Three essays on Informal Care, Health, and Education

by

Dörte Heger

A thesis submitted to the
Department of Economics
in conformity with the requirements for
the degree of Doctor of Philosophy

Queen's University
Kingston, Ontario, Canada
June 2014

Abstract

This dissertation is a collection of three essays that use economic tools to address policy-relevant issues related to ageing, population health, and education. The use of economic modelling and econometric analyses has the potential to provide information on the consequences and effectiveness of policy interventions in these areas and enables policymakers to make better informed decisions. Chapter 1 provides an introduction to these topics and is followed by the three essays.

In Chapter 2, I analyze how providing informal care to an elderly parent affects the caregiver's labour market outcomes, cognitive ability, and health; and study the influence of the institutional background on the caregiving decision and the effects of caregiving. My results show that negative effects on labour market outcomes can be avoided by the provision of formal care alternatives, but negative effects for caregivers' mental health persist. These findings give useful insights into the optimal provision of formal care in today's ageing societies.

Self-reported health measures are commonly collected in numerous surveys but might be influenced by respondents' definitions and frames of reference of health. In Chapter 3, I address the issue of response bias in population surveys by constructing an objective measure of health. I find that using a common definition of health nearly eliminates the reported health differences between the U.S. and Canada. Socioeconomic differences in health are stronger in the U.S., but remain an issue in Canada.

Chapter 4 studies the effect of post-secondary education on the continued development of reading proficiency during adolescence and young adulthood. Reading proficiency is essential for labour market success in a knowledge-based economy, but little is known about how advanced reading skills such as text interpretation and text evaluation are developed. The results show that university graduation increases students' reading proficiency relative to high school graduation, which demonstrates the importance of cognitive skill investments later in the life cycle.

Co-authorship

Chapter 4 is co-authored with Professor Casey Warman at Dalhousie University.

Dedication

To my parents, for their unwavering support and belief in me.

Acknowledgments

I am indebted to my supervisors Steven Lehrer and Vincent Pohl for their guidance, insight, and support in preparing this thesis. I am also grateful to Robin Boadway, Thor Koeppel, and Chris Ferrall for many helpful comments.

I benefited from the questions, comments and suggestions from my fellow students at Queen's University and the seminar participants at the CEA Conference (Calgary, Montréal, and Vancouver), the 4th SHARE User Conference, and the 4th International Workshop on Applied Economics of Education.

I also thank my family and friends for their patience, support, and encouragement.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Developed countries are experiencing rapid population ageing. Average life expectancy increased by ten years since 1970 and now exceeds 80 years across OECD countries (OECD, 2013). Over the same time period, the average effective retirement age in these countries decreased slightly and only recently this trend has been reversed.[1] Taken together, these trends have important implications for countries' social security systems and the provision of health care services. One of the main challenges policy makers will face over the next decades, is how to address the growing need for long-term care as people live longer. While publicly financed formal care makes up only a small share of total long-term care provided, the financial burden on public finances is non-negligible. In 2011, the latest year for which data is available, total public spending on long-term care amounted to 1.6% of GDP on average across OECD countries and this number is expected to double by 2060 (OECD, 2013).

An alternative to formal care is informal care provided by family members and

[1] OECD Labour Force Statistics database; the average effective retirement age across 34 OECD countries was 68.7 (66.6) years in 1970 and 64.2 (63.1) years in 2012 for men (women).

friends. Several countries are trying to encourage informal caregiving by creating incentives for informal care provision since it is hoped that this will reduce the financial burden on health care systems. Chapter 2 analyses the effects of informal caregiving on informal caregivers to predict possible consequences of such policies. The outcomes I consider - caregivers' work, cognitive ability, and health - influence the financial situation and the well-being of caregivers and are thus relevant for individual caregivers themselves and for the society as a whole. For one, reduced labour force participation of informal caregivers can lower the productivity of the workforce. For another, negative consequences for caregivers' mental and physical well-being will manifest themselves as an increase in health care expenditures in the future. My study assesses whether such concerns are warranted and in doing so, helps to inform the debate on the optimal provision of long-term care.

Many empirical studies in the area of health and labour economics, including the one presented in Chapter 2, depend on a meaningful measure of health, which is comparable between individuals both within and across countries. The reliability of the results of these studies depends to a large extent on the quality of that measure. A very common measure of health, which has been shown to be strongly correlated with various objective health outcomes, including, for example, future mortality, is individuals' self-assessed health status (Maddox and Douglass, 1973; Idler and Benyamini, 1997; Dowd and Zajacova, 2010). This measure has been widely used in national and international health comparisons as it is easy to collect from population surveys. However, it has been shown that a person's perception of her health status is influenced by her age, socioeconomic background, and even language use (Etilé and Milcent, 2006; Lindeboom and van Doorslaer, 2004; Jürges, 2007; Bago d'Uva et al.,

2008a).

In Chapter 3, I study differences in reporting styles of self-assessed health between the U.S. and Canada by constructing an objective health index based on a large set of health related information. Using an objective and thus comparable health measure has the advantage that findings will not be influenced by differences in reporting behaviour. Consequently, objective health measures are especially useful to compare health status or the influence of health on some outcome of interest across different populations. I use the objective health index to compare population health in the U.S. and Canada and to assess income and education related health differences between the two countries. International health comparisons are of great policy importance and are frequently used to evaluate the quality of health care systems. My study facilitates such comparisons by providing a more accurate analysis of true health differences.

Not only the comparison of health outcomes, but also international comparisons of educational outcomes are attracting substantial attention in the popular press. Comparisons of educational standards have been promoted by projects such as the Programme for International Student Assessment (PISA). PISA was introduced in 2000 when the reading, mathematics, and science performance of fifteen year old students in 28 OECD countries was assessed and thereupon compared. Canada followed the participants of the original PISA study in a biennial survey and reassessed their reading ability nine years later in the Youth in Transition Survey Reading Skills Reassessment (Statistics Canada, 2005; Cartwright, 2013).[2] The resulting data provides the unique opportunity to not only study reading ability at a single point in time,

[2]Recent results show that while Canadian youths perform well in a global context in reading, their math and science performance has declined in recent years with the exception of students in Québec (Knighton et al., 2010).

but to analyze the development of reading proficiency during adolescence and young adulthood, which is an area that is still understudied.

Chapter 4 aims to fill this gap in the literature by examining how post-secondary education affects reading proficiency. Reading proficiency provides social and cultural benefits, and is, in particular, an important determinant of labour market success as it is strongly associated with higher employment rates and higher wages (Knighton and Bussière, 2006; Finnie and Meng, 2006). Especially in a knowledge-based economy, reading proficiency has become increasingly important (Green and Riddell, 2001; Ingram and Neumann, 2006). Consequently, policy makers wish to increase the reading proficiency of the population (see, for example, Filmer et al. (2006)). Studying the impact of post-secondary education on reading proficiency helps to understand which policy tools may be used to achieve this goal. Moreover, studying the effect of post-secondary education on individuals' skill production extents our knowledge on the malleability of cognitive ability past early childhood (Hansen et al., 2004).

The main findings of the three studies are summarized in Chapter 5, which concludes this dissertation and provides an outlook of several future directions of research.

Chapter 2

Employment and well-being of informal caregivers

Informal caregivers provide valuable services to elderly persons with long-term care needs, but the consequences of caregiving on caregivers are not yet fully understood. In this paper, I introduce a theoretical framework to explain caregiving behaviour of mature daughters to their elderly parents and show how the availability of formal care alternatives influences the decision to provide care. Differentiating between countries with and without comprehensive formal care alternatives, I analyze the effects of informal caregiving on caregivers' labour force participation, cognitive ability, and health in thirteen European countries using longitudinal data from the Survey for Health, Ageing and Retirement in Europe (SHARE). I find that caregiving significantly reduces the probability of being employed in countries with few formal care alternatives but does not affect employment in countries with a more generous formal care system. In all countries, caregiving negatively affects mental health and slightly

increases cognitive ability.[1]

2.1 Introduction

A growing share of Europeans aged 50 and above still have living parents (Ogg and Renaut, 2006). When a parent's health declines, the adult child often volunteers his or her time to assist the parent with personal care or household chores. Such informal care far exceeds formal care in the number of hours of care provided, as well as in terms of monetary value (OECD, 2005). Moreover, the need for informal care is growing quickly due to the increase in life expectancy (Colombo et al., 2011). Using recent data from the U.S., Feinberg et al. (2013) estimate an economic value of informal caregiving of $450 billion (3.2% of GDP) in 2009, which signifies a 20% increase since 2007 alone. At the same time, fewer informal caregivers are available. For one, children tend to live further away from their parents, which makes caregiving more difficult. For another, increased female labour market participation requires women, who traditionally account for the vast majority of informal caregivers, to balance work and care commitments (Colombo et al., 2011). In response to the growing need for care, policymakers in several countries have aimed to encourage informal caregiving to reduce the financial pressure on public long term care (LTC) systems. However,

[1]This paper uses data from SHARE wave 4 release 1.1.1, as of March 28th 2013 or SHARE wave 1 and 2 release 2.5.0, as of May 24th 2011 or SHARELIFE release 1, as of November 24th 2010. The SHARE data collection has been primarily funded by the European Commission through the 5th Framework Programme (project QLK6-CT-2001-00360 in the thematic programme Quality of Life), through the 6th Framework Programme (projects SHARE-I3, RII-CT-2006-062193, COMPARE, CIT5- CT-2005-028857, and SHARELIFE, CIT4-CT-2006-028812) and through the 7th Framework Programme (SHARE-PREP, N 211909, SHARE-LEAP, N 227822 and SHARE M4, N 261982). Additional funding from the U.S. National Institute on Aging (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, R21 AG025169, Y1-AG-4553-01, IAG BSR06-11 and OGHA 04-064) and the German Ministry of Education and Research as well as from various national sources is gratefully acknowledged (see www.share-project.org for a full list of funding institutions).

the effect of caregiving on caregivers is not yet fully understood.

To better understand the impact of caregiving on caregivers, I analyze the effects of care provision to an elderly parent on women aged 50 to 75 in three important areas: labour force participation, cognitive ability, and health.[2] These outcomes represent valuable indicators of caregivers' quality of life, well-being, and risk of financial vulnerability and are of relevance both to the individual caregivers and to society. While mature caregivers make up a significant portion of informal caregivers in the European Union (Glendinning et al., 2009), they have a lower attachment to the labour force and may be more prone to negatively react to the stress of caregiving than younger caregivers. In addition, parental caregiving may be more emotionally challenging and thus lead to more severe consequences for the caregiver than care provision to a friend or neighbour due to close emotional ties between parents and children.

My paper proceeds in two steps. First, I present the trade off between leisure time, caregiving activities, and work in a theoretical framework and relate the daughter's decision to provide care to individual characteristics, the parent's need for care, and the availability of formal LTC options. Second, I estimate the causal effects of caregiving on labour force participation, cognitive ability, and health while carefully controlling for the endogeneity of caregiving using longitudinal data from the Survey of Health, Ageing, and Retirement (SHARE) collected in 2004/2005, 2006/2007, and 2010. This data set provides rich demographic and socio-economic information of the population 50+ in Europe and offers the unique opportunity to study the influence of institutional characteristics. My sample consists of 10,313 women from thirteen

[2] Due to data limitations, I only consider a person's own parents and do not include parents-in-law.

European countries, of which approximately one third still has a living parent. For these women, I additionally observe the distance to the parent and the parent's health status, which provides information about the time requirement of providing care and the parent's need for assistance.

As caregiving may depend on unobserved individual characteristics, such as the emotional closeness to the care receiver or the caregiver's sense of duty, I use fixed effects estimation to account for time-invariant endogeneity. Further, I consider time-varying endogeneity by using a change in parental health as an instrument for a change in caregiving behaviour. Considering time-variant and time-invariant endogeneity provides a more careful treatment of the endogeneity of caregiving than previous cross-sectional studies allow for. Furthermore, my empirical strategy accounts for the possibility that the effects of caregiving depend on the intensity of caregiving as measured by the frequency of care and by the institutional caregiving context. Specifically, I separately consider daily, weekly, and any frequency of caregiving and I test for heterogeneous effects of caregiving between formal care countries with relatively generous formal care alternatives and family care countries where elderly care is predominantly provided by the family. Comparing the effects of caregiving in family and formal care countries will increase our understanding of the influence of the institutional background on the effects of caregiving on informal caregivers and inform the policy debate about the optimal provision of LTC. In addition, my work contributes to the growing literature on the effects of caregiving on caregivers and adds to the still relatively limited European evidence. Moreover, to the best of my knowledge, this study is the first to look at cognitive ability of parental caregivers in a longitudinal, multi-country context.

To preview my results, I find a strong relationship between the institutional caregiving context and the effects of caregiving on caregivers. Caregiving decreases employment rates in family care countries by 26 to 47 percentage points depending on the frequency of care but has no impact on caregivers' employment probability in formal care countries. Similarly, the effects of caregiving on hours worked differ based on the caregiving context: Caregiving tends to increase hours worked in formal care countries but decreases hours worked in family care countries, though the estimates are mostly statistically insignificant. Further, I find caregiving leads to small positive or insignificant effects on cognitive ability with slight differences based on the institutional setting. While caregiving increases verbal fluency in family care countries, caregiving has positive effects on short-term memory functioning in formal care countries. With respect to health, I find caregiving increases the number of depressive symptoms by about three additional symptoms for daily caregiving independent of the caregiving context. While caregivers in formal care countries report slightly better self-perceived health, I find negative effects on physical health using a more objective measure of health. However, small positive effects on objective health are found for caregivers in family care countries.

The remainder of this paper is organized as follows. In Section 2.2, I lay out the possible pathways through which caregiving may affect caregivers by reviewing the existing literature and link the caregiving context to the intensity of caregiving. Section 2.3 introduces my theoretical framework explaining caregiving behaviour, Section 3.2 describes the SHARE data and the variables of interest, Section 2.5 discusses the estimation strategy, and Section 2.6 presents my results. Section 3.5 concludes.

2.2 Background and Literature

Given the often considerable time commitment of informal caregiving and the potential loss of productivity if caregivers reduce their labour supply, the effect of caregiving on labour market outcomes has received much attention (see Lilly et al. (2007) for an overview). Labour market outcomes for mature caregivers are especially important since caregivers close to the retirement age are less attached to the labour force than younger caregivers. As early retirement is generally linked to a reduction in pension benefits, understanding whether caregiving and work are compatible for elderly caregivers is important to evaluate caregivers' risk of financial vulnerability.

Theoretically, the direction of the effect of caregiving on labour market outcomes is ambiguous. Caregiving may reduce employment or hours worked if caregivers substitute work time with care time, or if increased absenteeism of caregivers to fulfill caregiving obligations leads to job loss (Heitmueller and Inglis, 2007). Yet, labour market participation may increase if work offers respite from caregiving or if caregivers require additional income (Carmichael and Charles, 1998, 2003). Empirically, determining causal effects, however, is challenging due to reverse causality. Persons with less attachment to the labour market or a low time cost are more likely to become caregivers (Ettner, 1995, 1996; Heitmueller, 2007; Carmichael et al. 2010). In addition, labour force participation and caregiving may be influenced by the same unobserved characteristics, which would produce a spurious correlation. For example, individuals with high (unobserved) ability may be more productive employees and may also be more productive at providing informal care (He and McHenry, 2013).[3] As a result, cross-sectional studies may not always be able to detect causal relationships.

[3]Such a relationship may, for example, explain the finding by Pavalko and Artis (1997) that an individual's employment history does not influence the likelihood of becoming a caregiver.

In a review of the literature, Lilly et al. (2007) conclude that caregiving generally is associated with a moderate reduction in hours worked, intensive caregiving leads to negative consequences for both hours worked and labour force participation. Infrequent caregivers, however, are sometimes found to be less likely to retire (Carmichael and Charles, 1998; Dentinger and Clarkberg, 2002). While most of the analyzed studies rely on cross-sectional data and are thus subject to above mentioned concerns, the results have been largely mirrored by recent longitudinal studies (Heitmueller, 2007; King and Pickard, 2013; van Houtven et al., 2013).

The vast majority of evidence relies on U.S. or U.K. data, two countries with very “liberal” welfare regimes (Haberkern and Szydlik, 2010). For continental Europe, the focus of my study, the effects of caregiving may not necessarily be the same since these countries have different welfare arrangements. Little is known about the effects of caregiving on labour supply under different institutional settings and the existing evidence is inconclusive. Crespo and Mira (2010) find a stronger negative effect of daily caregiving on employment in Southern European countries than in Northern European countries. Looking at cross-sectional data, Bolin et al. (2008b) analyze the effects of caregiving on employment, hours worked, and wages and find more adverse effects of caregiving on labour force participation in Central Europe than in Southern Europe; Crespo (2006) find negative effects on employment for caregivers in both Northern and Southern European countries. My paper extends these studies by considering two additional countries, Poland and the Czech Republic, and by also including data collected in 2010 from the most recent wave of the SHARE data, which allows me to observe individuals over a longer time period. Moreover, my paper contributes to a better understanding of the caregiving decision by carefully assessing

both time constant as well as time varying sources of endogeneity and allows for heterogeneous effects of caregiving by studying different dimensions of the intensity of care.

Similar to labour market outcomes, interest in the health effects of caregiving is not limited to the individual caregiver. While informal caregiving helps to reduce public expenditures in the short run, long run effects are uncertain as negative health outcomes for caregivers may increase their health care needs in the future. Whether to expect positive or negative effects of caregiving on caregivers' physical and mental health is again ambiguous in theory. For one, according to the stress process model proposed by Aneshensel (1995), caregiving may lead to a decline in health due to increased stress or a change in health behaviour or diet in response to a reduction in personal time. Physiological responses to stress consistent with this theory have been documented in both medical and psychological studies (see, for example, Vitaliano et al. (2007)). For another, caregiving may increase physical activity and lead to health improvements. This view is expressed by the "healthy caregiver hypothesis" (Fredman et al., 2008).

Likewise, both positive and negative effects are also possible regarding mental health. Caregiving can be rewarding as it conveys the feeling that the caregiver is needed, yet it might lead to loneliness if caregiving prevents other social activities. Which effect of caregiving on physical or mental health dominates is still debated (see Pinquart and Sörensen (2003, 2007) for an overview). While Coe and van Houtven (2009) find continued caregiving decreases self-rated health and increases depressive symptoms over time, Brown et al. (2009) find decreased mortality rates of caregivers compared to non-caregivers. Moreover, the sign and magnitude of health effects is

influenced by the intensity of caregiving and caregivers' characteristics, with intensive caregivers and caregivers from lower socioeconomic background generally experiencing worse outcomes (Schulz et al., 1997; Hirst, 2005; Pinquart and Sörensen, 2005).

Cognitive effects of caregiving have received little attention in studies using large representative population surveys. However, like health, cognitive ability is of significant personal and public interest since preserving cognitive ability is essential for a self-determined and independent life. Thus, cognitive ability should be considered as an outcome of interest. Higher levels of stress have been shown to impact cognitive ability negatively, but caregiving may also provide cognitive stimulation and hence have positive effects on cognitive ability (Vitaliano et al. 2005; Vitaliano et al., 2007; Vitaliano et al., 2011; Lee et al., 2004; Bertrand et al. 2012). The empirical evidence is still inconclusive and more research is needed.[4] Related, several studies look at the effects of social participation and volunteering on cognitive ability and generally find positive results (Glei et al., 2005; Green et al., 2008; Engelhardt et al., 2010; Dobrescu and Christelis, 2012). Though caregiving is sometimes included as one possible volunteering activity, caregiving differs substantially from participating in a political, cultural or religious organization or club (Hsu, 2007). In particular, caregiving is usually motivated by the care receivers need for care and social norms rather than by a person's desire to provide care and is thus likely to be more stressful than other volunteering activities. Consequently, the effects of caregiving on cognitive ability may differ from the effects reported for social engagement.

While the frequency of care provides an intuitive measure of the care intensity and more frequent caregiving has been associated with worse outcomes for caregivers as

[4] Ambiguous results are found in studies focusing on broadly defined spousal caregiving, while mostly negative effects are found for caregivers to persons suffering from dementia; see, for example, Caswell et al. (2003), de Vugt et al. (2006), Leipold et al. (2008), and Mackenzie et al. (2009).

outlined above, institutional differences have been less studied due to data limitations but may be equally important in explaining differential effects of caregiving. For one, Matire and Schulz (2012) emphasize that the magnitude and direction of the effects of caregiving are influenced by caregivers' appraisal of the caregiving demands and adaptive capacities. When caregivers feel capable of dealing with the caregiving demands, positive outcomes on cognitive ability and health are more likely than when caregivers feel overwhelmed. Hence, the institutional setting may influence caregivers' appraisal of caregiving demands, for example, by determining the extent to which formal care options are available or by offering financial support for informal caregivers. For another, formal care alternatives may reduce negative consequences for caregivers by directly reducing the caregiving burden (Bass et al. 1996).

To analyze the influence of the intensity of caregiving, I separate countries into family and formal care countries based on their spending level on LTC. This grouping corresponds to separating countries into what are commonly thought of as countries with strong formal LTC, and countries with strong family based LTC. For example, Haberkern and Szydlik (2010) evaluate countries based on whether and to what extent there is a legal obligation to support relatives in need, the services that the state provides or funds, and public opinion about whether the state or the family should in general be responsible for the care and support of dependent elderly people. Countries with strong formal-care systems include the Scandinavian countries, the Netherlands, and Switzerland. Belgium and France are intermediate cases between family and formal based LTC; I group them with formal care countries based on their public LTC spending. The countries with family based care systems include the Mediterranean countries, Germany, and Austria. I group the Czech Republic and Poland, which are

not part of Haberkern and Szydlik's study, with the family care countries as both countries have very low public expenditures on LTC.[5] Consistent with the motivation for the group selection, the two country groups differ substantially with respect to several additional institutional characteristics and social values related to LTC as shown in Table 2.1. Compared to formal care countries, family care countries have on average a lower share of public health expenditure, fewer nursing home beds, less financial support for informal caregivers (such as, for example, caregiving allowances), and elderly care is generally considered to be the responsibility of the family rather than the state.[6]

The institutional characteristics and cultural values related to caregiving have been shaped over a long period of time and are influenced by the political process. But while the institutional setting is endogenous, changing a country's LTC care system or society's view on elderly care requires time. Consequently, the formal help available to a potential caregiver through the LTC system can be considered an exogenous factor in the decision whether to provide informal care or not. As formal care may be used to complement or substitute informal care, I expect caregiving to be more strenuous in countries with few formal care options compared to countries with more generous formal care alternatives (Bolin et al., 2008a; Bonsang, 2009).

[5] Based on their institutional setting, France and Germany might be considered to be in the middle between formal and family care countries (for example, Germany is the only formal care country offering a caregiver allowance, while France is the only formal care country, which does not offer a caregiver allowance). As a robustness check for my empirical analysis, I estimate all regressions for a reduced sample without France and Germany. While I lose some significance due to the reduced sample size, the main results of my study remain unchanged. Results are available upon request.

[6] A caregiver allowance is a public benefit paid directly to the caregiver (Columbo et al., 2011).

Table 2.1: Institutional characteristics and social values

	Public LTC exp. (% of GDP)[a]	Public health exp. (% of total)[b]	Nursing home beds (per 1000 pop 65+)[c]	Caregiver allowance[d]	Elderly care state responsibility[e]	Old age dep. ratio[f]
Countries with predominantly family based LTC						
Poland	0.70	70.28	17.10	0	2.33	19.40
Czechia	0.80	84.18	43.30	0	2.64	23.40
Spain	0.80	73.05	29.30	0	2.86	25.80
Germany	1.40	76.45	50.30	1	2.77	31.20
Greece	1.40	65.05	1.40	0	2.34	29.90
Austria	1.60	67.83	38.70	0	2.90	26.20
Italy	1.90	77.84	16.60	0	2.76	31.60
Average	1.23	73.53	28.10	0.14	2.71	26.79
Countries with strong formal LTC						
Switzerland	2.16	64.89	68.90	1	2.95	25.30
France	2.20	76.75	51.80	0	3.39	26.60
Belgium	2.30	75.91	70.80	1	3.25	26.40
Netherlands	3.80	85.60	68.40	1	3.55	24.40
Sweden	3.90	81.62	80.40	1	3.64	29.20
Denmark	4.50	85.31	52.30	1	4.12	26.70
Average	3.14	78.35	65.43	0.83	3.43	26.43

[a]European Commission (2012) and Eurostat; 2010 values. [b]OECD Health Data 2013; 2011 or nearest year available. [c]OECD Health Data 2012 - Long-Term Care Data; 2009; Greece: 2000. [d]Colombo et al. 2011. [e]SHARE; Average of the SHARE respondents' answers to the questions who - the family (1) or the state (5) - should bear the responsibility for help with household chores/personal care for older persons. [f]Eurostat; 2012 values.

2.3 Theoretical Framework

This section presents a simple theoretical model to describe the interrelation between work, caregiving, cognitive ability, and health, and points out how this interrelation may be influenced by the caregiving context. The sole decision maker in my model is the adult daughter who is altruistic towards her elderly parent. The daughter derives utility from consumption (C), leisure (L), and parental well-being (U^P). The daughter's utility is given by

$$U^D = U^D(C, L, U^P), \tag{2.1}$$

while the parent's utility is given by

$$U^P = U^P(CG, OC), \tag{2.2}$$

where CG represents informal care provided by the daughter, and OC indicates other care or consumption.[7] Hence, the model does not rule out the possibility of formal care or informal care provided by another family member. However, I assume that informal care received from the daughter has a special value to the parent and care provided by someone else is not a perfect substitute, that is the additional utility the parent receives is specific to informal care provided by the daughter in my model. For this reason, I also do not allow for the option of substituting informal caregiving with a monetary transfer to the parent. The daughter's altruism may be rooted in the desire to support her parent but may also arise from family obligations or social and

[7]Since this paper focuses on the daughter as the sole decision maker, I abstract from possible altruism of the parent towards the daughter.

cultural norms to provide care, which are probably more predominant in family care countries. I assume that both the daughter's and the parent's utility are increasing in each input with decreasing marginal utility.

The daughter's budget and time constraints are given by

$$C = Y^{*} + w(HC)H - e\,CG, \tag{2.3}$$

and

$$T = H + L + CG, \tag{2.4}$$

where Y^{*} represents outside income, $w(HC)$ denotes the wage, which depends on the daughter's level of health capital and cognitive ability HC and $\partial w(HC)/\partial HC \geq 0$, H denotes hours worked, e represents the cost of caregiving, which can be interpreted as a combination of monetary effort cost and additional expenditures for medication or equipment for the care recipient, and T is the individual's total time endowment. If a daughter wants to increase the amount of time devoted to informal care, she will incur additional expenditures and must either reduce consumption or increase her work hours. The extent of these additional expenditures will depend, among other factors, on the generosity of the formal care system such as the existence of a caregiver allowance, which will reduce the cost of caregiving. However, as care provision is time consuming, the time constraint forces the daughter to either reduce her leisure time or her work hours.

Lastly, health capital and cognitive ability evolve according to

$$HC = HC_0 + \gamma CG, \tag{2.5}$$

where HC_0 represents the initial level of health capital and cognitive ability and γ represent the effectiveness of caregiving as an investment in the production of health capital and cognitive ability (Grossman, 1972).[8] If γ is positive, caregiving increases the stock of health capital and cognitive ability, otherwise caregiving is detrimental to the caregiver's health or cognitive ability.[9] As discussed above, positive effects seem more plausible for less intensive caregiving or when the caregiving activity leads to physical activity or cognitive stimulation, while negative effects are to be expected for very stressful caregiving.

The daughter chooses the amount of consumption, leisure, hours work, and hours providing care in order to maximize her utility subject to her time and budget constraints and the law of motion for health capital and cognitive ability. Her maximization problem is equivalent to maximizing the Lagrangian:

$$\begin{aligned}\mathcal{L} = &U^D(C, L, U^P(CG, OC)) \\ &+ \lambda_1(Y^* + w(HC)H - eCG - C) \\ &+ \lambda_2(T - H - L - CG) \\ &+ \lambda_3(HC - HC_0 - \gamma CG). \end{aligned} \tag{2.6}$$

[8]For simplicity, I abstract from other possible investments in health and cognitive ability and use a combined measure of both health and cognitive ability. Differential affects of caregiving on these outcomes can easily be incorporated by allowing for different γ coefficients for each outcome and I consider different aspects of health and human capital in my empirical analysis.

[9]Extending the model to a multi-period set-up or allowing for the depreciation of health capital and cognitive ability may affect the magnitude of health and cognitive benefits derived from caregiving if investments in health or cognitive capital persist (caregiving in one period will increase - or decrease - utility in future periods by increasing - or decreasing - the stock of health or cognitive capital) but does not change the qualitative findings of the model.

where λ_1 to λ_3 are the Lagrange multipliers for the budget constraint, the time constraint and the law of motions for health capital and cognitive ability, respectively. Assuming the daughter chooses a strictly positive amount of consumption and leisure and defining $U_X \equiv \frac{\partial U}{\partial X}$, an optimal solution to her maximization problem is characterized by the following necessary first-order conditions (FOCs):

$$\begin{aligned}
i)&\ \ U^D_C - \lambda_1 = 0,\\
ii)&\ \ U^D_L - \lambda_2 = 0,\\
iii)&\ \ w(HC)\lambda_1 - \lambda_2 \leq 0, \quad H \geq 0,\\
iv)&\ \ U^D_{U^P} U^P_{CG} - e\lambda_1 - \lambda_2 - \gamma\lambda_3 \leq 0, \quad CG \geq 0,\\
v)&\ \ \frac{\partial w(HC)}{\partial HC} H\lambda_1 + \lambda_3 = 0,
\end{aligned}$$

where in $iii)$ and $iv)$ at most one inequality can be strict.

After substituting and rearranging terms, the FOCs can be expressed as:

$$w(HC)U^D_C \leq U^D_L, \quad H \geq 0, \tag{2.7}$$

which characterizes the typical trade-off between the returns and costs to work from increased consumption and foregone leisure, respectively, and determines whether the individual will work or not, and

$$U^D_{U^P} U^P_{CG} + U^D_C \frac{\partial w(HC)}{\partial HC} H\gamma \leq eU^D_C + U^D_L, \quad CG \geq 0, \tag{2.8}$$

which determines whether the individual will provide care or not.[10] Again both

[10] If the daughter is employed, her FOC with respect to caregiving can be explicitly expressed in

inequalities in the same line cannot be strict. The daughter's caregiving decision will be determined by the costs and benefits of providing care. Specifically, the daughter will not provide care if $U^D_{U_P} U^P_{CG} + U^D_C \frac{\partial w(HC)}{\partial HC} H\gamma < eU^D_C + U^D_L$, as in this situation the total marginal benefit of providing care, which consists of the marginal benefit derived from the parent's well-being ($U^D_{U^P}$) times the parent's marginal utility from caregiving (U^P_{CG}), and the marginal benefit of an increase in health and cognitive ability ($U^D_C \frac{\partial w(HC)}{\partial HC} H$) times the effectiveness of caregiving as an investment in health and cognitive ability (γ) is less than the the total marginal cost of caregiving, which is the sum of the cost of providing care (e) times the marginal utility of consumption (U^D_C) and the marginal utility of leisure (U^D_L). If the optimal decision is to provide a positive amount of care, the daughter will provide informal care up to the point where her marginal costs and benefits of providing an additional hour of care are equalized. Ceteris paribus caregiving increases with the daughter's level of altruism and the parent's need for care, and decreases in the costs associated with caregiving and in the time cost. Besides, a positive effect of caregiving on health capital and cognitive ability will encourage caregiving through the positive effect of health and cognitive ability on wages, whereas the opposite is true if caregiving leads to a reduction in health and cognitive ability.

---

terms of the wage: $U^D_{U^P} U^P_{CG} + U^D_C \frac{\partial w(HC)}{\partial HC} H\gamma \leq (e + w(HC))U^D_C, \quad CG \geq 0$, where again at most one inequality can be strict.

2.4 Data

This paper uses data from wave 1, 2 and 4 of the Survey for Health, Ageing and Retirement in Europe (SHARE) collected in 2004/2005, 2006/2007, and 2010.[11] SHARE is modeled after the English Longitudinal Study of Ageing (ELSA) and the U.S. Health and Retirement Study (HRS) and is the first data set to include information on health, cognitive ability, socio-economic status, and social engagement of the elderly at a pan-European level, which provides the unique opportunity to study the effect of institutional characteristics on caregivers' outcomes. The survey includes individuals aged 50 years or older living in residential households and their partners independent of age.[12] Partners younger than 50 are dropped from the analysis.[13].

2.4.1 Sample Selection and Caregiving Behaviour

My sample consists of women who participated in at least two interviews, which restricts my sample to thirteen European countries from Scandinavia to the Mediterraneans. I limit my sample to individuals aged 75 or younger since for older individuals the labour market outcomes are largely irrelevant, the importance of investments in health and human capital is declining, and the number of respondents with a

[11]Wave 3 of SHARE consists of a special retrospective survey (SHARELIFE), which does not ask about caregiving activities. SHARE provides five imputed data sets to deal with missing information; the results in this study are based on the first of these data sets. Item non-response is negligible for non-financial questions and the results are not sensitive to which imputed data set is used (see Christelis (2011) for more information on item non-response and the imputation process).

[12]Data is collected using a computer-assisted personal interviewing technique (CAPI). The main questionnaire is complemented by a self-selection paper and pencil drop off questionnaire. Sample selection varies across countries from simple random selection of households to multi-stage designs due to varying institutional conditions regarding sampling. All results presented in this paper are weighted using population weights based on individuals' first observation. For more detailed information see Börsch-Supan and Jürges (2005). The SHARE questionnaires and data are available at www.share-project.org.

[13]For respondents younger than age 50 no sample weights are available.

living parent is very small. Countries participating in all three waves are Austria, Germany, Sweden, the Netherlands, Spain, Italy, France, Denmark, Switzerland, and Belgium. Greece participated in wave 1 and 2; Czechia and Poland are part of SHARE since wave 2. After deleting observations with missing information, my sample includes 10,313 individuals. 3,017 women are observed three times, resulting in 23,643 person-wave observations. 3,588 women had at least one living parent when they first participated in SHARE.

Sample size and caregiving rates by country are shown in Table 2.2 for the total sample and for individuals with a living parent (the "parent sample").[14] Caregiving activities include help with personal care such as dressing, bathing or showering, eating, getting in or out of bed, or using the toilet; as well as practical household help such as help with home repairs, gardening, transportation, shopping, and household chores.[15] Help can be provided to a parent outside the household, or to a parent living in the same household. To distinguish between help during a short-term sickness from caregiving, a person providing help to another household member is only considered a caregiver if this help occurs regularly, defined as almost daily during at least three months. Caregiving is measured by an indicator variable for whether a person provided help to a parent either outside or inside the household during the 12 months prior to the interview.[16]

Caregiving rates show large geographic variation.[17] In Denmark and Sweden,

[14] Table A.1 in the Appendix shows changes in caregiving activities over time.

[15] The definition of caregiving in wave 1 and wave 2 also includes help with paperwork such as filling out forms, or settling financial or legal matters. Less than 1% of caregivers state that they solely provide help with paperwork. For consistency, these individuals are not considered to be caregivers.

[16] In wave 2, respondents who have been interviewed in wave 1 are asked whether they have provided care since the last interview, which corresponds to approximately 24 months. First time respondents are asked about caregiving activities during the last 12 months.

[17] Equality of caregiving rates in all countries can be rejected with a p-value of 0.000 for all

Table 2.2: Sample size and caregiving rates by country

		Caregiving rates		
Country	Sample size	Any	Weekly	Daily
Total sample				
Poland	528	0.07	0.05	0.04
Austria	557	0.09	0.07	0.04
Greece	866	0.09	0.07	0.05
Italy	988	0.10	0.08	0.06
Spain	717	0.11	0.08	0.08
France	1,014	0.13	0.08	0.04
Czechia	480	0.13	0.08	0.03
Switzerland	445	0.16	0.09	0.03
Germany	824	0.16	0.11	0.04
Denmark	729	0.16	0.08	0.01
Netherlands	995	0.16	0.11	0.03
Belgium	1,161	0.17	0.14	0.06
Sweden	1,009	0.20	0.08	0.02
Total	10,313	0.13	0.09	0.05

		Caregiving rates		
Country	Sample size	Any	Weekly	Daily
Parent sample				
Poland	141	0.23	0.16	0.11
Italy	324	0.26	0.22	0.14
Greece	285	0.26	0.20	0.12
Austria	158	0.26	0.21	0.13
Spain	244	0.28	0.20	0.18
France	430	0.29	0.17	0.07
Switzerland	186	0.36	0.19	0.06
Germany	302	0.40	0.27	0.09
Denmark	258	0.41	0.20	0.03
Czechia	131	0.42	0.27	0.07
Sweden	393	0.44	0.17	0.03
Netherlands	331	0.44	0.29	0.06
Belgium	405	0.46	0.36	0.14
Total	3,588	0.33	0.22	0.10

Countries ordered by caregiving rates (any frequency).
Weighted values based on individuals' first observation.

respectively 41% and 44% of respondents with living parents provide some positive amount of care, though only 3% of the respondents in either country provide care on an almost daily basis. In contrast, in Spain and Italy only 28% and 26% of daughters with a living parent provide care, however, over half of caregivers provide daily care.

