UNIVERSITÀ
DEGLI STUDI
DI PADOVA

PhD in ECONOMICS AND MANAGEMENT
ECONOMICS CURRICULUM
Department of Economics and Management
University of Padova

Women's Labor Supply –
Competitive Preferences and Informal Caregiving

Head of the PhD Program: Prof. Antonio Nicolò
Supervisor: Prof. Guglielmo Weber

PhD Candidate: Lotte Kofoed Jørgensen

PhD Thesis

Women's Labor Supply – Competitive Preferences and Informal Caregiving

by

Lotte Kofoed Jørgensen

September 2019
Dipartimento di Scienze Economiche e Aziendali "Marco Fanno"
Università degli Studi di Padova
Via del Santo 33 - 35123 Padova
www.economia.unipd.it

Preface

This set of essays constitutes my PhD thesis on "Women's Labor Supply – Competitive Preferences and Informal Caregiving". The research for the thesis was carried out while I was a PhD student at the University of Padova. I would like to take this opportunity to thank my professors, colleagues, friends, and other staff members at the Department of Economics and Management for their support and encouragement throughout this process.

My collaboration with Prof. Marco Piovesan and Helene Willadsen at the University of Copenhagen has enabled me to co-create the first paper (Chapter 1), and I am thankful for the opportunity to experimentally explore with them the competitive preferences of children. I am greatly indebted to my supervisor, Prof. Guglielmo Weber, and Prof. Martina Celidoni for their valuable help, comments and feedback on numerous aspects of my second paper (Chapter 2). I thank the Fondazione Cassa di Risparmio di Padova e Rovigo for making my PhD studies at the University of Padova possible with their generous scholarship.

Lotte Kofoed Jørgensen

Padova, September 2019

Contents

Introduction

Gender differences in the labor market persist despite recent advances towards equality (Blau and Kahn, 2017). While women in many countries are pursuing higher education at similar rates to men, they continue to have lower earnings, and compared to men they are less likely to hold top positions in firms (OECD, 2012). Possible explanations for the gender gap include preferences for family, gender differences in abilities, and preferences over jobs (Niederle and Vesterlund, 2011). Over the last two decades economists have become increasingly interested in investigating whether gender differences in economic preferences can help explain the observed gender gap in labor market outcomes. Previous experimental studies find women to be less likely than men to self-select into competitive settings (Niederle and Vesterlund, 2007). Why do young men and women sort into different careers? If women on average are more reluctant to compete, they may be less likely to choose educations in male dominated, competitive fields and less willing to seek promotions in competitive environments.

A concern raised in the economic literature is that high performing women appear to opt out of competitive environments. In a societal perspective, the loss of workers with high abilities is costly. Understanding which factors that may drive these gender differences, is important when designing equality-promoting policies. Is the gender gap in competitive preferences fixed by nature, or do gender differences in competitiveness develop with age, as boys and girls grow up? Fairly little is known about the development of gender differences in competitiveness. Studying children of different age groups and in different environments may give us a better picture of gender differences in economic preferences. The aim of my first paper in Chapter 1 is to investigate whether boys and girls differ in their preferences for competition. The experimental design greatly follows that of Niederle and Vesterlund (2007), and this allows for comparisons of the results with those of previous studies.

Chapter 1

Gender Differences in Competitiveness: Evidence from Children in Denmark

We study gender differences in competitive behavior, running an experiment with Danish schoolchildren. In previous experimental studies the gender-stereotype associated with the task has been shown to influence results (Dreber et al., 2014). A new Lego task was therefore created to measure performance in a competitive environment, where participants build with Lego bricks on time and against each other. The experimental analysis further explores other factors that may influence or drive gender differences in competitiveness. We control for individual performance, risk preferences, confidence, stereotypes, and interactions with the opposite gender, and we find that Danish girls are 8.4 % less likely to choose a competitive scheme. Moreover, we elicit the network of friends within each class and find that a girl is 9.9 % more likely to enter a tournament, if one more of her friends competes. Finally, the gender gap increases to 12.7 % when we focus our analysis on children performing above the average. When we compute the optimal decision for entering the tournament, we discover that with age, girls avoid competition when from a profit maximizing aspect, they instead should go for it.

With an ageing population in Europe, there is a rising need for elderly care. Due to increasing costs of formal care provision, governments are expected to shift some of the responsibility for long-term care of the elderly onto family members, thereby intensifying the demand for informal caregiving. Informal care is the primary type of long-term care for older individuals, often provided by their adult children, and especially the daughters (Bettio and Verashchagina, 2010). Providing intensive care for a parent or parent-in-law can be time consuming, and caregivers may as a result decrease their working hours, drop out of the labor market, or retire early if close to pension age. Gender roles and family responsibilities

clearly affect women's labor market participation at different points in their working lives. Taking time off or retiring early to provide informal care imply a lost potential because of today's substantial investments in women's education and careers. This is an important issue as women, if they continue to increase their labor supply in their fifties and early sixties, have a real opportunity to decrease the gender wage gap and financial inequality in later years (Goldin and Katz, 2018).

On the other hand, the supply of informal care could be shrinking due to the increasing labor market participation of women, as a stronger labor force attachment effectively translates into a higher opportunity cost of informal caregiving. While recent studies mostly aim at identifying the effect of informal caregiving on the employment status and work hours of individuals of retirement-age, I instead explore the other direction; the effect of retirement on informal caregiving. The dominant reform strategy to increase pension eligibility ages and tighten the generosity of pension benefits may have negative consequences for the availability of informal care.

Chapter 2

Does retirement increase informal caregiving for parents and parents-in-law?

Using SHARE data, I examine the effect of retirement on daily caregiving to parents and parents-in-law. I focus on individuals aged 50-70 in 18 European countries. I account for the possible simultaneity bias of the retirement decision and the choice to provide informal care, by instrumenting the retirement status with having reached country-specific pension eligibility ages for early and statutory retirement. I find a significant and positive causal effect of retirement on daily caregiving. Moreover, I find evidence of heterogeneous effects across European regions. The effect of retirement on daily caregiving seems stronger for individuals in Scandinavia and The Netherlands. The policy ambition to extend the working lives of retirement-aged individuals and increase women's labor force participation, may diminish the supply of

informal care. The negative effects of later retirement should be taken into consideration, if government are to rely more on informal care provision to meet the increased demand for the long-term care of the elderly.

References

Bettio, F., and Verashchagina, A. (2010). Long-term care for the elderly: Provisions and providers in 33 European countries. *European Commission*.

Blau, Francine D., and Lawrence M. Kahn. "The gender wage gap: Extent, trends, and explanations." *Journal of Economic Literature* 55.3 (2017): 789-865.

Dreber, A., von Essen, E., and Ranehill, E. (2014). Gender and competition in adolescence: task matters. *Experimental Economics*, *17*(1), 154-172.

Goldin, C. and Katz, L. (2018). *Women Working Longer* (p. 11). University of Chicago Press.

OECD, *Gender Equality in Education, Employment and Entrepreneurship: Final Report to the MCM 2012*. OECD Publishing (NJ1), 2012.

Niederle, Muriel and Lise Vesterlund. (2007). Do Women Shy Away from Competition? Do Men Compete Too Much? *Quarterly Journal of Economics 122*(3), 1067–1101.

Niederle, Muriel, Carmit Segal, and Lise Vesterlund. (2013). How costly is diversity? Affirmative action in light of gender differences in competitiveness. *Management Science 59*(1), 1-16.

Chapter 1

Gender Differences in Competitiveness: Evidence from Children in Denmark

Gender Differences in Competitiveness: Evidence from Children in Denmark

Lotte Kofoed Jørgensen

University of Padua

Marco Piovesan[1]

CEBI, University of Copenhagen

Helene Willadsen

CEBI, University of Copenhagen

July 2019

This study has benefited from comments made by participants in various conferences: the 2nd LABEL - IEPR Conference in Los Angeles; the 2017 Spring School in Behavioral Economics, in San Diego; the 2016 Economic Science Association conference in Bergen; the 2016 IMEBESS conference in Rome. We thank the Carlsberg Foundation for their generous financial support (Project: 108579). We also acknowledge the assistance received from research assistants Janek Bligaard Eskildsen and Benjamin Mathiasen. Last, we thank the participating schools for their cooperation and help with making this research possible.

[1] Corresponding Author: Marco Piovesan, Center for Economic Behavior and Inequality, Department of Economics, University of Copenhagen. Address: Øster Farimagsgade 5, DK-1353 Copenhagen. Phone: +45 353 36872. Email: marco.piovesan@econ.ku.dk.

Abstract

We study gender differences in competitive behavior, running an experiment with Danish schoolchildren. We control for individual performance, risk preferences, confidence, stereotypes, and interactions with the opposite gender and we find Danish girls are 8.4 % less likely to choose a competitive scheme. Moreover, we elicit the network of friends within each class and we find that girls are 9.9 % more likely to enter a tournament if one more of her friends competes. Finally, the gender gap increases to 12.7 % when we focus our analysis on children performing above the average. When we compute the payoff maximizing decision, we discover that with age, girls avoid the competition when they should choose it.

Keywords: competition; gender; stereotypes; children; experiment.

JEL code: C70, C91, J16, J24, J31, M52

1. Introduction

Gender differences in career choices and labor market outcomes are often associated with the different attitudes that men and women have towards competition. The pioneering work of Muriel Niederle and Lise Vesterlund suggests that men are twice as more likely to self-select into competition (Niederle and Vesterlund, 2007). Dozens of subsequent studies have confirmed the general tendency of women to shy away from competition, but the estimated size of this gap varies enormously from one study to another. Individual risk preferences, confidence, and gender stereotypes all seem to be relevant factors that can help explain the heterogeneity of the observed gender gap (Dreber et al., 2011). To study the emergence of this gender gap, a growing body of literature has focused on different age groups within various cultures and social environments (Andersen et al., 2013, Cárdenas et al. 2012). Most of the experiments conducted with adults found a significant gender gap in competitiveness. However, the evidence with children is more mixed: some papers report a gender gap in competitiveness to be present already in early years (Sutter and Glätzle-Rützler, 2015), while others find no significant gender differences (Khachatryan et. al. 2015).

In this paper we report evidence from our experiment with 341 Danish children aged 7 to 16 using the Niederle and Vesterlund paradigm. Since previous studies have shown that the heterogeneity of the estimated gender gap can depend on the task used (Flory et al., 2010; Dreber et al. 2014), we propose a novel real-effort task designed specifically for children. Moreover, we elicit children's beliefs about their (relative) performance, and we ask them to guess the performance of two "real" children; a boy and a girl. Finally, we measure children's risk preferences, and we elicit the network of friends in each class. To the best of our knowledge, our study is one of the first that control for all these different measures at the same time.

Our results show that Danish girls are 8.4 % less likely to choose a competitive scheme, when we control for individual ability, confidence, risk aversion, stereotypical beliefs, and interaction with the op-

posite gender. In addition, we investigate if friends influence the decision to compete; we find that a girl is 9.9 % more likely to enter the tournament if one more friend of hers competes. We also investigate the difference in competitiveness for boys and girls performing above average, and we find that in this case a girl's probability of selecting into competition is 12.7 % lower. Finally, we compute the payoff maximizing decisions whether to enter the competition (or not), and we compare it with the observed decisions in our experiment. Our results show that with age, boys become better in making the payoff maximizing decision. while girls avoid the competition when they should choose it.

We decided to study Danish children, because like other Scandinavian countries, Denmark has a long history of implementing gender-equalizing policies, and gender equality is generally a hot topic in the public debate. Danish children spend a large part of their childhood and adolescence in childcare and educational institutions: from 6-12 months, when children enter nursery, to 16 years old, when they leave compulsory school. During these school years, norms and styles of social interactions are established. The Danish Ministry of Education aim to remove gender stereotypical expectations to allow boys and girls to have equal opportunities and unfold their dreams and potential (Ministry of Education, 2017). This ambition is pursued and supported by teachers in removing gender stereotypical expectations. Moreover, the Gender Equality Act (Ligestillingsloven) promotes gender equality, including equal integration, equal influence and equal opportunities in all functions of society based on women and men's equal status. In addition, the Act aims to remove direct and indirect discrimination on the grounds of gender.

Even so, women and men continue to choose education and occupations according to traditional gender stereotypes[2] with consequences for equal pay and career prospects. Gender segregation in the Danish labor market is widespread: women work primarily in the public sector, while men work in the private sector; women are overrepresented in part-time work and take all or most of the parental leave.

[2] Statistics Denmark, http://www.statistikbanken.dk/u16 and /u14

This hampers flexibility and causes congestions in some trades (COWI, 2014); for example, not enough young women pursue education in STEM fields (Science, Technology, Engineering and Math)[3]. Stereotypes that men naturally are more interested and perform better in math and science are thought to influence the STEM aspirations and achievements of boys and girls, men and women (Nosek et al., 2009). Countries with high levels of gender equality have some of the largest STEM gaps in higher education (Stoet et al., 2018). The setting of Denmark is especially interesting to study because of this contradiction.

