitudinal study on aging. *The Journals of Gerontology: Series B*, 77(9):1740–1757.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

World Health Organization (2001). The world health report 2001: Mental health: new understanding, new hope.

Yang, T.-C., Shoff, C., Noah, A. J., Black, N., and Sparks, C. S. (2014). Racial segregation and maternal smoking during pregnancy: A multilevel analysis using the racial segregation interaction index. *Social Science & Medicine*, 107:26–36.

Zhang, X. and Kramer, M. S. (2009). Variations in mortality and morbidity by gestational age among infants born at term. *The Journal of pediatrics*, 154(3):358–362.

Zhou, M. and Kan, M.-Y. (2021). The varying impacts of covid-19 and its related measures in the uk: A year in review. *PLoS One*, 16(9):e0257286.