2.4.2 Summary Statistics

Descriptive statistics for the remaining variables are presented in Table 2.3. To highlight the differences between caregivers and non-caregivers, Table 2.4 reports values separately for individuals who never provide care (never caregivers) and individuals who state they are caregivers in at least one interview (ever caregivers) for the parent sample. 31% of the respondents still have a mother, 13% still have a father. Ever caregivers are slightly less likely than never caregivers to have reached the official retirement age and are thus likely to face financial losses if they drop out of the labour force.[18]

Table 2.3: Descriptive statistics based on first observations (Total sample)

	Mean	St. dev.	Min	Max	N
Caregiver status					
Caregiver (any frequency)	0.13	0.33	0	1	10,313
Weekly caregiver	0.09	0.29	0	1	10,313
Daily caregiver	0.05	0.21	0	1	10,313
Labour force participation					
Employed	0.29	0.45	0	1	10,297
Weekly hours worked (main job)	10.27	17.20	0	168	10,238
Weekly hours if hours > 0	32.45	14.67	1	168	3,452

Continued on next page

frequencies of care both in the total sample and in the parent sample.

[18]The difference between never caregivers and ever caregivers just fails to be significant at the 10% level.

Table 2.3 – continued from previous page

	Mean	St. dev.	Min	Max	N
Cognitive ability					
Verbal fluency	19.06	7.07	0	67	10,239
Short-term memory	5.08	1.74	0	10	10,262
Long-term memory	3.67	1.97	0	10	10,264
Numeracy	3.19	1.09	1	5	10,295
Health					
Depression (EURO-D)	2.89	2.39	0	12	10,216
Self-perceived health	3.07	1.02	1	5	10,313
Grip stregth	27.88	6.84	2	65	9,799
Socioeconomic and demographic information					
Age	60.17	6.46	50	74	10,313
Age squared/100	36.62	7.88	25	54.76	10,313
Reached off. Retirement age	0.39	0.49	0	1	10,313
2 years to off. Retirement age	0.50	0.50	0	1	10,313
5 years to off. Retirement age	0.65	0.48	0	1	10,313
Married	0.65	0.48	0	1	10,313
Number of chronic conditions	1.55	1.45	0	10	10,313
Number of limitations with ADL	0.13	0.56	0	6	10,313
Number of limitations with IADL	0.20	0.69	0	7	10,313
Other activities	0.35	0.48	0	1	10,268
Financial distress	0.47	0.50	0	1	10,313
Primary education of less	0.32	0.47	0	1	10,292
Secondary education	0.48	0.50	0	1	10,292
Post-secondary education	0.19	0.40	0	1	10,292
Household size	2.25	1.11	1	12	10,313
Has sister(s)	0.65	0.48	0	1	10,313
Sister information missing	0.10	0.30	0	1	10,313
Has living parent(s)	0.35	0.48	0	1	10,182
Mother alive	0.31	0.46	0	1	10,219
Father alive	0.13	0.34	0	1	10,234
Only one parent alive	0.26	0.44	0	1	10,157
Parental information					
Parent in poor health	0.10	0.31	0	1	10,056
Parent in fair health	0.12	0.33	0	1	10,056
Parent in good health	0.09	0.29	0	1	10,056
Parent in very good health	0.02	0.15	0	1	10,056
Parent in same household	0.02	0.15	0	1	10,048
Parent within 5 km	0.12	0.32	0	1	10,048
Parent between 5 and 25 km	0.07	0.26	0	1	10,048
Parent between 25 and 100 km	0.06	0.23	0	1	10,048
Parent within >100 km	0.07	0.26	0	1	10,048

Weighted values based on individuals' first observation.

Table 2.4: Descriptive statistics based on first observations (Parent sample)

	Never caregivers			Ever caregivers		
	Mean	St. dev.	N	Mean	St. dev.	N
Labour force participation						
Employed	0.45	0.50	1,803	0.49	0.50	1,778
Weekly hours worked (main job)	15.72	19.08	1,792	16.90	19.80	1,756
Weekly hours if hours > 0	33.96	12.92	859	33.49	14.86	964
Cognitive ability						
Verbal fluency**	19.43	7.17	1,789	21.74	6.82	1,773
Short-term memory**	5.25	1.68	1,795	5.67	1.70	1,777
Long-term memory**	3.84	2.00	1,796	4.34	1.90	1,777
Numeracy**	3.27	1.08	1,805	3.45	1.01	1,779
Health						
Depression (EURO-D)**	2.99	2.46	1,785	2.68	2.27	1,771
Self-perceived health**	2.95	1.02	1,807	2.71	1.00	1,781
Grip stregth**	28.85	7.12	1,697	30.04	6.37	1,737
Socioeconomic and demographic information						
Age	56.03	4.97	1,807	56.02	4.65	1,781
Age squared/100	31.64	5.82	1,807	31.60	5.41	1,781
Reached off. Retirement age	0.16	0.37	1,807	0.13	0.34	1,781
2 years to off. Retirement age	0.23	0.42	1,807	0.22	0.42	1,781
5 years to off. Retirement age	0.42	0.49	1,807	0.39	0.49	1,781
Married	0.71	0.46	1,807	0.72	0.45	1,781
Number of chronic conditions	1.24	1.34	1,807	1.14	1.22	1,781
Number of limitations with ADL**	0.11	0.49	1,807	0.04	0.25	1,781
Number of limitations with IADL**	0.19	0.72	1,807	0.08	0.29	1,781
Other social activities**	0.31	0.46	1,796	0.45	0.50	1,776
Financial distress**	0.47	0.50	1,807	0.39	0.49	1,781
Primary education of less**	0.27	0.44	1,802	0.17	0.38	1,778
Secondary education	0.50	0.50	1,802	0.51	0.50	1,778
Post-secondary education**	0.23	0.42	1,802	0.32	0.47	1,778
Household size	2.47	1.12	1,807	2.45	1.09	1,781
Has sister(s)**	0.71	0.45	1,807	0.59	0.49	1,781
Sister information missing**	0.08	0.27	1,807	0.12	0.33	1,781
Mother alive*	0.87	0.33	1,806	0.91	0.29	1,778
Father alive	0.36	0.48	1,794	0.38	0.49	1,773
Only one parent alive*	0.77	0.42	1,793	0.72	0.45	1,770
Parental information						
Parent in poor health**	0.28	0.45	1,739	0.33	0.47	1,723
Parent in fair health	0.35	0.48	1,739	0.36	0.48	1,723
Parent in good health*	0.29	0.45	1,739	0.23	0.42	1,723

Continued on next page

Table 2.4 – continued from previous page

	Never caregivers			Ever caregivers		
	Mean	St. dev.	N	Mean	St. dev.	N
Parent in very good health	0.07	0.26	1,739	0.07	0.25	1,723
Parent in same household**	0.03	0.17	1,737	0.10	0.30	1,717
Parent within 5 km**	0.28	0.45	1,737	0.43	0.49	1,717
Parent between 5 and 25 km	0.21	0.41	1,737	0.21	0.41	1,717
Parent between 25 and 100 km**	0.19	0.39	1,737	0.13	0.34	1,717
Parent within > 100 km**	0.29	0.45	1,737	0.13	0.34	1,717

**; *; $^{+}$ indicates means are statistically different at the 1%, 5%, and 10% level, respectively, based on a two sided t-test. Weighted values based on individuals' first observation.

Labour Force Participation

Labour force participation is measured by an indicator variable for employment, which takes on the value of 1 if the respondent is employed or self-employed and 0 otherwise.[19] As shown in Figure 2.1, employment rates fall steadily with age.[20] Ever caregivers below the age of 63 have slightly higher employment rates than never caregivers. No apparent difference exists between ever caregivers and never caregivers for older individuals.

As an alternative to dropping out of the labour force completely, I consider changes in labour force participation on the intensive margin by looking at the number of hours a respondent worked in her main job per week. Weekly hours worked, conditional on that the individual is working, decline from about 35 hours per week to 20 hours per week with a distinct drop after the age of 65 for both ever and never caregivers.

[19] I group anyone not employed or self-employed as being non-employed.

[20] The graphs in Figure 2.1 to 2.3 pool all observations and are not meant to be interpreted in a causal fashion but merely serve illustrative purposes. In particular, note that the graphs do not distinguish between age and cohort effects and may overstate the effect of age.

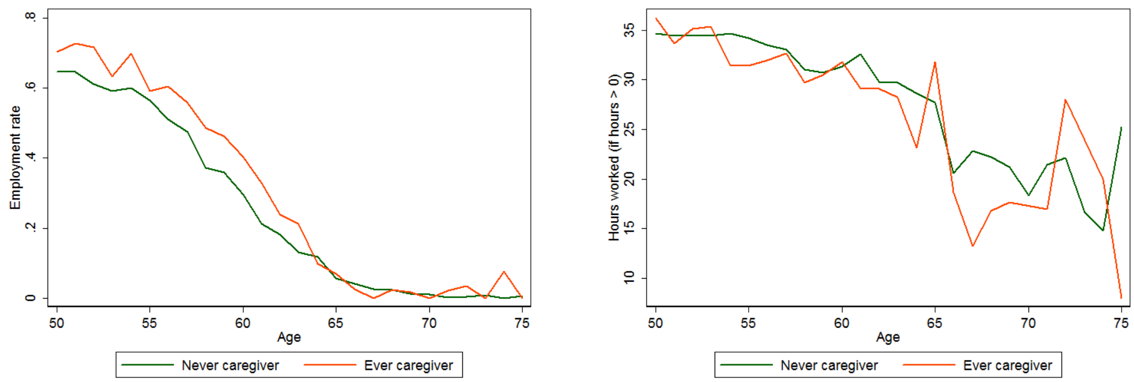


Figure 2.1: Mean employment rates and hours worked (conditional on that the individual is working) of ever and never caregivers by age

Cognitive Ability

SHARE includes several measures of cognitive ability, which differ in the domains they test. First, verbal fluency probably represents the most basic measure of cognitive ability and is measured by the number of animals the respondent can name. Each acceptable animal counts as one point up to a maximum score of 100. Second, short term memory is tested by the respondent's ability to remember a list of ten words the interviewer reads to her. The score is given by the number of correctly recalled words.[21] At the end of the cognitive function module, the respondent is asked to state as many of the words as she remembers from the list read to her earlier to measure long term memory. The score is again given by the number of correctly recalled words. Finally, numeracy skills are assessed based on a series of four math questions taken from everyday life situations and probably represent the most difficult measure

[21] In wave 1 and 2, the list of words consists of: butter, arm, letter, queen, ticket, grass, corner, stone, book, and stick. In wave 4, respondents are randomly assigned to either of the four following lists: 1) hotel, river, tree, skin, gold, market, paper, child, king, and book; 2) Sky, ocean, flag, dollar, wife, machine, home, earth, college, and butter; 3) Woman, rock, blood, corner, shoes, letter, girl, house, valley, and engine; 4) Water, church, doctor, palace, fire, garden, sea, village, baby, and table.

of cognitive ability included in SHARE. The numeracy score ranges from 1 (poor numeracy skills) to 5 (good numeracy skills).[22]

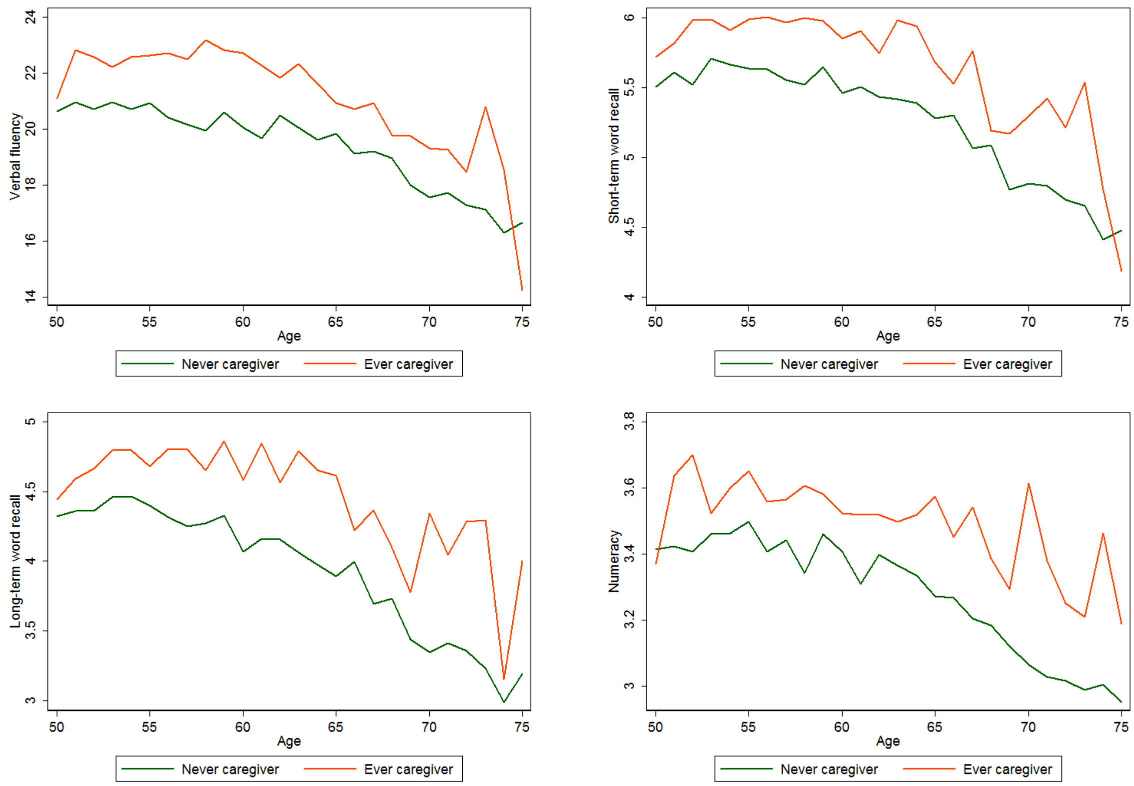


Figure 2.2: Mean cognitive ability of ever and never caregivers by age

Figure 2.2 shows the mean scores of the cognitive outcomes by age separately for ever caregivers and never caregivers. For all measures of cognitive ability, ever caregivers have higher average scores. Furthermore, while relatively constant at first, the measures of cognitive ability show a progressive decline with age, highlighting the relevance of efforts to increase or preserve cognitive ability in later life.

[22]SHARE also includes a fifth measure of cognitive ability, which tests respondents' knowledge of the year, month, and day of the interview. However, most respondents perform well on this test and the measure includes too little variation for an informative analysis. The time limit for each cognitive test is one minute.

Health

I consider both mental and physical health outcomes of caregivers. Mental health is measured by the EURO-D depression scale ranging from 0 (not depressed) to 12 (very depressed). This measure has been developed to capture symptoms of depression (including, for example, feelings of guilt, loss of interest, trouble sleeping, fatigue, or tearfulness).[23] Next, I analyze changes in self-perceived general health. Respondents are asked to evaluate their general health based on the categories excellent (1), very good (2), good (3), fair (4), and poor (5). While self-perceived health is a strong predictor of future morbidity and mortality (Idler and Benyamini, 1997; Maddox and Douglass, 1973), respondents' answering behaviour may be influenced by individuals' characteristics and their economic situation (Etilé and Milcent, 2006; Lindeboom and van Doorslaer, 2004). Especially if caregivers are faced with the declining health of a parent, caregiving may change caregivers' perception of their own health. To complement self-perceived health, I use grip strength as an objective indicator of physical health (Ziebarth, 2010; Jürges, 2007). Grip strength is measured in a gripping exercise where respondents are asked to squeeze their hand as hard as they can for a couple of seconds.[24]

Figure 2.3 shows mean health outcomes for ever caregivers and never caregiver by age. Depression scores are relatively flat at first and increase for high ages; ever caregivers have slightly fewer mental health problems. Self-perceived health and grip strength both show a clear decline of health with age. Ever caregivers have somewhat better health outcomes, though the difference between caregivers and never caregivers is small.

[23] For more information see Castro-Costa et al. (2008).

[24] Two measures are taken for each hand; I use the maximum result of the four measures.

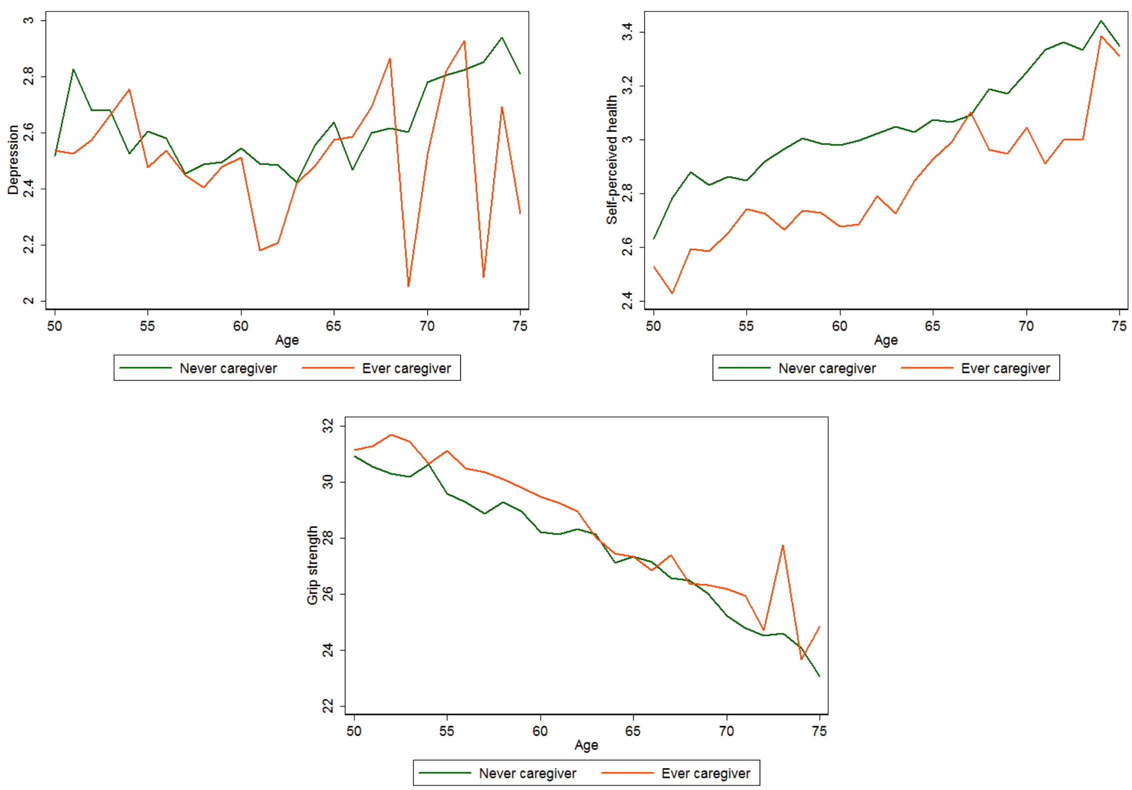


Figure 2.3: Mean health outcomes of ever and never caregivers by age

2.5 Empirical Strategy

2.5.1 Estimating the Effects of Caregiving

Using y_{it} to denote the individual i's labour market, cognitive, or health outcome of interest at time t, the FOCs from Equations 2.7 and 2.8 can be described in terms of the outcomes as follows:

$$y_{it} = f(CG_{it}, X_{it}, c_i, u_{it}), \tag{2.9}$$

where y_{it} is a function of caregiving activity, CG_{it}, individual demographic and socioeconomic characteristics, X_{it}, an individual specific error term, c_i, and an idiosyncratic error term, u_{it}, and $i = 1, ..., N$, and $t = 1, ..., T$.

Approximating Equation 2.9 by a linear function, my estimation equation is given by

$$y_{it} = \alpha_1 CG_{it} + \alpha_2 X_{it} + c_i + u_{it}, \tag{2.10}$$

$$E(u_{it}|CG_i, X_i, c_i) = 0,\ t = 1, 2, ..., T;$$

$$CG_i \equiv (CG_{i1}, CG_{i2}, ..., CG_{iT}),\ X_i \equiv (X_{i1}, X_{i2}, ..., X_{iT}).$$

As unobserved time-invariant individual characteristics captured by the individual specific error term c_i may be correlated with caregiving behaviour, I estimate Equation 2.10 using fixed effects. Fixed effects estimation assumes strict exogeneity of the explanatory variables conditional on the individual fixed effects, hence it is assumed that $E(u_{it}CG_i) = 0$ and $E(u_{it}X_i) = 0$, but allows for arbitrary correlation between the observed independent variables and c_i, that is $E(CG_i'c_i) \neq 0$ and $E(X_i'c_i) \neq 0$ (Wooldridge, 2002).[25] Allowing for unobserved, time-invariant heterogeneity in this

[25]While the linear specification technically requires a continuous dependent variable, with fixed

way appropriately describes the data because individual characteristics and personality traits, such as closeness to the parent, selflessness, or upbringing, may influence caregiving behaviour but can reasonably be assumed to be constant over time for individuals aged 50 and above. However, time-varying endogeneity may still be an issue. For example, as work and caregiving activities both compete for the individual's time endowment, an unobserved shock may change the daughter's time cost and impact both labour supply and caregiving activities. To account for such time-varying endogeneity, I use an indicator for whether at least one parent suffers from poor health as an instrument for caregiving behaviour. The measure is based on the daughter's assessment of her mother's and father's health.[26] Poor parental health increases the parent's need for care and thus encourages the daughter to provide care. Likewise, an improvement of the parent's health status or the death of the parent reduces or eliminates the need for informal care.[27] In order for parental health

effects estimation it is standard in the labour economics literature to use a linear probability model for binary dependent variables such as employment status.

[26]The possible categories for parental health changed from "very good", "good", "fair", "poor", and "very poor" in wave 1 to "excellent", "very good", "good", "fair", and "poor" in waves 2 and 4. Jürges et al. (2008) show that it is still possible to obtain a consistent measure by combining the categories "poor" and "very poor" and "very good" and "excellent". Poor parental health indicates poor health based on the newly constructed health variable. If two parents are alive parental health is defined as the health status of the parent in worse health.

[27]While self-perceived individual health is not necessarily a good objective measure of individual health (see Chapter 3 of this dissertation), the daughter's decision to provide care depends on her perception of her parent's health rather than on the parent's objective health level. Hence, potential subjectivity of the measure is less of a concern in this study. It could, however, be the case that daughters misreport their parents' health to justify not providing informal care. As daughters living close to their parents may feel more pressured to justify not providing care to their parents, I can test for such behaviour by comparing parental health for parents living close to their daughters and parents living further away from their daughters. Results are presented in Table A.2 in the Appendix. I find no significant differences in parental health based on the distance to the daughter. Besides, the question about parental health is asked at the beginning of the SHARE interview, whereas the questions about caregiving behaviour are asked later on, which further limits the possibility for strategic answering behaviour.

(PH) to be a valid instrument, parental health must influence the daughter's outcome under consideration only by affecting caregiving behaviour. Thus, in addition to $corr(CG_{it}, PH_{it}) \neq 0$, $E(u_{it}|PH_{it}) = 0$ must hold. While the level of parental health does not represent a valid instrument, fixed effects estimation only considers within individual variation, which means a change in the parent's health status serves as an instrument for a change in caregiving behaviour. Arguably, after controlling for all other covariates, a worsening of the parent's health status only influences the adult daughter's labour supply, cognitive ability or health through its effect on caregiving behaviour.[28] To estimate separate effects for family and formal care countries, I use a change in parental health interacted with the indicators for each country group as instruments, as a change in parental health may have a differential effect on daughters' caregiving decision based on the availability of formal care alternatives.[29]

Caregiving activities are measured by daily, at least weekly, or any frequency of caregiving. Which variables are included in X_{it} depends on the outcome of interest. In all regressions I control for the number of chronic conditions and limitations with activities of daily living (ADL) and instrumental ADL (IADL). Changes in these health measures capture more severe changes in health than what is expected from caregiving based on previous findings in the literature, and I thus treat them as exogenous. As the death of a partner may impact the surviving spouse's outcomes under consideration, I control for whether a person is married or living in a registered partnership. To capture changes over time, I control for a linear and quadratic effect of the respondent's age. Time-invariant covariates cannot be included in fixed effects estimation as they are part of the individual specific fixed effects. However, life

[28]This argument has also been made by Van Houtven et al. (2013) in their analysis of the effect of caregiving on work for mature men and women in the U.S..

[29]Estimation is performed using the Stata command xtivreg2 (Schaffer, 2012).

expectancy, population health, and the general education level differ across countries. As a result, the natural deterioration rate of cognitive ability or health might also differ across countries. To capture this effect, I include country dummies interacted with age (both linear and quadratic terms). To pick up additional variation over time due to a change in the economic situation or as respondents become familiar with the type of cognitive skills questions, I include two wave dummies for wave 2 and wave 4. The dummy variable for wave 2 remains one for wave 4 in order to only pick up any changes between wave 1 and wave 2.

For labour market outcomes, I also include indicator variables equaling one if the individual is above the country specific official retirement age and I control for household size defined as the number of people living in the same household.[30] As few individuals actually work until the official retirement age, I further include two indicators for when an individual is within two and within five years of the country specific official retirement age, respectively.

Labour market participation may affect both cognitive ability and health by providing cognitive stimulation or affecting physical activity and stress levels. When analyzing cognitive and health outcomes, I control for whether a person is employed (including self-employed) or not. Similarly, social participation may serve as an investment in cognitive ability or health. I capture such activities by an indicator variable which equals one if the respondent participated in any of the following activities: voluntary or charity work; attended an educational or training course; gone to a sport, social or other kind of club; taken part in activities of a religious organization; or taken part in a political or community-related organization. Further, the financial

[30]Information for country specific retirement ages is based on the official retirement age in 2010 reported by the OECD. Country specific retirement ages are listed in Table A.3 in the Appendix.

situation may impact cognitive ability and health. I measure financial distress by the household's ability to make ends meet and include an indicator variable for whether the respondent has great or some difficulties to make ends meet. Moreover, when studying daughters' mental health, I include an indicator for parental loss, which equals one if a parent passed away since the last interview. As I focus on mature daughters it seems reasonable to assume that the declining health of a parent is seen as a natural occurrence, which does not directly affect the daughter's mental health after controlling for all other covariates including the loss of a parent.

2.5.2 Evidence on the Caregiving Decision

In the following, I first provide suggestive evidence that caregivers respond to the costs and benefits of caregiving as predicted by my model and then link differences in the motivation to differences in the intensity of caregiving across countries. I estimate the probability of being a caregiver for the sample of daughters who have at least one living parent using a simple probit model for cross sectional data from wave 2. In addition, for the subsample of caregivers, I estimate an ordered probit model with the frequency of care provided as the dependent variable. Since the frequency of care is measured in discrete intervals, that is $y_i = 1, 2$, or 3 for less than weekly care, weekly (but not daily) care, and daily care provision, respectively, I estimate the effect of the explanatory variables x on caregiving frequency using the latent variable model

$$y_i^o = x_i\beta + u_i, \quad u_i \sim NID(0,1), \tag{2.11}$$

Table 2.5: Probability of care provision and frequency of care

Dependent variable:	Caregiver	Caregiving frequency
Only one parent alive	0.191**	0.355**
	(0.097)	(0.139)
Parent in poor health	0.702***	0.616***
	(0.130)	(0.193)
Parent in fair health	0.329***	0.334*
	(0.125)	(0.194)
Parent in good health	0.178	-0.031
	(0.134)	(0.200)
Parent in same household	1.401***	2.688***
	(0.168)	(0.443)
Parent within 5 km	1.005***	1.994***
	(0.117)	(0.369)
Parent between 5 and 25 km	0.621***	1.129***
	(0.127)	(0.377)
Parent between 25 and 100 km	0.231*	0.746**
	(0.132)	(0.371)
Has sister(s)	-0.202**	0.098
	(0.094)	(0.132)
Sister information missing	-0.175	0.151
	(0.129)	(0.151)
LTC exp. (% of GDP)	0.155***	-0.283***
	(0.058)	(0.079)
Caregiver allowance	0.253**	0.012
	(0.121)	(0.161)
Employed	0.031	-0.319***
	(0.095)	(0.119)
Financial distress	-0.012	0.287**
	(0.093)	(0.126)
Married	0.010	-0.013
	(0.092)	(0.142)
Age	0.169	0.310
	(0.150)	(0.225)
Age squared/100	-0.138	-0.229
	(0.125)	(0.188)
Chronic conditions	0.027	-0.025
	(0.032)	(0.047)
Limitations with ADL	-0.187	0.227
	(0.119)	(0.185)
Limitations with IADL	-0.166*	-0.373**
	(0.094)	(0.189)
Secondary education	0.116	0.176
	(0.111)	(0.163)
Post-secondary education	0.471***	0.341*
	(0.127)	(0.177)
Constant	-7.097	
	(4.514)	
Cut 1		11.391*
		(6.731)
Cut 2		12.882*
		(6.732)
Observations	2718	957
Pseudo R-squared	0.12	0.25

Standard errors clustered at household level in parentheses. **p<0.01, *p<0.05, $^{+}$p<0.1.

where y_i^o represents the underlaying continuous care frequency. For cut points $\alpha_1 < \alpha_2$, I observe

$$y_i = 1 \text{ if } y_i^o \leq \alpha_1,$$
$$y_i = 2 \text{ if } \alpha_1 < y^o \leq \alpha_2, \text{ and}$$
$$y_i = 3 \text{ if } y_i^o > \alpha_2.$$

The estimates for the ordered probit model corresponding to Equation 2.11 are obtained by maximizing the likelihood function

$$l(\beta, \alpha_1, \alpha_2) = \sum_{y_i=1} log\left(\Phi(\alpha_1 - x_i\beta)\right) + \sum_{y_i=2} log\left(\Phi(\alpha_2 - x_i\beta) - \Phi(\alpha_1 - x_i\beta)\right) + \sum_{y_i=3} log\left(\Phi(x_i\beta - \alpha_2)\right), \quad (2.12)$$

where $\Phi()$ is the cumulative standard normal distribution function.[31]

The covariates can be grouped into variables related to the parent's need for informal care, characteristics of the daughter, and institutional characteristics influencing either the parent's need for informal care or the daughter's cost of providing care. Need related variables include most importantly the parent's health, whether only one parent is alive, and whether the respondent has a sister.[32] The latter two variables provide information on whether informal care from other family members may be available and thus influences the parent's need for informal care provided by the

[31] For the special case of two discrete outcomes, the ordered probit model reduces to the regular probit model.

[32] Due to the large number of missing information, I also include an indicator for whether having a sister is not known.

daughter in my sample. Variables related to the cost of caregiving include the daughter's distance to the parent, which determines the cost of getting to the parent, and variables affecting the daughter's time cost or ability to provide care: being employed, being in financial distress, being older, being married, and having health problems. The daughter's highest level of education may influence the cost of caregiving as persons with higher education levels may find it easier to fulfill caregiving tasks. Besides, I control for whether the daughter's country of residence offers a caregiving allowance to caregivers providing financial assistance and thus reducing the cost of caregiving. In addition, I include public spending on LTC as a percentage of GDP, which serves as a proxy of the availability of formal care alternatives.[33] Formal care options may either reduce the need for informal care or reduce the cost of informal caregiving depending on whether formal care is used as a substitute or as a complement to informal care. As policies regarding LTC as well as observed caregiving rates may be influenced by social values regarding informal care, my estimates should only be interpreted as correlations but are nevertheless informative for explaining different caregiving motivations.

Results for the probit model and the ordered probit model are shown in Table 2.5. Parental need for care is strongly related to caregiving patterns. Daughters who have a parent in poor or fair health are much more likely to provide informal care and, conditional on providing informal care, are more likely to provide more frequent care. Likewise, having only one living parent increases both the probability of caregiving and the frequency of care provided; having a sister reduces the probability of providing care. Further, the cost of caregiving influences caregiving behaviour as expected.

[33] I use values for 2007, which is when most of the data for wave 2 was collected (OECD Health Data 2013). Using more recent values does not change my findings.

Proximity to one's parent is closely related to a higher likelihood of providing care and more frequent care provision. Employment does not affect the probability of being a caregiver, but employed daughters are less likely to provide frequent care. Both a higher level of spending on LTC and offering a caregiving allowance increase the probability a daughter provides care. However, high public spending on LTC is strongly related to a lower frequency of care. This pattern corresponds with Bonsang's (2009) finding that formal care options are used to complement infrequent informal care but are used as substitutes for frequent or more involved informal care and is analyzed in more detail in Figure 2.4, which shows the correlation between public expenditure on LTC (as % of GDP) and caregiving rates across countries. Countries with high public expenditures have higher weekly or less than weekly caregiving rates (as shown in the left graph), whereas countries with high LTC expenditures have lower daily caregiving rates (as shown in the right graph). As caregivers in countries with few formal care options have fewer possibilities to substitute burdensome caregiving tasks with formal care, this suggests a negative correlation between the generosity of the LTC system and the caregiving intensity.

2.6 Results

In this section, I present estimation results for labour force participation, cognitive ability, and health in turn. For the sake of brevity, I limit my discussion to the effects of caregiving on these outcomes and only address additional covariates if the estimated parameters are of special interest. Results are presented for the fixed effects estimation with parental health as an instrument for caregiving activities (FE-IV) and

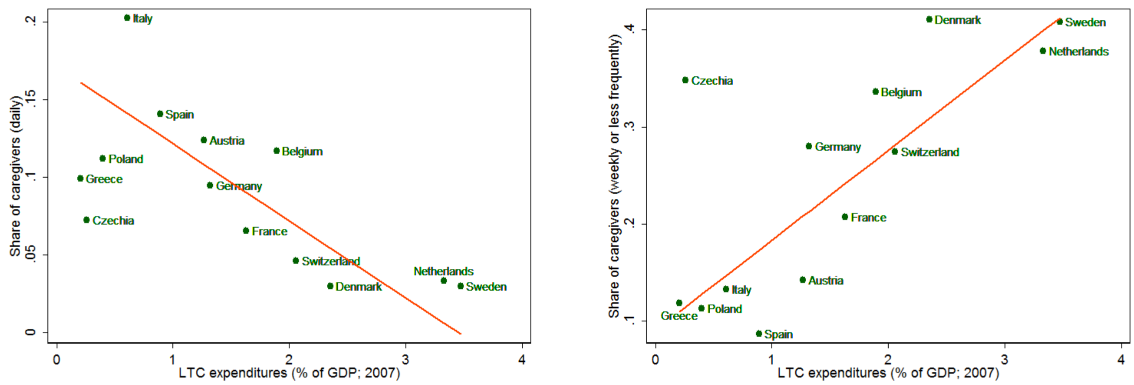


Figure 2.4: Caregiving rates and generosity of LTC systems (Weighted caregiving rates of individuals with at least one living parent in wave 2.)

without (FE).[34] My instruments are highly predictive of caregiving behaviour and satisfy the requirement for a strong instrument with all first stage F statistics greater than 10.[35] Following Fletcher and Lehrer (2009), I test whether individual fixed effects are necessary to account for time constant endogeneity by comparing OLS-IV and FE-IV estimates. Tests of joint significance of the individual specific fixed effects are statistically significant at the 1% significance level for all specifications.

Table 2.6 and 2.7 show the estimated effects of caregiving on employment for all countries combined and separately for family and formal care countries, respectively. Exogeneity of caregiving is rejected at the 10% level. In the FE-IV regression, any frequency of caregiving reduces the probability that the daughter is employed by 15 percentage points, daily caregiving reduces the probability by 28 percentage points. The point estimates in the FE regressions without the instrument are close to zero and statistically insignificant. The FE regressions do not show a negative effect of

[34]Results for the first stage regression are shown in Table A.4 to ?? in the Appendix.

[35]If exogeneity of caregiving cannot be rejected, both the FE and FE-IV estimates can be interpreted as causal effects. Results may differ slightly due to the smaller sample size when using the instrument and as the FE-IV specification captures the effect of caregiving as a result of poor parental health and such caregiving may involve more physical or emotionally demanding tasks.