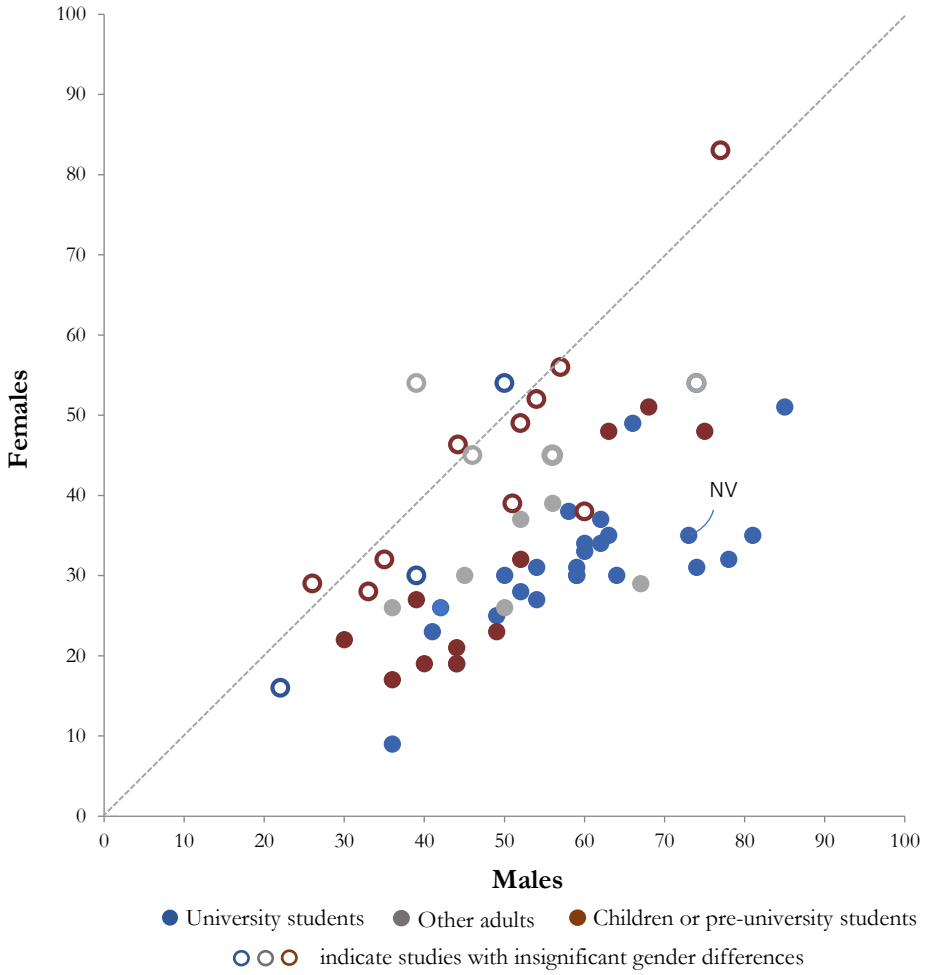
2. Related literature

Two pioneering studies have experimentally investigated competitiveness as one of the many explanations behind the gender gap. Gneezy et al. (2003) find that as an environment becomes more competitive, men increase their performance, while women generally do not. Women under-perform, when competing against men, but increase their performance opposite other women. Performing in mixed gendered group may heighten participants' awareness of their own gender and induce them to behave in accordance with stereotypes about men being more competitive and performing better in tournaments (Steele, 1997). Niederle and Vesterlund (2007) show that gender differences in competitiveness persist even when controlling for gender differences in risk aversion, confidence, and aversion towards feedback about one's relative performance. Various experimental studies have replicated these findings, and researchers have more recently focused their attention to children. For instance, a study by Sutter and Glätzle-Rützler (2015) show that gender differences in competitiveness emerge early in life. They examine the competitive choices of 1,570 Austrian children from early childhood to late teenage years and find that boys at the age of three are already significantly more likely than girls to choose competition.

[3] Final admissions 2015 - Engineering courses. Ministry of Higher Education and Science, 2015.

Figure 1 compares the willingness to compete for males and females and boys and girls in 47 experimental studies that use the Niederle-Vesterlund (2007) paradigm[4]: the majority of these studies report that males are more competitive than females, but there is a substantial variation in the results[5].

Figure 1. Willingness to compete in studies using the Niederle-Vesterlund paradigm (in percentage)



Note: The figure shows the ratio of males (x-axis) and females (y-axis) that chose to compete in previous studies on competitiveness. We have included experiments that were run with University students (blue), Other adults (grey) and Children or

[4] The paradigm consists of stages that differ in how performance translates into earnings. In the first stage, participants are paid for each correct answer using a piece rate. In the second stage, each participant competes in a tournament against one (or more) randomly selected anonymous participants. Winner takes all: the participant with the highest performance receives the total payoff; the other(s) receive zero. As the main measure of competitiveness, often in the third stage, the participants must choose between the piece rate and the tournament.

[5] 70 % of all studies find significant gender differences, while for studies with children it is 52 %. The full list of studies used in Figure 1 is included in Table A.1. in Appendix A.

pre-university students (red). The filled dots indicate significant gender differences, while unfilled dots indicate insignificant gender differences. For the full list of included studies, we refer to Table A.1. in Appendix A. If men and women are equally competitive then studies will report results along the 45-degree line. If women are less competitive then studies will report results below the 45-degree line. The dot marked "NV" highlights the study of Niederle and Vesterlund (2007) that reported a significant gender gap. The figure shows that across studies it is a robust finding that adult females are less competitive than adult males. For children, results are mixed.

If men are more (over)confident about their relative performance, one would expect them to enter competitive environments at a higher rate than women of equal performance. Kamas and Preston (2012) examine whether gender differences in willingness to compete are due to gap in overconfidence. They let students solve both an addition task and a word search task: when asked about their beliefs, participants consider men better at the math task, and women better at the word task. Yet, the authors find that women perform significantly better in the math task and marginally better in the word task, while men believe that they outperform women in both tasks. Interestingly, controlling for overconfidence eliminates the gender gap in the willingness to compete in the math task, where there is a significant gender gap in competitiveness. Dreber et al. (2014) ask Swedish high school students to perform both a math task and a word search task. They find girls and boys to be equally likely to self-select into the tournament in the verbal task, while boys are much more willing to enter the competitive setting in the math task. Controlling for actual performances and relative performance beliefs diminishes the gender gap in competitiveness, since girls are more under-confident about their performance in the math task than boys are. Wozniak et al. (2014) study how confidence affects tournament-entry and they find that when subjects are provided with information about the performances of others, the gender gap in competitive preferences disappears, which seems to suggest that lack of information creates confidence errors that may explain gender differences in economic outcomes.

Gender differences in competitiveness may also depend on the type of the task used. Samek (2013) investigate gender differences in competitive preferences among preschool children, who are matched randomly into pairs and take part in a gender-neutral fishing task, performing under piece rate

and tournament incentives, before choosing their preferred incentive scheme in stage 3. Contrary to the results reported in Sutter and Glätzle-Rützler (2015), the young boys and girls are equally willing to compete. Yet, the performances of girls under competitive incentives are significantly lower than that of boys. The children are very confident; 84 % believe they are the winner and being paired with either a boy or a girl does not seem to matter. With children in Armenia aged 8 to 16, Khachatryan et al. (2015) study the willingness to compete both in an addition task and a word search task and find no gender differences at any age.

In the labor market, gender stereotypes associated with a job or task could explain the under-representation of women in competitive environments (Akerlof and Kranton, 2000). If women perceive competitiveness as something more of a male trait, they may feel subject to a conflict in their gender identity and feel pressured to act in a more feminine way to meet the expectations of their peers (Steele, 1997). Implicit stereotypes need not be explicitly endorsed to influence individual behavior. Andersen et al. (2013) compare the competitive choices of children in matrilineal and patriarchal societies in India using a bucket toss task. Before performing the task, participants choose between piece rate pay and tournament pay. While the authors find no gender gap in competitiveness at any age in the matrilineal society, in the patriarchal society, boys become more competitive and girls less competitive around the age of puberty. Cárdenas et al. (2012) study the competitive performance and the willingness to compete of 1,200 children in Sweden and Colombia who perform in running, skipping rope, in a math task, and a word search task. Their results show no gender differences in tournament performance across the four tasks, and in Colombia, girls and boys are equally competitive in all tasks and measures, while, surprisingly, boys are more willing to compete than girls in Sweden. Almås et al. (2016) examines the role of family background combining results from a lab experiment with register data, and find that Norwegian adolescents of parents with a low socioeconomic background tend to be less willing to compete than other children. Results also show significant gender differences in competitiveness with the largest differences

among children with well-educated parents. They find no gender differences in self-confidence, risk preferences, or personality traits.

3. The experiment

One week before the experiment parents signed a consent form and filled in a questionnaire with some demographic variables (age, gender, siblings and their gender). For every questionnaire returned, we donated 25 DKK to the class account for social activities. We contacted 413 parents, and 353 parents gave us their consent (85.5 %). However, our final sample comprises 341 children aged 7-16 (grades 1 to 9) because 10 questionnaires were not properly filled out and two children chose not to participate. On the day of the experiment, we placed cardboard dividers on the tables and then randomly allocated children to seats. Once seated we informed children that during the experiment, they had the opportunity to win *points* that later would be converted into gifts such as small toys, games and schools materials, desirable for all participants across gender and age groups. We explained the children that more points corresponded to more gifts, but we did not reveal the exact exchange rate (1 point = 0.40 DKK). [6]

In our experiment, children receive a box with 120 Lego bricks. They have 3 minutes to build as many rows as they can following specific instructions in terms of colors and sizes of the bricks (we use bricks in 2 different sizes and in 3 different colors). Following Niederle and Vesterlund (2007), children have a trial run and they then solve the task in three different stages. In the first stage, children receive 1 point per correct row (*piece-rate* scheme). In the second stage, they compete against an unknown child in their class; this means that if they build more rows than their opponent does, they receive 2 points per

[6] Information about the experimental procedures was revealed sequentially, immediately before the students were performing each task. The instructions were given verbally as suggested by the school leaders, because of the level of literary among younger students and the time constraint of the sessions. The instructions were interactive and involved several examples and control questions to make sure that all participants understood the details of the different payment schemes.

correct row, and 0 points otherwise (*competitive* scheme). In the third stage, they can choose to perform the task under *piece-rate* or *competitive* scheme. We do not inform children about their performance or payoff until the end of the experiment.

We carefully explained the children that under the *competitive* scheme they are matched randomly with an anonymous child in class: participant A competes against participant B, participant B against participant C, and so on. We randomly matched each child with the same opponent in stage 2 and 3, regardless of the other child's decision to compete. We follow the experimental design of Dreber et al. (2011), and the circular matching allows for a child's decision to compete not to affect the payment of another student.

After the Lego task, we asked children to assess their (relative) performance. Children must indicate both the total number of children participating in that session and then their own rank within the class. For instance, in a class with 17 children participating, children must write "17" and then indicate what they believe to be their own ranking position, i.e. a number between 1 (best performer) and 17 (worst performer). We asked children to rank their performance for stage 2. In a different decision sheet, we asked them to guess the performance of two "real" children, a boy and a girl, depicted while building a Lego wall.[7] We took these pictures during the pre-test of our experiment and both these children built 11 correct rows within 3 minutes. This task allows us to elicit children's stereotypes, i.e. their belief about the performances of boys and girls.

Moreover, we elicited children's risk preferences using a modified (and simpler) version of the Bomb Risk Elicitation Task (hereafter BRET) developed by Crosetto and Filippin (2013). Children must decide how many "steps" of a winding road (with 100 numbered steps) they want to walk across, knowing that one step contains a "bomb". For each step they select, they receive 1 point; however, if one of the

[7] The two children were both in the 5th grade.

steps they selected contains the bomb, they receive 0 points. After all children made their decision, we randomly determined the location of the bomb in front of the class. Finally, we elicited the network of friends in the class using a network elicitation task developed by Chen et al. (2016). We distributed a drawing of a table with six chairs. We asked children to write their own name above the chair at the head of the table and indicate, in order of preference, maximum five names of classmates (also participating in the experiment) that they would like to have seated at their table. We emphasized that nobody but the experimenters would see their answers.

At the end of the session, we called children individually to receive their *points* and to choose their gifts. During the experiment, we naturally assisted the children who needed help with the instructions. Appendix B and C include more detailed information about our experimental procedures, the script of our oral instructions, our decision sheets, and some pictures taken during the experiment.

For our power calculations, we take the unweighted average across all previous studies using children or pre-university subjects[8], resulting in 49 % of the boys and 37 % of the girls choosing to compete (there obviously is a lot of variability between the studies). Our sample size (166 boys and 175 girls) will allow us to detect a difference in the competitive behavior of boys and girls 99.99 % of the time, when comparing and testing their average willingness to compete at the 5 % level (McCrum-Gardner, 2010). If 37 % of the girls choose to compete (based on previous results in the literature), we will be able to detect a difference 80 % of the time at a 5 % level as long as boys are 4.9 %-points more likely to compete.[9] It should be noted that in our regression results in Table 2, we also split the sample by gender and average performance in class, and naturally the power of the estimates obtained in these subsamples is expected to be much smaller.

[8] See full list of studies in Table A.1. in Appendix A.
[9] The power calculation was performed using the software GPower.

4. Results

In our analysis, we divide our sample into three age groups: Grades 1-3 (aged 7-10), Grades 4-6 (aged 9-13), and Grades 7-9 (aged 12-16) according to the administrative division of grades in the Danish school system.[10] To test our hypotheses, we use the Mann-Whitney U test (hereafter MWW) and the Wilcoxon matched pairs Signed-Rank test (hereafter S-R).

Table 1 gives an overview of the average performances in the Lego task under the piece-rate (stage 1) and competitive payment scheme (stage 2), across age groups and gender.

Table 1. Performance in the Lego task in stage 1 and 2 and decision to compete in stage 3

	Gender	N	Piece-rate - Stage 1 *(average #rows)*	Competitive - Stage 2 *(average #rows)*	Performance increase *(average #rows)*	Choosing to Compete – Stage 3 *(mean percentage)*
Grades 1-3	Boys	49	7.76	8.18	0.43	0.53
	Girls	61	8.62	9.64	1.02***	0.51
	Total	110	8.24	8.99	0.75***	0.52
Grades 4-6	Boys	68	9.97	11.44	1.47***	0.71
	Girls	54	10.19	11.83	1.65***	0.56
	Total	122	10.07	11.61	1.55***	0.64
Grades 7-9	Boys	49	10.78	12.73	1.96***	0.59
	Girls	60	12.32	13.87	1.55***	0.62
	Total	109	11.62	13.36	1.73***	0.61
All	Boys	166	9.55	10.86	1.31***	0.62
	Girls	175	10.37	11.77	1.39***	0.56
	Total	341	9.97	11.33	1.35***	0.59

Note: Stage 1 and 2 average performance, the performance increase (*** $p<0.001$), and the decision to compete in Stage 3 for boys and girls, and together in the age groups and for all.