Table 2.6: The effect of informal caregiving on employment

	FE-IV			FE		
Dependent variable: employed						
Frequency of care:	any	weekly	daily	any	weekly	daily
Caregiver	-0.154^{+}	-0.178^{+}	-0.278^{+}	0.005	0.008	-0.018
	(0.087)	(0.104)	(0.159)	(0.015)	(0.017)	(0.018)
Chronic conditions	0.004	0.004	0.004	0.002	0.002	0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Limitations with ADL	-0.003	-0.003	-0.004	-0.002	-0.002	-0.002
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Limitations with IADL	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)
Age	-0.143**	-0.134**	-0.127**	-0.115**	-0.116**	-0.115**
	(0.033)	(0.032)	(0.032)	(0.030)	(0.030)	(0.030)
Age squared/100	0.123**	0.117**	0.111**	0.103**	0.103**	0.103**
	(0.023)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Wave 2	-0.052**	-0.052**	-0.054**	-0.052**	-0.052**	-0.052**
	(0.019)	(0.019)	(0.019)	(0.018)	(0.018)	(0.018)
Wave 4	-0.085**	-0.085**	-0.083**	-0.077**	-0.077**	-0.077**
	(0.027)	(0.027)	(0.027)	(0.025)	(0.025)	(0.025)
Married	-0.011	-0.010	-0.005	-0.010	-0.010	-0.009
	(0.021)	(0.021)	(0.021)	(0.020)	(0.020)	(0.020)
Household size	-0.018**	-0.020**	-0.018**	-0.018**	-0.018**	-0.018**
	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)
Reached off. retirement age	-0.084**	-0.086**	-0.084**	-0.087**	-0.087**	-0.087**
	(0.014)	(0.015)	(0.015)	(0.014)	(0.014)	(0.014)
2 years to off. retirement age	-0.081**	-0.080**	-0.083**	-0.080**	-0.080**	-0.080**
	(0.016)	(0.016)	(0.016)	(0.015)	(0.015)	(0.015)
5 years to off. retirement age	-0.064**	-0.069**	-0.068**	-0.071**	-0.071**	-0.071**
	(0.018)	(0.018)	(0.018)	(0.017)	(0.017)	(0.017)
Observations	22,379	22,379	22,379	23,525	23,525	23,525
Unique individuals	9,809	9,809	9,809	10,265	10,265	10,265
R-squared	0.08	0.07	0.07	0.09	0.09	0.09
First stage F statistic	71.3	59.18	42.65			
Endogeneity test (p-value)	0.057	0.057	0.095			

Standard errors clustered at household level in parentheses. $^{**}p<0.01$, $^{*}p<0.05$, $^{+}p<0.1$
Regressions also include age*country and age squared/100*country interaction terms.

Table 2.7: The effect of informal caregiving on employment by caregiving context

	FE-IV			FE		
Dependent variable: employed						
Frequency of care:	any	weekly	daily	any	weekly	daily
Caregiver x	-0.258*	-0.285*	-0.465*	-0.014	-0.003	-0.011
family care country	(0.115)	(0.137)	(0.221)	(0.022)	(0.024)	(0.023)
Caregiver x	0.055	0.070	0.098	0.033*	0.027	-0.033
formal care country	(0.129)	(0.165)	(0.235)	(0.017)	(0.019)	(0.029)
Chronic conditions	0.004	0.005	0.005	0.002	0.002	0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Limitations with ADL	-0.003	-0.003	-0.003	-0.002	-0.002	-0.002
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Limitations with IADL	-0.001	-0.002	-0.002	-0.001	-0.001	-0.001
	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)
Age	-0.125**	-0.126**	-0.130**	-0.114**	-0.115**	-0.115**
	(0.033)	(0.032)	(0.031)	(0.030)	(0.030)	(0.030)
Age squared/100	0.111**	0.113**	0.115**	0.102**	0.103**	0.103**
	(0.023)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Wave 2	-0.052**	-0.054**	-0.051**	-0.052**	-0.052**	-0.053**
	(0.019)	(0.019)	(0.019)	(0.018)	(0.018)	(0.018)
Wave 4	-0.085**	-0.090**	-0.083**	-0.077**	-0.077**	-0.077**
	(0.027)	(0.028)	(0.027)	(0.025)	(0.025)	(0.025)
Married	-0.011	-0.010	-0.009	-0.010	-0.010	-0.009
	(0.020)	(0.020)	(0.021)	(0.020)	(0.020)	(0.020)
Household size	-0.018**	-0.019**	-0.020**	-0.018**	-0.018**	-0.018**
	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)
Reached off. retirement age	-0.082**	-0.083**	-0.085**	-0.087**	-0.087**	-0.087**
	(0.015)	(0.015)	(0.015)	(0.014)	(0.014)	(0.014)
2 years to off. retirement age	-0.078**	-0.081**	-0.081**	-0.080**	-0.080**	-0.081**
	(0.016)	(0.016)	(0.016)	(0.015)	(0.015)	(0.015)
5 years to off. retirement age	-0.066**	-0.069**	-0.071**	-0.072**	-0.071**	-0.071**
	(0.018)	(0.018)	(0.018)	(0.016)	(0.017)	(0.017)
Observations	22,379	22,379	22,379	23,525	23,525	23,525
Unique individuals	9,809	9,809	9,809	10,265	10,265	10,265
R-squared	0.07	0.06	0.04	0.09	0.09	0.09
Significant difference (p-value)	0.070	0.097	0.081	0.087	0.331	0.553
First stage F statistic (family)	23.44	18.96	11.88			
First stage F statistic (formal)	12.63	11.82	12.05			
Endogeneity test (p-value)	0.071	0.062	0.064			

Standard errors clustered at household level in parentheses. **p<0.01, *p<0.05, +p<0.1
Regressions also include age*country and age squared/100*country interaction terms.

Table 2.8: The effect of informal caregiving on hours worked

	FE-IV			FE		
Dependent variable: hours worked						
Frequency of care:	any	weekly	daily	any	weekly	daily
Caregiver	-1.217	-1.398	-2.191	0.463	0.326	-0.894
	(3.194)	(3.686)	(5.729)	(0.593)	(0.690)	(0.740)
Chronic conditions	0.096	0.099	0.098	0.010	0.010	0.014
	(0.124)	(0.125)	(0.125)	(0.121)	(0.120)	(0.121)
Limitations with ADL	0.012	0.013	0.005	-0.053	-0.055	-0.059
	(0.165)	(0.165)	(0.165)	(0.175)	(0.176)	(0.176)
Limitations with IADL	-0.138	-0.141	-0.135	-0.175	-0.174	-0.175
	(0.293)	(0.294)	(0.288)	(0.284)	(0.283)	(0.285)
Age	-4.622**	-4.544**	-4.489**	-3.537**	-3.565**	-3.540**
	(1.150)	(1.122)	(1.122)	(1.230)	(1.231)	(1.232)
Age squared/100	3.862**	3.804**	3.766**	3.080**	3.101**	3.083**
	(0.813)	(0.789)	(0.787)	(0.881)	(0.882)	(0.882)
Wave 2	-1.317+	-1.309+	-1.327+	-1.357+	-1.362+	-1.372+
	(0.757)	(0.755)	(0.758)	(0.740)	(0.740)	(0.740)
Wave 4	-2.305*	-2.298*	-2.283*	-2.139*	-2.150*	-2.170*
	(1.062)	(1.059)	(1.053)	(1.018)	(1.019)	(1.021)
Married	-0.554	-0.549	-0.506	-0.561	-0.559	-0.542
	(0.800)	(0.801)	(0.811)	(0.778)	(0.778)	(0.777)
Household size	-0.153	-0.162	-0.149	-0.133	-0.130	-0.122
	(0.295)	(0.295)	(0.295)	(0.272)	(0.272)	(0.273)
Reached off. retirement age	-1.846**	-1.860**	-1.842**	-2.054**	-2.052**	-2.051**
	(0.518)	(0.523)	(0.519)	(0.496)	(0.496)	(0.497)
2 years to off. retirement age	-1.833**	-1.833**	-1.851**	-1.968**	-1.971**	-1.989**
	(0.564)	(0.565)	(0.573)	(0.544)	(0.544)	(0.544)
5 years to off. retirement age	-1.170+	-1.211+	-1.205+	-1.563*	-1.551*	-1.551*
	(0.657)	(0.657)	(0.655)	(0.620)	(0.620)	(0.620)
Observations	22,228	22,228	22,228	23,375	23,375	23,375
Unique individuals	9,748	9,748	9,748	10,207	10,207	10,207
R-squared	0.07	0.07	0.07	0.07	0.07	0.07
First stage F statistic	68.7	57.64	41.18			
Endogeneity test (p-value)	0.565	0.586	0.822			

Standard errors clustered at household level in parentheses. $^{**}p<0.01$, $^{*}p<0.05$, $^{+}p<0.1$
Regressions also include age*country and age squared/100*country interaction terms.

Table 2.9: The effect of informal caregiving on hours worked by caregiving context

	FE-IV			FE		
Dependent variable: hours worked						
Frequency of care:	any	weekly	daily	any	weekly	daily
Caregiver x	-3.013	-3.328	-5.462	0.073	0.078	-0.990
family care country	(4.220)	(4.732)	(7.639)	(0.923)	(1.025)	(0.955)
Caregiver x	2.491	3.095	4.335	1.035^{+}	0.758	-0.653
formal care country	(4.537)	(5.656)	(8.027)	(0.552)	(0.630)	(1.004)
Chronic conditions	0.102	0.106	0.107	0.011	0.010	0.015
	(0.125)	(0.126)	(0.126)	(0.120)	(0.120)	(0.121)
Limitations with ADL	0.004	0.014	0.021	-0.055	-0.055	-0.059
	(0.165)	(0.165)	(0.167)	(0.175)	(0.176)	(0.176)
Limitations with IADL	-0.144	-0.158	-0.155	-0.176	-0.176	-0.176
	(0.300)	(0.303)	(0.289)	(0.285)	(0.283)	(0.286)
Age	-4.308**	-4.414**	-4.558**	-3.502**	-3.563**	-3.547**
	(1.175)	(1.126)	(1.127)	(1.228)	(1.230)	(1.232)
Age squared/100	3.649**	3.751**	3.827**	3.057**	3.104**	3.087**
	(0.820)	(0.784)	(0.793)	(0.879)	(0.881)	(0.882)
Wave 2	-1.305^{+}	-1.360^{+}	-1.275^{+}	-1.352^{+}	-1.365^{+}	-1.368^{+}
	(0.758)	(0.767)	(0.760)	(0.740)	(0.740)	(0.741)
Wave 4	-2.298*	-2.403*	-2.288*	-2.139*	-2.161*	-2.169*
	(1.065)	(1.082)	(1.059)	(1.019)	(1.018)	(1.021)
Married	-0.564	-0.546	-0.581	-0.563	-0.560	-0.544
	(0.796)	(0.797)	(0.807)	(0.777)	(0.777)	(0.777)
Household size	-0.135	-0.156	-0.172	-0.131	-0.130	-0.124
	(0.295)	(0.295)	(0.298)	(0.273)	(0.272)	(0.273)
Reached off. retirement age	-1.812**	-1.798**	-1.865**	-2.048**	-2.046**	-2.051**
	(0.518)	(0.524)	(0.522)	(0.497)	(0.497)	(0.497)
2 years to off. retirement age	-1.791**	-1.839**	-1.824**	-1.962**	-1.972**	-1.988**
	(0.563)	(0.566)	(0.572)	(0.543)	(0.544)	(0.543)
5 years to off. retirement age	-1.208^{+}	-1.214^{+}	-1.243^{+}	-1.563*	-1.548*	-1.550*
	(0.658)	(0.659)	(0.659)	(0.620)	(0.620)	(0.620)
Observations	22,228	22,228	22,228	23,375	23,375	23,375
Unique individuals	9,748	9,748	9,748	10,207	10,207	10,207
R-squared	0.07	0.07	0.07	0.07	0.07	0.07
Significant difference (p-value)	0.375	0.383	0.375	0.371	0.572	0.808
First stage F statistic (family)	22.88	18.46	11.40			
First stage F statistic (formal)	11.84	11.55	11.73			
Endogeneity test (p-value)	0.674	0.649	0.703			

Standard errors clustered at household level in parentheses. $^{**}p<0.01$, $^{*}p<0.05$, $^{+}p<0.1$
Regressions also include age*country and age squared/100*country interaction terms.

caregiving as the self-selection into caregiving is not controlled for. Daughters may choose to provide care when they know they can combine work with caregiving, or when they are already retired and thus caregiving does not influence their employment probability. However, when caregiving is triggered by a decline in parental health, daughters respond to the parent's need for care even if this requires them to quit working.

Separating the countries into family and formal care countries reveals heterogeneous effects based on the caregiving context. Exogeneity is again rejected at the 10% level. Caregiving in family care countries reduces employment rates by 26, 29, and 47 percentage points for any frequency of caregiving, weekly, and daily caregiving, respectively. The effect of caregiving on employment is insignificant in formal care countries and the difference between the estimates for family and formal care countries are significant at the 10% level. Negative effects of caregiving on employment are thus more severe if the intensity of caregiving is high, that is if care is provided more frequently or if few formal care options are available. For comparison, the FE specification shows any frequency of caregiving slightly increases the employment probability in formal care countries, the estimates for weekly or daily caregiving and for caregiving in family care countries are statistically insignificant.

I do not find significant effects of caregiving on hours worked when I restrict the effect to be the same in all countries, though the point estimates show a decline of hours worked for all caregiving frequencies in the FE-IV and for daily caregiving in the FE specification as shown in Table 2.8. Exogeneity of caregiving cannot be rejected. When I allow for differential effects for family and formal care countries, the point estimates for the IV-FE specification become increasingly negative for family

care countries while they become increasingly positive in formal care countries as the frequency of care increases. However, the effects are statistically insignificant (Table 2.9). Mirroring the result for employment, I find any frequency of caregiving increases the number of hours worked per week by one hour in formal care countries. Given the high standard errors of the estimates, I cannot reject that the effects of caregiving on hours worked are the same in family and formal care countries. Again, exogeneity cannot be rejected.

Tables 2.10 to 2.17 show the results for verbal fluency, short- and long-term memory, and numeracy. Exogeneity of caregiving cannot be rejected for any cognitive measure. I find caregiving, like other social activities, increases individuals' verbal fluency scores. The effects of caregiving range from 1.4 to 2.5 additional words in the FE-IV regressions but are statistically insignificant; the effects in the FE regressions are smaller in magnitude but statistically significant. Separating between family and formal care countries, large positive effects are found for caregiving in family care countries and negative (but insignificant) effects are found in formal care countries using the FE-IV specification. In the FE specification, caregiving increases verbal fluency by close to one additional word in family care countries for all caregiving frequencies; the effects are statistically significant at the 1% or 5% level. Contrary, caregiving in countries with strong formal care does not lead to a significant change in verbal fluency.

Short-term memory functioning is considered a more difficult cognitive skill than verbal fluency, and caregiving is not found to influence respondents' test score assuming the same effect in all countries. However, separating between family and formal

Table 2.10: The effect of informal caregiving on verbal fluency

	FE-IV			FE		
Dependent variable: verbal fluency						
Frequency of care:	any	weekly	daily	any	weekly	daily
Caregiver	1.385	1.620	2.501	0.568*	0.671**	0.595+
	(1.259)	(1.477)	(2.313)	(0.232)	(0.260)	(0.344)
Chronic conditions	0.156*	0.153*	0.154*	0.129*	0.128*	0.129*
	(0.066)	(0.066)	(0.066)	(0.064)	(0.064)	(0.064)
Limitations with ADL	-0.220+	-0.220+	-0.210+	-0.183	-0.184	-0.182
	(0.124)	(0.124)	(0.125)	(0.122)	(0.122)	(0.122)
Limitations with IADL	-0.192	-0.189	-0.197	-0.218+	-0.216+	-0.219+
	(0.123)	(0.123)	(0.121)	(0.119)	(0.119)	(0.118)
Age	1.061+	0.991+	0.937+	1.057*	1.026*	1.017*
	(0.552)	(0.548)	(0.546)	(0.516)	(0.516)	(0.516)
Age squared/100	-0.577	-0.524	-0.486	-0.592	-0.568	-0.562
	(0.404)	(0.399)	(0.398)	(0.377)	(0.378)	(0.377)
Wave 2	-0.240	-0.249	-0.227	-0.223	-0.226	-0.223
	(0.439)	(0.438)	(0.439)	(0.430)	(0.429)	(0.429)
Wave 4	-1.262*	-1.271*	-1.288*	-1.257*	-1.262*	-1.277*
	(0.582)	(0.580)	(0.578)	(0.561)	(0.560)	(0.560)
Married	-0.429	-0.427	-0.495	-0.400	-0.395	-0.412
	(0.345)	(0.343)	(0.353)	(0.356)	(0.354)	(0.355)
Employed	0.354	0.345	0.375	0.363	0.361	0.371
	(0.276)	(0.276)	(0.276)	(0.266)	(0.266)	(0.265)
Other activities	0.478**	0.485**	0.494**	0.428*	0.431*	0.434*
	(0.179)	(0.179)	(0.178)	(0.175)	(0.175)	(0.174)
Financial distress	-0.093	-0.085	-0.114	-0.110	-0.108	-0.113
	(0.173)	(0.173)	(0.175)	(0.169)	(0.169)	(0.169)
Observations	22,161	22,161	22,161	23,299	23,299	23,299
Unique individuals	9,711	9,711	9,711	10,164	10,164	10,164
R-squared	0.03	0.03	0.02	0.03	0.03	0.03
First stage F statistic	72.06	59.74	42.47			
Endogeneity test (p-value)	0.579	0.553	0.425			

Standard errors clustered at household level in parentheses. **p<0.01, *p<0.05, +p<0.1
Regressions also include age*country and age squared/100*country interaction terms.

Table 2.11: The effect of informal caregiving on verbal fluency by caregiving context

	FE-IV			FE		
Dependent variable: verbal fluency						
Frequency of care:	any	weekly	daily	any	weekly	daily
Caregiver x	2.666^{+}	3.026^{+}	4.763	1.022**	0.874*	1.003*
family care country	(1.557)	(1.810)	(2.946)	(0.331)	(0.353)	(0.416)
Caregiver x	-1.187	-1.472	-2.123	-0.095	0.323	-0.423
formal care country	(2.109)	(2.637)	(3.773)	(0.298)	(0.363)	(0.615)
Chronic conditions	0.152*	0.149*	0.149*	0.127*	0.127*	0.128*
	(0.066)	(0.067)	(0.067)	(0.063)	(0.064)	(0.064)
Limitations with ADL	-0.214^{+}	-0.221^{+}	-0.222^{+}	-0.180	-0.184	-0.184
	(0.124)	(0.124)	(0.125)	(0.122)	(0.122)	(0.122)
Limitations with IADL	-0.189	-0.178	-0.183	-0.218^{+}	-0.215^{+}	-0.217^{+}
	(0.125)	(0.126)	(0.121)	(0.119)	(0.119)	(0.118)
Age	0.862	0.896	0.976^{+}	1.018*	1.023*	1.038*
	(0.573)	(0.554)	(0.552)	(0.517)	(0.516)	(0.517)
Age squared/100	-0.447	-0.487	-0.521	-0.569	-0.570	-0.577
	(0.417)	(0.402)	(0.402)	(0.378)	(0.378)	(0.378)
Wave 2	-0.239	-0.205	-0.267	-0.225	-0.223	-0.239
	(0.438)	(0.439)	(0.442)	(0.429)	(0.429)	(0.429)
Wave 4	-1.256*	-1.187*	-1.290*	-1.253*	-1.252*	-1.286*
	(0.581)	(0.585)	(0.580)	(0.560)	(0.560)	(0.560)
Married	-0.438	-0.445	-0.434	-0.402	-0.395	-0.400
	(0.342)	(0.344)	(0.350)	(0.356)	(0.355)	(0.354)
Employed	0.408	0.380	0.364	0.377	0.365	0.367
	(0.281)	(0.280)	(0.276)	(0.266)	(0.266)	(0.265)
Other activities	0.464**	0.463*	0.474**	0.420*	0.427*	0.427*
	(0.180)	(0.181)	(0.179)	(0.175)	(0.175)	(0.174)
Financial distress	-0.074	-0.049	-0.109	-0.107	-0.104	-0.111
	(0.174)	(0.176)	(0.175)	(0.169)	(0.169)	(0.169)
Observations	22,161	22,161	22,161	23,299	23,299	23,299
Unique individuals	9,711	9,711	9,711	10,164	10,164	10,164
R-squared	0.03	0.02	0.02	0.03	0.03	0.03
Significant difference (p-value)	0.144	0.161	0.150	0.012	0.276	0.054
First stage F statistic (family)	23.82	18.78	12.06			
First stage F statistic (formal)	14.98	13.92	11.81			
Endogeneity test (p-value)	0.618	0.443	0.442			

Standard errors clustered at household level in parentheses. $^{**}p<0.01$, $^{*}p<0.05$, $^{+}p<0.1$
Regressions also include age*country and age squared/100*country interaction terms.

Table 2.12: The effect of informal caregiving on short-term memory

	FE-IV			FE		
Dependent variable: short-term memory						
Frequency of care:	any	weekly	daily	any	weekly	daily
Caregiver	0.008	0.010	0.015	0.068	0.054	-0.007
	(0.383)	(0.451)	(0.695)	(0.066)	(0.075)	(0.087)
Chronic conditions	-0.026	-0.026	-0.026	-0.031$^+$	-0.031$^+$	-0.031$^+$
	(0.019)	(0.019)	(0.019)	(0.018)	(0.018)	(0.018)
Limitations with ADL	-0.006	-0.006	-0.006	-0.012	-0.012	-0.012
	(0.045)	(0.045)	(0.046)	(0.043)	(0.043)	(0.043)
Limitations with IADL	-0.069$^+$	-0.069$^+$	-0.069$^+$	-0.073*	-0.073*	-0.073*
	(0.037)	(0.037)	(0.037)	(0.036)	(0.036)	(0.036)
Age	0.204	0.203	0.203	0.187	0.184	0.184
	(0.183)	(0.181)	(0.181)	(0.170)	(0.170)	(0.170)
Age squared/100	-0.167	-0.167	-0.167	-0.175	-0.173	-0.173
	(0.135)	(0.134)	(0.134)	(0.126)	(0.126)	(0.126)
Wave 2	0.125	0.125	0.125	0.165	0.164	0.164
	(0.108)	(0.108)	(0.108)	(0.106)	(0.105)	(0.106)
Wave 4	0.102	0.102	0.102	0.201	0.200	0.198
	(0.149)	(0.148)	(0.148)	(0.143)	(0.143)	(0.143)
Married	0.307*	0.307*	0.306*	0.302*	0.302*	0.302*
	(0.128)	(0.128)	(0.127)	(0.123)	(0.123)	(0.123)
Employed	0.145*	0.145*	0.145*	0.167**	0.167**	0.168**
	(0.067)	(0.067)	(0.067)	(0.063)	(0.063)	(0.064)
Other activities	0.011	0.011	0.011	-0.001	0.000	0.000
	(0.047)	(0.047)	(0.047)	(0.045)	(0.045)	(0.045)
Financial distress	-0.006	-0.006	-0.006	-0.026	-0.026	-0.025
	(0.048)	(0.048)	(0.049)	(0.047)	(0.047)	(0.047)
Observations	22,202	22,202	22,202	23,344	23,344	23,344
Unique individuals	9,730	9,730	9,730	10,185	10,185	10,185
R-squared	0.03	0.03	0.03	0.03	0.03	0.03
First stage F statistic	72.96	59.88	42.64			
Endogeneity test (p-value)	0.852	0.889	0.961			

Standard errors clustered at household level in parentheses. **p<0.01, *p<0.05, $^+$p<0.1
Regressions also include age*country and age squared/100*country interaction terms.

Table 2.13: The effect of informal caregiving on short-term memory by caregiving context

	FE-IV			FE		
Dependent variable: short-term memory						
Frequency of care:	any	weekly	daily	any	weekly	daily
Caregiver x	-0.283	-0.320	-0.499	0.045	-0.028	-0.138
family care country	(0.498)	(0.563)	(0.900)	(0.096)	(0.106)	(0.110)
Caregiver x	0.580	0.732	1.056	0.100	0.194*	0.321*
formal care country	(0.595)	(0.719)	(1.054)	(0.081)	(0.089)	(0.128)
Chronic conditions	-0.025	-0.025	0.293*	-0.031+	-0.031+	-0.031+
	(0.019)	(0.019)	(0.128)	(0.018)	(0.018)	(0.018)
Limitations with ADL	-0.008	-0.006	-0.025	-0.012	-0.012	-0.012
	(0.045)	(0.046)	(0.019)	(0.043)	(0.043)	(0.043)
Limitations with IADL	-0.069+	-0.071+	-0.004	-0.073*	-0.074*	-0.074*
	(0.037)	(0.037)	(0.046)	(0.036)	(0.036)	(0.036)
Age	0.248	0.226	-0.072+	0.189	0.185	0.177
	(0.185)	(0.183)	(0.037)	(0.170)	(0.170)	(0.170)
Age squared/100	-0.196	-0.175	0.195	-0.176	-0.172	-0.168
	(0.137)	(0.135)	(0.183)	(0.127)	(0.127)	(0.126)
Wave 2	0.124	0.115	-0.159	0.165	0.163	0.169
	(0.109)	(0.108)	(0.135)	(0.106)	(0.105)	(0.106)
Wave 4	0.100	0.082	0.134	0.201	0.196	0.201
	(0.149)	(0.149)	(0.110)	(0.143)	(0.143)	(0.143)
Married	0.309*	0.311*	0.102	0.302*	0.302*	0.298*
	(0.128)	(0.128)	(0.149)	(0.123)	(0.123)	(0.123)
Employed	0.133+	0.137*	0.148*	0.166**	0.166**	0.169**
	(0.068)	(0.068)	(0.067)	(0.063)	(0.063)	(0.063)
Other activities	0.014	0.016	0.015	0.000	0.002	0.003
	(0.048)	(0.048)	(0.047)	(0.045)	(0.045)	(0.045)
Financial distress	-0.010	-0.014	-0.007	-0.026	-0.027	-0.026
	(0.049)	(0.049)	(0.049)	(0.047)	(0.047)	(0.047)
Observations	22,202	22,202	22,202	23,344	23,344	23,344
Unique individuals	9,730	9,730	9,730	10,185	10,185	10,185
R-squared	0.03	0.03	0.03	0.03	0.03	0.03
Significant difference (p-value)	0.268	0.251	0.262	0.661	0.110	0.006
First stage F statistic (family)	23.88	18.86	12.05			
First stage F statistic (formal)	15.45	13.96	12.02			
Endogeneity test (p-value)	0.538	0.649	0.706			

Standard errors clustered at household level in parentheses. **p<0.01, *p<0.05, +p<0.1
Regressions also include age*country and age squared/100*country interaction terms.

Table 2.14: The effect of informal caregiving on long-term memory

	FE-IV			FE		
Dependent variable: long-term memory						
Frequency of care:	any	weekly	daily	any	weekly	daily
Caregiver	-0.240	-0.283	-0.436	-0.005	-0.060	-0.028
	(0.397)	(0.468)	(0.722)	(0.065)	(0.074)	(0.100)
Chronic conditions	-0.019	-0.018	-0.018	-0.025	-0.025	-0.025
	(0.021)	(0.021)	(0.021)	(0.020)	(0.021)	(0.020)
Limitations with ADL	-0.057	-0.057	-0.059	-0.061	-0.061	-0.061
	(0.043)	(0.043)	(0.044)	(0.041)	(0.041)	(0.041)
Limitations with IADL	-0.099**	-0.099**	-0.098**	-0.094**	-0.094**	-0.094**
	(0.037)	(0.037)	(0.037)	(0.036)	(0.036)	(0.036)
Age	0.657**	0.669**	0.678**	0.624**	0.624**	0.624**
	(0.191)	(0.190)	(0.190)	(0.186)	(0.186)	(0.186)
Age squared/100	-0.365*	-0.374**	-0.380**	-0.363**	-0.363**	-0.364**
	(0.144)	(0.143)	(0.142)	(0.141)	(0.141)	(0.141)
Wave 2	-0.101	-0.099	-0.103	-0.053	-0.054	-0.054
	(0.125)	(0.124)	(0.125)	(0.121)	(0.121)	(0.121)
Wave 4	-0.225	-0.224	-0.221	-0.123	-0.125	-0.123
	(0.173)	(0.173)	(0.173)	(0.166)	(0.166)	(0.166)
Married	0.304+	0.303+	0.315*	0.281+	0.281+	0.281+
	(0.158)	(0.158)	(0.154)	(0.151)	(0.151)	(0.151)
Employed	-0.099	-0.097	-0.103	-0.061	-0.060	-0.061
	(0.078)	(0.078)	(0.078)	(0.075)	(0.075)	(0.075)
Other activities	0.012	0.010	0.009	-0.005	-0.004	-0.005
	(0.055)	(0.055)	(0.055)	(0.053)	(0.053)	(0.053)
Financial distress	0.011	0.010	0.015	-0.010	-0.010	-0.010
	(0.054)	(0.054)	(0.054)	(0.052)	(0.052)	(0.052)
Observations	22,207	22,207	22,207	23,349	23,349	23,349
Unique individuals	9,732	9,732	9,732	10,187	10,187	10,187
R-squared	0.05	0.05	0.05	0.05	0.05	0.05
First stage F statistic	72.96	59.88	42.64			
Endogeneity test (p-value)	0.529	0.607	0.593			

Standard errors clustered at household level in parentheses. **p<0.01, *p<0.05, +p<0.1
Regressions also include age*country and age squared/100*country interaction terms.

Table 2.15: The effect of informal caregiving on long-term memory by caregiving context

	FE-IV			FE		
Dependent variable: long-term memory						
Frequency of care:	any	weekly	daily	any	weekly	daily
Caregiver x	-0.117	-0.134	-0.217	-0.063	-0.137	-0.134
family care country	(0.515)	(0.585)	(0.925)	(0.093)	(0.104)	(0.120)
Caregiver x	-0.483	-0.609	-0.881	0.080	0.073	0.236
formal care country	(0.612)	(0.782)	(1.131)	(0.085)	(0.093)	(0.174)
Chronic conditions	-0.019	-0.019	-0.019	-0.025	-0.025	-0.025
	(0.021)	(0.021)	(0.021)	(0.020)	(0.021)	(0.020)
Limitations with ADL	-0.057	-0.057	-0.060	-0.061	-0.061	-0.060
	(0.043)	(0.043)	(0.044)	(0.041)	(0.041)	(0.041)
Limitations with IADL	-0.099**	-0.098**	-0.097**	-0.094**	-0.094**	-0.094**
	(0.037)	(0.037)	(0.037)	(0.036)	(0.036)	(0.036)
Age	0.638**	0.659**	0.682**	0.629**	0.625**	0.619**
	(0.195)	(0.192)	(0.190)	(0.187)	(0.186)	(0.186)
Age squared/100	-0.352*	-0.370**	-0.384**	-0.366**	-0.363**	-0.360*
	(0.146)	(0.143)	(0.143)	(0.141)	(0.141)	(0.141)
Wave 2	-0.101	-0.095	-0.107	-0.053	-0.055	-0.050
	(0.125)	(0.125)	(0.126)	(0.121)	(0.121)	(0.120)
Wave 4	-0.225	-0.215	-0.221	-0.123	-0.128	-0.121
	(0.173)	(0.174)	(0.174)	(0.166)	(0.166)	(0.165)
Married	0.303+	0.302+	0.321*	0.281+	0.281+	0.278+
	(0.158)	(0.157)	(0.156)	(0.151)	(0.151)	(0.151)
Employed	-0.093	-0.094	-0.104	-0.063	-0.062	-0.060
	(0.079)	(0.079)	(0.078)	(0.075)	(0.075)	(0.075)
Other activities	0.010	0.008	0.007	-0.004	-0.003	-0.003
	(0.055)	(0.055)	(0.055)	(0.053)	(0.053)	(0.053)
Financial distress	0.013	0.013	0.015	-0.010	-0.011	-0.010
	(0.054)	(0.054)	(0.054)	(0.052)	(0.052)	(0.052)
Observations	22,207	22,207	22,207	23,349	23,349	23,349
Unique individuals	9,732	9,732	9,732	10,187	10,187	10,187
R-squared	0.05	0.05	0.05	0.05	0.05	0.05
Significant difference (p-value)	0.648	0.627	0.648	0.258	0.134	0.080
First stage F statistic (family)	23.88	18.87	12.05			
First stage F statistic (formal)	15.45	13.96	12.02			
Endogeneity test (p-value)	0.630	0.658	0.593			

Standard errors clustered at household level in parentheses. $^{**}p<0.01$, $^{*}p<0.05$, $^{+}p<0.1$
Regressions also include age*country and age squared/100*country interaction terms.

Table 2.16: The effect of informal caregiving on numeracy

	FE-IV			FE		
Dependent variable: numeracy						
Frequency of care:	any	weekly	daily	any	weekly	daily
Caregiver	-0.135	-0.157	-0.242	0.005	0.006	-0.033
	(0.164)	(0.192)	(0.295)	(0.025)	(0.028)	(0.038)
Chronic conditions	-0.008	-0.008	-0.008	-0.009	-0.009	-0.009
	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)
Limitations with ADL	0.034^{+}	0.034^{+}	0.033^{+}	0.030^{+}	0.030^{+}	0.030^{+}
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Limitations with IADL	-0.044^{*}	-0.044^{*}	-0.043^{*}	-0.042^{*}	-0.042^{*}	-0.042^{*}
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Age	0.141^{+}	0.148^{+}	0.153^{+}	0.126^{+}	0.126^{+}	0.127^{+}
	(0.079)	(0.078)	(0.078)	(0.075)	(0.075)	(0.075)
Age squared/100	-0.105^{+}	-0.110^{+}	-0.113^{+}	-0.099^{+}	-0.099^{+}	-0.100^{+}
	(0.059)	(0.058)	(0.058)	(0.056)	(0.056)	(0.056)
Wave 2	0.055	0.056	0.054	0.066	0.066	0.066
	(0.056)	(0.056)	(0.056)	(0.054)	(0.054)	(0.054)
Wave 4	-0.020	-0.019	-0.017	0.010	0.010	0.010
	(0.071)	(0.071)	(0.070)	(0.068)	(0.068)	(0.068)
Married	0.016	0.016	0.023	0.025	0.025	0.026
	(0.052)	(0.052)	(0.053)	(0.050)	(0.050)	(0.050)
Employed	-0.006	-0.005	-0.008	0.002	0.002	0.002
	(0.034)	(0.034)	(0.034)	(0.032)	(0.032)	(0.032)
Other activities	0.026	0.025	0.025	0.024	0.024	0.024
	(0.020)	(0.020)	(0.020)	(0.019)	(0.019)	(0.019)
Financial distress	-0.035	-0.036	-0.033	-0.039^{+}	-0.039^{+}	-0.039^{+}
	(0.024)	(0.024)	(0.024)	(0.023)	(0.023)	(0.023)
Observations	22,226	22,226	22,226	23,370	23,370	23,370
Unique individuals	9,744	9,744	9,744	10,200	10,200	10,200
R-squared	0.01	0.01	0.01	0.01	0.01	0.01
First stage F statistic	71.76	59.88	42.64			
Endogeneity test (p-value)	0.398	0.375	0.487			

Standard errors clustered at household level in parentheses. $^{**}p<0.01$, $^{*}p<0.05$, $^{+}p<0.1$
Regressions also include age*country and age squared/100*country interaction terms.

Table 2.17: The effect of informal caregiving on numeracy by caregiving context

	FE-IV			FE		
Dependent variable: numeracy						
Frequency of care:	any	weekly	daily	any	weekly	daily
Caregiver x family care country	-0.221	-0.251	-0.396	0.003	0.020	-0.044
	(0.220)	(0.252)	(0.395)	(0.037)	(0.040)	(0.049)
Caregiver x formal care country	0.041	0.050	0.071	0.007	-0.018	-0.006
	(0.235)	(0.287)	(0.415)	(0.030)	(0.034)	(0.055)
Chronic conditions	-0.008	-0.007	-0.007	-0.009	-0.009	-0.009
	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)
Limitations with ADL	0.034^{+}	0.034^{+}	0.034^{+}	0.030^{+}	0.030^{+}	0.030^{+}
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Limitations with IADL	-0.044^{*}	-0.045^{**}	-0.044^{**}	-0.042^{*}	-0.042^{*}	-0.042^{*}
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Age	0.154^{+}	0.154^{*}	0.150^{+}	0.127^{+}	0.126^{+}	0.126^{+}
	(0.080)	(0.078)	(0.078)	(0.075)	(0.075)	(0.075)
Age squared/100	-0.113^{+}	-0.112^{+}	-0.111^{+}	-0.099^{+}	-0.099^{+}	-0.099^{+}
	(0.059)	(0.058)	(0.058)	(0.056)	(0.056)	(0.056)
Wave 2	0.055	0.053	0.057	0.066	0.067	0.066
	(0.056)	(0.056)	(0.056)	(0.054)	(0.054)	(0.054)
Wave 4	-0.020	-0.024	-0.017	0.010	0.011	0.010
	(0.071)	(0.071)	(0.071)	(0.068)	(0.068)	(0.068)
Married	0.017	0.017	0.019	0.025	0.025	0.025
	(0.052)	(0.052)	(0.052)	(0.050)	(0.050)	(0.050)
Employed	-0.010	-0.008	-0.008	0.002	0.002	0.002
	(0.035)	(0.035)	(0.034)	(0.032)	(0.032)	(0.032)
Other activities	0.027	0.027	0.026	0.024	0.023	0.024
	(0.020)	(0.020)	(0.020)	(0.019)	(0.019)	(0.019)
Financial distress	-0.036	-0.038	-0.033	-0.039^{+}	-0.039^{+}	-0.039^{+}
	(0.024)	(0.024)	(0.024)	(0.023)	(0.023)	(0.023)
Observations	22,226	22,226	22,226	23,370	23,370	23,370
Unique individuals	9,744	9,744	9,744	10,200	10,200	10,200
R-squared	0.01	0	0	0.01	0.01	0.01
Significant difference (p-value)	0.419	0.433	0.415	0.944	0.462	0.609
First stage F statistic (family)	23.69	18.67	11.97			
First stage F statistic (formal)	14.93	14.10	12.02			
Endogeneity test (p-value)	0.604	0.536	0.678			

Standard errors clustered at household level in parentheses. $^{**}p<0.01$, $^{*}p<0.05$, $^{+}p<0.1$
Regressions also include age*country and age squared/100*country interaction terms.

care countries shows significant positive effects for weekly and daily caregiving in formal care countries in the FE specification. Weekly and daily caregiving increases the memory score by 0.2 and 0.3 words, respectively, which is larger than, for example, the effect found for being employed.

Long-term memory and numeracy are even more difficult cognitive skills than short-term memory and I do not find significant effects of caregiving on these measures. Possibly, the stress of caregiving has more negative consequences for difficult cognitive skills, in which case positive and negative effects may cancel each other out (Arpino and Bordone, 2012).

Results for health outcomes are reported in Tables 2.18 to 2.23. Caregiving has severe negative effects on caregivers' mental health. Using the FE-IV specification, I find any frequency of caregiving leads to 2, daily caregiving leads to 3 additional depressive symptoms. In comparison, parental loss only slightly reduces mental health, which makes it unlikely that parental health would have a significant effect on mental health other than through its effect on the daughter's caregiving behaviour. Other social activities lead to a small but only marginally significant or insignificant improvement in mental health. Exogeneity of caregiving is rejected at the 1% level; the FE estimates for weekly and any frequency of caregiving are still statistically significant but smaller in magnitude. Similar effects are found for caregiving in family and formal care countries.

The effects of caregiving on self-perceived health for all countries are statistically insignificant and exogeneity cannot be rejected. The point estimates show a small reduction in self-perceived health for the FE-IV specification and are close to zero for the FE specification. Likewise, differentiating between family and formal care

Table 2.18: The effect of informal caregiving on mental health

	FE-IV			FE		
Dependent variable: mental health						
Frequency of care:	any	weekly	daily	any	weekly	daily
Caregiver	2.076**	2.225**	3.123**	0.223**	0.202*	0.245+
	(0.693)	(0.744)	(1.059)	(0.082)	(0.095)	(0.137)
Parental loss	0.286*	0.205+	0.093	0.049	0.037	0.025
	(0.126)	(0.113)	(0.106)	(0.095)	(0.094)	(0.094)
Chronic conditions	0.196**	0.192**	0.193**	0.199**	0.198**	0.199**
	(0.028)	(0.028)	(0.028)	(0.027)	(0.027)	(0.027)
Limitations with ADL	0.165*	0.164*	0.175**	0.160*	0.160*	0.161*
	(0.065)	(0.065)	(0.065)	(0.063)	(0.063)	(0.063)
Limitations with IADL	0.368**	0.372**	0.361**	0.374**	0.374**	0.373**
	(0.059)	(0.059)	(0.060)	(0.058)	(0.059)	(0.059)
Age	-0.278	-0.372	-0.432+	-0.346	-0.358	-0.365+
	(0.237)	(0.235)	(0.232)	(0.217)	(0.218)	(0.218)
Age squared/100	0.243	0.315+	0.357*	0.304+	0.313+	0.317+
	(0.179)	(0.177)	(0.175)	(0.165)	(0.165)	(0.165)
Wave 2	-0.247+	-0.256+	-0.223	-0.266+	-0.267+	-0.265+
	(0.150)	(0.148)	(0.149)	(0.142)	(0.142)	(0.142)
Wave 4	0.168	0.151	0.119	0.060	0.057	0.053
	(0.201)	(0.199)	(0.200)	(0.189)	(0.189)	(0.189)
Married	-0.877**	-0.877**	-0.965**	-0.886**	-0.885**	-0.893**
	(0.199)	(0.201)	(0.200)	(0.193)	(0.193)	(0.193)
Employed	0.009	-0.006	0.031	0.009	0.008	0.011
	(0.099)	(0.099)	(0.098)	(0.092)	(0.092)	(0.092)
Other activities	-0.113+	-0.096	-0.087	-0.093	-0.091	-0.090
	(0.063)	(0.062)	(0.061)	(0.058)	(0.058)	(0.058)
Financial distress	0.201**	0.210**	0.174*	0.199**	0.200**	0.198**
	(0.071)	(0.070)	(0.074)	(0.068)	(0.068)	(0.068)
Observations	21,911	21,911	21,911	22,287	22,287	22,287
Unique individuals	9,611	9,611	9,611	9,758	9,758	9,758
R-squared	-0.01	-0.01	-0.02	0.05	0.05	0.05
First stage F statistic	49.83	50.89	42.71			
Endogeneity test (p-value)	0.004	0.003	0.003			

Standard errors clustered at household level in parentheses. **p<0.01, *p<0.05, +p<0.1
Regressions also include age*country and age squared/100*country interaction terms.

Table 2.19: The effect of informal caregiving on mental health by caregiving context

	FE-IV			FE		
Dependent variable: mental health						
Frequency of care:	any	weekly	daily	any	weekly	daily
Caregiver x	1.919*	1.958*	2.915*	0.262*	0.280*	0.168
family care country	(0.838)	(0.913)	(1.448)	(0.120)	(0.135)	(0.158)
Caregiver x	2.416*	2.797*	3.516*	0.168+	0.073	0.421
formal care country	(0.976)	(1.104)	(1.405)	(0.100)	(0.117)	(0.259)
Parental loss	0.292*	0.202+	0.093	0.049	0.039	0.025
	(0.127)	(0.114)	(0.105)	(0.095)	(0.094)	(0.094)
Chronic conditions	0.197**	0.193**	0.194**	0.199**	0.198**	0.199**
	(0.028)	(0.028)	(0.028)	(0.027)	(0.027)	(0.027)
Limitations with ADL	0.164*	0.164*	0.176**	0.161*	0.160*	0.162*
	(0.065)	(0.065)	(0.065)	(0.063)	(0.063)	(0.063)
Limitations with IADL	0.368**	0.370**	0.360**	0.374**	0.374**	0.372**
	(0.059)	(0.059)	(0.060)	(0.058)	(0.058)	(0.059)
Age	-0.250	-0.349	-0.434+	-0.351	-0.363+	-0.368+
	(0.248)	(0.243)	(0.234)	(0.217)	(0.218)	(0.218)
Age squared/100	0.225	0.304+	0.359*	0.307+	0.315+	0.320+
	(0.185)	(0.182)	(0.176)	(0.165)	(0.165)	(0.165)
Wave 2	-0.247	-0.264+	-0.219	-0.266+	-0.265+	-0.263+
	(0.151)	(0.150)	(0.149)	(0.142)	(0.142)	(0.142)
Wave 4	0.168	0.135	0.120	0.061	0.062	0.054
	(0.202)	(0.202)	(0.200)	(0.189)	(0.189)	(0.188)
Married	-0.875**	-0.874**	-0.970**	-0.886**	-0.886**	-0.895**
	(0.199)	(0.202)	(0.200)	(0.193)	(0.193)	(0.193)
Employed	0.002	-0.012	0.032	0.010	0.009	0.012
	(0.101)	(0.100)	(0.098)	(0.092)	(0.092)	(0.092)
Other activities	-0.112+	-0.093	-0.085	-0.094	-0.092	-0.089
	(0.063)	(0.062)	(0.062)	(0.059)	(0.058)	(0.058)
Financial distress	0.198**	0.203**	0.174*	0.200**	0.201**	0.198**
	(0.071)	(0.071)	(0.074)	(0.068)	(0.068)	(0.068)
Observations	21,911	21,911	21,911	22,287	22,287	22,287
Unique individuals	9,611	9,611	9,611	9,758	9,758	9,758
R-squared	-0.01	-0.01	-0.02	0.05	0.05	0.05
Significant difference (p-value)	0.673	0.537	0.769	0.547	0.248	0.405
First stage F statistic (family)	23.45	17.94	12.03			
First stage F statistic (formal)	13.51	11.78	12.17			
Endogeneity test (p-value)	0.006	0.004	0.006			

Standard errors clustered at household level in parentheses. **p<0.01, *p<0.05, +p<0.1
Regressions also include age*country and age squared/100*country interaction terms.

Table 2.20: The effect of informal caregiving on self-perceived health

	FE-IV			FE		
Dependent variable: self-perceived health						
Frequency of care:	any	weekly	daily	any	weekly	daily
Caregiver	0.215	0.248	0.391	-0.016	0.005	0.024
	(0.201)	(0.232)	(0.370)	(0.030)	(0.035)	(0.047)
Chronic conditions	0.121**	0.120**	0.120**	0.125**	0.125**	0.125**
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Limitations with ADL	0.055**	0.055**	0.056**	0.057**	0.058**	0.058**
	(0.019)	(0.019)	(0.019)	(0.018)	(0.018)	(0.018)
Limitations with IADL	0.095**	0.096**	0.094**	0.094**	0.094**	0.094**
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Age	-0.153+	-0.164+	-0.172*	-0.182*	-0.181*	-0.182*
	(0.085)	(0.085)	(0.085)	(0.082)	(0.082)	(0.082)
Age squared/100	0.127*	0.135*	0.141*	0.147*	0.146*	0.146*
	(0.065)	(0.064)	(0.064)	(0.062)	(0.062)	(0.062)
Wave 2	0.159**	0.158**	0.161**	0.153**	0.153**	0.153**
	(0.055)	(0.055)	(0.056)	(0.054)	(0.054)	(0.054)
Wave 4	0.161*	0.160*	0.157*	0.166*	0.167*	0.168*
	(0.076)	(0.076)	(0.075)	(0.073)	(0.073)	(0.073)
Married	0.071	0.071	0.060	0.085	0.085	0.084
	(0.057)	(0.058)	(0.058)	(0.057)	(0.057)	(0.057)
Employed	-0.008	-0.009	-0.004	-0.020	-0.020	-0.020
	(0.034)	(0.034)	(0.034)	(0.032)	(0.032)	(0.032)
Other activities	-0.098**	-0.097**	-0.096**	-0.096**	-0.096**	-0.096**
	(0.022)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
Financial distress	0.048*	0.049*	0.045+	0.041+	0.041+	0.040+
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Observations	22,271	22,271	22,271	23,414	23,414	23,414
Unique individuals	9,761	9,761	9,761	10,216	10,216	10,216
R-squared	0.07	0.07	0.07	0.08	0.08	0.08
First stage F statistic	73.10	61.22	42.58			
Endogeneity test (p-value)	0.250	0.325	0.288			

Standard errors clustered at household level in parentheses. **p<0.01, *p<0.05, +p<0.1
Regressions also include age*country and age squared/100*country interaction terms.

Table 2.21: The effect of informal caregiving on self-perceived healthy by caregiving context

	FE-IV			FE		
Dependent variable: self-perceived health						
Frequency of care:	any	weekly	daily	any	weekly	daily
Caregiver x	0.156	0.173	0.285	0.024	0.052	0.080
family care country	(0.246)	(0.271)	(0.446)	(0.044)	(0.048)	(0.058)
Caregiver x	0.332	0.418	0.605	-0.075*	-0.075+	-0.114
formal care country	(0.357)	(0.450)	(0.660)	(0.038)	(0.045)	(0.071)
Chronic conditions	0.121**	0.121**	0.121**	0.125**	0.125**	0.125**
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Limitations with ADL	0.055**	0.055**	0.057**	0.058**	0.058**	0.057**
	(0.019)	(0.019)	(0.019)	(0.018)	(0.018)	(0.018)
Limitations with IADL	0.095**	0.095**	0.094**	0.094**	0.094**	0.094**
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Age	-0.144	-0.158+	-0.173*	-0.185*	-0.182*	-0.179*
	(0.088)	(0.086)	(0.086)	(0.082)	(0.082)	(0.082)
Age squared/100	0.121+	0.133*	0.142*	0.149*	0.146*	0.144*
	(0.066)	(0.064)	(0.065)	(0.062)	(0.062)	(0.062)
Wave 2	0.159**	0.155**	0.163**	0.153**	0.154**	0.151**
	(0.055)	(0.055)	(0.057)	(0.054)	(0.054)	(0.054)
Wave 4	0.160*	0.155*	0.157*	0.167*	0.170*	0.166*
	(0.076)	(0.076)	(0.076)	(0.073)	(0.073)	(0.073)
Married	0.071	0.072	0.058	0.084	0.084	0.086
	(0.057)	(0.058)	(0.059)	(0.057)	(0.057)	(0.057)
Employed	-0.010	-0.011	-0.004	-0.019	-0.019	-0.021
	(0.034)	(0.034)	(0.034)	(0.032)	(0.032)	(0.032)
Other activities	-0.097**	-0.095**	-0.095**	-0.097**	-0.097**	-0.097**
	(0.022)	(0.022)	(0.022)	(0.021)	(0.021)	(0.021)
Financial distress	0.047*	0.047*	0.044+	0.041+	0.041+	0.040+
	(0.023)	(0.024)	(0.024)	(0.023)	(0.023)	(0.023)
Observations	22,271	22,271	22,271	23,414	23,414	23,414
Unique individuals	9,761	9,761	9,761	10,216	10,216	10,216
R-squared	0.07	0.07	0.06	0.08	0.08	0.08
Significant difference (p-value)	0.687	0.642	0.688	0.089	0.054	0.034
First stage F statistic (family)	23.93	19.55	12.05			
First stage F statistic (formal)	15.32	13.71	12.07			
Endogeneity test (p-value)	0.421	0.490	0.440			

Standard errors clustered at household level in parentheses. **p<0.01, *p<0.05, +p<0.1
Regressions also include age*country and age squared/100*country interaction terms.

Table 2.22: The effect of informal caregiving on grip strength

	FE-IV			FE		
Dependent variable: grip strength						
Frequency of care:	any	weekly	daily	any	weekly	daily
Caregiver	0.456	0.530	0.834	0.279	0.271	0.126
	(0.980)	(1.142)	(1.800)	(0.179)	(0.201)	(0.236)
Chronic conditions	-0.362**	-0.363**	-0.362**	-0.344**	-0.345**	-0.344**
	(0.064)	(0.064)	(0.064)	(0.063)	(0.063)	(0.063)
Limitations with ADL	-0.531**	-0.532**	-0.529**	-0.540**	-0.541**	-0.540**
	(0.171)	(0.171)	(0.170)	(0.166)	(0.166)	(0.166)
Limitations with IADL	-0.543**	-0.543**	-0.546**	-0.555**	-0.554**	-0.556**
	(0.131)	(0.131)	(0.130)	(0.128)	(0.128)	(0.128)
Age	1.751**	1.730**	1.712**	1.605**	1.591**	1.590**
	(0.575)	(0.573)	(0.573)	(0.541)	(0.542)	(0.541)
Age squared/100	-1.484**	-1.467**	-1.455**	-1.305**	-1.294**	-1.293**
	(0.441)	(0.439)	(0.438)	(0.416)	(0.417)	(0.416)
Wave 2	0.177	0.171	0.182	-0.034	-0.037	-0.036
	(0.340)	(0.340)	(0.341)	(0.335)	(0.335)	(0.335)
Wave 4	-1.003*	-1.010*	-1.014*	-1.266**	-1.272**	-1.280**
	(0.469)	(0.467)	(0.466)	(0.457)	(0.457)	(0.457)
Married	0.009	0.010	-0.012	0.037	0.039	0.034
	(0.301)	(0.301)	(0.304)	(0.291)	(0.291)	(0.292)
Employed	-0.074	-0.077	-0.068	-0.170	-0.171	-0.168
	(0.218)	(0.219)	(0.219)	(0.228)	(0.229)	(0.229)
Other activities	0.138	0.142	0.144	0.164	0.167	0.168
	(0.145)	(0.144)	(0.144)	(0.143)	(0.143)	(0.143)
Financial distress	-0.092	-0.089	-0.099	-0.001	0.001	0.000
	(0.151)	(0.150)	(0.152)	(0.145)	(0.145)	(0.145)
Observations	20,482	20,482	20,482	21,554	21,554	21,554
Unique individuals	9,017	9,017	9,017	9,452	9,452	9,452
R-squared	0.06	0.06	0.06	0.06	0.06	0.06
First stage F statistic	65.51	55.01	38.28			
Endogeneity test (p-value)	0.905	0.854	0.741			

Standard errors clustered at household level in parentheses. **p<0.01, *p<0.05, +p<0.1
Regressions also include age*country and age squared/100*country interaction terms.

Table 2.23: The effect of informal caregiving on grip strength by caregiving context

	FE-IV			FE		
Dependent variable: grip strength						
Frequency of care:	any	weekly	daily	any	weekly	daily
Caregiver x	1.869	2.143	3.515	0.487^{+}	0.510^{+}	0.240
family care country	(1.303)	(1.513)	(2.552)	(0.259)	(0.278)	(0.293)
Caregiver x	-2.427^{+}	-2.877^{+}	-4.043^{+}	-0.027	-0.142	-0.166
formal care country	(1.460)	(1.659)	(2.359)	(0.221)	(0.260)	(0.381)
Chronic conditions	-0.364**	-0.366**	-0.370**	-0.345**	-0.345**	-0.344**
	(0.064)	(0.064)	(0.065)	(0.063)	(0.063)	(0.063)
Limitations with ADL	-0.518**	-0.527**	-0.532**	-0.539**	-0.540**	-0.540**
	(0.172)	(0.172)	(0.171)	(0.166)	(0.166)	(0.166)
Limitations with IADL	-0.538**	-0.527**	-0.538**	-0.554**	-0.552**	-0.556**
	(0.131)	(0.132)	(0.130)	(0.128)	(0.128)	(0.128)
Age	1.552**	1.637**	1.778**	1.588**	1.587**	1.596**
	(0.587)	(0.576)	(0.577)	(0.541)	(0.541)	(0.541)
Age squared/100	-1.351**	-1.434**	-1.499**	-1.294**	-1.295**	-1.297**
	(0.448)	(0.440)	(0.441)	(0.416)	(0.416)	(0.416)
Wave 2	0.162	0.216	0.103	-0.036	-0.033	-0.043
	(0.342)	(0.343)	(0.345)	(0.335)	(0.335)	(0.335)
Wave 4	-1.015*	-0.920^{+}	-1.056*	-1.265**	-1.262**	-1.285**
	(0.471)	(0.473)	(0.470)	(0.457)	(0.457)	(0.457)
Married	-0.020	-0.033	0.022	0.034	0.036	0.037
	(0.301)	(0.305)	(0.312)	(0.291)	(0.291)	(0.292)
Employed	-0.008	-0.036	-0.075	-0.163	-0.166	-0.168
	(0.220)	(0.220)	(0.221)	(0.228)	(0.229)	(0.229)
Other activities	0.127	0.117	0.122	0.161	0.163	0.167
	(0.145)	(0.145)	(0.144)	(0.143)	(0.143)	(0.143)
Financial distress	-0.065	-0.046	-0.091	0.001	0.005	0.001
	(0.153)	(0.153)	(0.155)	(0.145)	(0.145)	(0.145)
Observations	20,482	20,482	20,482	21,554	21,554	21,554
Unique individuals	9,017	9,017	9,017	9,452	9,452	9,452
R-squared	0.05	0.05	0.05	0.06	0.06	0.06
Significant difference (p-value)	0.029	0.026	0.030	0.131	0.087	0.399
First stage F statistic (family)	22.02	17.07	10.30			
First stage F statistic (formal)	13.50	13.59	12.31			
Endogeneity test (p-value)	0.163	0.187	0.107			

Standard errors clustered at household level in parentheses. **p<0.01, *p<0.05, $^{+}$p<0.1
Regressions also include age*country and age squared/100*country interaction terms.

countries indicates a small, statistically insignificant reduction in self-perceived health as a result of caregiving in both family and formal care countries for the FE-IV specification. Without instruments, I find a slight improvement in self-perceived health as a result of weekly and any frequency of caregiving in formal care countries.

However, these findings are not reflected by the objective grip strength measure, which casts doubt on the reliability of self-perceived health as a valid health measure in this context because caregiving may change the caregivers point of reference for good health. While the results are statistically insignificant, the point estimates show an increase in grip strength as a result of caregiving for both the FE-IV and FE specification assuming a common effect in all countries. Exogeneity cannot be rejected. The effects differ somewhat across family and formal care countries. Caregiving in family care countries is found to increase grip strength consistent with the healthy-caregiver hypothesis, while caregiving in formal care countries is found to reduce grip strength for the FE-IV specification. Only the latter effects are statistically significant. For the FE specification, I find small positive effects for weekly and any frequency of caregiving in family care countries and negligible effects in formal care countries. The overall effects of caregiving on health are thus more difficult to evaluate. While caregiving is detrimental to caregivers' mental health in all countries, both positive and negative effects are found with respect to physical health.

2.7 Conclusion

Over 50% of European women aged between 50 and 75 with a living parent assist their parent with household work or personal care at some point in their lives.[36]

[36] The share of ever caregivers in the parent sample is 56% (Table A.1 in the Appendix).

While care receivers usually prefer informal care provided by a family member to formal care, care provision requires caregivers to reduce leisure time or work time and exposes them to additional stress, which may affect their work and well-being. This paper provides a comprehensive analysis of the effects of parental caregiving on caregivers by studying labour force participation, cognitive ability, and health outcomes of mature female caregivers in thirteen European countries and shows that caregiving can lead to both positive and negative effects. To understand the role of the intensity of caregiving as an important determinant of whether positive or negative effects prevail, I separately estimate the effects of daily, at least weekly, and any frequency of caregiving. In addition, I allow the effects of caregiving to differ between countries with strong formal LTC and countries with predominantly family based LTC to analyze the influence of the caregiving context on caregivers' outcomes and inform the policy debate on the optimal provision of elderly care. My results show that both dimensions of intensity influence the effects of caregiving on the caregiver. Contradictory effects of caregiving reported in earlier studies might therefore be driven by differences in the definition of caregiving and by the country of residence of the caregivers.

In particular, I find caregiving decreases the probability of being employed in family care countries, while caregiving does not affect labour market outcomes in formal care countries. Policies aimed at increasing informal caregiving rates will thus likely lead to a further reduction in employment rates of mature women in family care countries and additional formal care options would increase the compatibility of work and caregiving obligations in these countries. Further, my analysis shows small but positive effects of informal caregiving on cognitive ability with more pronounced effects

for more frequent caregiving. The different findings for verbal skills and memory skills may be caused by differences in the stress level of caregiving across family and formal care countries, though further research is needed to determine the direct pathways. While caregiving is similar to other social activities in that it preserves cognitive ability, caregiving has the opposite effect of other social activities with respect to mental health. Imposing a common effect across all countries, my results show a strong negative relationship between caregiving and mental health problems, with larger negative effects found for more frequent caregiving. An important policy goal in both family and formal care countries would thus be to reduce the mental burden of caregivers. As mental health problems may reduce the ability of informal caregivers to provide adequate care, such policies would not only help caregivers but would also benefit care receivers. While I find a small increase in self-perceived health for caregivers in formal care countries, using the objective grip strength measure shows a slight reduction in physical health for these countries; caregiving improves physical health in family care countries. However, the effects of caregiving on caregivers' physical health are small relative to its effects on caregivers' mental health.

This study focuses on daughters providing care to a parent because the close emotional connection creates a situation in which caregivers are more willing to provide very demanding caregiving services. In addition, cultural values can impose a strong obligation on children to provide care to their parents. While women tend to provide care more often than men, a growing number of men are becoming caregivers as the traditional gender roles soften and extending the analysis to sons providing care to a parent hence represents a natural extension of this study. In particular, the social pressure to provide informal care may be weaker for men than for women, which could

provide additional insight about the influence of the caregiving context on the effects of caregiving.

Chapter 3

Measuring Health and Inequality in Canada and the United States using comparable Health Indices

People in Canada and the U.S. often make claims regarding whose country has a better health system. Several researchers have attempted to address this question by comparing subjective health measures between the two countries, thus assuming a common definition of "good" health. Using data from the Joint Canada/U.S. Survey of Health, I generate objective health indices and show that Canadians and Americans define "good" health differently. The results indicate that by controlling for reporting heterogeneity, health differences between Americans and Canadians are largely eliminated, with the sole exception of Americans having a higher propensity to have "fair" or "poor" health. I then use these indices to calculate income and education gradients for different levels of health. I find socioeconomic inequality increases steeply with poor health and is stronger in the U.S..

3.1 Introduction

Being able to compare individuals' health statuses is crucial to assessing the quality of health systems and to evaluating health outcomes across different countries or population subgroups. Particular attention has been given to the comparison of health in the U.S. and Canada because of differences in their health systems and their geographic and cultural proximity (see, for example, Eng and Feeny, 2007; Guyatt et al., 2007; O'Neill and O'Neill, 2007; and Sanmartin et al., 2004).

Two important differences between the Canadian and American health systems are the extent of coverage and the method of funding.[1] The Canadian health system is publicly funded and provides universal, though somewhat incomplete, coverage for all Canadians.[2] Similar systems - Medicare and Medicaid - exist in the U.S., but only for certain subgroups: the population aged 65 and above, individuals under a certain income level, and disabled individuals.[3] Other Americans often receive health insurance benefits for themselves and their families through their employers. If employer-sponsored plans are not offered, one must purchase private insurance or go uninsured. These differences have yielded significant differences in health spending. In 2011, private health expenditures represented 52% of total health expenditures in the U.S., but only 30% in Canada; per capita total health expenditure amounted to US$8,508 (17.7% of GDP) in the U.S. compared to US$4,522 (11.2% of GDP) in Canada (OECD Health Data 2014).[4] Whether such differences in spending and

[1] Health care is administered by states in the U.S. and by provinces and territories in Canada, leading to slight differences in the extent of coverage within each country. Due to data limitations, this paper only focuses on cross-country differences in coverage.

[2] Dental care and prescription drugs are not covered universally, though supplemental coverage is often provided by employers.

[3] The current U.S. health-care reform aims to introduce nearly universal coverage by 2014.

[4] Values adjusted for purchasing power.

in the structure of the system translate into differences in health outcomes is still debated. While general health measures such as life expectancy at birth (81.0 years for Canadians, 78.7 years for Americans) and infant mortality (4.9 deaths per 1,000 live births in Canada, 6.1 in the U.S.) slightly favour Canada, O'Neill and O'Neill (2007) argue that these measures are misleading as they reflect the higher proportion of immigrants and higher number of pre-term births in the U.S., rather than differences in general health or in the quality of the health systems.[5]

In addition, the choice of health measure and empirical method might influence the results of health comparisons (Makdissi et al., 2011). A prevalent measure of individuals' health is self-assessed health (SAH), which is simple to collect from population surveys and can predict various objective health outcomes (Maddox and Douglass, 1973; Idler and Benyamini, 1997; Dowd and Zajacova, 2010).[6] However, using subjective health measures implicitly assumes that everybody shares the same understanding of "good" health, which has been challenged by recent research (Etilé and Milcent, 2006; Lindeboom and van Doorslaer, 2004; Jürges, 2007; and Bago d'Uva et al., 2008b). A violation of this assumption leads to reporting heterogeneity which is a concern to researchers as it affects the reliability of models where health influences individuals' decisions (Lindeboom and Kerkhofs, 2009).

Reporting styles have been shown to differ within countries by income (Etilé and Milcent, 2006), demographic characteristics such as age and gender (Lindeboom and van Doorslaer, 2004) and employment status (Kerkhofs and Lindeboom, 1995), as well

[5]All figures are from OECD Health Data 2014. Latest available years are 2009 for Canada, 2011 for the U.S..

[6]A commonly used measure is respondents' self-assessment of their general health status based on a five-point scale poor, fair, good, very good, or excellent in the "American version", and very bad, bad, fair, good, and very good in the "European version" (Bolin et al. 2010).

as across countries (Jürges, 2007; Bago d'Uva et al., 2008b). As a result, much effort has gone into obtaining internationally comparable data on health outcomes and comparable health measures to increase the comparability of cross-country studies (Sadana et al., 2000, 2002). One approach is to use vignettes, where survey respondents not only have to rate their own health but also the health of a hypothetical person described to them (Bago d'Uva et al., 2008a). Using this information, one can correct for reporting differences between individuals. However, not all surveys provide this information and the vignette approach has been criticized for its strong underlying assumptions.[7]

An alternative approach controls for reporting heterogeneity by using objective health measures such as biomarkers (Dowd and Zajacova, 2007) and grip strength (Ziebarth, 2010; Jürges, 2007), or composite health measures such as the Health Utility Index Mark 3 (HUI3) (van Doorslaer and Jones, 2003; Eng and Feeny, 2007) and the Short Form (36) Health Survey (SF-36) (Shmueli, 2003). These latter two measures are sometimes referred to as "quasi-objective": While they are based on more detailed information, they still rely on self-reports of physical or mental health problems (Ziebarth, 2010). Individual's responses on more detailed health information are generally considered to be more objective than general SAH as they refer to conditions diagnosed by a health professional or address a very specific aspect of an individual's health (Lindeboom and Kerkhofs, 2009). However, Baker et al. (2004) show that self-reports of specific ailments still suffer from misreporting. In particular,

[7]Using the English Longitudinal Survey of Ageing (ELSA) Bago d'Uva et al. (2011) show that the two necessary assumptions that form the foundation of the vignette approach are often not satisfied: "Vignette equivalence", meaning the described health state is perceived the same by all respondents, and "response consistency", requiring respondents to use the same scale to rate their own health and the health status described in the vignette, are violated in the authors' sample in all but one tests.

they find evidence that suggests people overreport health conditions to justify their absence from the labour market. As a consequence, analyses of labour market outcomes using any single specific health condition as explanatory variable may suffer from equally large bias than when SAH is used.

Nevertheless, there are several arguments in favour of their use as a measure of health. First, self-reports might provide information that is unobtainable from other sources and may be used if other sources are unavailable or too expensive to collect (Maddox and Douglass, 1973). Given that medical diagnoses are usually unavailable to researchers, using individuals' stated diagnoses is widely accepted (Pfarr et al., 2011). Second, information from a large number of self-reported health conditions and limitations can be used to derive detailed measures of health directly from survey data. Since many different health outcomes are considered and combined, over- or unerreporting of some of the conditions would only slightly misstate an individual's health and would not lead to a healthy person being classified as unhealthy or vice versa.

Using a large number of self-reported, detailed health information, Jürges (2007) and Pfarr et al. (2011) show that SAH measures are not comparable across European countries and construct health indices to correct for reporting heterogeneity. Both studies find smaller health differences between European countries when health is measured by such health indices. Jürges (2007) uses SAH as the dependent variable in an ordered probit regression on a number of quasi-objective health conditions to construct a health index that he then uses to adjust SAH for reporting heterogeneity. Pfarr et al. (2011) use a regular probit model with a binary health indicator as the dependent variable in the first step. The authors argue their method makes the index

less likely to suffer from the same reporting bias as SAH.

I use this latter approach to construct a health index using data from the Joint Canada/U.S. Survey of Health (JCUSH) from 2002-03, which has been specially designed to compare health outcomes between the two countries. For comparison and to increase the robustness of my results, I construct a second health index using principal component analysis. Principal component analysis has, for example, been used by Poterba et al. (2010), who find that an index designed using this method can be a good predictor of mortality. Principal component analysis relies on the quasi-objective health information without imposing any specific structure. It thus differs from Pfarr et al.'s (2011) method, where an empirical model is needed to compute the effect of quasi-objective health conditions on an underlying measure of health. Both methods yield very similar results, which makes it less likely that my findings depend on the empirical specification.

With the help of the constructed health indices, this paper addresses two questions. First, how does general health status compare between Canada and the U.S. once reporting heterogeneity is controlled for? Second, are the income-health and education-health gradients different for the two countries? In addressing these questions, the paper contributes to the existing literature on comparable health indices and on health differences between the U.S. and Canada in several ways. For one, it contributes a North American perspective to cross-country reporting heterogeneity, which has, for the most part, focused on European countries.[8] To my knowledge, this is the first study that tries to control for reporting heterogeneity in self-assessed health between the U.S. and Canada. Further, my health indices provide comparable health measures that enable a detailed analysis of health differences between the

[8]A notable exception is Bago d'Uva et al.(2008), who examine Asian countries.

two countries. In particular, I study the effect of socioeconomic status on health at different points of the health distribution. In doing so, I go beyond previous studies by Eng and Feeny (2007) and Sanmartin et al. (2004) by explicitly addressing the problem of heterogeneity in reporting behaviour between countries. After controlling for reporting heterogeneity, I find Americans and Canadians to be in very similar health, with the exception that more Americans suffer from "fair" or "poor" health relative to Canadians. I also extend the work of O'Neill and O'Neill (2007), by comparing health gradients between the U.S. and Canada using more detailed health measures and by focusing on the effect of socioeconomic status at different points of the health distribution. I show that gradients increase steeply with poor health. While the socioeconomic health-gradients are steeper in the U.S., socioeconomic inequality remains an issue in Canada despite universal coverage and public financing of the health system.

The remainder of the paper is organized as follows. Section 3.2 describes the data. Section 3.3 presents the construction of the health indices and compares objective health status between the U.S., and Canada. Socioeconomic inequality in health and how this inequality compares between the two countries is assessed in Section 3.4. Section 3.5 concludes.

3.2 The Joint Canada/U.S. Survey of Health

I use data from the Joint Canada/U.S. Survey of Health (JCUSH) 2002-03. The survey is specifically designed to obtain comparable data between the U.S. and Canada

and contains information on physical and mental health, income, education, and demographic characteristics.[9] Data was collected by computer-assisted telephone interviews from 3,505 Canadians and 5,183 Americans aged 18 or older living in households to obtain reliable national estimates for three age groups (18 to 44, 45 to 64 and 65 years and older), by sex. Overall response rates are 65.5% for the Canadian sample and 50.2% for the U.S. sample.[10]

Table 3.1: Summary statistics

		Canada		U.S.			
		Mean	SD	Mean	SD	Min	Max
Health variables	has health limitation	0.30	0.46	0.29	0.45	0	1
	self-assessed health	2.69	1.03	2.67	1.10	0	4
Socio-economic variables	high school graduate**	0.31	0.46	0.38	0.48	0	1
	college graduate**	0.21	0.41	0.14	0.34	0	1
	university graduate**	0.28	0.45	0.37	0.48	0	1
	household income**	4.51	2.82	5.13	3.10	0	10.6
Demographic variables	age	45.11	17.12	45.35	16.98	18	85
	female	0.51	0.50	0.52	0.50	0	1
	immigrant**	0.20	0.40	0.16	0.37	0	1
	living with partner+	0.65	0.48	0.63	0.48	0	1
	widowed	0.06	0.23	0.06	0.25	0	1
	separated/divorced**	0.08	0.27	0.11	0.31	0	1
	white**	0.82	0.38	0.72	0.45	0	1
Health care variables	doctor consultations	3.60	3.65	3.53	3.59	0	12
	hospital nights	0.64	3.41	0.63	3.09	0	31
	has health insurance**	1.00	0.00	0.89	0.32	0	1

Household income is measured in US$10,000 and right censored at US$106,000. Sample size is 3,505 for Canada and 5,183 for the U.S.. ** ; * ; + indicates means are statistically different at the 1%, 5%, and 10% level, respectively, based on a two sided t-test.

[9]The questionnaire is the same for both countries except for questions on health insurance coverage and race, owing to differences in the health systems and population composition, respectively.

[10]Sample weights are used to obtain representative results of the survey population. For more information, see the user guide published by Statistics Canada and the United States National Center for Health Statistics (2004).

Table 3.1 presents summary statistics of the socioeconomic, demographic and health-related variables relevant to my empirical analysis for Canada and the U.S.. I make use of two different health measures. First, I use a value for health limitations (from now on simply "limitations"), which is based on four questions asking to what degree health problems reduce the respondent's ability to partake in activities at home, at work, at school and in other activities such as transportation or leisure.[11] The variable equals 1 if health problems limit any kind of activity at least sometimes and 0 otherwise.[12] The second health measure is general SAH, with five categories ranging from "poor" (0) to "excellent" (4).

To estimate health gradients, I measure socioeconomic status by highest educational attainment and income. Education is grouped into four categories: less than high school, high school graduation or equivalent, college graduation (trades certificate, vocational school, community college, or CEGEP[13]) and university graduation (university or college certificate, including below bachelor). "Less than high school" serves as reference group. Income is measured by household income (in US$10,000, corrected for purchasing power) and is adjusted for household size by dividing by the square root of the number of household members (Figini, 1998).[14] Demographic control variables include age (both age and age squared are included) and gender. Lastly, I control for immigrant status, marital status (living with a partner, widowed,

[11]While overreporting of health problems that limit the ability to work of non-workers may cause so called justification bias, the size of this effect would be negligible as only 1% of respondents state that health problems only limit their ability to work.