In stage 1, not surprisingly, we find that older children perform significantly better than younger ones (MWW: Grade 1-3 vs. Grade 4-6, $p<0.001$; Grades 4-6 vs. Grades 7-9, $p<0.001$). Moreover, on average girls perform 8 % better than boys (MWW, $p=0.009$), but there is some variation with age (MWW:

[10] Note that dividing the sample in groups by age instead of grade does not significantly change our results.

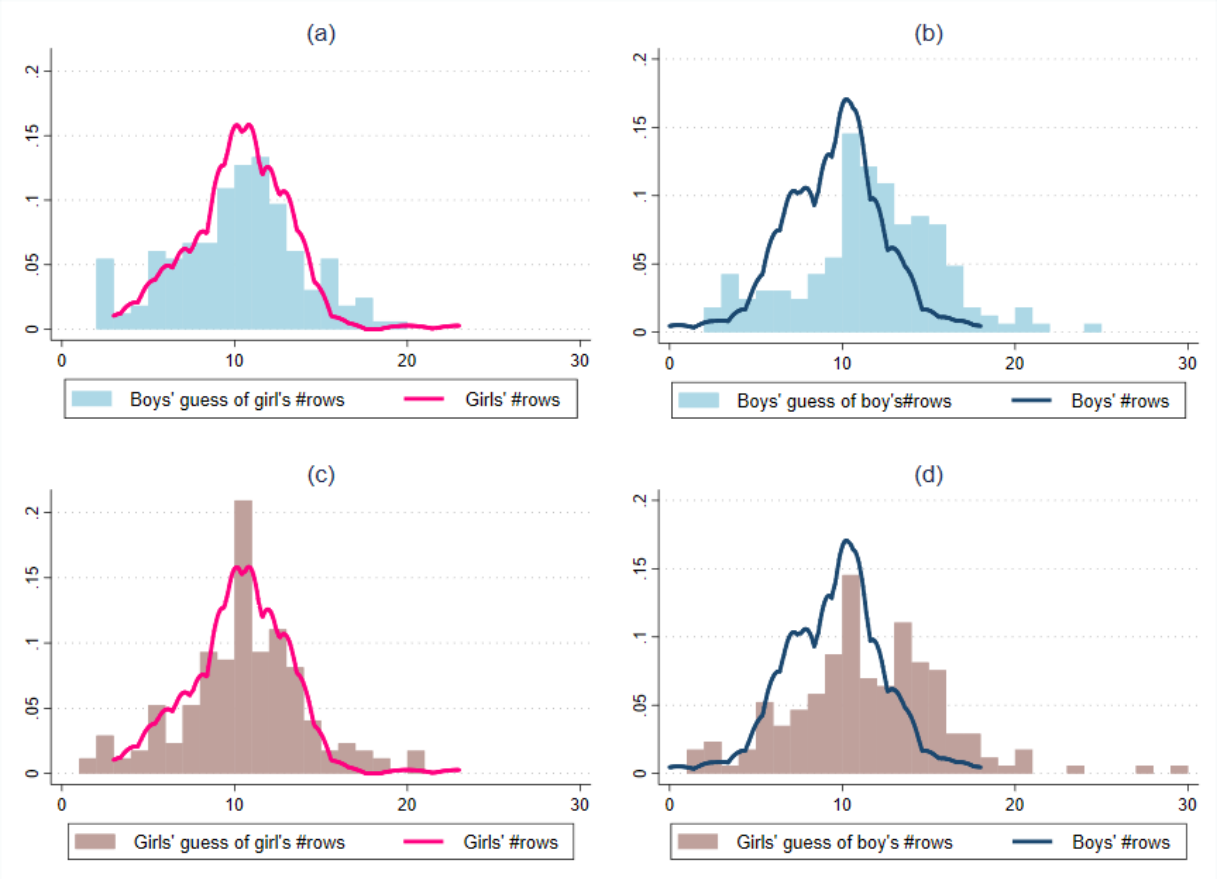
Grade1-3, +11 %, p=0.049; Grades 4-6, +2 %, p=0.543; Grades 7-9, +14 %; p=0.006). We observe a similar pattern for performance in stage 2 (MWW: All: +8 %, p=0.012; Grades 1-3, +17 %, p=0.003; Grades 4-6, +3 %, p= 0.501; Grades 7-9, +8 %; p=0.043). Overall, both girls and boys increase their performance under the competitive scheme (S-R: p<0.001). Individually, all groups significantly raise their average performance between the two stages (S-R: p<0.001) except for the youngest boys (S-R: p=0.143). Girls improve more in the younger grades, while boys improve slightly more in the older grades. Yet, there is no significant difference between the performance increase of boys and girls (MWW: p=0.374). In stage 3, boys are more likely to choose to compete than girls are (62 % vs. 56 %), but this difference is not significant (MWW: p=0.257). The difference is also not significant for younger compared to older children (MWW: Grades 1-3, p=0.816; Grades 7-9, p=0.793), and it is only marginally significant for the middle age group (MWW: Grades 4-6, p=0.087).

In our regression analysis, we need to control for the fact that girls are more "productive" than boys; we use their performance in stage 2. Moreover, three other factors may affect the willingness to compete: risk preferences, confidence, and stereotypical beliefs. In the Bomb Risk Elicitation Task boys take significantly more steps than girls (37.1 vs. 33.6, MWW: p=0.018), and comparing younger with older children, risk taking significantly increases with age for boys (MWW: p=0.026) and marginally significantly for girls (MWW: p=0.054). To control for children self-confidence, we used the children's guess about their relative performance ranks. We characterize a child as confident (confident=1) if the child believes that his or her performance in stage two lies in the top half of the distribution.[11]

[11] In Figure A.1. in the Appendix A we use another definition of self-confidence and term a child as being over (under-) confident if the child believes that his or her performance lies in the top (bottom) half of the performance distribution in class when it does not. The figure shows the distribution of under- and overconfidence both by gender and age group.

Finally, we study children's stereotypical beliefs, i.e. their guesses about the performance of a depicted boy and girl. Figure 2 shows boys' (top panels; (a) and (b)) and girls' (at the bottom; (c) and (d)) guesses in combination with the smoothed distribution of the actual performances of the boys and girls.

Figure 2. Beliefs about the performance vs actual performance



Note: The figure shows the number of Lego rows that boys (166) and girls (175), respectively, expect a boy and a girl to have built, compared with the distribution of the number of rows the boys and girls actual built.

Both boys' and girls' guesses of the girl's performance (on the left; (a) and (c)) are precise. Whereas, their guesses of the boy's performance (on the right; (b) and (d)) are more right-skewed, meaning that guesses of the boy's performance are much higher than the actual performance of boys. The percentage of girls who believe that the boy is better increases with age (Grade 1-3, 54 % vs. Grade 7-9, 72 %, MWW: p=0. 047) while the opposite is true for boys (Grade 1-3, 78 % vs. Grade 7-9, 57 %, MWW: p=0.032).

This result shows that for girls the stereotype of a high male performance in the task gets more pronounced with age and suggests that for girls the stereotypical belief is developing with age[12].

In Table 2, we present the results from a probit regression on the decision to enter the tournament in stage 3. To ease the interpretation of the regression results, we report the marginal effects estimated at the mean. Our regression analysis follows that of Niederle and Vesterlund (2007), and similar to previous studies, we gradually include more determinants of selecting into competition. In model (1)-(3), the "Female" variable is always negative and (marginally) significant (p=0.062; p=0.075; p=0.091). This means that girls are less willing to enter in the tournament compared to boys. Moreover, in Model (1) children with a high performance (number of correctly built Lego rows) in stage 2 are 4.8 % more likely to choose competitive incentives in stage 3 (p <0.001). Risk aversion negatively affects the probability to enter the tournament (p =0.083).[13] Risk averse individuals are 12.4 % less likely to compete.

Compared to Model (1), Model (2) adds variables for confidence about one's absolute performance and beliefs about stereotypes (i.e. boy is best); both are insignificant. Model (3) adds number of friends of opposite gender, number of siblings, and number of friends who self-select into competition.[14] When we calculate the marginal effects estimated at the mean, the result tells us that for two otherwise average children, a girl would be 8.4 % less likely than a boy to select the competitive incentive scheme. Both the number of siblings (with 5.4 %) and number of competing friends (with 5.3 %) increase the probability of entering the tournament in stage 3.

[12] We find no significant difference in the stereotypical beliefs for girls above and below average performance, a division of the observations into subgroups that we will use in Table 2's model (6) and (7).
[13] We define risk aversion as dummy variable (=1) for choosing less than 50 steps in the BRET.
[14] For both the variables, number of friends of opposite gender and number of friends who compete, the variables take a value from 0 to 5, based on characteristics of the five friends reported in the network elicitation task.

Table 2. Marginal Effects on the Probability of the Choice of Compensation Scheme

	Dependent variable: Choice of competitive payment scheme						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female (=1)	-0.105*	-0.099*	-0.084*	-	-	-0.127**	-0.039
	(0.056)	(0.056)	(0.050)	-	-	(0.062)	(0.089)
Performance in stage 2	0.048***	0.041**	0.041**	0.044*	0.034	0.043	-0.001
	(0.013)	(0.018)	(0.018)	(0.024)	(0.027)	(0.036)	(0.032)
Risk averse (=1)	-0.124*	-0.114	-0.095	-0.149	0.011	-0.076	-0.075
	(0.071)	(0.074)	(0.073)	(0.102)	(0.092)	(0.101)	(0.081)
Grade	0.006	0.006	0.009	0.007	0.016	-0.029	0.050
	(0.050)	(0.051)	(0.044)	(0.065)	(0.056)	(0.039)	(0.059)
Average performance in class	-0.041	-0.031	-0.039	-0.038	-0.036	-0.002	-0.039
	(0.067)	(0.070)	(0.060)	(0.102)	(0.060)	(0.049)	(0.086)
Confident in stage 2 (=1)		0.055	0.064	0.072	0.061	-0.097	0.172**
		(0.068)	(0.070)	(0.109)	(0.078)	(0.112)	(0.087)
Boy is best (=1)		-0.003	-0.005	0.011	-0.025	-0.006	-0.004
		(0.012)	(0.012)	(0.013)	(0.015)	(0.015)	(0.014)
Number of opposite gender friends			0.019	0.021	0.015	0.017	0.027
			(0.024)	(0.042)	(0.041)	(0.051)	(0.033)
Number of siblings			0.054***	0.069**	0.046	0.031	0.056**
			(0.020)	(0.032)	(0.038)	(0.037)	(0.027)
Number of competing friends			0.053**	0.002	0.099***	0.061***	0.037
			(0.020)	(0.051)	(0.033)	(0.022)	(0.041)
Sample	All	All	All	Males	Females	Above average	Below average
Observations	341	339	338	164	174	168	170
R-squared	0.047	0.046	0.07	0.079	0.088	0.07	0.053

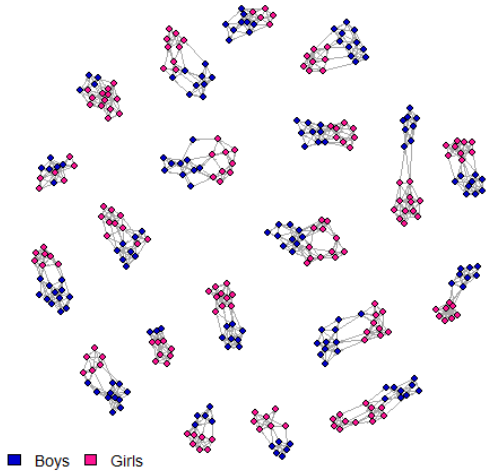
Note: The dependent variable of the probit regression is equal to 1 if the participant chose tournament incentives in Stage 3. Robust standard errors clustered by class ID are in parentheses. The participants are divided into two groups, males and females, with the corresponding regression results in column (4) for males and in column (5) for females. In column (6) and (7) the participants are split by their *actual* average performance in stage 2; we run a regression for the children performing above average in column (6) and below in column (7). 'Grade' and 'Number of siblings' are continuous variables, while 'Risk averse', 'Confident in stage 2', and 'Boy is best' are dummies. The number of girls above average is 98 (out of 168) and the number of girls below average is 76 (out of 170). *p<0.1, **p<0.5, ***p<0.01

In model (4) and (5), we run our model separately for boys and girls. For boys, a higher number of siblings increases the willingness to compete with 6.9 % (p= 0.032). There is a positive correlation between tournament entry and the number of siblings for boys (r=0.11, p <0.001). For girls, all things being equal, if one more of her friend competes, the girl is 9.9 % more likely to enter the tournament (p=0.003). There is a positive correlation between tournament entry and the number of friends who

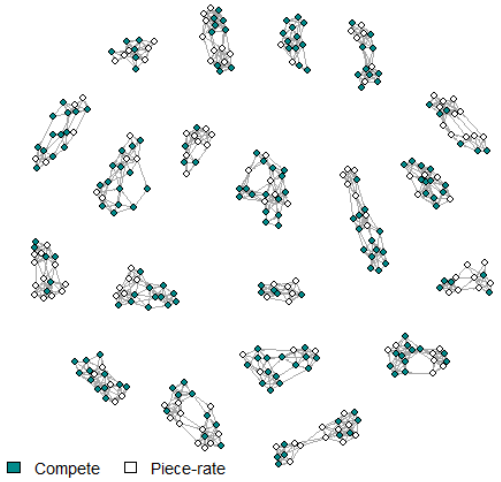
compete for girls (r=0.23, p <0.001). To understand if this pattern reflects that children cluster together based on a certain trait, we analyze the friendship network in class.[15] Panel (a) in Figure 3 shows that children cluster based on gender, i.e. boys (girls) are mainly friends with boys (girls). On the contrary, panel (b) and panel (c) show that children neither cluster based on "competitiveness" nor "performance".[16]

Figure 3. Children's connections within each class

a) Gender



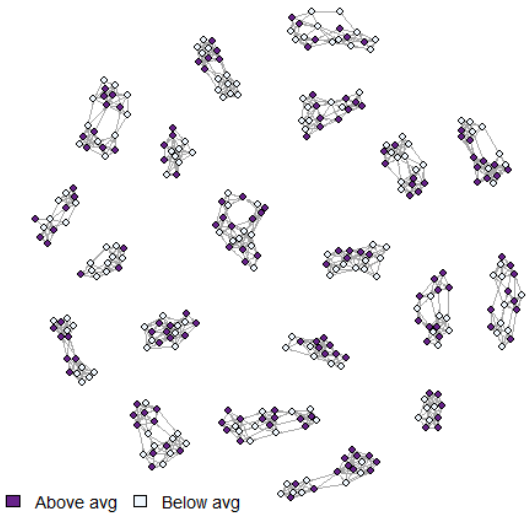
b) Competitiveness



[15] Other individual measures of the network structure, such as the centrality of the child in the network or the number of reciprocated links with friends of the opposite gender, are insignificant and only adds white noise to the estimation results in the regression analysis.