[12]For a detailed description, see the documentation for derived variables, which is available at ftp://ftp.cdc.gov/pub/Health_Statistics/nchs/Dataset_Documentation/JCUSH/questionaires/.

[13]CEGEPs (Collge d'enseignement gnral et professionnel) are public general and vocational colleges in Quebec, Canada.

[14]Canadian dollars are converted using the average purchasing power parity of actual individual consumption in 2002 and 2003 (US$1=CAD$0.816, OECD.Stat (2012)).

separated/divorced, or single (the reference group)) and race (white or non-white)[15].

As my health indices might not capture every aspect of health, I control for health care utilization using visits to a health care professional in the past year (either 0 consultations (the reference group), 1-3, 4-11, or 12 or more consultations) and whether the person spent time in hospital (either no overnight visits (the reference group), 1-4 nights, or five or more nights) during the 12 months prior to the interview. For U.S. respondents, I include a dummy variable that takes a value of 1 if the person has insurance coverage and 0 otherwise.[16]

To avoid losing information due to item non-response, I impute missing values using a simple hot deck procedure where missing data points are stochastically replaced by draws from the non-missing data; it is assumed that information is missing at random, conditional on specified categorical control variables. I follow Reilly (1993) and use a multiple imputation approach and draw missing values from the conditional distribution of the data.[17]

3.3 Objective Health Indices

I consider two empirical strategies to provide an objective and comprehensive measure of health based on diagnosed conditions and stated health limitations, including chronic diseases, physical and mental health conditions, body weight and behavioural

[15] Information on race differs between the two countries, which prevents a more detailed analysis.

[16] All Canadians are insured under universal insurance coverage.

[17] I use the HOTDECK command in STATA (Mander and Clayton, 1999). For SAH, educational attainment, immigrant status, insurance coverage, race and marital status I condition on country, age and gender. For all detailed health questions and health limitations, I additionally condition on educational attainment and SAH and replace age by age group (ages 18-44, 45-64, or 65 and older). For income, the number of missing values is quite large, so I include a dummy variable indicating missing income in my regressions.

considerations such as smoking and physical activity. The construction of the first health index uses a probit model (the Probit-Index), and the second index uses principal component analysis (the PCA-Index). I describe these indices in turn and then use them for statistical analysis.

3.3.1 The Probit-Index

Following Pfarr et al. (2011), I use a binary health indicator, y_i, measured by the variable "limitations" (described in the data section), as a measure of the underlying latent health status of individual i. I estimate the effect of various health problems on overall health with a probit model, which defines the probability of the dependent variable equalling 1 as

$$Pr(y_i = 1|X = X_i) = \Phi(X_i'\beta), \tag{3.1}$$

where X is a set of exogenous variables and $\Phi()$ is the cumulative standard normal distribution function. Equation 3.1 can be derived from a latent variable model

$$y_i^* = X_i'\beta + u_i, \tag{3.2}$$

where u_i is assumed to be normally distributed with mean 0 and y_i^* denotes latent health. I observe

$$y_i = 1 \text{ if } y_i^* > 0,$$

$$y_i = 0 \text{ if } y_i^* \leq 0.$$

In my analysis, the explanatory variables consist of a set of health variables as well as a country and gender dummy, to allow for country and gender-specific differences in severity and impact. Categorical health variables are transformed into a set of dummy variables, indicating different levels of severity. The prevalence of each condition in Canada and the U.S. is shown in Table 3.2 (displayed from low to high severity for categorical variables), columns two and three, respectively. For example, more Americans than Canadians report high blood pressure (18% vs. 14%), but more Canadians smoke daily (35% vs. 29%) Estimating the model for the entire sample forces the impact of a certain condition to be the same in each country. This assumption ultimately allows one to isolate the effect of reporting heterogeneity, since objective health becomes comparable across individuals.[18] While it is possible for the impact of a disease to vary across countries (e.g. due to treatment options), this difference is arguably very small between the U.S. and Canada (O'Neill and O'Neill, 2007).

The coefficients of the probit regression indicate how much each health problem increases the probability of suffering from a health limitation while all other variables are held constant at their means. Thus, positive values increase the probability of bad health, whereas a negative value decreases this probability. A coefficient's value states the "disability weight" of a health problem, that is it states how much the condition reduces health. The larger the disability weight of a certain health problem, the greater is its adverse health effect.

I construct the Probit-Index by weighting each health condition by the linear prediction from the probit results and summing these raw disability weights over all health conditions for each individual. To simplify interpretation, I reverse signs

[18] I formally address the validity of this assumption in Subsection 3.3.3.

and normalize the disability weights so that the final health index lies between 0 and 1, such that 0 refers to the worst observed health state and 1 signifies perfect health. The estimation results and disability weights are shown in Table 3.2, columns four to six.[19] The disability weight of a health condition states the contribution of this condition to the health index and indicates how much this condition impacts an individual's health. Being in severe pain and having great difficulties participating in social activities is associated with the largest disability weight for the Probit-Index, while suffering from diabetes or high blood pressure only has a low disability weight. Some conditions are assigned a negative disability weight, suggesting they might lead to a health benefit, however, the probit estimates are not statistically different from zero in these cases.

3.3.2 The Principal Component Analysis-Index

The second health measure is based on a principal component analysis (PCA) of the same set of health variables. PCA is a statistical technique used to reduce the dimensionality of data, to identify patterns while retaining the largest amount of information possible. The principal components of some zero-mean variables x_1, x_2,...,x_n, is a set of linear combinations of theses variables, $a_1'x$, that solves the following problem:[20]

$$\begin{aligned} a_1 &= \arg \max_{||a||=1} Var[a'x] \\ &\vdots \\ a_k &= \arg \max_{\substack{||a||=1, \\ a \perp a_1,,...,a_{k-1}}} Var[a'x]. \end{aligned} \tag{3.3}$$

[19]McFadden's pseudo R-squared for the probit regression is 0.37.
[20]Variables with non-zero means need to be mean centred before PCA can be performed.

Table 3.2: Prevalence of conditions and disability weights

Health limiting	Prevalence		Probit-Index			PCA-Index
condition or activity$^{\mp}$	Canada N=3,505	U.S. N=5,183	. Coef.	Std. Err.	Disability weight	Disability weight
bending	0.100	0.086	0.131^{+}	(0.079)	0.007	0.010
	0.063	0.074	0.411**	(0.095)	0.020	0.012
	0.040	0.052	0.388**	(0.136)	0.019	0.016
	0.020	0.028	0.334^{+}	(0.194)	0.017	0.021
carrying	0.043	0.032	-0.141	(0.127)	-0.007	0.012
	0.029	0.036	0.093	(0.147)	0.005	0.014
	0.022	0.021	0.044	(0.247)	0.002	0.016
	0.017	0.020	-0.299	(0.241)	-0.015	0.020
grasping	0.038	0.031	-0.007	(0.145)	0.000	0.012
	0.018	0.026	-0.045	(0.197)	-0.002	0.014
	0.009	0.014	0.678*	(0.304)	0.034	0.018
pushing	0.057	0.046	0.207*	(0.105)	0.010	0.011
	0.043	0.044	0.536**	(0.127)	0.027	0.013
	0.024	0.026	0.285	(0.203)	0.014	0.015
	0.027	0.034	0.666**	(0.189)	0.033	0.019
reaching	0.033	0.033	0.203	(0.126)	0.010	0.012
	0.025	0.031	0.228	(0.193)	0.011	0.014
	0.021	0.025	0.040	(0.205)	0.002	0.018
relaxing	0.018	0.018	-0.321	(0.210)	-0.016	0.013
	0.020	0.020	0.160	(0.241)	0.008	0.016
shopping	0.034	0.029	0.327*	(0.164)	0.016	0.013
	0.032	0.031	-0.242	(0.183)	-0.012	0.015
	0.017	0.017	0.011	(0.250)	0.001	0.017
	0.013	0.018	-0.157	(0.335)	-0.008	0.021
sitting	0.056	0.049	0.173^{+}	(0.103)	0.009	0.010
	0.036	0.036	0.317*	(0.140)	0.016	0.012
	0.028	0.030	0.333^{+}	(0.193)	0.017	0.016
socializing	0.028	0.029	0.377*	(0.167)	0.019	0.013
	0.022	0.023	0.409^{+}	(0.214)	0.020	0.015
	0.011	0.015	0.824*	(0.331)	0.041	0.017
	0.012	0.012	0.905*	(0.429)	0.045	0.021
climbing stairs	0.047	0.049	0.149	(0.115)	0.007	0.012
	0.031	0.038	0.040	(0.153)	0.002	0.014
	0.020	0.024	0.400	(0.251)	0.020	0.017
	0.012	0.020	0.266	(0.247)	0.013	0.021
standing	0.071	0.067	0.275**	(0.086)	0.014	0.010
	0.046	0.053	0.374**	(0.107)	0.019	0.012
	0.039	0.042	0.266^{+}	(0.146)	0.013	0.015
	0.033	0.048	0.214	(0.179)	0.011	0.019
walking	0.050	0.057	0.249*	(0.104)	0.012	0.011
	0.041	0.044	0.142	(0.129)	0.007	0.013
	0.021	0.031	-0.022	(0.182)	-0.001	0.016
	0.021	0.035	0.143	(0.203)	0.007	0.020

Health limiting condition or activity	Prevalence Canada N=3,505	Prevalence U.S. N=5,183	Probit-Index . Coef.	Probit-Index Std. Err.	Probit-Index Disability weight	PCA-Index Disability weight
needs equipment	0.045	0.058	0.261*	(0.129)	0.013	0.015
fracture	0.049	0.053	0.263**	(0.099)	0.013	0.006
injury	0.015	0.021	-0.051	(0.152)	-0.003	0.005
pain	0.050	0.050	0.257**	(0.095)	0.013	0.009
	0.056	0.052	0.672**	(0.103)	0.033	0.010
	0.037	0.048	0.794**	(0.127)	0.040	0.013
	0.031	0.033	0.450*	(0.224)	0.022	0.017
arthritis	0.173	0.194	0.315**	(0.063)	0.016	0.011
asthma	0.067	0.076	0.399**	(0.093)	0.020	0.005
angina	0.024	0.019	0.193	(0.199)	0.010	0.014
coronary disease	0.015	0.020	0.102	(0.239)	0.005	0.013
back problems	0.106	0.093	0.272**	(0.086)	0.014	0.008
diabetes	0.048	0.068	0.057	(0.092)	0.003	0.009
had heart attack	0.036	0.035	0.073	(0.158)	0.004	0.013
high blood pressure	0.139	0.184	0.068	(0.061)	0.003	0.009
heart disease	0.049	0.057	0.183	(0.120)	0.009	0.012
lung problems	0.022	0.020	0.336+	(0.172)	0.017	0.012
other problems	0.142	0.129	0.219**	(0.073)	0.011	0.008
pulmonary disease	0.010	0.019	0.382*	(0.186)	0.019	0.013
cognitive problems	0.038	0.035	0.269*	(0.112)	0.013	0.005
	0.166	0.176	0.244**	(0.060)	0.012	0.007
	0.048	0.065	0.237*	(0.103)	0.012	0.010
	0.022	0.026	0.551**	(0.174)	0.027	0.013
depressed	0.012	0.015	0.208	(0.199)	0.010	0.006
	0.015	0.015	0.242	(0.198)	0.012	0.006
	0.021	0.021	0.239	(0.171)	0.012	0.006
	0.029	0.032	0.629**	(0.133)	0.031	0.007
	0.032	0.035	0.528**	(0.150)	0.026	0.009
emotional problems	0.154	0.164	0.287**	(0.063)	0.014	0.006
	0.030	0.046	0.176	(0.132)	0.009	0.010
hearing problems	0.018	0.021	0.266+	(0.141)	0.013	0.008
	0.012	0.016	0.687**	(0.213)	0.034	0.010
vision problems	0.538	0.518	0.098*	(0.048)	0.005	0.006
	0.014	0.016	0.508**	(0.195)	0.025	0.013
weight problems	0.482	0.434	-0.130	(0.149)	-0.006	0.003
	0.334	0.336	-0.076	(0.151)	-0.004	0.006
	0.113	0.138	-0.097	(0.158)	-0.005	0.007
	0.028	0.045	0.032	(0.174)	0.002	0.008
	0.013	0.023	0.412+	(0.212)	0.021	0.010
is moderatly active	0.257	0.206	0.004	(0.058)	0.000	-0.004
is active	0.267	0.217	0.023	(0.058)	0.001	-0.007
smokes daily	0.348	0.292	0.033	(0.051)	0.002	0.002
smokes occasionally	0.190	0.168	0.092	(0.064)	0.005	0.004

$^{+}$, *, ** indicate significance at 10%, 5%, 1% respectively. $^{\mp}$Variables are equal to 1 if the individual has trouble with the stated activity/is suffering from specified condition, 0 otherwise. For variables with several rows, the severity of the health problem increases with each row.

Solving the eigenproblem for the covariance matrix $\Sigma = Var[x]$ gives:

$$\Sigma a = \lambda \, a, \tag{3.4}$$

which yields the principal component weights, a, scores, $a'x$, and eigenvalues $\lambda_1 \geq ... \geq \lambda_n$ (Kolenikov and Angeles, 2009).

PCA assumes normality of the included variables. Thus, a key issue is how to deal with binary or ordinal variables. I use a method proposed by Kolenikov and Angeles (2009) based on polychoric and polyserial correlations. In a comparison, the authors find that their method is preferable to the commonly used alternative, outlined in Filmer and Pritchett (2001), of transforming ordinal variables into a set of dummy variables. To estimate polychoric correlation matrices, it is assumed that discrete variables are obtained from an underlying latent normally distributed variable. In a first step, the categorizing thresholds are computed. In a second step, the likelihood function is maximized with respect to the categorizing thresholds and polychoric correlation factors (see Kolenikov and Angeles (2009)). The computed correlation matrix is then used in the eigenvalue problem of Equation 3.4.

I use the first principal component, $\Sigma a_1 = \lambda_1 a_1$, as the PCA-Index.[21] As before, I reverse the sign and normalize the index to lie between 0 and 1, such that 0 refers to the worst observed health state and 1 signifies perfect health. The PCA disability weights are shown in Table 3.2, column seven. Similar to the Probit-Index, severe pain and having difficulties participating in social activities is associated with a high disability weight in the PCA-Index. In addition, the PCA-Index rates mobility

[21] The first principal component gives the direction of the greatest variability in the data, i.e. it minimizes the residual sum of squares of a regression of the data onto this line (Kolenikov and Angeles, 2009).

limitations such as having difficulties climbing stairs or walking as severely reducing well-being, while being active leads to a health improvement.

By construction, PCA disability weights increase with the severity of the condition, whereas probit disability weights may increase or decrease. Relatively low disability weights for severe health problems are plausible if comorbidities between different conditions reduce the negative health effect of an additional health problem.

3.3.3 Oaxaca-Blinder Decomposition

To address the issue of heterogeneous health effects of certain health problems across countries, I perform an Oaxaca-Blinder decomposition of my health indices to test for possible differences in the disability weights between Canada and the U.S. (Oaxaca, 1973; Blinder, 1973). I estimate health index HI for country i by the linear model

$$HI_i = X_i'\beta_i + \epsilon_i,\ E(\epsilon_i) = 0,\ i \in \{Can, US\} \tag{3.5}$$

separately for the U.S. and Canada. X contains the health information (as a set of dummy variables) used to construct the health indices and a constant. The decomposition makes use of the property of the linear model that

$$E(HI_i) = E(HI'\beta_i + \epsilon_i) = E(HI'\beta_i) + E(\epsilon_i) = E(X_i)'\beta_i, \tag{3.6}$$

since $E(\beta_i) = \beta_i$ and $E(\epsilon_i) = 0$ by assumption.

Equation 3.6 allows to decompose the predicted mean outcome difference D,

$$D = E(HI_{Can}) - E(HI_{US}), \tag{3.7}$$

into three components: the endowment effect (E), the difference in coefficients (C), and the interaction term (I). Formally,

$$D = \underbrace{[E(X_{Can}) - E(X_{US})]'\beta_{US}}_{E} + \underbrace{E(X_{US})'(\beta_{Can} - \beta_{US})}_{C} + \underbrace{[E(X_{Can}) - E(X_{US})]'(\beta_{Can} - \beta_{US})}_{I}. \tag{3.8}$$

Equation 3.8 can be estimated using OLS estimates of the parameters and replacing the expected values by the respective sample averages (Firpo et al. 2007).[22]

Differences in the prevalence of conditions are captured by the endowment effect, whereas differences in the impact of conditions on health are explained by the difference in coefficients.[23] Finally, the interaction term states the difference in means caused by differences in the prevalence of conditions and difference in coefficients occurring together. The validity of the health indices discussed above relies on the assumption that the difference in coefficients is negligible.

The results of the Oaxaca-Blinder decompositions are shown in Table 3.3. For each health index, the cross-country difference in health can almost entirely be explained by endowment differences. Both the differences in coefficients and the differences caused by the interaction effect are negligible, thus my assumption, that the impact of a disease is the same for each country, is supported. On average, health is slightly higher in Canada relative to the U.S. with a difference of 0.011 and 0.012 for the Probit-Index and PCA-Index, respectively.

A difficulty that arises with the use of health indices is how to interpret score

[22]Estimation of the three components is performed using the Stata command oaxaca introduced in Jann (2008).

[23]Equation 3.8 uses the U.S. as a reference point; that is it measures the resulting difference if the U.S. had Canada's prevalence of conditions and coefficients (Jann, 2008).

Table 3.3: Oaxaca-Blinder decomposition

	Probit-Index		PCA-Index	
	Coef.	Std. Err.	Coef.	Std. Err.
Mean health status in Canada	0.843	0.003	0.871	0.003
Mean health status in the U.S.	0.832	0.002	0.859	0.002
Difference in health statuses	0.011	0.004	0.012	0.004
Difference in health statuses explained by:				
Endowment effect	0.011	0.004	0.012	0.004
Coefficient effect	1.12e-09	5.11e-10	1.98e-10	4.77e-10
Interaction effect	5.56e-10	2.66e-10	5.93e-11	2.26e-10

differences. As a benchmark, I apply the criteria used for the HUI3 index and consider a score difference of 0.01 as potentially clinically important and a difference of 0.03 as clearly clinically important (Horsman et al., 2003).[24] Hence, average health in the U.S. and Canada is potentially different from a clinical perspective.

3.3.4 Cross-country Reporting Heterogeneity and Differences in Objective Health

This section compares the health indices with SAH to analyze reporting differences between the U.S. and Canada. Cross-country reporting heterogeneity is present if comparisons based on SAH yield different results than comparisons based on the objective health indices. To test for country-specific reporting styles, I construct (for each index) an adjusted general health variable (adjusted SAH) by dividing the index distribution into five intervals according to the categories of SAH. The threshold

[24] I implicitly assume here that changes in the HUI3 and in our health indices are comparable. As the HUI3 and my indices have similar distributions and are highly correlated (80% and 77% for the Probit-Index and the PCA-Index, respectively) this assumption does not seem unreasonably strong.

values for each interval can be arbitrary, but need to be the same across respondents. I compute the cut-off points such that if 5% of all respondents reported to be in poor health, the cut-off point between poor and fair is the fifth percentile of the (unweighted) health index distribution of both countries combined. This way, individuals are assigned a health category using the average respondent's definition of health (Jürges, 2007).

A correct self-assessment of health is strongly linked with higher educational attainment and higher household income. The self-reported and adjusted health category matches for 45% (46%) of respondents with household income in the fifth income quintile but only for 36% (40%) of respondents in the lowest household income quintile for the Probit-Index (the PCA-Index). For education, correct predictions make up 43% (46%) for respondents with a university degree and 33% (35%) for respondents with less than a high school degree for the Probit-Index (the PCA-Index).

Table 3.4 shows the (weighted) percentages of people reporting the different categories of SAH and adjusted SAH by country and standardized by age and sex. In line with the findings by Jürges (2007), cross-country differences in self-reported health become smaller for adjusted SAH. Unlike the studies by Sanmartin et al. (2004) and Eng and Feeny (2007), I do not find a significant difference in health status between Canadians and Americans for health statuses "good", "very good" or "excellent", once reporting heterogeneity is controlled for. However, both indices indicate that Americans are more often in "poor" or "fair" health than Canadians. Without adjustment, Americans report "fair" and "excellent" health more frequently, but "very good" health less frequently than Canadians. In general, Canadians rate their health

too negatively compared to the average JCUSH respondent. For the U.S., the pattern is less clear. Americans more often use extreme categories such as "poor" or "excellent" rather than the more moderate category "very good".

As a sensitivity analysis, I repeat my analysis but exclude uninsured Americans; it does not change my general findings. When I restrict the sample to whites only (to account for the different proportions of non-whites in the two countries), self-reported differences between the U.S. and Canada remain for the categories "good" and "excellent". When reporting heterogeneity is controlled for, the Probit-Index and PCA-Index find slightly more Americans to be in "fair" and "poor" health, respectively, while all other categories are not significantly different across countries. The differences are significant at the 10% level (see Table 3.4).

Above results show that Americans and Canadians differ in their response behaviour when it comes to SAH. As a consequence, health comparisons based on self-reports misstate actual differences in population health between the two countries. Moreover, whether SAH and objective health measures predict the same health category, depends on an individual's education and income. For my further empirical analysis, I rely on the objective health indices introduced above to be able to make unbiased cross-country comparisons.

3.4 Socioeconomic Inequality

In this section, I first analyze socioeconomic inequality within each country and present graphical results, then I test whether inequality differs between countries. Each time, I focus on the effect of socioeconomic factors across the health distribution.

Table 3.4: SAH and adjusted SAH categories (in %)

	Canada			U.S.		
	SAH	Probit-Index	PCA-Index	SAH	Probit-Index	PCA-Index
Full Sample						
poor	3.4^{+}	3.1*	2.8**	4.2^{+}	4.1*	4.1**
	(0.182)	(0.172)	(0.166)	(0.201)	(0.198)	(0.197)
fair	8.1**	7.7*	8.0^{+}	10.3**	9.0*	9.0^{+}
	(0.273)	(0.267)	(0.271)	(0.304)	(0.286)	(0.287)
good	28.1	26.8	26.2	26.6	26.4	26.2
	(0.449)	(0.443)	(0.440)	(0.442)	(0.441)	(0.440)
very good	36.4**	36.6	35.3	32.6**	34.7	34.6
	(0.481)	(0.482)	(0.478)	(0.469)	(0.476)	(0.476)
excellent	24.0*	25.8	27.7	26.4*	25.8	26.1
	(0.427)	(0.438)	(0.448)	(0.441)	(0.437)	(0.439)
White individuals						
poor	3.5	3.2	2.8^{+}	3.9	3.7	3.8^{+}
	(0.183)	(0.177)	(0.165)	(0.195)	(0.189)	(0.192)
fair	8.3	8.0^{+}	8.5	8.0	9.3^{+}	8.7
	(0.276)	(0.272)	(0.279)	(0.271)	(0.290)	(0.283)
good	26.3	27.8	27.1	25.7	27.0	27.2
	(0.440)	(0.448)	(0.444)	(0.437)	(0.444)	(0.445)
very good	37.3^{+}	36.3	35.3	34.8^{+}	34.8	34.7
	(0.484)	(0.481)	(0.478)	(0.476)	(0.477)	(0.476)
excellent	24.6*	24.6	26.4	27.6*	25.2	25.5
	(0.431)	(0.431)	(0.441)	(0.447)	(0.434)	(0.436)
Individuals with health insurance						
poor	3.4*	3.1*	2.8**	4.4*	4.1*	4.1**
	(0.182)	(0.172)	(0.166)	(0.204)	(0.198)	(0.199)
fair	8.1*	7.7*	8.0^{+}	9.6*	9.2*	9.1^{+}
	(0.273)	(0.267)	(0.271)	(0.295)	(0.289)	(0.287)
good	28.1*	26.8	26.2	25.6*	26.3	26.6
	(0.449)	(0.443)	(0.440)	(0.436)	(0.440)	(0.442)
very good	36.4*	36.6	35.3	33.4*	34.7	33.8
	(0.481)	(0.482)	(0.478)	(0.472)	(0.476)	(0.473)
excellent	24.0**	25.8	27.7	27.0**	25.8	26.4
	(0.427)	(0.438)	(0.448)	(0.444)	(0.437)	(0.441)

Sample size for the full sample/white individuals/individuals with health insurance is 3,505/2,913/3,505 for Canada; 5,183/3,998/4,653 for the U.S.. $^{+}$, *, ** indicate significant difference between the U.S. and Canada at 10%, 5%, 1% respectively. Age and sex adjusted percentages. Standard deviation in parentheses.

There are several reasons to expect income and education to be positively correlated with health (see, for example, Cutler et al. 2008). Highly educated people may find it easier to navigate through the health system, communicate their medical needs or follow a treatment plans. In addition, healthy behaviours such as physical activity, healthy body weight and non-smoking are strongly linked with higher education (O'Neill and O'Neill, 2007; Jürges, 2009). With respect to income, more affluent people can better afford insurance coverage or treatments that are not covered under their insurance plan. Similarly, costly services inaccessible to lower-income individuals, such as personal training or nutritional advice, may facilitate a healthy lifestyle. While a public health system tries to eliminate socioeconomic health inequality as much as possible, interconnections between health and socioeconomic status make it difficult for any health system to overcome socioeconomic differences in health outcomes. For this reason, and because my analysis relies on a cross-section of individuals, I do not attempt a causal interpretation of the relationship between socioeconomic status and health.

3.4.1 Health Gradients

Using the health indices introduced above, I now turn to analyzing socioeconomic inequality in the U.S. and Canada by estimating education-health and income-health gradients for the two countries. Such socioeconomic gradients in health describe the universally found link between a lower socioeconomic position and worse health outcomes (Feinstein, 1993). I estimate socioeconomic gradients by the unconditional quantile regression estimator proposed by Firpo et al. (2009), which allows me to

analyze the effect of income and education on different percentiles of the (unconditional) distribution of the health indices. Contrary to ordinary least squares (OLS) regression analysis, which looks at average effects, this approach does not force the effect of socioeconomic status on health to be the same for individuals in poor and in good health. Hence, I can determine whether socioeconomic status affects people with different health levels differently.

In an additive model of health,

$$HI_i = X_i'\beta_i + \epsilon_i,\ E(\epsilon_i) = 0,\ i \in \{Can, US\}, \tag{3.9}$$

where X stands for the socioeconomic, demographic and health care utilization-related variables described in Section 3.2, the gradients are given by the unconditional quantile partial effect (UQPE)

$$UQPE = \frac{\partial q_\tau}{\partial x}. \tag{3.10}$$

q_τ denotes the τth quantile of the health index, HI, and x denotes socioeconomic status, which is measured by household income and educational attainment to compute the income-health and education-health gradients, respectively.[25] The UQPE for income, for example, states the effect on the health index at the τth quantile from a small increase in household income, while all other variables are held constant.

Let F_{HI} stand for the distribution of the health index. The estimation of the UQPEs relies on the influence function (IF), which gives the influence of an individual

[25]The results in this section are based on individuals with household income less than US$106,000 (higher incomes are right censored and not comparable across countries). The reduced sample size is 3,327 and 4,861 for Canada and the U.S., respectively.

observation on the quantile τ, and the recentred influence function (RIF). For quantile τ, the IF is given by

$$IF(HI; q_\tau, F_{HI}) = \frac{\tau - I(HI \leq q_\tau)}{f_{HI}(q_\tau)}, \tag{3.11}$$

where f_{HI} denotes the density function of the health index and $I()$ is an indicator function which equals 1 if its argument is true and 0 otherwise. The recentred influence function is defined by

$$RIF(HI; q_\tau, F_{HI}) = IF(HI; q_\tau, F_{HI}) + q_\tau, \tag{3.12}$$

which in expectation is equal to q_τ.

Firpo et al. (2009) show that the UQPEs are given by the parameter estimates of the "unconditional quantile regression", an ordinary least squares regression of the estimated RIF on the control variables.[26] The estimates of this unconditional quantile regression state the effect of a small change in the explanatory variables on the unconditional quantile of the health index (Ding and Lehrer, 2011.

The density functions required for the estimation of the RIF are estimated using Gaussian kernels. As the distributions of the health indices are strongly skewed to the left with most people being in relatively good health, I follow Silverman (1986) to avoid oversmoothing and calculate the optimal bandwidth, h, as

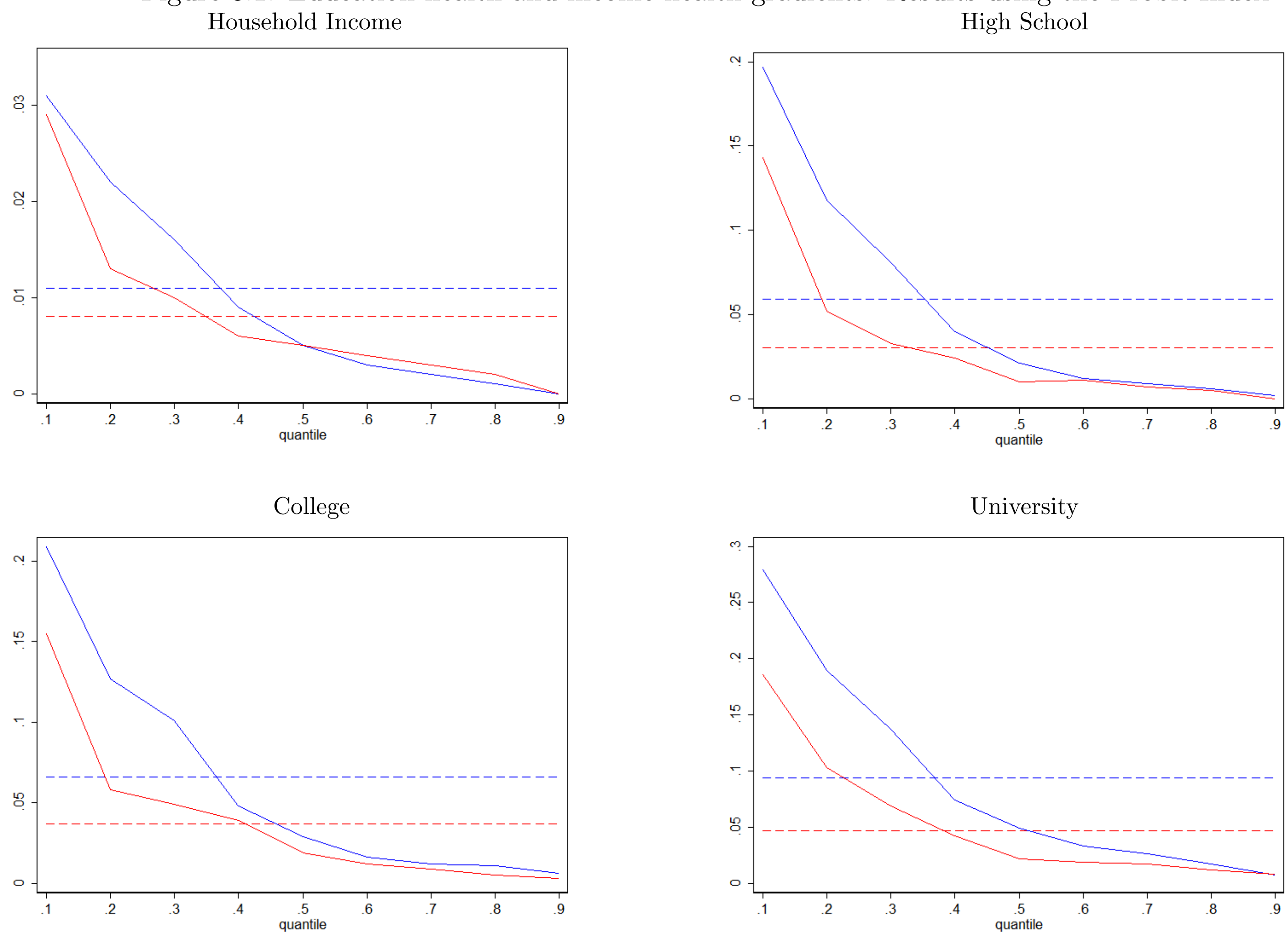
$$h = 0.9\, min(\hat{\sigma}, IQR/1.349)\, n^{-1/5},$$

[26] I use the Stata program rifreg to estimate the unconditional quantiles. The program is downloadable from Nicole M. Fortini's homepage http://www.econ.ubc.ca/nfortin/hmpgfort.htm. Their method yields consistent estimates if $Pr[HI > q_\tau | X = X_i]$ is linear in X_i, where X denotes the control variables (Firpo et al., 2009).

where $\hat{\sigma}$ is the standard deviation, IQR is the interquartile range defined by the difference between the third and first quantiles of the health index, and n is the sample size. The formula yields optimal bandwidths between 0.01 and 0.02, though the results are not very sensitive to small changes of the bandwidth. As I am interested in comparing socioeconomic inequality across the U.S. and Canada, I run separate regressions for each country.

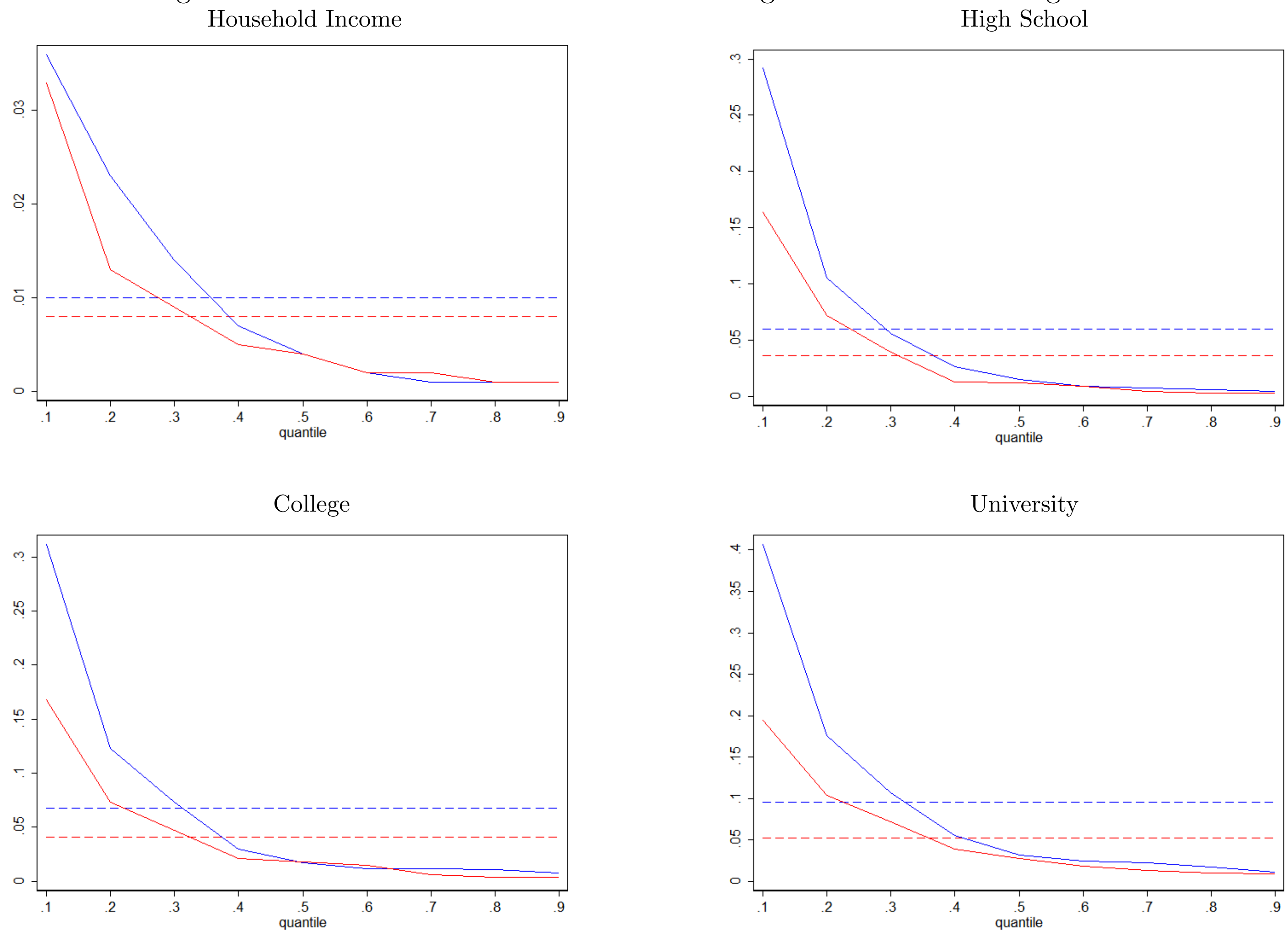
The results for the income and education variables for both OLS and unconditional quantile regressions are shown in Figure 3.1 for the Probit-Index and in Figure 3.2 for the PCA-Index. In both Canada and the U.S., higher household income is associated with better health at all levels of the health distribution, though the effect is much higher for lower levels of health. An increase in household income of US$10,000 for all households would increase the health status (as measured by the Probit-Index) of people in the lowest health decile by approximately 0.03 in both countries, which is considered clinically important. The median health level would only improve by approximately 0.005, which even though it is still statistically different from zero, it is no longer clinically important. When health is measured by the PCA-Index, the results are similar. The lowest health decile would increase by between 0.03 and 0.04, whereas the median health level would increase by 0.004 in both Canada and the U.S.. The OLS estimates of the income-health gradients are statistically different from zero but only marginally important in a clinical sense. As OLS estimates an average effect, the estimate cannot capture the influence of the initial health level on the gradient. The cross country difference is not statistically significant.