[16] The networks are visualized with the software R.

c) Performance



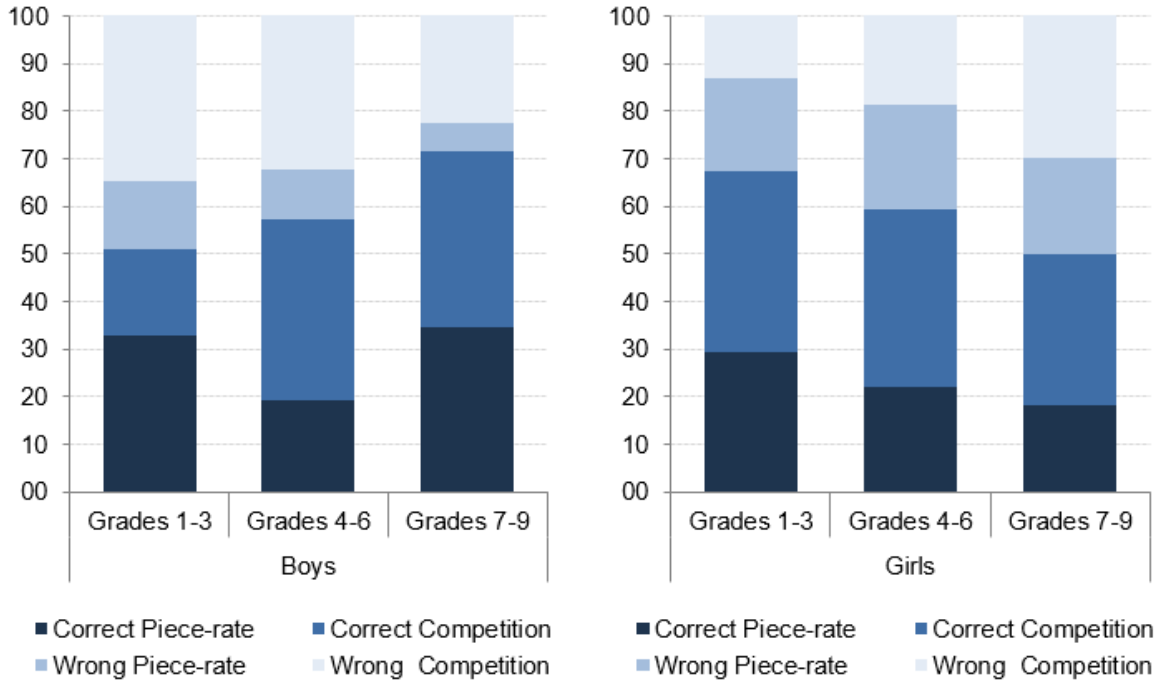
Note: The figures above show ties between children as reported in the friend elicitation task. Each class is a separate cluster (20 classes) because children can only indicate peers from their class as friends. In the three panels, we show children's network of friendship according to gender (panel a), their choice to compete (panel b) and whether they perform above or below average (panel c).

In model (6) and (7), we split the sample by the children's average performance in stage 2. In model (6), i.e. with the above average performers, being female has a significantly negative effect on choosing to compete (p= 0.040). On average, a girl's probability of selecting into competition is 12.7 % lower. The number of friends who compete has a significantly positive effect on the willingness to compete (p= 0.006). Having one more friend who selects competition, raises the probability of entering the competition with 6.1 %. All things being equal, above average performers should select into competition to maximize their payoff, but girls are less likely to do so. As we are already looking at the top performers, performance in stage 2 has no significant effect on the choice to compete. Risk aversion also plays a smaller role. The risk of no payoff is lower for top performers, when entering the tournament, since top performers perform relatively better in class and are surer to win. In model (7), i.e. with below average performers, the female variable is insignificant, while being confident and the number of siblings matters. For two otherwise average children, who perform below average, being confident increases the willingness to compete with 17.2 % (p= 0.047). The number of siblings increases the probability to compete by

5.6 % (p= 0.034).[17] In Table A.2. in Appendix A we replicate the estimations reported in Table 2 by including a dummy that specifies the age group rather than using grade as a continuous variable, and we get similar results with estimates of almost identical size.

Finally, we investigate whether children make the payoff maximizing decision in stage 3. Since we randomly matched the children, a child should choose to compete if his or her performance is above the average performance in the class. Figure 4 shows that older boys make significantly more correct (i.e. payoff maximizing) decisions than younger boys (51 % vs. 71 %, MWW: p=0.039), while older girls make fewer correct decisions than younger girls (67 % vs. 50 %, MWW: p=0.056).[18]

Figure 4. Correct (payoff maximizing) and wrong decisions in stage 3 (in percentage)



Note: A child should enter competition if his or her performance is above the average performance in class. The figure shows the fraction of children who correctly (the payoff maximizing decision) or wrongly chose piece-rate or competitive incentives by gender and age group.

[17] We get similar results when dividing the sample by *above* or *below median performance* in class.
[18] We compare the mean between the youngest and oldest age group.

Among younger children, girls make better decisions than boys (MWW: p=0.086), because younger boys choose to compete when they should not (MWW: p= 0.0075). For the older children, boys are significantly better than girls in making the profit maximizing decision (MWW: p=0.024), and many girls avoid the competition when they should choose it (MWW: p=0.037).

5. Conclusion

We conducted an experiment using an intuitive and novel real-effort task designed specifically for children. We find that Danish boys are 8.4 % more willing to compete than Danish girls. This gap increases to 12.7 % when we only focus on children performing above the average. In our regression analysis, we control for individual ability (girls in our task are more productive than boys), confidence (boys, especially if young, are more overconfident than girls), risk aversion (girls are more risk averse than boys), stereotypical beliefs (both boys and girls think boys are more productive), interaction with the opposite gender (having a brother or a sister, or having friends of the opposite gender) and with competitive peers (number of friends competing). Moreover, we find that girls are 9.9 % more likely to enter the tournament if one more of her friends competes, and the gender gap in competitiveness increase to 12.7 % when we focus our analysis on children performing above the average.

Interestingly, when we compute the payoff maximizing decision of whether to enter the competetion, we see that boys with age significantly improve their number of payoff maximizing decisions by reducing their "competitiveness". Girls also reduce their "competitiveness" with age, but in their case, it implies that they make more wrong decisions as they become older. The extent to which gender differences manifest themselves seem to vary with a wider set of social factors (Stoet et al., 2018). Our findings suggest that boys and girls from an early age receive different "social signals" about how to behave from

their social environment; i.e. their school, family, and friends. These behavioral social cues may negatively affect girls' confidence and contribute to enlarge the gender gap observed in studies with adults.

Our results demonstrate that even in a society with active gender policies and where women have a high labor market participation, we still find significant gender differences in competitive preferences. This raises a question about what kind of gender policies that manages to equalize competitive behavior. As our results suggest, high-performing girls may especially benefit to STEM-related interventions, if they can increase the girls' confidence about own abilities and overcome implicit stereotypical beliefs (Nosek et al., 2009), and at the same time deal with the cultural reality that fewer women pursue scientific careers. Identifying policies that can help equalize the competitiveness of young girls would be a significant advance in the literature on gender differences in competitive behavior and provide a unique insight into the formation of these gender differences. Finally, at this stage it is probably necessary to develop an ambitious research project that, using the same paradigm and procedures, can elicit men and women's willingness to compete and all the related measures with participants of different ages, cultures and social backgrounds. Only a comprehensive and standardized experiment can provide definitive evidence on a gender gap in competitiveness.

References

Akerlof, G., and Kranton, R. (2000). Economics and identity. *Quarterly journal of Economics*, 715–753.

Almås, I., Cappelen, A. W., Salvanes, K. G., Sørensen, E. Ø., and Tungodden, B. (2015). Willingness to compete: Family matters. Management Science, *62*(8), 2149-2162.

Andersen, S., Ertac, S., Gneezy, U., List, J. A., and Maximiano, S. (2013). Gender, competitiveness, and socialization at a young age: Evidence from a matrilineal and a patriarchal society. *Review of Economics and Statistics*, *95*(4), 1438-1443.

Apicella, C. L., and Dreber, A. (2015). Sex differences in competitiveness: Hunter-gatherer women and girls compete less in gender-neutral and male-centric tasks. *Adaptive Human Behavior and Physiology*, *1*(3), 247-269.

Apicella, C. L., Demiral, E. E., and Mollerstrom, J. (2017). No gender difference in willingness to compete when competing against self. *American Economic Review*, *107*(5), 136-40.

Ashton, M. C., and Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and social psychology review*, *11*(2), 150-166.

Balafoutas, L., Kerschbamer, R., and Sutter, M. (2012). Distributional preferences and competitive behavior. *Journal of economic behavior and organization*, *83*(1), 125-135.

Banerjee, R., Gupta, N. D., and Villeval, M. C. (2018). The spillover effects of affirmative action on competitiveness and unethical behavior. *European Economic Review*, *101*, 567-604.

Barber, B. M., and Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The quarterly journal of economics*, *116*(1), 261-292.

Berlin, N., and Dargnies, M. P. (2016). Gender differences in reactions to feedback and willingness to compete. *Journal of Economic Behavior and Organization*, *130*, 320-336.

Bönte, W., Procher, V., and Urbig, D. (2018). Gender differences in selection into self-competition. *Applied Economics Letters*, *25*(8), 539-543.

Brandts, J., Groenert, V., and Rott, C. (2014). The impact of advice on women's and men's selection into competition. *Management Science*, *61*(5), 1018-1035.

Buser, T., Niederle, M., and Oosterbeek, H. (2014). Gender, competitiveness, and career choices. *The Quarterly Journal of Economics*, *129*(3), 1409-1447.

Buser, T., Dreber, A., and Mollerstrom, J. (2017). The impact of stress on tournament entry. *Experimental economics*, *20*(2), 506-530.

Buser, T., Peter, N., and Wolter, S. C. (2017). Gender, competitiveness, and study choices in high school: Evidence from Switzerland. *American economic review*, *107*(5), 125-30.

Buser, T., Gerhards, L., and Van Der Weele, J. (2018). Responsiveness to feedback as a personal trait. *Journal of Risk and Uncertainty*, *56*(2), 165-192.

Cadsby, C. B., Servátka, M., and Song, F. (2013). How competitive are female professionals? A tale of identity conflict. *Journal of Economic Behavior and Organization*, *92*, 284-303.

Cárdenas, J. C., Dreber, A., Von Essen, E., and Ranehill, E. (2012). Gender differences in competitiveness and risk taking: Comparing children in Colombia and Sweden. *Journal of Economic Behavior and Organization*, *83*(1), 11-23.

Cason, T. N., Masters, W. A., and Sheremeta, R. M. (2010). Entry into winner-take-all and proportional-prize contests: An experimental study. *Journal of Public Economics*, *94*(9-10), 604-611.

Cassar, A., Wordofa, F., and Zhang, Y. J. (2016). Competing for the benefit of offspring eliminates the gender gap in competitiveness. *Proceedings of the National Academy of Sciences*, *113*(19), 5201-5205.

Chen, J., Houser, D., Montinari, N., and Piovesan, M. (2016). Beware of popular kids bearing gifts: A framed field experiment. *Journal of Economic Behavior and Organization*, *132*, 104-120.

COWI. Recruitment of engineers for the infrastructure sector. VEJ-EU, 2014. http: //www.vej-eu.dk/Nyheder.aspx?ID=65.

Crosetto, P., and Filippin, A. (2013). The "bomb" risk elicitation task. *Journal of Risk and Uncertainty*, *47*(1), 31-65.

Dargnies, M. P. (2012). Men too sometimes shy away from competition: The case of team competition. *Management Science*, *58*(11), 1982-2000.

Dariel, A., Kephart, C., Nikiforakis, N., and Zenker, C. (2017). Emirati women do not shy away from competition: Evidence from a patriarchal society in transition. *Journal of the Economic Science Association*, *3*(2), 121-136.

Datta Gupta, N., Poulsen, A., and Villeval, M. C. (2013). Gender matching and competitiveness: Experimental evidence. *Economic Inquiry*, *51*(1), 816-835.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, *9*(3), 522-550.

Dreber, A., von Essen, E., and Ranehill, E. (2014). Gender and competition in adolescence: task matters. *Experimental Economics*, *17*(1), 154-172.

Flory, J. A., Gneezy, U., Leonard, K. L., and List, J. A. (2018). Gender, age, and competition: A disappearing gap?. *Journal of Economic Behavior and Organization*, *150*, 256-276.

Gebreyes, Rahel. 'LEGO' Documentary Directors On The Gender Imbalance Of Toys. 2015. http://www.huffingtonpost.com/entry/lego-gender-toy-documentary.

Gneezy, U., Niederle, M., and Rustichini, A. (2003). Performance in competitive environments: Gender differences. *The Quarterly Journal of Economics*, *118*(3), 1049-1074.

Gneezy, U., Leonard, K. L., and List, J. A. (2009). Gender differences in competition: Evidence from a matrilineal and a patriarchal society. *Econometrica*, *77*(5), 1637-1664.

Halko, M. L., and Sääksvuori, L. (2017). Competitive behavior, stress, and gender. *Journal of Economic Behavior and Organization*, *141*, 96-109.

Healy, A., and Pate, J. (2011). Can teams help to close the gender competition gap? *The Economic Journal*, *121*(555), 1192-1204.

Kamas, L., and Preston, A. (2012). The importance of being confident; gender, career choice, and willingness to compete. *Journal of Economic Behavior and Organization*, *83*(1), 82-97.

Khachatryan, K., Dreber, A., Von Essen, E., and Ranehill, E. (2015). Gender and preferences at a young age: Evidence from Armenia. *Journal of Economic Behavior and Organization*, *118*, 318-332.

Khadjavi, M., and Nicklisch, A. (2018). Parents' ambitions and children's competitiveness. *Journal of Economic Psychology*, *67*, 87-102.

Krause, K., and Harbaugh, W. T. (1999). Economic experiments that you can perform at home on your children. *University of Oregon Department of Economics Working paper*, (1999-1).

Lee, S., Niederle, M., and Kang, N. (2014). Do single-sex schools make girls more competitive? *Economics Letters*, *124*(3), 474-477.

McCrum-Gardner, E. (2010). *Sample size and power calculations made simple. International Journal of Therapy and Rehabilitation, 17*(1), 10-14.

Masclet, D., Peterle, E., and Larribeau, S. (2015). Gender differences in tournament and flat-wage schemes: An experimental study. *Journal of Economic Psychology*, *47*, 103-115.

Mayr, U., Wozniak, D., Davidson, C., Kuhns, D., and Harbaugh, W. T. (2012). Competitiveness across the life span: the feisty fifties. *Psychology and aging*, *27*(2), 278.

Müller, J., and Schwieren, C. (2012). Can personality explain what is underlying women's unwillingness to compete? *Journal of Economic Psychology*, *33*(3), 448-460.

Niederle, M., Segal, C., and Vesterlund, L. (2013). How costly is diversity? Affirmative action in light of gender differences in competitiveness. *Management Science*, *59*(1), 1-16.

Niederle, Muriel, Carmit Segal, and Lise Vesterlund. "How costly is diversity? Affirmative action in light of gender differences in competitiveness." *Management Science* 59.1 (2013): 1-16.

Nosek, B.A., Smyth, F.L., Sriram, N., Lindner, N., Devos, T., Ayala, A., Bar-Anan, Y., Bergh, R., Cai, H., Gonsalkorale, K. and Kesebir, S. (2009). National differences in gender–science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences, 106*(26), 10593-10597.

Price, C. R. (2012). Gender, competition, and managerial decisions. *Management Science*, *58*(1), 114-122.

Rapport fra Udvalget om ligestilling i dagtilbud og uddannelse 2017. Ministry of Education, 2017

Reuben, E., Sapienza, P., and Zingales, L. (2015). *Taste for competition and the gender gap among young business professionals* (No. w21695). National Bureau of Economic Research.

Reuben, E., Wiswall, M., and Zafar, B. (2017). Preferences and biases in educational choices and labour market expectations: Shrinking the black box of gender. *The Economic Journal*, *127*(604), 2153-2186.

Saccardo, S., Pietrasz, A., and Gneezy, U. (2017). On the size of the gender difference in competitiveness. *Management Science*, *64*(4), 1541-1554.

Samak, A. C. (2013). Is there a gender gap in preschoolers' competitiveness? An experiment in the US. *Journal of Economic Behavior and Organization*, *92*, 22-31.

Shurchkov, O. (2012). Under pressure: gender differences in output quality and quantity under competition and time constraints. *Journal of the European Economic Association*, *10*(5), 1189-1213.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American psychologist*, *52*(6), 613.

Stoet, G. and Geary, D. C. (2018). The gender-equality paradox in science, technology, engineering, and mathematics education. *Psychological science, 29*(4), 581-593.

Sutter, M., and Glätzle-Rützler, D. (2014). Gender differences in the willingness to compete emerge early in life and persist. *Management Science*, *61*(10), 2339-2354.

Sutter, M., Glätzle-Rützler, D., Balafoutas, L., and Czermak, S. (2016). Cancelling out early age gender differences in competition: an analysis of policy interventions. *Experimental Economics*, *19*(2), 412-432.

Wozniak, D., Harbaugh, W. T., and Mayr, U. (2014). The menstrual cycle and performance feedback alter gender differences in competitive choices. *Journal of Labor Economics*, *32*(1), 161-198.

Zhang, Y. J. (2018). Culture, institutions and the gender gap in competitive inclination: Evidence from the communist experiment in China. *The Economic Journal*, *129*(617), 509-552.

Zhong, S., Shalev, I., Koh, D., Ebstein, R. P., and Chew, S. H. (2018). Competitiveness and stress. *International Economic Review*, *59*(3), 1263-1281.

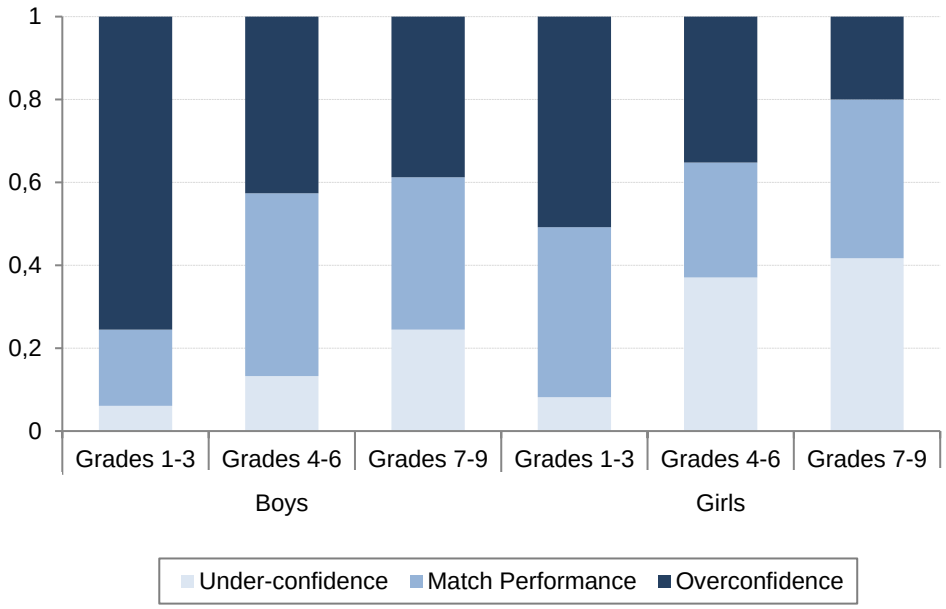
Appendix A. Additional tables and figures

Table A.1. List of studies using the Niederle-Vesterlund (2007) design in figure 1

Study	Sample size	Task	% selecting competition		
			Male	Female	Significance
Niederle and Vesterlund (2007)	80	Addition	73	35	Yes
Gneezy Leonard, and List (2009)	74	Bucket toss	50	26	Yes
	74	Bucket toss	39	54	No
Cason, Masters and Sheremeta (2010)	69	Addition	56	45	Yes
Dohmen et al. (2011)	240	Addition	62	37	Yes
Healy and Pate (2011)	192	Addition	81	35	Yes
Balafoutas and Sutter (2012)	72	Addition	64	30	Yes
Balafoutas, Kerschbamer, and Sutter (2012)	132	Addition	59	31	Yes
Cárdenas et al. (2012)	304	Math search	44	19	Yes
	304	Word search	39	27	Yes
	315	Math search	35	32	No
	316	Word search	26	29	No
Dargnies (2012)	76	Addition	85	51	Yes
Kamas and Preston (2012)	310	Addition	41	23	Yes
Mayr et al. (2012)	543	Addition	56	39	Yes
Mueller and Schwieren (2012)	127	Addition	42	26	No
Price (2012)	310	Addition	66	49	Yes
Shurchkov (2012)	128	Verbal	39	30	No
	84	Match puzzles	44	19	Yes
Andersen et al. (2013)	172	Bucket toss	52	49	No
	146	Bucket toss	51	39	No
Cadsby Servátka, and Song (2013)	132	Addition	36	9	Yes
Datta Gupta, Poulsen, and Villeval (2013)	100	Mazes	60	34	Yes
Niederle, Segal, and Vesterlund (2013)	84	Addition	74	31	Yes
Samak (2013)	123	Fishing task	77	83	No
Brandts, Groenert, and Rott (2014)	89	Addition	59	30	Yes
Buser, Niederle, and Oosterbeek (2014)	362	Addition	49	23	yes
Dreber, von Essen, and Ranehill (2014)	216	Addition	36	17	Yes
	216	Word search	33	28	No
Lee Niederle, and Kang (2014)	640	Addition	30	22	Yes
Wozniak, Harbaugh, and Mayr (2014)	128	Addition	54	31	Yes
	128	Verbal	50	30	Yes
Apicella and Dreber (2015)	191	Skipping rope	45	30	Yes
	88	Bead collection	52	37	Yes
	70	Handgrip strength	67	29	Yes

Khachatryan et al. (2015)	824	Addition	54	52	No
	824	Word search	57	56	No
Masclet, Peterle, and Larribeau (2015)	202	Decoding	62	34	Yes
Reuben, Sapienza, and Zingales (2015)	409	Addition	60	33	Yes
Sutter and Glätzle-Rützler (2015)	717	Addition	40	19	Yes
Zhang (2015)	96	Addition	60	38	No
	80	Addition	75	48	Yes
	96	Addition	63	48	Yes
Almås et al. (2016)	483	Addition	52	32	Yes
Berlin and Dargnies (2016)	228	Addition	63	35	Yes
Cassar, Wordofa, and Zhang (2016)	358	Addition	36	26	Yes
Sutter et al. (2016)	246	Addition	44	21	Yes
Apicella, Demiral, and Mollerstrom (2017)	100	Addition	58	38	Yes
Buser, Dreber, and Möllerström (2017)	104	Addition	52	28	Yes
Buser, Peter, and Wolter (2017)	249	Addition	68	51	Yes
Dariel et al. (2017)	147	Addition	50	54	No
Halko and Sääksvuori (2017)	80	Addition	74	54	No
Reuben, Sapienza, and Zingales (2017)	257	Addition	54	27	Yes
Banerjee, Gupta, and Villeval (2018)	168	Memory task	22	16	No
Buser, Gerhards, and van der Weele (2018)	297	Mix	42	26	Yes
Bönte, Procher, and Urbig (2018)	225	Math	56	45	No
Flory et al. (2018)	730	Arranging shapes	46	45	No
Khadjani and Nicklish (2018)	84	Running task	44	46	No
Saccardo, Pietrasz, and Gneezy (2018)	126	Ball-tossing	78	32	Yes
Zhong et al. (2018)	197	Addition	49	25	Yes

Figure A.1. Under- and overconfidence by gender and age group



Note: Comparing children's guessed rank with their actual performance, the figure shows the proportion of children who are overconfident, under-confident, or match their performance by age group and gender.

In Figure A.1. the columns are ordered by age groups first and then gender to illustrate the development in confidence as the children get older. A very large percentage of the boys in Grades 1-3 are overconfident but this decreases for boys in the older age groups. The percentage of girls who are under-confident increases with age (when comparing the oldest age group for girls with the younger age groups).