Figure 3.1: Education-health and income-health gradients: Results using the Probit-Index



Red lines show values for Canada and blue lines for the U.S. estimated at each decile and connected. Solid lines show UQPEs and dashed lines show OLS estimates. OLS estimates differ across countries at the 1% for High School, and University, at the 5% level for College and are not statistically significant for Household Inocme.

Figure 3.2: Education-health and income-health gradients: Results using the PCA-Index



Red lines show values for Canada and blue lines for the U.S. estimated at each decile and connected. Solid lines show UQPEs and dashed lines show OLS estimates. OLS estimates differ across countries at the 5% for High School, and College, at the 1% level for University and are not statistically significant for Household Inocme.

The education-health gradients are generally steeper in the U.S. than in Canada for all levels of educational attainment, with a larger effect of education on health for individuals with higher levels of education. The effect of educational attainment on health is clearly clinically important at low levels of health in both countries, ranging from 0.14 to 0.19 in Canada and from 0.20 to 0.41 in the U.S. at the first decile. As health increases, education becomes a less important predictor of health. For individuals with very good health, I find no longer a clinically important influence of educational attainment on health. A clear ordering exists, with the highest level of education (university degree and above) having the largest effect on health, followed by the second-highest level (college degree) and so on. For each level of education, the OLS estimates predict a significantly stronger effect of education on health for the U.S. than for Canada (the differences are significant at the 1% or 5% level) and all effects are statistically different from zero at the 1% level. Again, the simple regression approach can only provide an average effect, whereas my results show that education has a changing impact on health across different health levels.

I repeat the sensitivity analysis from above and estimate both the OLS and the unconditional quantile regressions for whites only and insured people only. The general pattern does not change by restricting the sample in any major way.

3.4.2 Differences in Socioeconomic Inequality

Above results present socioeconomic health gradients separately for each country. A comparison of the magnitude of socioeconomic inequality between the two countries based on these health gradients is difficult, however, as the population in the U.S. differs from the one in Canada in the distribution of educational attainment and

income as shown in Table 3.1. Such endowment differences will have consequences for health outcomes even with equal levels of socioeconomic inequality.

In this subsection, I quantify cross-country differences in socioeconomic inequality by assessing its effect on health outcomes while controlling for differences in population characteristics. Specifically, I test whether health outcomes between Canada and the U.S. differ because of different population characteristics or because the effects of these characteristics on health differ. The methodology relies on an extension of the Oaxaca-Blinder decomposition by Firpo et al. (2007). In their paper, the authors introduce a two-stage procedure based on a reweighting method and RIFs, which extends the Oaxaca-Blinder method from the mean to any distributional statistic of interest.

In the first stage, the distribution of the health index in country i, is constructed under counterfactual endowments (i.e. using country j's population characteristics) by reweighting the data. Under the assumption of ignorability and common support, reweighting identifies the difference between the distributions due to endowment differences and due to differences in the coefficients.[27]

In the second stage, the distribution of interest is expressed in terms of the expected value of its RIF. Using the law of iterated expectations, the unconditional quantile q_τ can be expressed as

$$q_\tau = E(RIF(HI; q_\tau, F_{HI})) = E_X[E(RIF(HI; q_\tau, F_{HI})|X)] = E[X]\beta, \qquad (3.13)$$

[27]Ignorability assumes that the error term is independent across groups (Canada and the U.S.) given the covariates. Overlapping support requires the probability of belonging to group j strictly lies between zero and one, conditional on the covariates (Firpo et al. 2007). Overlapping support does not hold for the health insurance dummy as all Canadian respondents have health insurance. However, this does not affect the interpretation of the overall decomposition or the detailed decomposition with respect to the socioeconomic variables.

which generalizes the Oaxaca-Blinder decomposition to any distributional statistic and allows the detailed decomposition by each variable of interest.

As before, I focus on different quantiles of the health distribution. Using Firpo et al.'s (2007) reweighting method, I construct counterfactual health distributions that would occur had Canada and the U.S. the same population composition and decompose the observed difference in health levels between the U.S. and Canada in differences caused by the endowment effect (differences in the population composition) and differences caused by the coefficient effect (differences in the effect of the explanatory variables) at each decile. While differences in endowments due to differences in the population composition may lead to different health levels given the same level of inequality, differences in the coefficients of the socioeconomic variables reflect differences in the degree of inequality between the two countries. The interaction effect captures differences in health outcomes as a result of differences in endowments and coefficients occurring together.

The decomposition results are shown in Table 3.5. I report results for the first four deciles of the health distribution (for relatively unhealthy individuals); for higher deciles, neither the difference between health levels across countries nor the differential effect of socioeconomic status are significant in a clinical sense.[28] The average health of individuals in the lower deciles is higher in Canada than in the U.S., though this effect decreases with better health and becomes clinically insignificant beginning at the fourth decile for both the Probit- and the PCA-Index. For low health status, the endowment effect is negative, which indicates that health status in the U.S. would decrease in the counterfactual situation that the U.S. had Canada's endowments. The coefficient effect is large and positive, suggesting that differences in coefficients lead to

[28]The full results are available upon request.

a significant difference in health outcomes. The interaction effect is small and neither statistically nor clinically significant.

In addition to the average effect, Table 3.5 also presents the coefficient effects for educational attainment and household income. As Oaxaca-Blinder decomposition for categorical variables is sensitive to the omitted category, I normalize the effect for educational attainment and present coefficient effects for all four levels of educational attainment, including the reference category: less than high school completion (Jann 2008). I focus my discussion on the Probit-Index; the results for the PCA-Index are very similar. I further restrict the discussion to results that are statistically significant at least at the 5% level. If the U.S. had the same effect of less than high school compeletion on health than Canada, the health level in the U.S. would increase by 0.012 at the 10th quantile of the health disribution. However, as higher educational attainment is more strongly associated with good health outcomes in the U.S., the health level in the U.S. would decrease by 0.019 if university graduates faced the same coefficients than Canadian university graduates. These results show that differences in socioeconomic inequaltity between the two countries lead to clinically and statistically significant differences in population health. For individuals with better health, the effects become smaller and eventually are no longer clinically important.

Differences in household income contribute even more to differences in health outcomes between the U.S. and Canada. At the 10th, 20th, and 30th quantile, assuming the same (Canadian) effect of income on health would reduce health in the U.S. by 0.022, 0.034, and 0.022, respectively, though the estimated coefficient at the 10th quantile of the health distribution is not statistically significant.

Table 3.5: Differences in socioeconomic inequality between the U.S. and Canada

	Probit-Index				PCA-Index			
Decile	1	2	3	4	1	2	3	4
Health status in Canada	0.578**	0.729**	0.812**	0.858**	0.605**	0.774**	0.853**	0.898**
	(0.013)	(0.007)	(0.004)	(0.003)	(0.015)	(0.006)	(0.004)	(0.002)
Health status in the U.S.	0.546**	0.694**	0.791**	0.850**	0.548**	0.748**	0.839**	0.889**
	(0.009)	(0.006)	(0.004)	(0.003)	(0.013)	(0.006)	(0.004)	(0.002)
Difference in health statuses	0.032*	0.034**	0.021**	0.008*	0.057**	0.026**	0.014**	0.009**
	(0.016)	(0.009)	(0.006)	(0.004)	(0.020)	(0.009)	(0.005)	(0.003)
Difference in health statuses explained by:								
Endowment effect	-0.037**	-0.014*	-0.009*	-0.005^{+}	-0.045**	-0.011	-0.007^{+}	-0.003
	(0.010)	(0.007)	(0.004)	(0.003)	(0.014)	(0.007)	(0.004)	(0.002)
Coefficient effect	0.062**	0.047**	0.029**	0.012**	0.089**	0.041**	0.021**	0.012**
	(0.016)	(0.009)	(0.006)	(0.004)	(0.019)	(0.008)	(0.005)	(0.003)
Interaction effect	0.007	0.001	0.001	0.002	0.013	-0.004	-0.000	-0.000
	(0.009)	(0.006)	(0.004)	(0.002)	(0.012)	(0.005)	(0.003)	(0.002)
Coefficient effects for:								
less than high school completion	0.012**	0.007**	0.005**	0.002*	0.020**	0.006**	0.002^{+}	0.001*
	(0.004)	(0.002)	(0.001)	(0.001)	(0.005)	(0.002)	(0.001)	(0.001)
high school graduation	-0.010	-0.004	-0.001	0.000	-0.008	0.001	0.002	0.000
	(0.009)	(0.005)	(0.003)	(0.002)	(0.011)	(0.005)	(0.003)	(0.002)
college graduation	-0.002	-0.002	-0.002	-0.000	-0.006	-0.003	-0.002	-0.001
	(0.004)	(0.002)	(0.002)	(0.001)	(0.005)	(0.002)	(0.001)	(0.001)
university graduation	-0.019*	-0.010*	-0.007*	-0.005*	-0.033**	-0.010*	-0.004	-0.002
	(0.009)	(0.005)	(0.003)	(0.002)	(0.011)	(0.005)	(0.003)	(0.002)
household income	-0.022	-0.034**	-0.022**	-0.010+	-0.025	-0.034**	-0.016*	-0.006
	(0.022)	(0.012)	(0.008)	(0.005)	(0.027)	(0.011)	(0.007)	(0.004)

$^{+}$, *, ** indicate significance at 10%, 5%, 1% respectively. N=8,188. Standard errors in parantheses.

Even though the employed method takes differences in endowments into account, I repeat the analysis for the restricted samples of only white individuals and only individuals with health insurance. These restrictions do not change my findings in any major way.

3.5 Conclusion

In this paper I construct two objective health indices to control for heterogeneous reporting behaviour and compare health status in the U.S. and Canada. The results show that differences in SAH overstate objective health differences and can largely be explained by different response behaviours rather than by differences in health between Americans and Canadians. The differences in response behaviour are possibly caused by slight differences in language use between the two countries. Even though reporting differences between the U.S. and Canada are relatively small, the problems related to the use of subjective health measures could easily be magnified for countries that are less similar.

Despite fundamental differences in their health systems, I find smaller health differences than self-reported measures suggest once reporting heterogeneity is controlled for. In particular, the much higher per capita health expenditures in the U.S. do not lead to better general health for Americans compared to Canadians. One notable difference is the higher proportion of Americans in "poor" or "fair" health. This finding is not driven by lower health status of minorities or lack of health insurance at the time of the JCUSH interview. Canada's publicly provided health insurance remains nevertheless a possible explanation for this finding, as I cannot observe past periods of being uninsured or the cost of private insurance. Future research could explore

whether the recent health care reform in the U.S., which introduced nearly universal health insurance coverage – so called "Obamacare" – will further reduce differences between the U.S. and Canada.

In both countries, income and educational attainment are positively related to health outcomes with much larger effects for individuals with poor health. Compared to simple OLS estimates, my results predict a larger effect of socioeconomic variables at the bottom of the health distribution and a smaller effect at the top. This finding is important for the design of policies that target socioeconomic health inequalities, as they need to take the differential impact into account in order to attain their goal and to allocate resources efficiently. Though socioeconomic inequality is stronger in the U.S., my findings indicate that socioeconomic inequality with respect to health remains a concern in Canada. However, some care has to be taken in interpreting these results as causal effects due to the complex relationship between socioeconomic status and health.

Chapter 4

The Impact of Post-Secondary Education on Language Ability

In a knowledge-based economy, language ability is essential for labour market success. Most research on language ability focuses on developments during childhood and early adolescence, though language ability continues to improve during late adolescence and early adulthood. To better understand language development during these later stages, this paper analyses gains in reading proficiency during post-secondary education using reading test scores just prior to the age of post-secondary education and just after from the Youth in Transition Survey (YITS). We find university graduation increases reading proficiency by almost half a proficiency level, which corresponds to over half of the gap in respondents' reading test scores between ages 15 and 24. Community college graduation does not increase reading skills relative to high school completion. Finally, females outperform males on both reading tests, but the gender gap is smaller at age 24. Differences in the effects of education on reading test scores are not statistically significant across genders.

4.1 Introduction

Language ability greatly influences economic and social well-being. In particular, language ability is essential for labour market success. Individuals with good language skills are less likely to be unemployed, are more likely to complete higher education, and earn on average higher wages than individuals with poor language skills (Knighton and Bussière, 2006; Finnie and Meng, 2006). Moreover, a large part of the positive effect of schooling on wages can directly be attributed to literacy skills; and this effect has increased in more recent years (Green and Riddell, 2001; Ingram and Neumann, 2006). In addition to these positive labour market outcomes, good language ability also brings non-monetary individual, social, and cultural benefits. Reading literacy, which is a major component of language ability, includes the "understanding, using, reflecting on and engaging with written texts, in order to achieve ones goals, to develop ones knowledge and potential, and to participate in society" (OECD, 2010). This definition of reading literacy, which is used by the Programme for International Student Assessment's (PISA), also comprises, for example, health and financial literacy, which describes the ability to understand medical and financial information, respectively. The former has been found to be linked to lower mortality rates in old age (see, for example, Baker et al. 2007), while the latter helps people to make good investment decisions or plan for retirement (Lusardi and Mitchell, 2011). Because of the many benefits of language ability, policy makers aim to increase individuals' language ability, and the question whether post-secondary education increases language ability, and can thus be used to achieve this goal, is of great interest.

Although language development occurs primarily during childhood and early adolescence, late adolescence and early adulthood are still fundamental periods for the

acquisition of language skills (for an overview, see Reder (2000)). For example, vocabulary knowledge and the ability to understand complex narratives continue to increase until well into adulthood (Alwin and McCammon, 2001; Mascolo and Fischer, 2010). The importance of post-secondary education for the continued development of language ability is documented in several studies. Desjardins (2003) looks at literacy proficiency throughout the life-cycle in a cross-sectional sample of 18 countries. After controlling for other determinants such as home background, labour force status and age, tertiary education (measured by a one standard deviation of education above the country mean) remains the most important predictor of literacy proficiency. Likewise, Bray et al. (2004) assess reading comprehension and attitudes towards literacy activities among college students in the U.S. and find a positive association between improvements in reading comprehension and exposure to post-secondary education. However, evidence on the contribution of different kinds of post-secondary education to language ability is rare. Moreover, measuring the causal impact of post-secondary education is difficult since data with measures of ability after the completion of post-secondary education generally do not contain a comparable measure prior to the period in which post-secondary education is undertaken. Therefore, it is not possible to control for pre post-secondary language ability. This lack of information makes the identification of the causal effect of post-secondary education on language ability difficult, as attendance of post-secondary education is in turn influenced by a student's initial level of language ability (Dougherty, 2003).

To fill this gap in the literature, this paper uses a unique data set from the integrated Youth in Transition Survey/Programme for International Student Assessment (YITS/PISA) and the YITS Reading Skills Reassessment, which contains measures

of language ability before and after post-secondary enrolment. Doing so allows us to extend previous work on the effect of post-secondary education on language ability by focusing on different forms of post-secondary education using reading test scores of young Canadians at ages 15 and 24. We estimate a reading skills production function to analyze the gains in reading proficiency, using students' initial reading test scores and parental education to control for a student's innate ability and to address the endogeneity of post-secondary enrolment. As a robustness check, we include math and science test scores at age 15 as additional measures for ability. All specifications yield very similar results. We find university graduation increases reading proficiency by about 33 points on the PISA scale, which corresponds to over half of the average gain in respondents' reading test scores between ages 15 and 24. Community college graduation does not increase reading proficiency relative to high school completion. Since gender gaps in language proficiency have been observed in numerous studies (see Logan (2009) for an overview), we also study whether the effect of post-secondary education on reading proficiency differs across genders. While females outperform males on both reading tests, men experience a greater increase in reading proficiency between ages 15 and 24, causing the gender gap to diminish. The differences in the effect of education on reading test scores, however, are not statistically significant across genders.

The remainder of this paper is organized as follows. Section 4.2 describes the contemporaneous model and the value-added model as methods to estimate a knowledge production function. Section 4.3 introduces the data and key variables used in this study, and Section 4.4 presents our estimation results and robustness checks. Section 4.5 discusses the results and concludes.

4.2 Methodology

4.2.1 Estimating the Knowledge Production Function

Growth in cognitive achievement is commonly described by a knowledge production function analogous to a firm's production function, where achievement (the output) is a function of initial endowments and investments (Ben-Porath, 1967). In the most general conceptual framework, cognitive development is described by a production technology R_a, which captures how each input contributes to growth in cognitive achievement at age a:

$$R_{ia} = R_a[S_i(a), NS_i(a), \mu_{i0}, \epsilon_{ia}]. \tag{4.1}$$

Cognitive achievement is measured by a test score R_{ia} of student i at age a, $S(a)$ and $NS(a)$ denote the histories of school and non-school (most importantly family) inputs up to age a, respectively, μ_{i0} denotes unobserved innate ability determined at conception, and ϵ_{ia} is a random error term (Todd and Wolpin, 2003).

Since we do not have data on the entire input history of school and non-school inputs starting at the time of conception up to time a, we need to make several simplifying assumptions in order to estimate the achievement production function. One simplification is to disregard past inputs entirely. This "contemporaneous" specification of cognitive achievement requires the very strict assumption that only current inputs matter for cognitive development, or that inputs are unchanging over time, in which case current input measures also capture the history of inputs (Todd and Wolpin, 2003). In addition, the contemporaneous inputs must be unrelated to (unobserved) innate ability. The longitudinal data available in the YITS allows us to relax

these assumptions somewhat. Using a baseline test score at time $a-1$ and a follow up test score at time a, we estimate the so called "value-added model". Instead of assuming past inputs are irrelevant to cognitive development, the value-added model assumes that the history of inputs up to age a can be captured by the baseline test score, R_{a-1}. Generally, the value-added model is found to be more appropriate to model cognitive achievement (Hanushek, 2003).

With additively separable inputs and under the assumption that the effects of all inputs, including innate ability, decay at a constant rate over time, we can estimate the impact of education on reading proficiency with the following estimation equation:

$$R_{24,i} = \alpha_0 + \alpha_1 R_{15,i} + \alpha_2 R_{15,i}^2 + \sum_{j=1}^{J} \delta_j Edu_j + X_i' \beta + \eta_i, \tag{4.2}$$

where $R_{24,i}$ and $R_{15,i}$ denote reading test score at age 24 and 15, respectively, Edu_j are indicators for different educational pathways, X_i includes additional background variables, and η_i is a random error term. The square of the initial reading score is included to model the increasing difficulty of improving language proficiency the higher is the initial proficiency level. To measure whether education has a differential effect across genders, we interact the education variables with a male-dummy M_i, in which case our estimation equation is given by:

$$R_{24,i} = \tilde{\alpha}_0 + \tilde{\alpha}_1 R_{15,i} + \tilde{\alpha}_2 R_{15,i}^2 + \sum_{j=1}^{J} \tilde{\delta}_j Edu_j + M_i \left(\tilde{\gamma}_0 + \sum_{j=1}^{J} \tilde{\gamma}_j Edu_j \right) + X_i' \tilde{\beta} + \tilde{\eta}_i, \tag{4.3}$$

The validity of the assumption of a constant decay rate arguably depends on the age range under consideration. The assumption might be overly strong to describe

different stages of early childhood development; however, as our focus lies on individuals aged 15 to 24, it seems more appropriate to assume age invariant effects of inputs. With $\alpha_1 = 0$ ($\tilde{\alpha}_1 = 0$) and $\alpha_2 = 0$ ($\tilde{\alpha}_2 = 0$), Equation 4.2 (Equation 4.3) reduces to the contemporaneous model and we present results from the contemporaneous model in addition to the value-added model for reference. In general, the OLS estimate for the effect of the initial test score will be downward biased due to the positive correlation between the measure of lagged reading test score and its measurement error (Todd and Wolpin, 2003). Thus, we cannot test whether the contemporaneous model or the value-added model is appropriate.

Whether to attend post-secondary education and what kind of post-secondary institution to attend is determined by individuals' decisions. Different sorting characteristics include, among many others, financial incentives, family effects, tastes, or ability (Willis and Rosen, 1979). Due to data limitations, we do not formally model the education decision in this paper. Hence, the endogeneity of educational choices could be a concern if we are unable to control for self-selection into post-secondary education. In particular, if unobserved ability is correlated with inputs, our estimation equation suffers from omitted variable bias.[1] The value-added model addresses this problem by controlling for unobserved ability by using the initial reading test score of an individual as a proxy for innate ability: Students with high innate ability will already have achieved high reading ability at age 15. By including the initial test score as a control variable, our results only capture the effect of post-secondary education on a further gain in reading proficiency after the age of 15. Additionally, we include dummy variables for different levels of parental education as further control variables

[1]Instrument variables for educational attainment could address this problem, however, due to the relatively small sample size, possible instruments suffer from the weak instrument problem.

as parental education is correlated with innate ability and a strong predictor of educational attainment (Caponi and Plesca, 2009). Moreover, we use students' math and science test scores as additional measures of ability as a robustness check. If including the initial reading test score successfully controls for innate ability, we would expect our estimates to be unaffected by adding these additional variables. Furthermore, we exclude Québec from the sample to verify that institutional or language differences do not influence our results.[2]

4.3 Data and Variables

4.3.1 The Youth in Transition Survey

The Youth in Transition Survey (YITS) is a longitudinal study designed to provide insight into the educational and labour market pathways of Canadian youths and factors influencing their transition from school to work. In 2000, the first wave of YITS collected data from two age groups, one aged 15 and one aged 18-20 years. Our study focuses on the survey of the younger cohort - the Reading Cohort - which is integrated with PISA 2000, and includes a special reading skills assessment test. Almost 30,000 students at age 15 (as of December 31st 1999) attending any form of schooling in one of the ten provinces were interviewed for YITS/PISA. The paper and pencil interview consisted of a two hour skills assessment and a 40 minutes questionnaire for PISA followed by a 30 minutes questionnaire for YITS. Information was also collected from the students' schools (PISA/YITS) to learn about high school characteristics and

[2] In Québec, secondary school goes to grade 11 (rather than to grade 12 or 13 as in other Canadian provinces) after which students can choose to attend a Collège d'enseignement général et professionnel (CEGEP) where they can pursue technical degree programs or pre-university programs.

from their parents (YITS) to obtain family background information in a telephone interview. After the initial questionnaire, respondents were followed biennially by the YITS.[3] Due to attrition, the sample size dropped to 14,650 for Cycle 5, from which 2,820 were selected to participate in a special YITS Cycle 5.5 Reading Skills Reassessment, conducted in 2009. Attrition may impact the representativeness of our results if respondents do not drop out of the sample at random. This problem is somewhat reduced by the sample selection procedure. To select the sample for the Reading Skills Reassessment Survey, the sample of cycle 5 respondents was stratified by (1) reading proficiency level in 2000, (2) level of education in cycle 5, and (3) gender (Cartwright 2012). The response rate for the reading skills reassessment survey was 46%. Our final sample consists of 1,296 respondents from the original sample, who were by then aged 24.[4] Table 4.1 provides summary statistics for the variables used in this study.[5]

[3]For more information, see the YITS 2000 Cycle 1 User Guide (Statistics Canada, 2005).

[4]To assess the extent of attrition and non-response with respect to individuals' reading test scores, Figure B in the Appendix shows the distribution of scores of all test takers at age 15 and compares them with the distribution of scores of the individuals who participated in the reading reassessment. On average, the individuals in our sample have somewhat higher reading test scores at age 15 than the average YITS/PISA respondent but the tails of the distributions of reading scores between participants and non-participants are very similar.

[5]Summary statistics on provinces are grouped for the Atlantic provinces and the Prairies due to Statistics Canada disclosure rules. Our regressions include dummy variables for all provinces separately (Ontario serves as the reference group). All numbers presented in this paper are weighted using the population weight provided in the YITS data.

Table 4.1: Summary statistics

	Mean	St. dev.
Test scores		
Reading score at 15	541.35	(91.82)
Reading score at 24	598.85	(74.69)
Difference between reading scores at 24 and 15	57.51	(62.18)
Math score at 15[a]	537.67	(77.95)
Science score at 15[b]	535.55	(83.87)
Education at age 24		
Less than high school	0.06	(0.23)
High school degree	0.29	(0.45)
College degree	0.34	(0.47)
University degree	0.32	(0.47)
Demographic information		
Male	0.51	(0.50)
Age in months (PISA-15)	189.51	(3.38)
English	0.76	(0.43)
French	0.21	(0.40)
First language not English or French	0.12	(0.32)
Visible minority	0.11	(0.31)
Missing parent interview	0.14	(0.35)
Has siblings	0.92	(0.27)
Geographic information		
Ontario	0.32	(0.47)
Atlantic provinces (N.L., P.E.I., N.S., and N.B.)	0.09	(0.28)
Prairies (Sask. and Man.)	0.07	(0.25)
Lived in a urban place at age 15	0.66	(0.47)
Québec	0.19	(0.39)
Alberta	0.09	(0.29)
British Columbia	0.11	(0.32)
Missing geographic information	0.14	(0.34)
Parental education		
Less than high school	0.08	(0.28)
High school degree	0.19	(0.39)
Some college	0.06	(0.24)
College	0.30	(0.46)
Bachelor degree	0.17	(0.38)
Graduate degree	0.09	(0.29)

N=1296. Only a subsample of the students completed math and science tests.
[a] N=723. [b] N=737.

4.3.2 Variables

Reading Proficiency

We measure reading proficiency using information from the PISA 2000 reading test, indicating reading proficiency at the age of fifteen (PISA-15), and the follow up reading test from the YITS reading reassessment conducted in 2009 at age 24 (PISA-24). PISA-15 is assessed based on 141 questions asking students to retrieve information, interpret texts, and reflect on and evaluate texts; the reassessment test in 2009 consists of 28 questions covering the same aspects of reading proficiency (see OECD (2012) for details). A student's reading proficiency is a latent variable and must be inferred from the observed item responses. To make this inference, PISA constructs plausible values, which are random draws from the distribution of likely scores for a student, given the student's item responses (Adams, 2002).[6] Five such plausible values are drawn for PISA-15, and ten plausible values are drawn for PISA-24 accounting for the greater uncertainty in the estimated distribution of likely scores as reading skills are assessed with fewer questions in the reassessment test.[7]

Though PISA was designed to assess skills of 15 year olds, it has been shown that the PISA test design remains valid to test the reading proficiency of 24 year olds;

[6]For a detailed discussion on the construction of the plausible values see Adams (2002) for the PISA 2000 reading test and Cartwright (2013) for the reassessment test. The construction of the plausible values for the reassessment test follows the original method as closely as possible to ensure the comparability of PISA-15 and PISA-24 plausible values.

[7]While plausible values are not exact individual scores, they are highly correlated with individual scores measured by weighted likelihood estimates (WLEs) (the correlation between WLEs and the average of the plausible values for PISA-15 is 0.97). As WLEs are unavailable for PISA-24 we use the average of the plausible values for our measure of reading scores and estimate Equation 4.2 and 4.3 using ordinary least squares. When using plausible values, we ideally would correct our standard errors to adjust for the uncertainty of the true value of reading proficiency. However, it remains unclear which strategy should be adopted when plausible values are included both as dependent and independent variable (see also Bratti and Checchi (2013)).

Table 4.2: Average reading test scores at age 15 and 24

	PISA-15		PISA-24		Difference	
	Mean	St. dev.	Mean	St. dev.	Mean	St. dev.
All participants	541.35	91.82	598.85	74.69	57.51	62.18
Male	525.43	93.22	589.09	78.94	63.66	64.11
Female	557.63	87.48	608.84	68.71	51.21	59.52
Education						
Less than high school	439.52	82.02	503.94	96.21	64.43	79.08
High school degree	513.82	92.42	576.10	72.35	62.28	61.91
College degree	532.45	80.88	585.70	57.31	53.25	61.30
University degree	592.89	73.69	649.41	53.96	56.52	59.75
Proficiency level at age 15						
Level 1	370.23	35.98	482.81	62.99	112.58	56.61
Level 2	450.57	18.97	549.58	63.66	99.01	64.20
Level 3	520.27	20.08	592.25	49.66	71.98	47.82
Level 4	590.87	20.07	630.10	46.62	39.22	45.65
Level 5	661.51	28.60	663.30	47.25	1.79	44.74

N=1296.

the difficulty of the questions remains challenging for older youths and test results do not suffer from a ceiling effect (OECD, 2012). The absence of a ceiling effect is also apparent in Figure 4.1, which shows the distribution of test scores at ages 15 and 24. The variation in test scores considerably decreases over time, and while fewer individuals have low test scores at the age of 24, very high test scores remain relatively rare.

Table 4.2 shows the average reading test scores at age 15 and 24 and the average score difference between ages 15 and 24 for all YITS Reading Skills Reassessment participants and by subgroups. The average score of 541 points for the YITS Reading Reassessment participants at age 15 is 41 points higher than the overall average score of 500 points of all the participating OECD countries (OECD, 2010). The average score increase between age 15 and 24 is 58 points. At age 15, females score on average

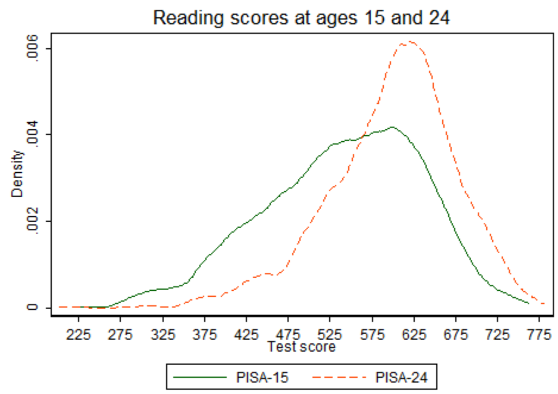


Figure 4.1: Reading test scores by age

32 points higher than males; the gender gap reduces to 20 points by age 24.[8]

PISA test scores and changes in test scores can be interpreted in terms of five "proficiency levels", with five being the highest level of reading proficiency. A score difference of 73 points is considered a difference in reading ability of one proficiency and represents a noticeable difference in the ability to understand, analyse, and interpret a written text (OECD, 2010). Proficiency levels are hierarchical, meaning students who score a certain reading proficiency level are also assumed to be proficient at lower levels. As shown in Table 4.2, students with low reading proficiency at age 15 increase their reading scores by more than students who have already reached a high level of reading proficiency by the age of 15, suggesting that the higher the initial level of language proficiency, the more difficult it becomes to further improve language skills.

[8]The difference in average reading scores between men and women is statistically significant at the 1% significance level for both PISA-15 and PISA-24; the difference in the score difference between the ages 15 and 24 is statistically significant at the 1% level.

Education

Educational attainment is measured by the highest level of education a respondent has completed by the age of 24, with the categories being less than high school, high school graduation (the reference category), college (community college or trades school) graduation and university graduation.

Figure 4.3.2 shows the kernel densities of reading test scores for individuals with high school, college, and university education at ages 15 and 24. Students who go on to attend university score on average higher in the PISA-15 reading test than students who graduate from college, and students who graduate from college tend to have higher scores than students who do not pursue post-secondary education. This pattern still holds at the age of 24: respondents with a university degree have on average the highest reading scores, followed by respondents with a college degree, followed by respondents with a high school diploma. The average gains in reading proficiency by education level, however, are very similar as shown in Table 4.2.[9]

Additional Covariates

We further control for demographic, family, and geographic characteristics. Demographic variables include a gender dummy (1 for males, 0 otherwise), and the respondent's age in months. In addition, while students might be fluent in both English and French, the language of the test can influence the obtained score. For this reason, we include two dummy variables that are equal to 1 if the test language at age 15 and 24

[9] The difference in average reading scores between university graduates and high school graduates and between high school dropouts and high school graduates is statistically significant at the 1% level for both PISA-15 and PISA-24. Neither the difference in scores between college graduates and high school graduates at any time nor the differences in scores between the ages 15 and 24 for all education levels are statistically different.

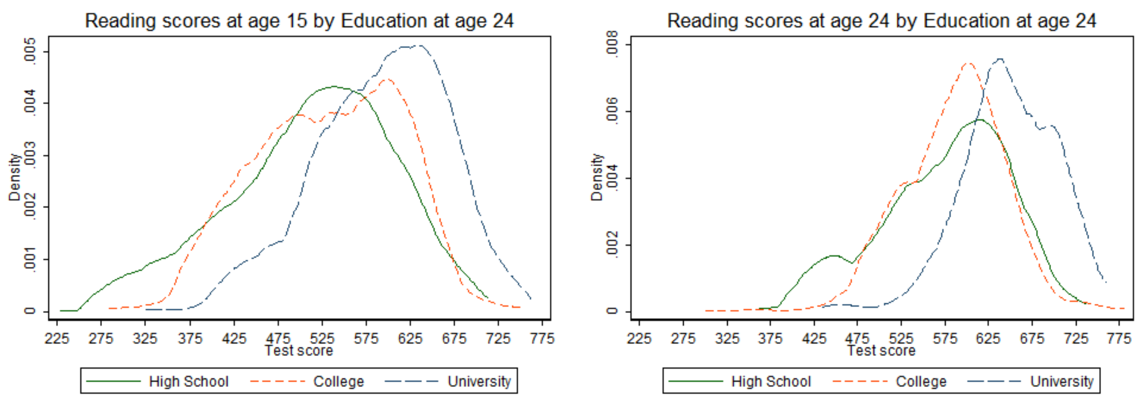


Figure 4.2: Reading test scores by post-secondary education level

is English or French, respectively. In addition, we include a foreign language dummy which equals 1 if the respondent's first language is different from English or French, and a dummy indicating whether the respondent is a visible minority.

The questions regarding minority status and the respondent's first language are asked in the parent interview. As several parent interviews are missing, we include a dummy variable indicating if the parent interview is missing to avoid losing observations.

Additional family background variables include information on whether the respondent has any siblings, and the highest level of education of either parent (both measured at age 15). The parental education categories are less than high school, high school, some college education, college graduation, a bachelor degree, or a graduate degree. High school education serves as the reference category.[10]

Geographical control variables include an indicator for living in an urban place at age 15, and a set of dummy variables indicating the province of residence at age

[10]Missing information on parental education is already captured by the variable indicating missing parent interview.

15. The reference category is Ontario. Missing geographic information is captured by a dummy variable. When available, missing time-invariant information is obtained from other YITS cycles. Observations with missing information for variables with only a few missing values are dropped from the analysis.

Lastly, while PISA 2000 focused on students' reading proficiency, subsamples of the PISA participants were also tested on their math and science proficiency. We use students' math and science scores as additional covariates in our robustness checks.

4.4 Results

In this section, we first present estimation results based on the contemporaneous model and on the value-added model with different sets of covariates included and discuss why the value-added model is more applicable to our data. We then present robustness checks for our preferred specification, the value-added specification with the full set of covariates, to test the reliability of our findings.

4.4.1 The Contemporaneous Model

Results for the contemporaneous specification for both women and men combined are presented in Panel A of Table 4.3 for different sets of covariates. In all regressions, high school graduation serves as the reference category. In column (1), we only control for educational attainment. Then we add additional control variables. Column (2) also includes demographic characteristics, column (3) adds geographic control variables, and column (4) adds family background variables.

Out of our post-secondary education variables, only the effect of university graduation on reading test scores is statistically significant.[11] The estimated coefficients are large, considering the average score increase for all respondents is 58 points: For the specifications with only educational control variables (column (1)), we estimate that the reading scores of individuals with a university degree are 73 points higher than reading scores of high school graduates, which corresponds to a difference in reading skills of one proficiency level. In the final specification with demographic, geographic, and parental education information included as control variables, university graduates have reading test scores that are 63 points higher.

The difference between university and high school graduates is slightly smaller but still surprisingly large when we allow the effects of post-secondary education to differ by gender as shown in Panel B of Table 4.3 for the same specifications as before. The reading test scores of university graduates are somewhere between 56 to 65 points higher. The point estimates indicate a slight additional benefit of post-secondary education for males, though none of the effects of the male-education interaction terms are statistically significant.

A concern with the contemporaneous specification is the significant drop in the coefficient for university graduation after including proxies for unobserved ability. After adding parental education, the university coefficients drop by about 10 points. This result is to be expected if, originally, the measured effects are biased upwards due to unobserved ability which is now (partly) picked up by parental education. However, it is impossible to tell whether the final estimate is still biased due to remaining unobserved ability. For this reason, including the initial test score at age

[11] Our main variables of interest are college and university graduation as well as these variables interacted with the male dummy. We include a dummy for high school dropouts (and, where applicable, its interaction with the male dummy) for completeness.