Table A.2. Marginal Effects on the Probability of the Choice of Compensation Scheme

	Dependent variable: Choice of competitive payment scheme						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female (=1)	-0.106*	-0.099*	-0.084*	-	-	-0.127**	-0.045
	(0.056)	(0.056)	(0.050)	-	-	(0.063)	(0.089)
Performance in stage 2	0.048***	0.041**	0.040**	0.043*	0.033	0.042	0.003
	(0.013)	(0.018)	(0.018)	(0.024)	(0.027)	(0.036)	(0.031)
Risk averse (=1)	-0.126*	-0.115	-0.097	-0.147	0.007	-0.077	-0.079
	(0.072)	(0.074)	(0.074)	(0.103)	(0.093)	(0.100)	(0.082)
Age group	0.077	0.085	0.079	0.074	0.128	-0.024	0.160
	(0.091)	(0.090)	(0.077)	(0.110)	(0.096)	(0.077)	(0.126)
Average performance in class	-0.062	-0.054	-0.056	-0.056	-0.061	-0.025	-0.042
	(0.049)	(0.052)	(0.045)	(0.070)	(0.041)	(0.038)	(0.078)
Confident in stage 2 (=1)		0.062	0.071	0.076	0.076	-0.099	0.190**
		(0.068)	(0.070)	(0.109)	(0.087)	(0.117)	(0.087)
Boy is best (=1)		-0.002	-0.005	0.012	-0.027*	-0.005	-0.002
		(0.012)	(0.012)	(0.013)	(0.016)	(0.015)	(0.014)
Number of opposite gender friends			0.018	0.020	0.013	0.014	0.032
			(0.022)	(0.041)	(0.039)	(0.050)	(0.034)
Number of siblings			0.055***	0.069**	0.047	0.033	0.055**
			(0.020)	(0.032)	(0.037)	(0.038)	(0.026)
Number of competing friends			0.051**	-0.001	0.100***	0.060***	0.033
			(0.021)	(0.051)	(0.034)	(0.022)	(0.042)
Sample	All	All	All	Males	Females	Above average	Below average
Observations	341	339	338	164	174	168	170
R-squared	0.049	0.048	0.072	0.081	0.094	0.069	0.057

Note: The dependent variable of the probit regression is equal to 1 if the participant chose tournament incentives in Stage 3. Robust standard errors clustered by class ID are in parentheses. The participants are divided into two groups, males and females, with the corresponding regression results in column (4) for males and in column (5) for females. In column (6) and (7) the participants are split by their *actual* average performance in stage 2; we run a regression for the children performing above average in column (6) and below in column (7). 'Number of siblings' is a continuous variable, while 'Age group', 'Risk averse', 'Confident in stage 2', and 'Boy is best' are dummies. The number of girls above average is 98 (out of 168) and the number of girls below average is 76 (out of 170). *p<0.1, **p<0.5, ***p<0.01

Appendix B. Preparations, prizes and task

Our experiment is a part of the project "KIDS" at the University of Copenhagen. The KIDS project has IRB approval from The Institutional Ethical Review Board at the Department of Psychology, University of Copenhagen (Approval numbers: 2015/12 and 2016/B04). We obtained consent from school leaders and parents (signed a consent form), agreeing to take part in our study aimed at investigating children's economic decision making. We informed the children that it was voluntary for them to participate, and that they could with-draw at any time. We never mentioned gender issues to schools, parents or children. The experiment took place during regular school hours in 20 classes in four different elementary schools; two private[19] and two public. The average number of children in a class who performed the experiment was 18. Concerning the gender composition in class in the three age groups; we had 45 % boys and 55 % girls in the youngest and oldest age group, and in the middle age group, we had 54 % boys and 46 % girls. Information about the experimental procedures was revealed sequentially, immediately before the students were performing each task. The instructions were given verbally as suggested by the school leaders, because of the level of literary among younger students and the time constraint of the sessions. The instructions were interactive and involved several examples and control questions to make sure that all participants understood the details of the different payment schemes.

Prizes

We used small gifts rather than money to make the comparisons of rewards less salient among the participants, as suggested in Krause and Harbaugh (1999). To find appropriate gifts and test the Lego task, we conducted a pilot with 30 children at the after-school program "Dronen" and at Experimentarium City in October, 2015 with children aged 8 to 11, which corresponds to grades 3 to 5. We carried out a

[19] The two private schools are Danish free schools (33 % of the schools in Denmark are private schools).

small "market research" when doing these pilot studies, and children assessed different items such as toys, small games and school supplies to investigate how popular the items were among the children their age. We selected the rewards for the actual experiment based on their opinions and suggestions to make sure the incentives of the participants were salient enough.

Lego task

We used a task based on Lego because most children in Denmark know how to build with Lego. Both boys and girls play with Lego, and from an early age, children know how to follow Lego manuals and build small constructions. Globally, Lego was proclaimed as one of the few gender-neutral toys on the market until recent years (Rahel Gebreyes, 2015). Thus, we created the Lego task anticipating boys' and girls' equal ability of building with Lego. Further, Lego is recognized as a toy making it fun for children to participate, which helps ensure the children's attention throughout the experiment. We kept the level of difficulty constant across ages, and older children were expected to perform better on average providing them with higher payoffs to match the age variability in children's pocket money and ensuring their motivation. We constructed Lego wall checker sheets from overhead projector plastic sheets. Instead of checking individually that every row of a Lego brick wall is built with the correct Lego bricks, it is possible to place the Lego checker sheets on top of the Lego construction and immediately tell if the rows are built according to the plan. In case of mistakes the bricks change colors on the sheets, as shown in the figures below.

Figure B.1. The Lego brick walls with laminated plans

Figure B.2. Children performing the Lego task

Figure B.3. Lego wall and transparent checker sheets

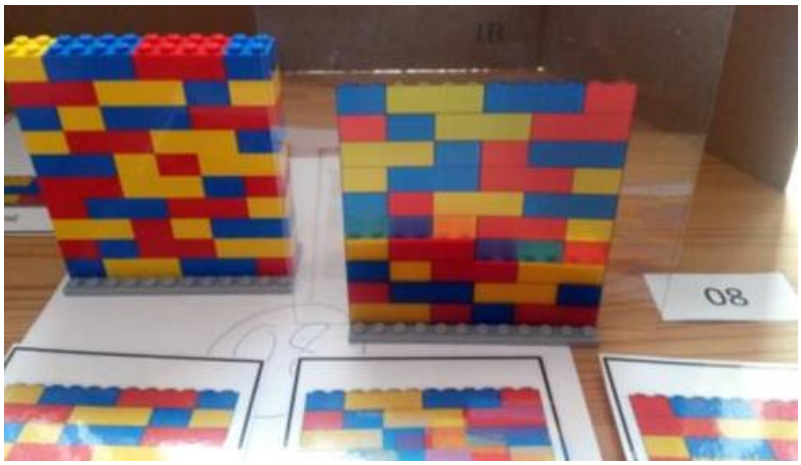


Figure B.4. Rewards - "The experiment shop"

Appendix C. Experimental procedures

Demographic questionnaire

A week before the experiment, parents receive a letter with an invitation to take part in the project along with a consent form. Parents of children participating in the experiment are asked to sign the consent form and fill out a demographic questionnaire. The demographic questionnaire include questions about the children's spare time activities, whether they receive pocket money (the amount and frequency), and the number of older and younger brothers and sisters.

Subjects and the settings

The participants in the experiment are children aged 7 to 16, corresponding to grades 1-9 of primary school. The experiment is carried out in the classroom, and the children are randomly assigned new seats, when they enter the room before the experiment begins. They draw a laminated number, and the experimenter directs the participant to his or her seat. Test dividers made of cardboard are placed on the tables,

so the participant can make his or her decisions in private. More space around the individual participant is created by rearranging the tables and the chairs in the classroom. On the table in front of each participant, there is a box of Lego bricks, a pen and four decision sheets tied together with paper clips and faced downwards. In order to motivate the participants' performance in the economic games, the experimenters will begin the sessions by showing them some of the toys, games, and school materials that they can buy with the points they earn in the games. The teacher is always present in the classroom to observe the experimental proceedings, and he or she is informed about the decision games beforehand. The teacher is asked not to discuss the games with the children; the experimenters will instead assist the participants, if they have any questions. While experimenter 1 begins to explain the instructions for the Lego game, experimenter 2 enters the names of the participants next to their drawn number in a prepared Excel sheet.

Script

"Welcome to our experiment. Together we will play three games. In every game, you have the opportunity to earn points that you can later spend in our Experiment Shop. With your points, you can buy small toys, games, and school materials. More points mean more prices. Your decisions in the games are private and no one will know how many points you will have earned in the games - not your classmates or your teacher - unless you tell them yourself. We will now explain the rules of the first game. Pay careful attention to these instructions, because the better you understand them, the more points you can earn. Please do not talk to each other from this point on. If you have a question, please raise your hand and ask. Otherwise, we urge you to be quiet and listen carefully just like you normally listen to your teacher in school."

Task 1. The Lego Task

"In front you on the table there are a pen and some sheets of paper. Do not touch that for now, we will explain about it later. Instead, focus on the box of Lego bricks. Inside you will find some bricks of different shapes and colors and three base plates. There is also a laminated drawing of a wall made of Lego bricks with three rows. Try to see if you can replicate the Lego wall. It has to be built exactly as it is on the plan. Once you have finished, place the Lego wall in front of you, so that we can check that it is built correctly." The participants now have 1-2 minutes to build the wall (Test). The two experimenters go around and check if the participants have understood the task. The experimenter praises the participants: "Good job! This was just a test. Now you have to separate all the bricks and put them back in the box. None of them must stick together." In the meantime, the two experimenters place a set of the first three laminated plans (1A, 1B, 1C) upside-down on the table next to each participant. "Don't look at this before we tell you to." Experimenter 1 addresses the whole class: "Now we begin the first part of the Lego game. In front of you, there are three plans with different Lego walls. Each wall now has ten rows, while before it was only three rows. We will give you three minutes to build as many rows as you can. For each row correctly built, you will receive 1 point. If you finish more than the ten rows of a wall, you just start a new wall with one of the other plans. Afterwards we will count the number of correctly built rows. Have you understood the task? Do you have any questions? Let us try an example: If Johan builds a Lego wall with five rows, how many points does he get? Five, yes that is correct. Now I will set the countdown clock. Are you ready? Ready, set, go build!"

"..3, 2, 1, now the time is up! Stop building and do not touch the Lego bricks. We will now come and check your Lego walls. The experimenters use the plastic checker sheets, and write down the number of correctly built rows on each participant's Lego decision sheet. "Good job all of you! Now you have to separate all the bricks and put them back in the box again. None of them must stick to together." The experimenters go around and replace the old Lego plans with new ones (2A, 2B, 2C) for the second part

of the Lego game. The plans are again placed upside down on the table. "Now we are ready for the second part of the Lego game. Again, you have plans with Lego walls in front of you, and as before, you will be given 3 minutes to build as many rows as you can, BUT this time it is a competition. Each of you is matched with another person in this room. This was done randomly, when you entered the room and picked a number. We will not tell you whom you are competing with. If you build more rows than your opponent does, you win the competition and you will receive 2 points per row you correctly build. However, if you finish fewer rows than your opponent does, you lose and you will get zero points. If you finish the same number of rows as your opponent, you get 1 point per row. Let us try an example. If Johan builds five rows and his opponent builds four rows, how many points does he get? Yes, he wins, and he receives 10 points. What if his opponent builds five rows? Then he gets 5 points. What if his opponent finishes six rows? Yes, he loses and gets zero points. Do you have any questions? Okay, I will set the countdown clock. Are you ready? Ready, set, go build!"

"..3, 2, 1, now the time is up! Stop building and do not touch the Lego bricks. We will now come and check your Lego walls." The experimenters use the plastic checker sheets, and write down the number of correctly built rows on each participant's Lego decision sheet. "Good job all of you! Now you have to separate all the bricks and put them back in the box as you did before. None of them must stick together." The experimenters go around and replace the old Lego plans with new ones (3A, 3B, 3C) for the third part of the Lego game. The Lego plans are again placed upside down on the table. "We are now at the third and last part of the Lego game. In front of you, we have placed three plans with Lego walls, and as before, you have 3 minutes to build as many rows of Lego that you can. However, this time you can decide which of the two payments schemes that you prefer. Do you want to enter the competition with a chance to earn 2 points per Lego row as you did in part 2, or do you want to stay out of the competition and earn 1 point per Lego row as in part one? You have tried both. We will give you one minute to choose, and then you must tick the box on the Lego decision sheet. Do not look to what your classmates are writing; it has to be your own decision. Please, raise your hand if you have any questions.

We will assist you now." The experimenters make sure that all participants indicate their decision on whether to enter the tournament or not on the Lego decision sheet. "Thank you for this. Now we are ready to begin. I will set the countdown clock. Are you ready? Ready, set, go build!"

"..3, 2, 1, and now the time is up! Stop building and do not touch the Lego bricks. We will now come and check your Lego walls." The experimenters use the plastic checker sheets, and write down the number of correctly built rows on each participant's Lego form. "Good job all of you! Now you have to separate all the bricks and put them back in the box. Again, none of them must stick together. Place the Test plan inside and close the box. We now ask you to answer two questions. Today you are # children participating in game. How fast do you think you were in building with Lego compared to your classmates in the first part of the game? If you think you did best of all, you write one. If you think you did second best, you write two, and so on. If you think you did worst, you write #. Do you understand? Similarly, in the second part of the game, how fast do you think you were in building with Lego compared to you classmates? Do not look to what your classmates are writing; it has to be your own decision. Please, raise your hand when you are done or if you have any questions. We will assist you now." The experimenters take the Lego decision sheets and the boxes of Lego, and experimenter 2 begins to enter the data in the Excel sheet. In front of each participant, the decision sheet with the bomb game now lies facing up.

Task 2. The Bomb Risk Elicitation Task with pen and paper

On the decision sheet of The Bomb Game, there is a drawing of a winding road with exactly one hundred steps and they are numbered one through 100.

"This is the second game, and it is called The Bomb Game. In this game you have to imagine that you are in a minefield, and on the winding road that you see on the paper sheet, exactly one bomb is hidden behind one of the 100 numbers. For every step you take, you gain 1 point. You have to start at step one, then 2, 3 and so on. You indicate every step you take by writing a cross over that number, and

you continue until you reach the step where you want to stop. You also have to write the number of steps that you finally decide you want to take in the box below on your left. (For the younger children: We will assist you with this.) You do not know the bomb's location. None of us does. You only know that it is equally likely to be behind any of the hundred steps on the winding road. After you have decided on the number of steps to take by writing a cross on each step, we will determine where the bomb is located by rolling two ten-sided dice (00 and 0=100). The dice look like this – see!"

If the number where the bomb is located is higher than the number of steps, you have taken, you have not stepped on the bomb and you earn 1 point for each of the steps you have taken. If the number where the bomb is located is lower than or equal to the number of steps you have decided to take, you have stepped on the bomb and lose all your points; it means that you get zero points. So the more steps you take, the more points you can make, but the risk that you will step on the bomb is also higher. How many steps will you take?

Let us try an example: Johan walks 25 steps and the bomb is hidden behind step number 21. Will he have stepped on the bomb? How many points will he get? What if Johan walks 75 steps and the bomb is hidden at step number 79, how many points will he get? Suppose Johan walks 48 steps and the bomb also is hidden at number 48, how many points will he get? Does the number of steps Johan decides to take have an influence on which step the bomb is located? You now have two minutes to make you decision. I will afterwards roll the dice and we will see where the bomb is located. You can begin now."

The participants have two minutes to make their decision, and they write it numerically below the winding road. The experimenters assist and check that this has been done for all the participants, before experimenter 1 roll the two ten-sided dice to determine the location of the bomb. The decision sheets are thereafter collected by the experimenters, and experimenter 2 starts to enter the data in the Excel sheet. On the table in front of each participant, the decision sheet with the Friends Elicitation Task now lies facing up.

Task 3. Eliciting the network of friends

The "A game with your friends" decision sheet has a drawing on it of a table with six chairs.

"The third game is a game with your friends. On this piece of paper a table and six chairs are drawn. Please write your own name above the chair at the head of the table. You also have to write the names of five friends in the class that you would like to have seated at the table next to you. No one will know what names you write, not your classmates or your teacher, we will not tell anyone. You will receive 10 points for completing this task. If you have two persons in this class, that have the same first name, try also to write their last names. You can begin now, but make sure that none of your classmates see what you write."

Experimenter 1 takes the decision sheets, while experimenter number 2 finishes the calculation of the points. "We have calculated your total points, and we will now open our Experiment Shop. Hold on to the laminated numbers that you received when you entered the room. We will call you up one by one, and we will start with the person with number 1. We will let you know how many points you got in total, and you can then choose your price."

Decision sheets and material

DET SAMFUNDSVIDENSKABELIGE FAKULTET
KØBENHAVNS UNIVERSITET

ID-nr.:_____

Lego

Number of rows in part 1: ☐

Number of rows in part 2: ☐

Do you choose to enter the competition: Yes ☐ No ☐

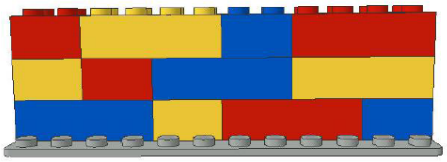
Number of rows in part 3: ☐

How fast do you think you were in building with Lego compared to your classmates?

We are ☐ that are participating in the game.

In part 1 I'm number ☐ from the top.

In part 2 I'm number ☐ from the top.



DET SAMFUNDSVIDENSKABELIGE FAKULTET
KØBENHAVNS UNIVERSITET

ID-nr.:_____

The bomb game

Imagine that you are in minefield and 1 bomb is hidden behind one of the 100 numbers below.

For every step you take you gain 1 point. If you step on the bomb you lose all your points.

How many steps will you take? Put a cross at every step you take.

The location of the bomb will afterwards be determined by a roll of two ten-sided dice.

Start here ↓

1									
2	3	4	5	6	7	8	9	10	11
									12
22	21	20	19	18	17	16	15	14	13
23									
24	25	26	27	28	29	30	31	32	33
									34
44	43	42	41	40	39	38	37	36	35
45									
46	47	48	49	50	51	52	53	54	55
									56
66	65	64	63	62	61	60	59	58	57
67									
68	69	70	71	72	73	74	75	76	77
									78
88	87	86	85	84	83	82	81	80	79
89									
90	91	92	93	94	95	96	97	98	99
									100

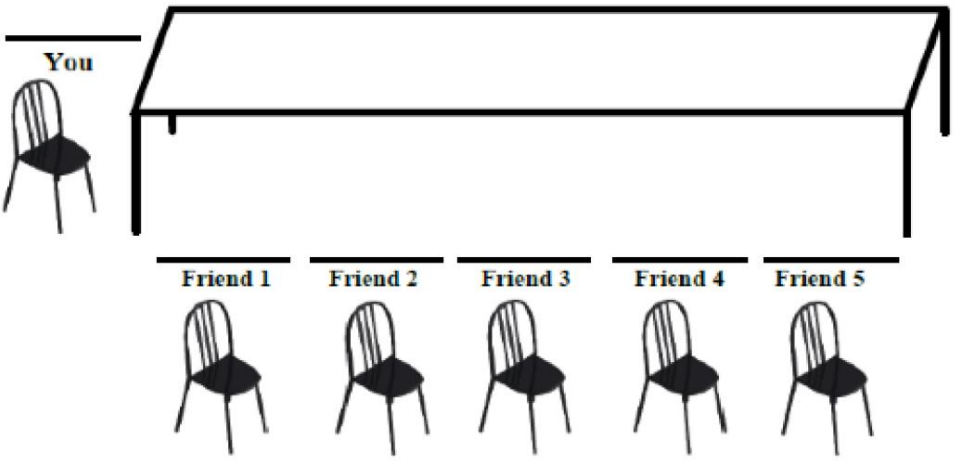
Number of steps: Points:

DET SAMFUNDSVIDENSKABELIGE FAKULTET
KØBENHAVNS UNIVERSITET

ID-nr.:_____

A game with your friends

Write your name above the chair at the head of the table. You have to write the names of up to five friends from your class above the remaining chairs at the long side of the table. Your best friend should be seated closest to you, your second-best friend second closest to you, and you go on like this.



DET SAMFUNDSVIDENSKABELIGE FAKULTET
KØBENHAVNS UNIVERSITET

ID-nr.:_____

Other questions

1. Here is a picture of Gertrud, and here is a picture of Victor. They are both in the 5th Grade. Here they are building with Lego bricks.

How many rows of Lego do you think Gertrud managed to build in 3 minutes? ☐

How many rows of Lego do you think Victor managed to build in 3 minutes? ☐

2. How difficult was it to understand and make your decisions in the Lego game? (please tick a box)

Very easy									Very difficult
1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

3. How difficult was it to understand and make your decision in the Bomb game? (please tick a box)

Very easy									Very difficult
1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Chapter 2

Does retirement increase informal caregiving for parents and parents-in-law?

Does retirement increase informal caregiving for parents and parents-in-law?

Lotte Kofoed Jørgensen

University of Padova

September 2019

Abstract

Using SHARE data, I examine the effect of retirement on daily caregiving to parents and parents-in-law. I focus on individuals aged 50-70 in 18 European countries. I account for the possible simultaneity bias of the retirement decision and the choice to provide informal care, by instrumenting the retirement status with having reached country-specific pension eligibility ages for early and statutory retirement. I find a significant and positive causal effect of retirement on daily caregiving. Moreover, I find evidence of heterogeneous effects across European regions.

Keywords: Informal care, retirement, instrumental variables, SHARE.

JEL classification: J22, J14, J18, J26, H55.

1. Introduction

Due to population ageing in Europe, the demand for long-term elderly care is expected to rise in the future. The increasing costs of providing formal care put welfare systems under pressure, with them greatly relying on families to take on more informal care. Traditional gender roles and the division of paid and un-paid work within families and partnerships may lead women to take on more of the caregiving role. Informal caregiving can be very time consuming, often at inflexible hours, which over time may impact the work hours and employment status of caregivers, along with the timing of their retirement decision. The time conflict between the supply of informal care and the labor supply in later work years may dissolve through retirement, as retirement typically offers more time, and so individuals engaged in more intensive care for a parent or parent-in-law may seize early retirement options.

To address concerns related to a shrinking workforce, official pension ages have been increasing across Europe, generally pushing people to extend their working lives by delaying retirement (OECD, 2017). Nevertheless, for decades social security programs themselves have provided strong incentives to leave the labor force through early retirement options, thereby increasing the financial burden caused by aging populations (Wise and Gruber, 2010). Tightening the generosity of pension benefits and increasing the pension eligibility ages, pension reforms aim to boost employment rates in later years and increase the labor market participation of underrepresented groups, especially women, with consequences for informal care provision. Individuals of retirement-age are likely to have a living parent or parent-in-law in need of care. If labor market activities crowd out informal caregiving, this may imply a declining supply of informal care. From a policy point of view, it is important to understand the correlation between the retirement status and the provision of informal care, and particularly whether policies increasing pension eligibility ages, and thereby advancing later retirement, negatively affect informal caregiving, especially in combination with programs that reduce the availability of formal care. Policies encouraging both informal care and later retirement may not be feasible, if they target the same population group without possibilities for more flexible work arrangements. Delaying retirement and subsequently increasing the labor supply

in later work years, may negatively affect the availability of individuals who can provide informal caregiving.

In this paper, I estimate the causal effects of retirement on informal caregiving to parents and parents-in-law. I use data drawn from the Survey of Health, Ageing and Retirement in Europe (SHARE), which offers rich longitudinal information on various life domains of European individuals aged fifty and above. Decisions about retirement and informal caregiving for a parent or parent-in-law may be joint endogenous, and this complicates the identification of a casual effect. While recent studies mostly aim at identifying the effect of informal caregiving on the employment status and work hours of individuals of retirement-age, I instead explore the other direction and estimate the effect of retirement on informal caregiving, applying an instrumental variables approach on a cross-country sample. Using country-specific pension age eligibility rules for both early and statutory retirement as instruments for the retirement status, I find that retirement significantly increases the probability of daily caregiving for a parent or parent-in-law. I furthermore find evidence of heterogeneous effects across European regions. My results contribute to the discussion on the potential impact of pension reforms delaying retirement on the future supply of informal care.

The remainder of the paper is organized as follows. In section 2, I survey the related literature. I present the theoretical foundation and the identification strategy in section 3, and I introduce the data set and describe the sample and key variables of interest in section 4. In section 5, I show and discuss the empirical results, and I check their robustness in section 6. Section 7 concludes.

2. Literature Review

Previous research on this topic mainly focuses on estimating the effect of informal caregiving on the labor market participation of retirement-aged caregivers without explicitly concentrating on the retirement outcome (Jacobs et al., 2014). More knowledge is needed about how informal caregiving is related to the retirement status and timing, and my study contributes to the empirical literature connecting the labor supply, the retirement decision, and the provision of informal care. Although I estimate the effect of retirement on the probability of informal caregiving, it is important to explore what the literature says about the opposite causal direction; the impact of informal caregiving on the employment of individuals who are close in years to retirement, as the two pathways seems interconnected. In fact, most empirical studies seem to confirm this two-way causality between caring responsibilities and the employment status (Carmichael et al., 2010). The estimation is potentially complicated by the presence of unobservable factors that may influence both the employment decision and the choice to provide informal care. For example, individuals who are less attached to the labor force and have fewer labor market opportunities may more likely choose to become caregivers, and this creates a selection bias. To deal with the endogeneity concerns, different research designs have been developed; panel data methods, first differences, and more recent, different instrumental variable approaches have been proposed by the literature with their relevance and validity being discussed (Ciani, 2011).

Somewhat mixed evidence has been reported, but in general authors find a small negative effect of intensive caregiving on paid work, even if the effect size and significance vary greatly with the applied estimation method and the specific data samples (Heitmueller, 2007; Bolin et al., 2008; Lilly et al., 2010; Van Houtven et al., 2013; Crespo and Mira, 2014; Ciccarelli and Van Soest, 2018). The caregiving intensity seems to be key, as previous studies find that caregivers tend to drop out of the labor force, if they control for more intensive care. An intensity threshold ranging from 10 to 20 hours per week must be reached

before caregiving significantly impacts the labor force participation (Lilly et al., 2011). Using IV and panel quasi-fixed effect methods on British data, Heitmueller (2007) finds a significant and negative effect of intensive caregiving on employment for co-residential caregivers, while Lilly et al. (2010) report that individuals involved in daily caregiving are considerably more likely to withdraw from the labor market than non-caregivers. Bolin et al. (2008) find that giving informal care to an elderly parent is associated with significant costs in terms of missed labor market opportunities. Caregiving also affects retirement; using data from the Health and Retirement Survey in the U.S., Van Houtven et al. (2013) report that caregiving leads to a higher probability of partial retirement through reduced weekly working hours.

A large subset of related previous studies focuses only on women. As primary caregivers, they more often provide informal care, even if the flow of women into the labor force in the past few decades has increased the size of the female labor participation in many countries (Gruber and Wise, 2010). Crespo and Mira (2014) use data from the first two waves of SHARE to study the link between parental health, daily caregiving, and the employment status of daughters, and they find that Southern European women, who engage in daily care, decrease their probability of being employed. Across Europe there is substantial heterogeneity with respect to the availability and generosity of public formal care for the elderly. Welfare systems in Southern Europe generally rely relatively more on the immediate family for support. Also using SHARE, Dobrescu (2015) finds evidence of a European north-south gradient in informal care provision and explain it by factors related to family norms and parental responsibility, arguing that because of weaker family ties in northern and central European countries, elderly parents rely less on their children. Stronger ties in southern countries imply that children take care of their parents in old age, often with several generations living together under the same roof. Cultural differences in norms, traditions and preferences for formal versus informal care are reflected in institutional differences. Adding cross-country variations in female labor participation rates to that, the effect size of retirement

on informal caregiving is expected to be heterogeneous across European regions, though general patterns may be similar.

While there is a substantial literature on the effect of caregiving on the labor supply, only a limited number of studies have investigated the effect of labor market participation on the willingness to undertake informal care (Carmichael et al., 2010; Michaud et al. 2010; Nizalova, 2012) with results in general suggesting a negative effect of employment on informal caregiving. Carmichael et al. (2010) use fifteen waves from the British Household Panel Survey to examine the relationship between the employment status and transitions into informal care using a discrete-time duration model employing binomial, multinomial and ordered logistic estimation procedures. Presenting their results separately for males and females, the authors find that the likelihood of becoming a caregiver is significantly and negatively related to previous participation in paid employment and to hourly earnings. Studying the employment and caregiving trajectories of respondents in the British Household Panel Study, Michaud et al. (2010) adopt a reduced-form dynamic approach to model care giving decisions and find a negative effect of employment on future caregiving both inside and outside the household. Nizalova (2012) estimates the wage elasticity of the informal care supply using the 1998 wave of the Health and Retirement Study. With an instrumental variables approach employing regional unemployment rates and information about the industry structure, the author finds a strong negative effect of wages on caregiving to elderly parents. Employment may also affect the demand for formal care; studies show that lower availability of informal care influences the use of formal care (Van Houtven and Norton, 2004).