Table 4.3: Estimation results from the contemporaneous model

Panel A

Dependent variable: PISA-24	(1)	(2)	(3)	(4)
University	73.312**	71.451**	73.225**	63.322**
	(9.186)	(8.078)	(7.973)	(7.481)
College	9.600	8.688	9.632	7.349
	(9.181)	(7.718)	(8.136)	(7.269)
Less than high school	-72.159**	-72.237**	-69.357**	-68.386**
	(16.160)	(16.430)	(16.090)	(15.829)
Male		-7.295	-7.430	-7.965
		(6.403)	(6.640)	(6.630)
Demographic information	no	yes	yes	yes
Geographic information	no	no	yes	yes
Parental education	no	no	no	yes
Constant	576.102**	848.062**	875.560**	875.252**
	(7.983)	(210.516)	(209.205)	(200.389)
Observations	1,296	1,296	1,296	1,296
R-squared	0.275	0.293	0.308	0.339

Panel B

Dependent variable: PISA-24	(1)	(2)	(3)	(4)
University	65.159**	65.013**	65.350**	56.530**
	(10.085)	(9.767)	(9.926)	(10.141)
University x male	12.160	12.580	15.101	13.180
	(15.370)	(14.931)	(13.952)	(13.573)
College	3.047	5.240	4.729	3.195
	(9.993)	(9.506)	(9.935)	(9.588)
College x male	8.371	5.444	7.705	6.403
	(15.622)	(14.547)	(14.174)	(13.966)
Less than high school	-53.246+	-55.197*	-55.473*	-58.477**
	(23.650)	(22.685)	(21.804)	(20.144)
Less than high school x male	-26.661	-24.239	-19.537	-13.932
	(30.028)	(29.820)	(29.286)	(27.904)
Male	-11.727	-11.971	-13.885	-13.669
	(13.370)	(12.651)	(12.309)	(11.541)
Demographic information	no	yes	yes	yes
Geographic information	no	no	yes	yes
Parental education	no	no	no	yes
Constant	584.081**	848.289**	876.476**	877.236**
	(7.748)	(207.364)	(205.142)	(197.382)
Observations	1,296	1,296	1,296	1,296
R-squared	0.280	0.296	0.311	0.340

Standard errors cluster at the school level in parentheses.
** $p<0.01$, * $p<0.05$, + $p<0.1$

15 seems warranted.

4.4.2 The Value-Added Model

Table 4.4 shows the results for the value-added model for both males and females combined. Again, we present several specifications: The first column controls for education and the initial test score, column (2) includes both the initial test score and the square of the initial test score (divided by 100), columns (3) to (5) add additional control variables as above.

We find that, relative to students who completed high school as their highest level of education, students who graduated from college have very similar gains in reading test scores between ages 15 and 24. The point estimates are close to zero and statistically insignificant across all specifications. In contrast, university graduation increases reading test scores by between 31 and 34 points compared to high school graduates across the different specifications, which corresponds to almost half a PISA proficiency level. Hence, including initial reading test scores noticeably reduces the estimates found by the contemporaneous model. Moreover, the estimated effect of educational attainment on reading proficiency is remarkably robust to adding additional control variables. In particular, adding parental education does not notably change the magnitude of the estimates. We take this finding as an indication that the initial test score at age 15 captures differences in unobserved innate ability and thus addresses the issue of self-selection into post-secondary education. Hence, the differential gains in reading test scores past the age of 15 for different education levels can be attributed to the different educational pathways.

The initial score contributes positively to gains in reading test scores, but at a

decreasing rate. However, as noted above, the coefficient for the initial test score is likely to be downward biased. Hence, the estimates serve as a lower bound and we cannot determine the exact magnitude of the effect.

Table 4.4: Estimation results from the value-added model

Dependent variable: PISA-24	(1)	(2)	(3)	(4)	(5)
University	32.652**	33.651**	33.719**	33.925**	32.607**
	(5.742)	(5.749)	(6.070)	(5.995)	(5.845)
College	0.019	-1.299	-0.770	-0.553	0.632
	(5.439)	(5.359)	(4.927)	(4.963)	(4.688)
Less than high school	-33.949*	-30.617*	-29.433*	-29.341*	-27.585*
	(13.908)	(13.699)	(13.627)	(13.959)	(13.428)
PISA-15	0.514**	1.196**	1.196**	1.169**	1.189**
	(0.031)	(0.227)	(0.213)	(0.208)	(0.214)
PISA-15 squared/100		-0.065**	-0.064**	-0.062**	-0.064**
		(0.021)	(0.020)	(0.020)	(0.020)
Male			2.347	1.772	2.614
			(4.374)	(4.332)	(4.207)
Demographic information	no	no	yes	yes	yes
Geographic information	no	no	no	yes	yes
Parental education	no	no	no	no	yes
	311.877**	138.973*	359.036**	375.944**	385.580**
	(17.612)	(59.630)	(138.120)	(135.402)	(132.005)
Observations	1,296	1,296	1,296	1,296	1,296
R-squared	0.595	0.603	0.609	0.612	0.623

Standard errors cluster at the school level in parentheses. ** p<0.01, * p<0.05, + p<0.1

Table 4.5 presents the same specifications again, but also includes male-education interaction terms. As before, college graduation does not significantly increase reading test scores for males and females. While the common effect is negative but insignificant, college education has a gender specific positive (but insignificant) effect for males across all specifications. The effect for university graduation for both males and females is slightly reduced compared to the specification without male-education interaction terms but remains highly significant. Males receive additional gains in

Table 4.5: Estimation results from the value-added model - Gender differences

Dependent variable: PISA-24	(1)	(2)	(3)	(4)	(5)
University	28.775**	30.910**	29.255**	29.705**	28.868**
	(7.654)	(7.749)	(7.996)	(8.050)	(8.064)
University x Male	7.259	5.002	7.356	6.896	6.214
	(9.903)	(9.799)	(9.801)	(9.704)	(9.658)
College	-6.230	-6.605	-6.493	-6.381	-4.131
	(6.689)	(6.703)	(6.604)	(6.639)	(6.184)
College x Male	12.874	11.352	10.508	10.737	8.551
	(9.398)	(9.211)	(9.177)	(9.090)	(8.582)
Less than high school	-18.128	-16.294	-17.656	-18.029	-20.296
	(18.204)	(17.021)	(15.774)	(16.619)	(12.662)
Less than high school x Male	-22.296	-20.368	-16.842	-16.071	-10.351
	(24.784)	(23.844)	(22.895)	(23.547)	(20.950)
PISA-15	0.516**	1.171**	1.175**	1.151**	1.175**
	(0.029)	(0.226)	(0.211)	(0.207)	(0.213)
PISA-15 squared/100		-0.063**	-0.062**	-0.060**	-0.063**
		(0.021)	(0.020)	(0.020)	(0.020)
Male	-2.906	-1.943	-2.834	-3.370	-1.876
	(7.343)	(7.249)	(7.323)	(7.405)	(7.244)
Demographic information	no	no	yes	yes	yes
Geographic information	no	no	no	yes	yes
Parental education	no	no	no	no	yes
Constant	312.858**	145.943*	354.466**	369.580**	381.942**
	(16.201)	(59.431)	(135.706)	(133.330)	(130.366)
Observations	1,296	1,296	1,296	1,296	1,296
R-squared	0.598	0.606	0.611	0.613	0.624

Standard errors cluster at the school level in parentheses. ** p<0.01, * p<0.05, + p<0.1

reading test scores through college and university graduation, but the effects are statistically insignificant. As before, the estimates remain unchanged when explanatory variables for parental education are added to control for unobserved ability.

4.4.3 Robustness Checks

In this subsection, we present the results of our robustness tests. As we believe that the value-added specification with the full set of covariates best controls for unobserved ability, we use this specification as our benchmark. Results are presented in Table 4.6. First, in column (1), we include PISA test scores for math and science as additional explanatory variables to control for unobserved ability.[12] Second, we exclude Québec from the sample (column (2)). Including math and science test scores as explanatory variables does not reduce the estimated effect of university graduation on reading proficiency. Similarly, excluding Québec from the sample does not significantly alter our results.

Columns (3) and (4) show the results of the robustness checks for the specification with male-education interaction terms. Including math and science test scores as covariates, or excluding Québec produces very similar results to our preferred specification with male-education interaction terms.

4.5 Discussion and Conclusion

Benefiting from the availability of reading test scores before and after post-secondary enrolment in the PISA/YITS and the YITS Reading Skills Reassessment data, this paper extends previous research by estimating the effect of different post-secondary

[12] The specification also includes indicators for whether a science or a math score is unavailable.

Table 4.6: Results from robustness checks

Dependent variable: PISA-24	(1)	(2)	(3)	(4)
	Math/Science information	Québec excluded	Math/Science information	Québec excluded
University	32.622**	30.712**	29.768**	25.832**
	(5.765)	(6.647)	(7.986)	(9.153)
University x Male			4.804	7.940
			(9.759)	(11.276)
College	0.776	-2.145	-3.334	-7.664
	(4.713)	(5.596)	(5.989)	(7.607)
College x Male			7.629	8.825
			(8.537)	(10.027)
Less than high school	-28.658*	-17.981	-18.632	-18.127
	(13.754)	(15.624)	(14.330)	(13.971)
Less than high school x Male			-14.370	-0.621
			(22.302)	(24.992)
PISA-15	1.217**	1.186**	1.203**	1.173**
	(0.216)	(0.250)	(0.216)	(0.251)
PISA-15 squared/100	-0.067**	-0.062**	-0.065**	-0.061**
	(0.020)	(0.023)	(0.020)	(0.023)
Male	3.102	3.906	-0.287	-1.984
	(4.508)	(4.974)	(7.532)	(8.839)
Demographic information	yes	yes	yes	yes
Geographic information	yes	yes	yes	yes
Parental education	yes	yes	yes	yes
Math/Science information	yes	no	yes	no
Constant	363.655**	436.398**	321.347**	164.748
	(130.271)	(153.278)	(114.481)	(154.548)
Observations	1,296	1,065	1,296	1,065
R-squared	0.626	0.635	0.627	0.635

Standard errors cluster at the school level in parentheses. ** $p<0.01$, * $p<0.05$, + $p<0.1$

education pathways on the development of reading proficiency of young Canadians. Moreover, we study whether the effect of post-secondary education on language ability differs across genders. We do not find a gender specific impact of education. The closing gender gap between ages 15 and 24 may simply be caused by differences in the age of when men and women acquire language skills.

Our finding that college graduation does not significantly improve reading proficiency relative to high school graduation could be explained by the more practical, labour market specific learning predominant in colleges. While colleges offer a more task specific education, university education focuses on the transfer of general knowledge and theoretical concepts, which are mostly taught out of textbooks. In addition, university students may be required to do more course readings, which would improve their reading proficiency. Our findings suggest that while the average reading proficiency still improves between the ages of 15 and 24, continued formal education only contributes to this improvement if it requires intensive reading tasks.

Due to the increasing importance of reading proficiency in a skill-based economy, policy makers are constantly looking for ways to increase language skills of the workforce. Our paper can inform decision makers by offering a better understanding of the impact of post-secondary education on reading proficiency, though more research is needed to confirm our hypothesis about which aspects of university education leads to the higher reading test scores of university graduates, and to analyze whether these aspects could also be implemented into the college curriculum.

Ultimately, post-secondary education aims to equip students with the necessary skills for a successful participation in the labour force. Green and Riddell (2001) estimate that approximately one-third of the return to education can be attributed to

literacy skills. Interestingly, looking at how different post-secondary education levels affect wages in Canada, Caponi and Plesca (2009) find that college graduates receive very similar wages on average compared to high school graduates, whereas workers with a university degree receive a significant wage premium. Thus, further research could explore whether our finding that only university graduation contributes to an improvement in reading proficiency, can help to explain the lack of a wage premium for college graduates.

Chapter 5

Conclusion

In this dissertation, I conduct three studies that can inform policy makers on contemporary policy debates related to the recent increase in longevity, the assessment and comparison of health outcomes and health care systems, and the importance of skill development in an increasingly skill oriented labour force. The chapters in this dissertation aim to extend our knowledge in these areas using economic tools to study individuals' decisions and to determine the causal relationship between two events.

In Chapter 2, I provide both an economic and econometric analysis of the effects of informal caregiving on mature caregivers. A better understanding of these effects is important in order to decide on the optimal mix of informal and formal provision of long-term care services. I show that some negative effects, such as reduced employment rates and hours worked of informal caregivers, are absent in countries with a strong formal care system. Public spending on long-term care is thus partly offset by a higher productivity of the workforce. However, the higher risk for mental health problems persists in all countries independent of their formal care systems. Knowledge of the risk of mental health problems allows health care providers to introduce

early intervention programs, which will benefit caregivers as well as reduce treatment costs. With the rapid ageing of our societies, elderly care will remain an extremely relevant topic for research. As more internationally comparable data on the health, socioeconomic situation, and social networks of middle-aged and older individuals is becoming available, it will become possible to study the effects of caregiving on caregivers in more detail. Given the impact of the institutional caregiving context and the intensity of care on the effect of caregiving, it would be especially interesting to focus on different policy interventions, or to distinguish between different forms of care provision to capture heterogeneous effects of different caregiving tasks. As an increasing number of men provide informal care, it would also be natural to extend the analysis of the effects of caregiving to include male caregivers and look at the caregiving decision in terms of a household model with intra-household bargaining.

Chapter 3 addresses a more technical issue and introduces an objective health measure that allows us to compare population health in the U.S. and Canada, and the extent of socioeconomic inequality in health outcomes. The comparison between the two countries is especially interesting as they differ dramatically in the organization of their health care systems and in their expenditure level for health care. However, health comparisons tend to find relatively similar levels of population health between the two countries. My study confirms these results. Moreover, I show that some of the reported differences in self-assessed health are due to differences in reporting behaviour rather than actual differences in individuals' health status. While socioeconomic health inequalities are stronger in the U.S., socioeconomic differences in health outcomes also exist in Canada, despite the country's universal health insurance system. As a comparable health measure is essential for many economic models and for

the evaluation of health care systems alike, this finding ought to encourage policy makers to invest in the collection of better health measures such as, for example, biomarkers.

Finally, Chapter 4 examines the effect of post-secondary education on reading proficiency and shows that only university graduation but not college graduation leads to an increase in reading proficiency. In the debate on whether cognitive ability is fixed early in a person's life or whether schooling can raise achievement measures, the results provide evidence for this latter view and are in line with the findings of, for example, Hansen et al. (2004) and Ding and Lehrer (2014). Given the growing importance of reading proficiency in the labour market, this study provides policymakers with valuable insights into how formal education can impact the further development of reading proficiency throughout young adulthood. Understanding which aspects of university education lead to the increase in reading proficiency would be a worthwhile undertaking for future research and would help to target reading proficiency in the development of college curricula and adult learning programs.

Bibliography

Adams, R. J. (2002): *Scaling PISA Cognitive Data*, PISA 2000 Technical Report. OECD Publications: Paris.

Alwin, D. F., and R. J. McCammon (2001): "Aging, cohorts, and verbal ability," *Journal of Gerontology*, 56(3), 151–161.

Aneshensel, C. S., L. Pearlin, J. T. Mullan, S. H. Zarit, and C. Whitlatch (1995): *Profiles in caregiving: The unexpected carer*. Academic Press, San Diego.

Arpino, B., and V. Bordone (2012): "Does grandparenting pay off? The effect of childcare on grandparents cognitive functioning," *European Demographic Research Papers No. 4*.

Bago dUva, T., M. Lindeboom, O. ODonnell, and E. Van Doorslaer (2011): "Slipping anchor? Testing the vignettes approach to identification and correction of reporting heterogeneity," *Journal of Human Resources*, 46(4), 875–906.

Bago d'Uva, T., O. O'Donnell, and E. van Doorslaer (2008a): "Differential health reporting by education level and its impact on the measurement of health

inequalities among older Europeans," *International Journal of Epidemiology*, 37(6), 1375–1383.

Bago d'Uva, T., E. van Doorslaer, M. Lindeboom, and O. O'Donnell (2008b): "Does reporting heterogeneity bias the measurement of health disparities?," *Health Economics*, 17(3), 351–375.

Baker, D. W., M. S. Wolf, J. Feinglass, J. A. Thompson, J. A. Gazmararian, and J. Huang (2007): "Health literacy and mortality among elderly persons," *Archives of Internal Medicine*, 167(14), 1503–1509.

Baker, M., M. Stabile, and C. Deri (2004): "What do self-reported, objective, measures of health measure?," *Journal of Human Resources*, 39(4), 1067–1093.

Bass, D. M., L. S. Noelker, and L. R. Rechlin (1996): "The moderating influence of service use on negative caregiving consequences," *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 51(3), 121–131.

Ben-Porath, Y. (1967): "The production of human capital and the life cycle of earnings," *The Journal of Political Economy*, pp. 352–365.

Bertrand, R. M., J. S. Saczynski, C. Mezzacappa, M. Hulse, K. Ensrud, and L. Fredman (2012): "Caregiving and Cognitive Function in Older Women: Evidence for the Healthy Caregiver Hypothesis," *Journal of Aging and Health*, 24(1), 48–66.

Blinder, A. S. (1973): "Wage Discrimination: Reduced Form and Structural Estimates," *The Journal of Human Resources*, 8(4), 436–455.

Bolin, K., D. Hedblom, A. Lindgren, and B. Lindgren (2010): "Asymmetric Information and the Demand for Voluntary Health Insurance in Europe," NBER Working Papers, National Bureau of Economic Research, Inc.

Bolin, K., B. Lindgren, and P. Lundborg (2008a): "Informal and formal care among single-living elderly in Europe," *Health economics*, 17(3), 393–409.

——— (2008b): "Your next of kin or your own career?: Caring and working among the 50+ of Europe," *Journal of Health Economics*, 27(3), 718–738.

Bonsang, E. (2009): "Does informal care from children to their elderly parents substitute for formal care in Europe?," *Journal of health economics*, 28(1), 143–154.

Bratti, M., and D. Checchi (2013): "Re-testing PISA Students One Year Later: On School Value Added Estimation Using OECD-PISA," IZA Discussion Paper.

Bray, G. B., E. T. Pascarella, and C. T. Pierson (2004): "Postsecondary education and some dimensions of literacy development: An exploration of longitudinal evidence," *Reading Research Quarterly*, 39(3), 306–330.

Brown, S. L., D. M. Smith, R. Schulz, M. U. Kabeto, P. A. Ubel, M. Poulin, J. Yi, C. Kim, and K. M. Langa (2009): "Caregiving behavior is associated with decreased mortality risk," *Psychological science*, 20(4), 488–494.

Caponi, V., and M. Plesca (2009): "Post-secondary education in Canada: can ability bias explain the earnings gap between college and university graduates?," *Canadian Journal of Economics/Revue canadienne d'Economique*, 42(3), 1100–1131.

Carmichael, F., and S. A. Charles (1998): "The labour market costs of community care," *Journal of Health Economics*, 17(6), 747–765.

——— (2003): "The opportunity costs of informal care: does gender matter?," *Journal of Health Economics*, 22(5), 781–803.

Carmichael, F., S. A. Charles, and C. Hulme (2010): "Who will care? Employment participation and willingness to supply informal care," *Journal of Health Economics*, 29(1), 182–190.

Cartwright, F. (2013): "Technical Feasibility of Reporting YITS 2010 Skill Assessment Results on the PISA 2000 Reading Scale," OECD Education Working Papers 69, OECD.

Castro-Costa, E., M. Dewey, R. Stewart, S. Banerjee, F. Huppert, C. Mendonca-Lima, C. Bula, F. Reisches, J. Wancata, K. Ritchie, et al. (2008): "Ascertaining late-life depressive symptoms in Europe: an evaluation of the survey version of the EURO-D scale in 10 nations. The SHARE project," *International Journal of Methods in Psychiatric Research*, 17(1), 12–29.

Caswell, L. W., P. P. Vitaliano, K. L. Croyle, J. M. Scanlan, J. Zhang, and A. Daruwala (2003): "Negative associations of chronic stress and cognitive performance in older adult spouse caregivers," *Experimental aging research*, 29(3), 303–318.

Christelis, D. (2011): "Imputation of Missing Data in Waves 1 and 2 of SHARE," *Available at SSRN 1788248*.

Coe, N. B., and C. H. Van Houtven (2009): "Caring for mom and neglecting yourself? The health effects of caring for an elderly parent," *Health Economics*, 18(9), 991–1010.

Colombo, F., A. Llena-Nozal, J. Mercier, and F. Tjadens (2011): "Help Wanted? Providing and Paying for Long-Term Care," OECD Health Policy Studies.

Crespo, L. (2006): "Caring for parents and employment status of European mid-life women," *Documentos de Trabajo (CEMFI)*, (15), 1.

Crespo, L., and P. Mira (2010): "Caregiving to elderly parents and employment status of European mature women," *Documentos de Trabajo (CEMFI)*, (7), 1.

Cutler, D. M., A. Lleras-Muney, and T. Vogl (2008): "Socioeconomic status and health: Dimensions and mechanisms," NBER Working Paper 14333, NBER.

de Vugt, M. E., J. Jolles, L. van Osch, F. Stevens, P. Aalten, R. Lousberg, and F. R. J. Verhey (2006): "Cognitive functioning in spousal caregivers of dementia patients: findings from the prospective MAASBED study," *Age and Ageing*, 35(2), 160–166.

Dentinger, E., and M. Clarkberg (2002): "Informal caregiving and retirement timing among men and women gender and caregiving relationships in late midlife," *Journal of Family Issues*, 23(7), 857–879.

Desjardins, R. (2003): "Determinants of literacy proficiency: a lifelong-lifewide learning perspective," *International Journal of Educational Research*, 39(3), 205 – 245.

Ding, W., and S. Lehrer (2011): "Experimental estimates of the impacts of class size on test scores: robustness and heterogeneity," *Education Economics*, 19(3), 229–252.

Ding, W., and S. F. Lehrer (2014): "Understanding the Role of Time-Varying Unobserved Ability Heterogeneity in Education Production," NBER Working Papers 19937, National Bureau of Economic Research, Inc.

Dobrescu, L., and D. Christelis (2012): "The Impact of Social Activities on Cognitive Ageing: Evidence From Eleven European Countries," Working Papers 201207, ARC Centre of Excellence in Population Ageing Research (CEPAR).

Dougherty, C. (2003): "Numeracy, literacy and earnings: evidence from the National Longitudinal Survey of Youth," *Economics of Education Review*, 22(5), 511 – 521.

Dowd, J. B., and A. Zajacova (2010): "Does Self-Rated Health Mean the Same Thing Across Socioeconomic Groups? Evidence From Biomarker Data," *Annals of Epidemiology*, 20(10), 743 – 749.

Eng, K., and D. Feeny (2007): "Comparing the health of low income and less well educated groups in the United States and Canada," *Population Health Metrics*, 5:10.

Engelhardt, H., I. Buber, V. Skirbekk, and A. Prskawetz (2010): "Social involvement, behavioural risks and cognitive functioning among older people," *Ageing and Society*, 30(5), 779.

Etilé, F., and C. Milcent (2006): "Income-related reporting heterogeneity in self-assessed health: evidence from France," *Health economics*, 15(9), 965–981.

Ettner, S. L. (1995): "The Impact of 'Parent Care' on Female Labor Supply Decisions," *Demography*, 32, 63–80.

——— (1996): "The opportunity costs of elder care," *Journal of Human Resources*, pp. 189–205.

European Commission (2012): "The 2012 Ageing Report-Economic and budgetary projections for the 27 EU Member States (2010-2060)," *European Economy (2)*, 2.

Feinberg, L., S. C. Reinhard, A. Hauser, and R. Choula (2013): "Valuing the Invaluable: 2011 Update The Growing Contributions and Costs of Family Caregiving," Insight on the issues 51.

Feinstein, J. S. (1993): "The relationship between socioeconomic status and health: a review of the literature," *The Milbank Quarterly*, pp. 279–322.

Figini, P. (1998): "Inequality Measures, Equivalence Scales and Adjustment for Household Size and Composition," Technical Paper 8, Trinity Economic Paper Series.

Filmer, D., A. Hasan, and L. Pritchett (2006): "A millennium learning goal: Measuring real progress in education," *Center for Global Development Working Paper*, 97.

Filmer, D., and L. Pritchett (2001): "Estimating wealth effects without expenditure data - or tears: An application to educational enrollments in states of India," *Demography*, 38(1), 115–132.

Finnie, R., and R. Meng (2006): "The Importance of Functional Literacy: Reading and Math Skills and Labour Market Outcomes of High School Drop-outs," Analytical studies branch research paper series, Statistics Canada, Analytical Studies Branch.

Firpo, S., N. Fortin, and T. Lemieux (2007): "Decomposing wage distributions using recentered influence function regressions," *University of British Columbia (June)*.

——— (2009): "Unconditional Quantiles Regressions," *Econometrica*, 77(3), 953–973.

Fletcher, J. M., and S. F. Lehrer (2009): "Using genetic lotteries within families to examine the causal impact of poor health on academic achievement," Discussion paper, National Bureau of Economic Research.

Fredman, L., J. A. Cauley, S. Satterfield, E. Simonsick, S. M. Spencer, H. N. Ayonayon, T. B. Harris, et al. (2008): "Caregiving, mortality, and mobility decline: the Health, Aging, and Body Composition (Health ABC) study," *Archives of internal medicine*, 168(19), 2154.

Glei, D. A., D. A. Landau, N. Goldman, Y.-L. Chuang, G. Rodríguez, and M. Weinstein (2005): "Participating in social activities helps preserve cognitive function: an analysis of a longitudinal, population-based study of the elderly," *International Journal of Epidemiology*, 34(4), 864–871.

Glendinning, C., H. Arksey, F. Tjadens, M. Morée, N. Moran, and H. Nies (2009): "Care provision within families and its socio-economic impact on care providers across the European Union," *Research Works No. 5.*

Green, A. F., G. Rebok, and C. G. Lyketsos (2008): "Influence of social network characteristics on cognition and functional status with aging," *International journal of geriatric psychiatry*, 23(9), 972–978.

Green, D. A., and W. C. Riddell (2001): "Literacy, Numeracy and Labour Market Outcomes in Canada," International adult literacy survey, Statistics Canada.

Grossman, M. (1972): "On the concept of health capital and the demand for health," *The Journal of Political Economy*, 80, 223 255.

Guyatt, G. H., P. F. Devereaux, J. Lexchin, S. B. Stone, A. Yalnizyan, D. Himmelstein, S. Woolhandler, Q. Zhou, L. J. Goldsmith, D. J. Cook, T. Haines, C. Lacchetti, J. N. Lavis, T. Sullivan, E. Mills, S. Kraus, and N. Bhatnagar (2007): "A systematic review of studies comparing health outcomes in Canada and the United States," *Open Medicine*, 1(1), 27–36.

Haberkern, K., and M. Szydlik (2010): "State care provision, societal opinion and children's care of older parents in 11 European countries," *Ageing and Society*, 30(2), 299.

Hansen, K. T., J. J. Heckman, and K. J. Mullen (2004): "The effect of schooling and ability on achievement test scores," *Journal of econometrics*, 121(1), 39–98.

Hanushek, E. A. (2003): "The Failure of Input-based Schooling Policies," *The Economic Journal*, 113(485), F64–F98.

He, D., and P. McHenry (2013): "Does Labor Force Participation Reduce Informal Caregiving?," *Available at SSRN 2292700.*

Heitmueller, A. (2007): "The chicken or the egg?: Endogeneity in labour market participation of informal carers in England," *Journal of Health Economics*, 26(3), 536–559.

Heitmueller, A., and K. Inglis (2007): "The earnings of informal carers: Wage differentials and opportunity costs," *Journal of health economics*, 26(4), 821–841.

Hirst, M. (2005): "Carer distress: a prospective, population-based study," *Social Science & Medicine*, 61(3), 697–708.

Horsman, J., W. Furlong, D. Feeny, and G. Torrance (2003): "The Health Utilities Index (HUI): concepts, measurement properties and applications," *Health Quality of Life Outcomes*, 1(45), 45.

Hsu, H. C. (2007): "Does social participation by the elderly reduce mortality and cognitive impairment?," *Aging & mental health*, 11(6), 699–707.

Idler, E. L., and Y. Benyamini (1997): "Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies," *Journal of Health and Social Behavior*, 38(1), 21–37.

Ingram, B. F., and G. R. Neumann (2006): "The returns to skill," *Labour Economics*, 13(1), 35–59.

Jann, B. (2008): "A Stata implementation of the Blinder-Oaxaca decomposition," *Stata Journal*, 8(4), 453–479.

Jürges, H. (2007): "True health vs response styles: Exploring cross-country differences in self-reported health," *Health Economics*, 16, 163–178.

——— (2009): "Healthy minds in healthy bodies," *Scottish Journal of Political Economy*, 56, 296–320.

Jürges, H., M. Avendano, and J. P. Mackenbach (2008): "Are different measures of self-rated health comparable? An assessment in five European countries," *European journal of epidemiology*, 23(12), 773–781.

Kerkhofs, M., and M. Lindeboom (1995): "Subjective Health Measures and State Dependent Reporting Errors," *Health Economics*, 4, 221–235.

King, D., and L. Pickard (2013): "When is a carer's employment at risk?: longitudinal analysis of unpaid care and employment in midlife in England," *Health and social care in the community*, 21(3), 303–314.

Knighton, T., P. Brochu, T. Gluszynski, et al. (2010): "Measuring up: Canadian results of the OECD PISA study: the performance of Canada's youth in reading, mathematics and science 2009: first results for Canadians aged 15," .

Knighton, T., and P. Bussière (2006): "Educational Outcomes at Age 19 Associated with Reading Ability at Age 15," Culture, tourism and the centre for education statistics research papers, Statistics Canada.

Kolenikov, S., and G. Angeles (2009): "Socioeconomic Status Measurement With Discrete Proxy Variables: Is Principal Component Analysis A Reliable Answer?," *Review of Income and Wealth*, 55(1), 128–165.

Lee, S., I. Kawachi, and F. Grodstein (2004): "Does caregiving stress affect cognitive function in older women?," *The Journal of nervous and mental disease*, 192(1), 51–57.

Leipold, B., C. Schacke, and S. Zank (2008): "Personal growth and cognitive complexity in caregivers of patients with dementia," *European Journal of Ageing*, 5(3), 203–214.

Lilly, M. B., A. Laporte, and P. C. Coyte (2007): "Labor market work and home care's unpaid caregivers: a systematic review of labor force participation rates, predictors of labor market withdrawal, and hours of work," *Milbank Quarterly*, 85(4), 641–690.

Lindeboom, M., and M. Kerkhofs (2009): "Health and Work of the Elderly: Subjective Health Measures, Reporting Errors and the Endogenous Relationship Between Health and Work," *Journal of Applied Econometrics*, 24, 1024–1046.

Lindeboom, M., and E. Van Doorslaer (2004): "Cut-point shift and index shift in self-reported health," *Journal of health economics*, 23(6), 1083–1099.

Logan, S., and R. Johnston (2009): "Gender differences in reading ability and attitudes: Examining where these differences lie," *Journal of Research in Reading*, 32(2), 199–214.

Lusardi, A., and O. S. Mitchell (2011): "Financial Literacy and Planning: Implications for Retirement Wellbeing," Working Paper 17078, National Bureau of Economic Research.

Mackenzie, C. S., U. J. Wiprzycka, L. Hasher, and D. Goldstein (2009): "Associations between psychological distress, learning, and memory in spouse caregivers of older adults," *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 64(6), 742–746.

Maddox, G. L., and E. B. Douglass (1973): "Self-Assessment of Health: A Longitudinal Study of Elderly Subjects," *Journal of Health and Social Behavior*, 14(1), 87–93.

Makdissi, P., M. Yazbeck, and H. Coldeboeuf (2011): "A Fuzzy Approach to the Measurement of Leakages for North American Health Systems," Cahiers de recherche 1107, CIRPEE.

Mander, A., and D. Clayton (1999): "HOTDECK: Stata module to impute missing values using the hotdeck method," Statistical Software Components, Boston College Department of Economics.

Mascolo, M. F., and K. W. Fischer (2010): *The Dynamic Development of Thinking, Feeling, and Acting over the Life Span*, The Handbook of Life-Span Development. John Wiley & Sons, Inc.

Matire, L. M., and R. Schulz (2012): "Informal caregiving to older adults: Health effects of providing and receiving care," in *Handbook of Health Psychology*, pp. 293–308. Psychology Press, 2 edn.

Oaxaca, R. (1973): "Male-Female Wage Differentials in Urban Labor Markets," *International Economic Review*, 14(3), 693–709.

OECD (2005): *Long-term care for older people*. OECD, Paris.

——— (2010): *PISA 2009 Results: What Students Know and Can Do Student Performance in Reading, Mathematics and Science*, vol. 1. OECD Publishing.

——— (2012): *Learning beyond Fifteen: Ten Years after PISA*. OECD Publishing.

——— (2013): *Health at a glance 2013: OECD indicators*. OECD, Paris.

Ogg, J., and S. Renaut (2006): "The support of parents in old age by those born during 1945-1954: A European perspective," *Ageing and Society*, 26(5), 723–744.

ONeill, J. E., and D. M. ONeill (2007): "Health status, health care and inequality: Canada vs. the US," *Frontiers in Health Policy Research*, 10(1), 3.

Pavalko, E. K., and J. E. Artis (1997): "Women's caregiving and paid work: Causal relationships in late midlife," *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 52(4), 170–S179.

Pfarr, C., A. Schmid, and U. Schneider (2011): "Reporting Heterogeneity in Self-Assessed Health among Elderly Europeans: The Impact of Mental and Physical Health Status," Discussion Paper 02-11, Universität Bayreuth.

Pinquart, M., and S. Sörensen (2003): "Associations of stressors and uplifts of caregiving with caregiver burden and depressive mood: a meta-analysis," *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 58(2), P112–P128.

——— (2005): "Ethnic differences in stressors, resources, and psychological outcomes of family caregiving: A meta-analysis," *The Gerontologist*, 45(1), 90–106.

——— (2007): "Correlates of physical health of informal caregivers: a meta-analysis," *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 62(2), P126–P137.

Poterba, J. M., S. F. Venti, and D. A. Wise (2010): "The Asset Cost of Poor health," NBER Working Papers 16389, National Bureau of Economic Research, Inc.

Reder, S. (2000): "Adult Literacy and Postsecondary Education Students Overlapping Populations and Learning Trajectories," vol. 1 of *Annual review of adult learning and literacy*, pp. 111 – 157. Jossey-Bass.

Reilly, M. (1993): "Data Analysis Using Hot Deck Multiple Imputation," *Journal of the Royal Statistical Society. Series D (The Statistician)*, 42(3), 307–313.

Sadana, R., C. D. Mathers, A. D. Lopez, C. J. L. Murray, and K. Iburg (2000): "Comparative analyses of more than 50 household surveys on health status," Global Programme on Evidence for Health Policy Discussion Paper 15, World Health Organization.

Sadana, R., A. Tandon, C. J. L. Murray, I. Serdobova, Y. Cao, W. J. Xie, S. Chatterji, and B. L. Ustün (2002): "Describing Population Health in Six Domains: Comparable Results From 66 Household Surveys," Global Programme on Evidence for Health Policy Discussion Paper 43, World Health Organization.

Sanmartin, C., E. Ng, D. Blackwell, J. Gentleman, M. Martinez, and C. Simile (2004): "Joint Canada/United States Survey of Health, 2002-03," Health Analysis and Measurement Group, Statistics Canada.

Schaffer, M. E. (2012): "xtivreg2: Stata module to perform extended IV/2SLS, GMM and AC/HAC, LIML and k-class regression for panel data models," *Statistical Software Components*.

Schulz, R., J. Newsom, M. Mittelmark, L. Burton, C. Hirsch, and S. Jackson (1997): "Health effects of caregiving: The caregiver health effects study: An ancillary study of the cardiovascular health study," *Annals of Behavioral Medicine*, 19(2), 110–116.

Shmueli, A. (2003): "Socio-economic and demographic variation in health and in its measures: the issue of reporting heterogeneity," *Social Science & Medicine*, 57, 125–134.

Silverman, B. W. (1986): *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London.

Statistics Canada (2005): "Youth in Transition Survey Cycle 1 YITS Reading Cohort User Guide," Discussion paper, Statistics Canada.

Statistics Canada and the National Center for Health Statistics, Centrers for Disease Control and Prevention (2004): "Joint Canada/United States Survey of Health: Public Use Microdata File User Guide," JCUSH Public Use Microdata File.

Todd, P. E., and I. W. Kenneth (2003): "On the Specification and Estimation of the Production Function for Cognitive Achievement," *Economic Journal*, 113(485), F3–F33.

van Doorslaer, E., and A. M. Jones (2003): "Inequalities in self-reported health: validation of a new approach to measurement," *Journal of Health Economics*, 22(1), 61–87.

Van Houtven, C. H., N. B. Coe, and M. Skira (2013): "The effect of informal care on work and wages," *Journal of Health Economics*, 32(1), 240 – 252.

Vitaliano, P., D. Echeverria, M. Shelkey, J. Zhang, and J. Scanlan (2007): "A cognitive psychophysiological model to predict functional decline in chronically stressed older adults," *Journal of Clinical Psychology in Medical Settings*, 14(3), 177–190.

Vitaliano, P. P., D. Echeverria, J. Yi, P. E. M. Phillips, H. Young, and I. C. Siegler (2005): "Psychophysiological mediators of caregiver stress and differential cognitive decline.," *Psychology and aging*, 20(3), 402.

Vitaliano, P. P., M. Murphy, H. M. Young, D. Echeverria, and S. Borson (2011): "Does Caring for a Spouse with Dementia Promote Cognitive Decline? A Hypothesis and Proposed Mechanisms," *Journal of the American Geriatrics Society*, 59(5), 900–908.

Willis, R. J., and S. Rosen (1979): "Education and self-selection," *Journal of Political Economy*, 87, 7–36.

Wooldridge, J. M. (2002): *Econometric analysis of cross section and panel data.* MIT press.

Ziebarth, N. (2010): "Measurement of Health, the Sensitivity of the Concentration Index, and Reporting Heterogeneity," *Social Science and Medicine*, 71, 116–124.

Appendix A

Additional Tables for Chapter 2

Table A.1: Changes in caregiving behaviour over time

		Number of individuals
Total sample		10,313
	Parent(s) alive at time of first interview	3,588
	Parent(s) alive at time of last interview	2,377
Ever caregiver		1,994
	Continuous caregivers	519
	Begin caregiving	592
	End caregiving	1,067
Ever weekly caregiver		1,437
	Continuous caregivers	320
	Begin caregiving	484
	End caregiving	775
Ever daily caregiver		737
	Continuous caregivers	141
	Begin caregiving	254
	End caregiving	410

Table A.2: Parental health and distance to the parent

	Distance to the parent					
Parental health	≤5 km	>5 km		≤25 km	>25 km	
poor	0.29	0.31	(0.407)	0.30	0.31	(0.826)
fair	0.34	0.37	(0.346)	0.35	0.37	(0.380)
good	0.29	0.25	(0.162)	0.28	0.25	(0.247)
very good	0.07	0.07	(0.504)	0.07	0.07	(0.943)

Weighted values for individuals with at least on living parent based on individuals' first observation. Two sided p-values of adjusted Wald test of equality of means in parentheses. N=3,323

Table A.3: Official retirement age for women in SHARE countries

Country	Retirement age	Comments
Austria	60	
Belgium	65	Workers can retire at age 60 with 40 years of contributions.
Czech Republic	61	
Denmark	65	
France	60	Workers can retire at age 60 with 40 years of contributions.
Germany	65	
Greece	62	Workers can retire at age 59 with 35 years of contributions.
Italy	60	Workers can retire at 57 (56 for manual workers) with 35 years of contributions.
Netherlands	65	
Poland	60	
Spain	65	
Switzerland	64	
Sweden	65	

Source: OECD Statistics on average effective age and official age of retirement in OECD countries (official retirement age for 2010)

Table A.4: First stage regressions for employed

Dependent variable:	Caregiving (any frequency)			Weekly caregiving			Daily caregiving		
Parent in poor health	0.198***			0.170***			0.109***		
	(0.023)			(0.022)			(0.017)		
Parent in poor health x		0.217***	0.001		0.197***	0.000		0.121***	0.000
family care country		(0.032)	(0.001)		(0.032)	0.000		(0.025)	0.000
Parent in poor health x		0.000	0.166***		0.000	0.130***		0.000	0.092***
formal care country		(0.001)	(0.033)		(0.001)	(0.027)		0.000	(0.019)
Chronic conditions	0.004	0.004*	0.000	0.006**	0.005**	0.001	0.004*	0.003*	0.000
	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.003)	(0.002)	(0.001)
Limitations with ADL	-0.001	-0.003	0.002	0.000	-0.001	0.000	-0.001	-0.001	-0.003
	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.002)	(0.002)
Limitations with IADL	-0.003	-0.002	-0.001	-0.004	-0.005	0.001	-0.003	-0.001	0.002
	(0.006)	(0.005)	(0.002)	(0.006)	(0.005)	(0.002)	(0.006)	(0.002)	(0.002)
Age	-0.061	0.005	-0.069*	-0.002	0.012	-0.018	-0.061	0.005	0.019
	(0.040)	(0.008)	(0.039)	(0.031)	(0.008)	(0.030)	(0.040)	(0.006)	(0.019)
Age squared/100	0.046	0.001	0.047	0.003	-0.001	0.006	0.046	-0.001	-0.015
	(0.029)	(0.004)	(0.029)	(0.022)	(0.003)	(0.022)	(0.029)	(0.003)	(0.014)
Wave 2	-0.004	-0.004	-0.001	0.000	-0.011	0.010	-0.008	-0.002	-0.006
	(0.019)	(0.011)	(0.016)	(0.017)	(0.010)	(0.014)	(0.015)	(0.007)	(0.013)
Wave 4	-0.041	-0.031*	-0.011	-0.031	-0.043**	0.010	-0.013	-0.012	-0.002
	(0.025)	(0.018)	(0.018)	(0.023)	(0.017)	(0.015)	(0.020)	(0.014)	(0.014)
Married	0.006	0.002	0.004	0.008	0.006	0.002	0.025*	0.009	0.016
	(0.021)	(0.015)	(0.015)	(0.020)	(0.014)	(0.015)	(0.014)	(0.011)	(0.010)
Household size	0.004	0.005	-0.001	-0.002	-0.002	-0.001	0.004	0.000	0.004
	(0.006)	(0.005)	(0.003)	(0.005)	(0.005)	(0.003)	(0.005)	(0.004)	(0.003)
Reached off. retirement age	-0.004	0.003	-0.007	-0.014	-0.001	-0.013**	-0.001	-0.004	0.002
	(0.011)	(0.009)	(0.006)	(0.011)	(0.009)	(0.006)	(0.011)	(0.008)	(0.008)
2 years to off. retirement age	-0.014	-0.003	-0.011	-0.011	-0.009	-0.002	-0.015	-0.008	-0.007
	(0.012)	(0.010)	(0.007)	(0.011)	(0.009)	(0.006)	(0.010)	(0.008)	(0.006)
5 years to off. retirement age	0.019	0.006	0.013*	-0.012	-0.008	-0.003	-0.005	-0.007	0.002
	(0.015)	(0.013)	(0.008)	(0.013)	(0.012)	(0.006)	(0.010)	(0.009)	(0.004)
Observations	22,379	22,379	22,379	22,379	22,379	22,379	22,379	22,379	22,379
Number of panelid	9,809	9,809	9,809	9,809	9,809	9,809	9,809	9,809	9,809
R-squared	0.06	0.07	0.04	0.04	0.06	0.03	0.03	0.03	0.02

Standard errors clustered at household level in parentheses. $^{**}p<0.01$, $^{*}p<0.05$, $^{+}p<0.1$
Regressions also include age*country and age squared/100*country interaction terms.

Table A.5: First stage regressions for hours worked

Dependent variable:	Caregiving (any frequency)			Weekly caregiving			Daily caregiving		
Parent in poor health	0.195***			0.170***			0.108***		
	(0.024)			(0.022)			(0.017)		
Parent in poor health x		0.215***	0.000		0.195***	0.000		0.118***	0.000
family care country		(0.032)	(0.001)		(0.032)	0.000		(0.025)	0.000
Parent in poor health x		0.000	0.162***		0.000	0.130***		0.000	0.093***
formal care country		(0.001)	(0.034)		(0.001)	(0.028)		0.000	(0.019)
Chronic conditions	0.004*	0.004*	0.000	0.006**	0.005**	0.001	0.003*	0.003*	0.000
	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)
Limitations with ADL	-0.001	-0.003	0.002	0.000	-0.001	0.000	-0.004	-0.001	-0.003
	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Limitations with IADL	-0.003	-0.002	-0.001	-0.005	-0.005	0.001	0.000	-0.002	0.002
	(0.006)	(0.005)	(0.002)	(0.006)	(0.005)	(0.002)	(0.003)	(0.002)	(0.002)
Age	-0.061	0.005	-0.069*	0.002	0.012	-0.013	0.027	0.005	0.020
	(0.040)	(0.008)	(0.039)	(0.031)	(0.008)	(0.030)	(0.020)	(0.006)	(0.020)
Age squared/100	0.046	0.001	0.048	-0.001	-0.001	0.003	-0.018	-0.001	-0.016
	(0.029)	(0.004)	(0.029)	(0.022)	(0.003)	(0.022)	(0.014)	(0.003)	(0.014)
Wave 2	-0.006	-0.005	-0.002	0.001	-0.010	0.010	-0.008	-0.002	-0.007
	(0.020)	(0.011)	(0.016)	(0.017)	(0.010)	(0.014)	(0.015)	(0.007)	(0.014)
Wave 4	-0.042*	-0.032*	-0.011	-0.031	-0.043**	0.010	-0.013	-0.012	-0.002
	(0.025)	(0.018)	(0.018)	(0.023)	(0.017)	(0.015)	(0.020)	(0.014)	(0.014)
Married	0.006	0.002	0.004	0.008	0.006	0.002	0.025*	0.009	0.016
	(0.021)	(0.015)	(0.015)	(0.020)	(0.014)	(0.015)	(0.014)	(0.011)	(0.010)
Household size	0.005	0.006	-0.001	-0.002	-0.001	-0.001	0.005	0.000	0.004
	(0.006)	(0.005)	(0.003)	(0.005)	(0.005)	(0.003)	(0.005)	(0.004)	(0.003)
Reached off. retirement age	-0.002	0.005	-0.007	-0.012	0.001	-0.013**	0.000	-0.002	0.002
	(0.011)	(0.009)	(0.006)	(0.011)	(0.009)	(0.006)	(0.011)	(0.008)	(0.008)
2 years to off. retirement age	-0.014	-0.002	-0.012	-0.012	-0.010	-0.002	-0.016	-0.008	-0.008
	(0.013)	(0.010)	(0.007)	(0.011)	(0.009)	(0.006)	(0.010)	(0.008)	(0.006)
5 years to off. retirement age	0.020	0.007	0.013	-0.012	-0.008	-0.003	-0.005	-0.007	0.002
	(0.015)	(0.013)	(0.008)	(0.013)	(0.012)	(0.006)	(0.010)	(0.009)	(0.004)
Observations	22,228	22,228	22,228	22,228	22,228	22,228	22,228	22,228	22,228
Number of panelid	9,748	9,748	9,748	9,748	9,748	9,748	9,748	9,748	9,748
R-squared	0.06	0.07	0.04	0.04	0.06	0.03	0.03	0.03	0.02

Standard errors clustered at household level in parentheses. $^{**}p<0.01$, $^{*}p<0.05$, $^{+}p<0.1$
Regressions also include age*country and age squared/100*country interaction terms.

Table A.6: First stage regressions for verbal fluency

Dependent variable:	Caregiving (any frequency)			Weekly caregiving			Daily caregiving		
Parent in poor health	0.201***			0.171***			0.111***		
	(0.024)			(0.022)			(0.017)		
Parent in poor health x		0.220***	0.001*		0.194***	0.001		0.123***	0.000
family care country		(0.032)	(0.001)		(0.032)	(0.001)		(0.025)	0.000
Parent in poor health x		0.001	0.168***		0.001	0.135***		0.000	0.093***
formal care country		(0.001)	(0.034)		(0.001)	(0.027)		0.000	(0.019)
Chronic conditions	0.004	0.004*	0.000	0.005**	0.005**	0.001	0.003*	0.003*	0.000
	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)
Limitations with ADL	0.000	-0.003	0.002	0.000	-0.001	0.000	-0.004	-0.002	-0.003
	(0.005)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Limitations with IADL	-0.003	-0.001	-0.001	-0.004	-0.004	0.001	0.000	-0.001	0.002
	(0.006)	(0.006)	(0.002)	(0.006)	(0.006)	(0.002)	(0.003)	(0.002)	(0.002)
Age	-0.055	0.005	-0.062	-0.004	0.009	-0.016	0.019	0.001	0.016
	(0.039)	(0.006)	(0.039)	(0.030)	(0.006)	(0.030)	(0.020)	(0.004)	(0.019)
Age squared/100	0.041	0.002	0.041	0.002	0.001	0.004	-0.013	0.001	-0.013
	(0.029)	(0.002)	(0.029)	(0.022)	(0.002)	(0.022)	(0.014)	(0.001)	(0.014)
Wave 2	-0.005	-0.007	0.001	0.001	-0.012	0.012	-0.008	-0.001	-0.007
	(0.020)	(0.011)	(0.017)	(0.017)	(0.010)	(0.014)	(0.016)	(0.007)	(0.014)
Wave 4	-0.043*	-0.035*	-0.009	-0.032	-0.044***	0.012	-0.013	-0.011	-0.003
	(0.026)	(0.018)	(0.018)	(0.023)	(0.017)	(0.015)	(0.020)	(0.014)	(0.014)
Married	0.003	0.005	-0.002	0.002	0.005	-0.003	0.028*	0.010	0.018*
	(0.021)	(0.015)	(0.015)	(0.019)	(0.014)	(0.014)	(0.015)	(0.011)	(0.010)
Eployed	0.006	-0.007	0.014**	0.011	0.002	0.010	-0.005	0.000	-0.004
	(0.014)	(0.012)	(0.007)	(0.013)	(0.012)	(0.006)	(0.009)	(0.008)	(0.004)
Other activities	0.016*	0.014*	0.001	0.009	0.011*	-0.002	0.002	0.004	-0.002
	(0.009)	(0.007)	(0.005)	(0.008)	(0.007)	(0.004)	(0.005)	(0.004)	(0.003)
Financial distress	0.004	-0.002	0.006	-0.001	-0.008	0.007*	0.011	0.007	0.004
	(0.008)	(0.007)	(0.004)	(0.008)	(0.007)	(0.004)	(0.007)	(0.006)	(0.004)
Observations	22,161	22,161	22,161	22,161	22,161	22,161	22,161	22,161	22,161
Number of panelid	9,711	9,711	9,711	9,711	9,711	9,711	9,711	9,711	9,711
R-squared	0.06	0.07	0.04	0.05	0.06	0.03	0.03	0.03	0.02

Standard errors clustered at household level in parentheses. **$p<0.01$, *$p<0.05$, $^{+}p<0.1$
Regressions also include age*country and age squared/100*country interaction terms.

Table A.7: First stage regressions for short-term memory

Dependent variable:	Caregiving (any frequency)			Weekly caregiving			Daily caregiving		
Parent in poor health	0.201***			0.171***			0.111***		
	(0.024)			(0.022)			(0.017)		
Parent in poor health x		0.219***	0.001*		0.193***	0.001		0.122***	0.000
family care country		(0.032)	(0.001)		(0.032)	(0.001)		(0.025)	0.000
Parent in poor health x		0.001	0.171***		0.001	0.136***		0.000	0.094***
formal care country		(0.001)	(0.034)		(0.001)	(0.027)		0.000	(0.019)
Chronic conditions	0.004	0.004*	0.000	0.005**	0.005**	0.001	0.003*	0.010	0.018*
	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.011)	(0.010)
Limitations with ADL	-0.001	-0.003	0.002	0.000	-0.001	0.000	-0.004	0.003*	0.000
	(0.005)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)	(0.001)
Limitations with IADL	-0.003	-0.001	-0.001	-0.004	-0.004	0.001	0.000	-0.002	-0.003
	(0.006)	(0.006)	(0.002)	(0.006)	(0.006)	(0.002)	(0.003)	(0.002)	(0.002)
Age	-0.055	0.005	-0.062	-0.004	0.009	-0.016	0.019	-0.001	0.002
	(0.039)	(0.006)	(0.039)	(0.030)	(0.006)	(0.030)	(0.020)	(0.002)	(0.002)
Age squared/100	0.041	0.002	0.041	0.002	0.001	0.004	-0.013	0.001	0.016
	(0.029)	(0.002)	(0.029)	(0.022)	(0.002)	(0.022)	(0.014)	(0.004)	(0.019)
Wave 2	-0.005	-0.007	0.001	0.001	-0.012	0.012	-0.008	-0.008	0.015
	(0.020)	(0.011)	(0.017)	(0.017)	(0.010)	(0.014)	(0.016)	(0.021)	(0.014)
Wave 4	-0.042*	-0.035*	-0.009	-0.031	-0.044***	0.012	-0.013	-0.001	-0.007
	(0.026)	(0.018)	(0.018)	(0.023)	(0.017)	(0.015)	(0.020)	(0.007)	(0.014)
Married	0.003	0.005	-0.002	0.002	0.005	-0.003	0.028*	-0.011	-0.002
	(0.021)	(0.015)	(0.015)	(0.019)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Eployed	0.007	-0.007	0.015**	0.011	0.002	0.009	-0.005	0.000	-0.005
	(0.014)	(0.012)	(0.007)	(0.013)	(0.012)	(0.006)	(0.009)	(0.008)	(0.004)
Other activities	0.016*	0.014*	0.001	0.009	0.011*	-0.002	0.002	0.004	-0.002
	(0.009)	(0.007)	(0.005)	(0.008)	(0.007)	(0.004)	(0.005)	(0.004)	(0.003)
Financial distress	0.004	-0.002	0.006	-0.001	-0.008	0.007*	0.011	0.007	0.004
	(0.008)	(0.007)	(0.004)	(0.008)	(0.006)	(0.004)	(0.007)	(0.006)	(0.004)
Observations	22,202	22,202	22,202	22,202	22,202	22,202	22,202	22,202	22,202
Number of panelid	9,730	9,730	9,730	9,730	9,730	9,730	9,730	9,730	9,730
R-squared	0.06	0.07	0.04	0.05	0.06	0.03	0.03	0.03	0.02

Standard errors clustered at household level in parentheses. $^{**}p<0.01$, $^{*}p<0.05$, $^{+}p<0.1$
Regressions also include age*country and age squared/100*country interaction terms.

Table A.8: First stage regressions for long-term memory

Dependent variable:	Caregiving (any frequency)			Weekly caregiving			Daily caregiving		
Parent in poor health	0.201***			0.171***			0.111***		
	(0.024)			(0.022)			(0.017)		
Parent in poor health x		0.219***	0.001*		0.193***	0.001		0.122***	0.000
family care country		(0.032)	(0.001)		(0.032)	(0.001)		(0.025)	0.000
Parent in poor health x		0.001	0.171***		0.001	0.136***		0.000	0.094***
formal care country		(0.001)	(0.034)		(0.001)	(0.027)		0.000	(0.019)
Chronic conditions	0.004	0.004*	0.000	0.005**	0.005**	0.001	0.003*	0.003*	0.000
	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)
Limitations with ADL	-0.001	-0.003	0.002	0.000	-0.001	0.000	-0.004	-0.002	-0.003
	(0.005)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Limitations with IADL	-0.003	-0.001	-0.001	-0.004	-0.004	0.001	0.000	-0.001	0.002
	(0.006)	(0.006)	(0.002)	(0.006)	(0.006)	(0.002)	(0.003)	(0.002)	(0.002)
Age	-0.055	0.005	-0.062	-0.004	0.009	-0.016	0.019	0.001	0.016
	(0.039)	(0.006)	(0.039)	(0.030)	(0.006)	(0.030)	(0.020)	(0.004)	(0.019)
Age squared/100	0.041	0.002	0.041	0.002	0.001	0.004	-0.013	0.001	-0.013
	(0.029)	(0.002)	(0.029)	(0.022)	(0.002)	(0.022)	(0.014)	(0.001)	(0.014)
Wave 2	-0.005	-0.007	0.001	0.001	-0.012	0.012	-0.008	-0.001	-0.007
	(0.020)	(0.011)	(0.017)	(0.017)	(0.010)	(0.014)	(0.016)	(0.007)	(0.014)
Wave 4	-0.042*	-0.035*	-0.009	-0.031	-0.044***	0.012	-0.013	-0.011	-0.002
	(0.026)	(0.018)	(0.018)	(0.023)	(0.017)	(0.015)	(0.020)	(0.014)	(0.014)
Married	0.003	0.005	-0.002	0.002	0.005	-0.003	0.028*	0.010	0.018*
	(0.021)	(0.015)	(0.015)	(0.019)	(0.014)	(0.014)	(0.014)	(0.011)	(0.010)
Eployed	0.007	-0.007	0.015**	0.011	0.002	0.009	-0.005	0.000	-0.005
	(0.014)	(0.012)	(0.007)	(0.013)	(0.012)	(0.006)	(0.009)	(0.008)	(0.004)
Other activities	0.016*	0.014*	0.001	0.009	0.011*	-0.002	0.002	0.004	-0.002
	(0.009)	(0.007)	(0.005)	(0.008)	(0.007)	(0.004)	(0.005)	(0.004)	(0.003)
Financial distress	0.004	-0.002	0.006	-0.001	-0.008	0.007*	0.011	0.007	0.004
	(0.008)	(0.007)	(0.004)	(0.008)	(0.006)	(0.004)	(0.007)	(0.006)	(0.004)
Observations	22,207	22,207	22,207	22,207	22,207	22,207	22,207	22,207	22,207
Number of panelid	9,732	9,732	9,732	9,732	9,732	9,732	9,732	9,732	9,732
R-squared	0.06	0.07	0.04	0.05	0.06	0.03	0.03	0.03	0.02

Standard errors clustered at household level in parentheses. $^{**}p<0.01$, $^{*}p<0.05$, $^{+}p<0.1$
Regressions also include age*country and age squared/100*country interaction terms.

Table A.9: First stage regressions for numeracy

Dependent variable:	Caregiving (any frequency)			Weekly caregiving			Daily caregiving		
Parent in poor health	0.200***			0.172***			0.112***		
	(0.024)			(0.022)			(0.017)		
Parent in poor health x		0.220***	0.001*		0.193***	0.001		0.123***	0.000
family care country		(0.032)	(0.001)		(0.032)	(0.001)		(0.025)	0.000
Parent in poor health x		0.001	0.167***		0.001	0.137***		0.000	0.095***
formal care country		(0.001)	(0.034)		(0.001)	(0.027)		0.000	(0.019)
Chronic conditions	0.004	0.004*	0.000	0.006**	0.005**	0.001	0.003*	0.003*	0.000
	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)
Limitations with ADL	0.000	-0.003	0.002	0.000	-0.001	0.000	-0.004	-0.002	-0.003
	(0.005)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Limitations with IADL	-0.003	-0.001	-0.001	-0.004	-0.004	0.001	0.000	-0.001	0.002
	(0.006)	(0.006)	(0.002)	(0.006)	(0.006)	(0.002)	(0.003)	(0.002)	(0.002)
Age	-0.053	0.005	-0.061	-0.004	0.010*	-0.016	0.019	0.002	0.016
	(0.039)	(0.006)	(0.039)	(0.030)	(0.006)	(0.030)	(0.019)	(0.004)	(0.019)
Age squared/100	0.040	0.002	0.040	0.002	0.001	0.003	-0.013	0.001	-0.013
	(0.028)	(0.002)	(0.028)	(0.022)	(0.002)	(0.022)	(0.014)	(0.001)	(0.014)
Wave 2	-0.006	-0.007	0.000	0.002	-0.012	0.013	-0.008	-0.002	-0.007
	(0.020)	(0.011)	(0.017)	(0.017)	(0.010)	(0.014)	(0.016)	(0.007)	(0.014)
Wave 4	-0.045*	-0.036**	-0.010	-0.032	-0.046***	0.012	-0.014	-0.012	-0.002
	(0.025)	(0.018)	(0.018)	(0.023)	(0.017)	(0.015)	(0.020)	(0.014)	(0.014)
Married	0.003	0.004	-0.002	0.002	0.005	-0.003	0.028*	0.010	0.018*
	(0.021)	(0.015)	(0.015)	(0.019)	(0.014)	(0.014)	(0.014)	(0.011)	(0.010)
Eployed	0.007	-0.007	0.015**	0.010	0.002	0.009	-0.005	0.000	-0.005
	(0.014)	(0.012)	(0.007)	(0.013)	(0.012)	(0.006)	(0.009)	(0.008)	(0.004)
Other activities	0.017*	0.015**	0.002	0.009	0.012*	-0.002	0.003	0.005	-0.002
	(0.009)	(0.007)	(0.005)	(0.008)	(0.007)	(0.004)	(0.005)	(0.004)	(0.003)
Financial distress	0.005	-0.002	0.007	-0.001	-0.008	0.007*	0.011	0.007	0.004
	(0.008)	(0.007)	(0.004)	(0.008)	(0.006)	(0.004)	(0.007)	(0.006)	(0.004)
Observations	22,226	22,226	22,226	22,226	22,226	22,226	22,226	22,226	22,226
Number of panelid	9,744	9,744	9,744	9,744	9,744	9,744	9,744	9,744	9,744
R-squared	0.06	0.07	0.04	0.05	0.06	0.03	0.03	0.03	0.02

Standard errors clustered at household level in parentheses. $^{**}p<0.01$, $^{*}p<0.05$, $^{+}p<0.1$
Regressions also include age*country and age squared/100*country interaction terms.

Table A.10: First stage regressions for mental health

Dependent variable:	Caregiving (any frequency)			Weekly caregiving			Daily caregiving		
Parent in poor health	0.178***						0.118***		
	(0.025)						(0.018)		
Parent in poor health x		0.206***	-0.011**	0.166***	0.189***	0.000		0.123***	0.003
family care country		(0.033)	(0.005)	(0.023)	(0.032)	(0.005)		(0.025)	(0.004)
Parent in poor health x		-0.011	0.161***		-0.007	0.136***		0.005	0.100***
formal care country		(0.007)	(0.035)		(0.007)	(0.029)		(0.006)	(0.020)
Parental loss	-0.061***	-0.031*	-0.031**	-0.021	-0.019	-0.002	0.021	0.012	0.009
	(0.022)	(0.017)	(0.014)	(0.020)	(0.016)	(0.012)	(0.018)	(0.014)	(0.010)
Chronic conditions	0.003	0.003	0.000	0.005**	0.004**	0.001	0.003	0.003	0.000
	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)
Limitations with ADL	-0.001	-0.003	0.002	0.000	-0.001	0.000	-0.004	-0.001	-0.003
	(0.005)	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
Limitations with IADL	-0.003	-0.001	-0.001	-0.004	-0.005	0.000	0.000	-0.001	0.001
	(0.007)	(0.006)	(0.002)	(0.007)	(0.006)	(0.002)	(0.003)	(0.002)	(0.002)
Age	-0.052	0.010	-0.064	-0.006	0.013**	-0.023	0.015	0.000	0.014
	(0.039)	(0.007)	(0.039)	(0.030)	(0.006)	(0.030)	(0.020)	(0.005)	(0.019)
Age squared/100	0.040	-0.002	0.043	0.005	-0.002	0.009	-0.010	0.002	-0.010
	(0.029)	(0.003)	(0.029)	(0.022)	(0.002)	(0.022)	(0.014)	(0.002)	(0.014)
Wave 2	-0.003	-0.005	0.002	0.001	-0.012	0.012	-0.010	-0.002	-0.009
	(0.020)	(0.011)	(0.017)	(0.017)	(0.010)	(0.014)	(0.016)	(0.007)	(0.014)
Wave 4	-0.045*	-0.035*	-0.011	-0.034	-0.046***	0.011	-0.014	-0.011	-0.004
	(0.026)	(0.018)	(0.018)	(0.023)	(0.017)	(0.015)	(0.020)	(0.014)	(0.015)
Married	0.002	0.003	-0.002	0.002	0.004	-0.003	0.029**	0.010	0.019*
	(0.021)	(0.015)	(0.015)	(0.020)	(0.014)	(0.014)	(0.015)	(0.011)	(0.010)
Eployed	0.008	-0.006	0.014**	0.014	0.004	0.010*	-0.002	0.002	-0.004
	(0.014)	(0.013)	(0.007)	(0.013)	(0.012)	(0.006)	(0.009)	(0.008)	(0.004)
Other activities	0.016*	0.014*	0.003	0.008	0.009	-0.001	0.003	0.004	-0.002
	(0.009)	(0.007)	(0.005)	(0.008)	(0.007)	(0.004)	(0.005)	(0.004)	(0.003)
Financial distress	0.005	-0.002	0.007*	0.000	-0.007	0.008**	0.012*	0.007	0.005
	(0.008)	(0.007)	(0.004)	(0.008)	(0.007)	(0.004)	(0.007)	(0.006)	(0.004)
Observations	21,911	21,911	21,911	21,911	21,911	21,911	21,911	21,911	21,911
Number of panelid	9,611	9,611	9,611	9,611	9,611	9,611	9,611	9,611	9,611
R-squared	0.06	0.07	0.04	0.05	0.06	0.03	0.03	0.03	0.02

Standard errors clustered at household level in parentheses. $^{**}p<0.01$, $^{*}p<0.05$, $^{+}p<0.1$
Regressions also include age*country and age squared/100*country interaction terms.

Table A.11: First stage regressions for self-perceived health

Dependent variable:	Caregiving (any frequency)			Weekly caregiving			Daily caregiving		
Parent in poor health	0.200***			0.174***			0.110***		
	(0.023)			(0.022)			(0.017)		
Parent in poor health x		0.217***	0.001*		0.197***	0.001		0.121***	0.000
family care country		(0.032)	(0.001)		(0.032)	(0.001)		(0.025)	0.000
Parent in poor health x		0.001	0.171***		0.001	0.136***		0.000	0.094***
formal care country		(0.001)	(0.034)		(0.001)	(0.027)		0.000	(0.019)
Chronic conditions	0.004	0.004*	0.000	0.006**	0.005**	0.001	0.003*	0.003*	0.000
	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)
Limitations with ADL	0.000	-0.003	0.002	0.000	-0.001	0.000	-0.004	-0.002	-0.003
	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Limitations with IADL	-0.003	-0.001	-0.001	-0.004	-0.004	0.001	0.000	-0.001	0.002
	(0.006)	(0.006)	(0.002)	(0.006)	(0.006)	(0.002)	(0.003)	(0.002)	(0.002)
Age	-0.054	0.005	-0.061	-0.003	0.010*	-0.016	0.019	0.002	0.016
	(0.039)	(0.006)	(0.039)	(0.030)	(0.006)	(0.030)	(0.019)	(0.004)	(0.019)
Age squared/100	0.040	0.002	0.040	0.002	0.001	0.004	-0.013	0.001	-0.013
	(0.028)	(0.002)	(0.028)	(0.022)	(0.002)	(0.022)	(0.014)	(0.001)	(0.014)
Wave 2	-0.005	-0.007	0.001	0.000	-0.013	0.012	-0.008	-0.002	-0.007
	(0.020)	(0.011)	(0.017)	(0.017)	(0.010)	(0.014)	(0.015)	(0.007)	(0.014)
Wave 4	-0.043*	-0.036**	-0.009	-0.034	-0.047***	0.012	-0.014	-0.012	-0.002
	(0.025)	(0.018)	(0.018)	(0.023)	(0.017)	(0.015)	(0.020)	(0.014)	(0.014)
Married	0.003	0.004	-0.002	0.002	0.005	-0.003	0.028*	0.010	0.018*
	(0.021)	(0.015)	(0.015)	(0.019)	(0.014)	(0.014)	(0.014)	(0.011)	(0.010)
Eployed	0.007	-0.007	0.015**	0.011	0.003	0.009	-0.005	0.000	-0.005
	(0.014)	(0.012)	(0.007)	(0.013)	(0.012)	(0.006)	(0.009)	(0.008)	(0.004)
Other activities	0.016*	0.015**	0.001	0.008	0.010	-0.002	0.003	0.005	-0.002
	(0.009)	(0.007)	(0.005)	(0.008)	(0.007)	(0.004)	(0.005)	(0.004)	(0.003)
Financial distress	0.004	-0.002	0.006	0.000	-0.007	0.007*	0.011	0.006	0.004
	(0.008)	(0.007)	(0.004)	(0.008)	(0.007)	(0.004)	(0.007)	(0.006)	(0.004)
Observations	22,271	22,271	22,271	22,271	22,271	22,271	22,271	22,271	22,271
Number of panelid	9,761	9,761	9,761	9,761	9,761	9,761	9,761	9,761	9,761
R-squared	0.06	0.07	0.04	0.05	0.06	0.03	0.03	0.03	0.02

Standard errors clustered at household level in parentheses. **p<0.01, *p<0.05, +p<0.1
Regressions also include age*country and age squared/100*country interaction terms.

Table A.12: First stage regressions for grip strength

Dependent variable:	Caregiving (any frequency)			Weekly caregiving			Daily caregiving		
Parent in poor health	0.204***			0.175***			0.111***		
	(0.025)			(0.024)			(0.018)		
Parent in poor health x		0.225***	0.001*		0.196***	0.001		0.118***	0.000
family care country		(0.034)	(0.001)		(0.034)	(0.001)		(0.027)	0.000
Parent in poor health x		0.001	0.168***		0.001	0.142***		0.000	0.101***
formal care country		(0.001)	(0.036)		(0.001)	(0.029)		(0.001)	(0.020)
Chronic conditions	0.005	0.004	0.001	0.006**	0.005**	0.001	0.003	0.003	0.000
	(0.003)	(0.003)	(0.001)	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)
Limitations with ADL	0.005	0.000	0.005*	0.005	0.002	0.003	0.001	0.001	0.000
	(0.005)	(0.004)	(0.003)	(0.005)	(0.004)	(0.002)	(0.003)	(0.003)	(0.002)
Limitations with IADL	-0.009	-0.006	-0.002	-0.008	-0.008	0.000	-0.002	-0.002	0.000
	(0.008)	(0.008)	(0.002)	(0.008)	(0.008)	(0.002)	(0.003)	(0.003)	(0.001)
Age	-0.046	0.006	-0.055	0.000	0.011	-0.013	0.022	0.001	0.021
	(0.041)	(0.007)	(0.041)	(0.031)	(0.006)	(0.031)	(0.020)	(0.005)	(0.019)
Age squared/100	0.037	0.002	0.037	0.000	0.001	0.001	-0.015	0.001	-0.015
	(0.030)	(0.002)	(0.030)	(0.023)	(0.002)	(0.023)	(0.014)	(0.001)	(0.014)
Wave 2	-0.011	-0.009	-0.003	0.002	-0.012	0.013	-0.012	0.000	-0.013
	(0.021)	(0.012)	(0.018)	(0.018)	(0.011)	(0.015)	(0.016)	(0.008)	(0.014)
Wave 4	-0.056**	-0.042**	-0.016	-0.036	-0.048***	0.011	-0.018	-0.010	-0.009
	(0.028)	(0.020)	(0.020)	(0.025)	(0.018)	(0.016)	(0.021)	(0.015)	(0.015)
Married	0.002	0.008	-0.006	-0.001	0.008	-0.009	0.027*	0.012	0.014
	(0.022)	(0.016)	(0.015)	(0.020)	(0.015)	(0.014)	(0.015)	(0.012)	(0.009)
Eployed	0.005	-0.010	0.015**	0.010	0.001	0.010	-0.005	-0.001	-0.004
	(0.015)	(0.013)	(0.007)	(0.014)	(0.013)	(0.006)	(0.009)	(0.008)	(0.005)
Other activities	0.016*	0.013*	0.003	0.008	0.010	-0.002	0.002	0.004	-0.002
	(0.010)	(0.008)	(0.005)	(0.008)	(0.007)	(0.004)	(0.006)	(0.005)	(0.003)
Financial distress	0.008	0.000	0.008*	0.001	-0.007	0.008*	0.012*	0.007	0.005
	(0.009)	(0.008)	(0.005)	(0.008)	(0.007)	(0.004)	(0.007)	(0.006)	(0.004)
Observations	20,482	20,482	20,482	20,482	20,482	20,482	20,482	20,482	20,482
Number of panelid	9,017	9,017	9,017	9,017	9,017	9,017	9,017	9,017	9,017
R-squared	0.06	0.07	0.04	0.05	0.06	0.03	0.03	0.03	0.03

Standard errors clustered at household level in parentheses. $^{**}p<0.01$, $^{*}p<0.05$, $^{+}p<0.1$
Regressions also include age*country and age squared/100*country interaction terms.

Appendix B

Additional Figure for Chapter 4

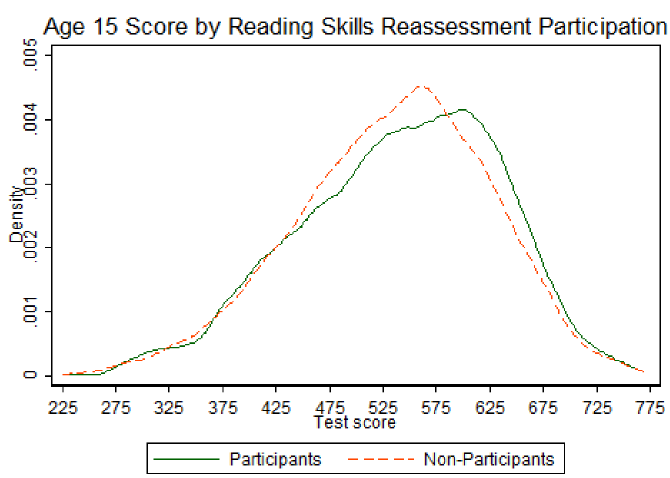


Figure B.1: The effect of attrition and non-response on the distribution of reading scores