Providing informal care is a costly activity for the caregiver, and retirement may lower that cost. Thoughts about the timing of retirement may factor in when choosing to provide care. In fact, informal caregivers tend to retire earlier (Fischer and Müller, 2019). Workers may decide on early retirement or

minimize weakly hours in their final years in the labor market to meet the care needs of elderly parents. Consequently, decisions about caregiving may coincide with retirement choices. There is some uncertainty in the existing literature as to whether caregiving actually leads to retirement (Jacobs et al., 2017). The closer the individual is in years to the pension eligibility age, the greater degree of choice he or she may possibly have in the provision of informal care. The intuition being that individuals, who are advancing in their careers in the labor market, are more likely to have a high opportunity cost of time and thereby less likely to substitute formal care for informal care. Individuals are perhaps giving up less at the end of their careers. Focusing on the retirement outcome, Jacobs et al. (2014) examine the relationship between the caregiving intensity and retirement status of older individuals. Using the Canadian 2007 General Social Survey, the authors apply multinomial logistic regressions to determine whether providing different intensities of informal care is significantly associated with the likelihood that the individual is retired, and they find that higher intensity caregiving is associated with being fully retired. Examining the opposite direction in a recent discussion paper, Fischer and Müller (2019) investigates whether women increase their informal care provision when retiring, applying a fuzzy regression discontinuity design and using early retirement thresholds for women in the German pension system as instruments for their retirement decision. The authors find a significant positive impact on informal care provided by women who retire from employment.

3. Theoretical foundation

To conceptualize the effect of retirement on an individual's decision to provide informal care, I use a standard microeconomic model of time allocation, starting from the static model of caregiving and labor supply presented by Wolf and Soldo (1994). The individual who is of retirement-age can divide his or her fixed, total amount of time, T, between work, rewarded by the wage rate, w, and other activities, such as leisure, L, and informal caregiving, IC. Time spent in the labor market is chosen to maximize the utility

function (1), which depends on both leisure, L, and consumption, C, subject to the budget constraint defined by the non-market income, Y_n, and the wage, w. Caregiving is inserted in the time allocation decision by a parent or parent-in-law's need for long-term care, as indicated by N. Satisfying the care need increases the utility of the individual. There are two possible care arrangements; either the adult child can rely on formal care, FC, or he or she can provide direct informal care, IC, using his or her own time. A third option, the no-care option, is implicitly available, although adult children in many countries have a legal obligation to take care of their parents in old age (Filial support laws).[1]

Individuals who are confronted with the demand for care face a time conflict between their labor supply and their supply of informal care. If the individual has a parent or parent-in-law in poor health and in need of care, the individual can reduce his or her working hours, exit the labor market completely, or reduce leisure time to provide informal care. Alternatively, the individual can maintain or increase his or her time in the labor market to pay for formal care with the wage income, w. I assume that formal and informal care are mutually exclusive, as they in this setup offer the same kind of services. The price of the care arrangements is built into the budget constraint, assuming a monetary cost of buying formal care, p^{FC}, while the cost of informal care is the opportunity cost of labor, w, for the individual in the labor market. Increased care duties may lead the individual to retire earlier, implying a positive correlation between caregiving and retirement. The relative price of providing informal care as opposed to purchasing formal care, may change with retirement. Becoming eligible for pension benefits, the individual may reevaluate his or her care decision. The opportunity cost of labor, w, falls, while the cost of formal care is kept constant, all else being equal, and depending on the relative price of the two care-arrangements, the individual may prefer the less costlier option to give informal care to the parent or parent-in-law; i.e.

[1] In Germany, for instance, persons who are related in a "direct line" (grandparents, parents, children, grandchildren) are required to support each other, and this includes adult children with parents in need of daily care (Gans et al., 2009).

the substitution effect dominates. On the other hand, the individual may delay retirement to pay for formal care, thereby suggesting a negative correlation between informal caregiving and retirement. A high wage income implies a high opportunity cost of labor, and thereby a high cost of providing informal care, which may be relatively higher than the price of formal care. In this case the income effect dominates; i.e. the individual remains in employment by delaying retirement to pay for formal care. To some extent there may be an element of choice about whether to undertake informal care or not, and the individual may be more willing to supply care, the lower the opportunity cost of doing so is. This reasoning is in line with the previous literature. For example, Carmichael et al. (2010) find that individuals are more likely to provide care, if they either are not in paid employment or their wages are relatively low.

If the individual is retired, $R = 1$, then providing informal care still decreases total time, T, but as $w(T - L - IC)$ falls to zero, the cost of informal care, which is equal to the marginal utility of leisure, drops. Here IC is a binary choice and not continuous, and it is formulated in the budget constraint as a fixed set of daily hours spent by giving informal care. The individual's problem is to maximize

$$max_{C,L,IC,FC,R} \; U(C, L, IC, FC, R; N) \tag{1}$$

s.t. $$C + p^{FC}FC \leq w(T - L - IC) + Y_n \quad if \quad R = 0 \tag{2}$$

$$C + p^{FC}FC \leq P + Y_n \quad if \quad R = 1 \tag{3}$$

$$N \leq FC + IC \tag{4}$$

$$T \leq IC + L + H \tag{5}$$

The retirement decision, R, alters the total income, as shown by the right-hand side of equation (2) and (3). Whether the individual is entitled to pension, P, depends on him or her having reached the pension eligibility age for early or statutory retirement. If the individual retires before the pension eligibility age

cut-off, $PE = 0$, the individual is left with a total income, Y, equal to the non-market income, Y_n. P is an independent fixed sum with $0 < P < w(T - L - IC)$. The maximization problem above leads to first-order conditions that summarize the relationships among the choice variables, C, L, IC, FC, and R, and they serve as the basis for the empirical specification and estimation method presented below. In equation (4) the amount of time devoted to care, N, needed by the parent or parent-in-law can be satisfied with either formal care, FC, or informal care, IC. In the equilibrium $N = FC + IC$, and so either FC or IC can be dropped from the utility maximization problem; i.e. when one is calculated, the other is known as they must sum to N. In equation (5), the possible time uses of the individual, such as leisure, L, hours of informal care, IC, and hours spent in the labor market, H, if not retired, must sum up to the overall time endowment, T.

It is not explicitly given that retirement automatically leads to more informal caregiving. Retirement increases the total time available, thereby lowering the opportunity cost of time for informal caregiving; the cost of leisure. The caregiving need of a parent or parent-in-law is often inflexible, and to accommodate that need while employed, the adult child may have reduced time spent on leisure instead of work hours. The effect of retirement on informal care could therefore be zero, if retirement induces the caregiver to increase leisure instead of time spent on informal care. The intuition behind the potential retirement effect on informal care giving can be thought of as a time conflict between the supply of labor and the supply of informal care, if the eligibility age for pension benefits is not yet reached. As the retirement decision with pension benefits becomes available, the time conflict between the labor supply and informal caregiving dissolves.

The decision to give informal care may also depend on the availability and cost of formal care. Labor market earnings can help individuals pay for formal care services to the care recipient instead of

providing informal care themselves. Formal care can be considered a substitute for long-term informal care, if it offers the same services (Bolin et al., 2008), and if the care needs of the elderly parent are relatively low and require an unskilled type of care (Bonsang, 2009). Regular types of formal home care may include paid domestic help, nursing care at home, and/or meals on wheels. Informal care is usually preferred over professional care by care recipients (Carmichael and Charles, 2003). There may be a learning effect involved when providing long-term care; with time the productivity increases, and care hours may decrease. On the other hand, the care intensity may also increase as the care need intensifies with the age and progression of the disability of the parent. Retirement often lowers total income, and combined with a lower opportunity cost of time, individuals may take on more informal care, if the two forms of care are close substitutes, and the relative cost between them changes, when the individual retire.

Based on the above framework of time allocation, I a priori expect the effect of retirement on informal caregiving to be positive, due to the combination of an increase in total time and a reduction in the total income. However, the estimation is complicated by the fact that individuals, for whom caregiving compared to their work in the labor market is relatively high-valued, are more likely to retire earlier. The retired individuals providing care for a parent or parent-in-law might therefore have been giving daily care even if not retired, and this leads to a bias. Factors related to the individuals' preferences, such as being more family-oriented or less ambitious, naturally affect all the choices of the individual considered in this model, including retirement and informal care. With the methods employed in this analysis, I do not control for the possibility that caregiving decisions made much earlier on in life determine future employment characteristics and the timing of retirement. The role of cultural norms and expectations also matters; some individuals may find it more difficult to turn down a request for informal care, particularly if their hourly earnings are relatively low or if they are retired.

3.1 Method

Caregiving and the retirement decision may be joint endogenous with unobservable factors influencing both. The characteristics and preferences of the individuals determine their behavior in terms of labor supply and retirement decisions as well as their provision of informal care. This motivates an instrumental variable approach with a two-stage least squares (2SLS) regression model, where I instrument the retirement status, R_i, with country-specific pension age eligibility criteria for early and statutory retirement, PE_i. Exploiting institutional changes in pension age eligibility rules as instruments for the retirement status has been widely used in the literature (Angelini et al., 2009; Mazzonna and Peracchi, 2014; and Celidoni and Rebba; 2017). The pension age thresholds generate exogenous variation in the labor supply. When individuals reach their pension eligibility age for early and statutory retirement and satisfy their contribution criteria, they become eligible for pension benefits. This changes their choice set and budget constraint because retirement with pension benefits becomes available. As instruments I define two zero–one dummy variables indicating whether the individual is eligible or not for either early retirement (EligibleER) or statutory (normal) retirement (EligibleSR).

The thresholds serve as instruments for the individual's retirement decision. For example, the dummy variable, EligibleER, is equal to 1, if the individual at the time of the interview has reached an age equal to or above the age threshold that allows the individual to retire early based on the current country-specific pension rules, and 0 otherwise (this is for individuals who are not yet retired). For the individual who is retired, the instrument considers the past pension eligibility age rule for early retirement that was in place at the time of the person's retirement. The instruments are not only country-specific, they are also characterized by within-country heterogeneity across the cohorts for each country in the sample. I conjecture that age eligibility for pension benefits allows individuals to resolve the dilemma between time spent in employment and on informal care. As the individual becomes eligible for retirement, he or she

may re-assess the care decision. Consequently, there could be a positive causal effect of reaching the age eligible retirement threshold on informal care provision. Because the crossing of the pension eligibility cut-off is solely determined by age, the treatment status is exogenous with respect to the caring decision. To estimate a 2SLS model I propose the following specification:

$$IC_i = x_i\beta + \delta_R R_i + v_i \tag{6}$$

$$R_i = x_i\varphi + PE_i\gamma + n_i \tag{7}$$

Equation (6) shows the second stage, in which the outcome of interest; daily caregiving, IC_i, is a function of the retirement status, R_i, a set of demographic variables and controls, x_i, and the error term, v_i, which is uncorrelated with x_i. Equation (7) represents the first stage, where the endogenous regressor, retirement, R_i, is determined by the instruments, PE_i, the same set of demographics and controls, x_i and the another error term, n_i, with $E[(n_i, PE_i)'] = 0$. Having reached the pension eligibility age for early or statutory retirement, must be correlated with the retirement status, if the instruments are able to strongly predict retirement in the first stage equation (7). In my analysis I am interested in δ_R; the coefficient associated with being retired. The coefficients β and δ_R are identified, if one can assume that the instruments, PE_i, are excluded from the daily caregiving equation (6); in other words, if PE_i are exogenous with respect to the unobserved heterogeneity in the second stage, v_i. For relevant and valid instruments, the predicted retirement probability carries only exogenous variation and is independent of the error term.

4. Data

In the empirical analysis, I use data from the Survey of Health, Ageing and Retirement in Europe (SHARE), which collects information on individuals aged fifty and above in 19 European countries (+ Israel). In my study, I restrict the sample to respondents in the age group 50 to 70 which covers the years just before and after most workers typically retire. This age also coincides with a period when the likelihood of having an elderly parent or parent-in-law still living and in need of long-term care is high. Since most caregivers provide help to elderly parents, I focus on individuals who had at least one living parent or parent-in-law at the time of the interview. I use five waves of the SHARE dataset, wave 1 to 6 (excluding wave 3; the life history survey) with comprehensive information on demographics, employment and pensions, health, and social support covering the period from 2004 to 2015. Moreover, I add information on past and current country-specific eligibility ages for early and statutory retirement. Appendix 2 includes the full list of pension age rules together with literature references.

The total sample of analysis includes 34,277 individuals from 18 European countries. Table 6 in Appendix 1 lists the countries from each of the five waves included in this study. The 34,277 individuals only appear once in the data set with one observation for each individual. The sample selection process was structured with all respondents in wave 6 and then I added respondents from wave 5 that had not already been included as part of wave 6, before adding individuals from wave 4 not already included from wave 6 and wave 5, and before continuing adding respondents from wave 2 and wave 1 in the same fashion.[2] In this way, I collect the most recent information, and each observation uniquely identify one

[2] Another approach would be to simply pool the five waves of SHARE and have repeated observations for the interviewed respondents in multiple waves. Valid inference of the econometric specification could still be conducted clustering standard errors at the individual level.

individual in the sample. The main variables of interest are those that measure the respondents' caregiving activities and their retirement status.

The geographical proximity to parents matters. I restrict the sample of analysis to include individuals who had at least one living parent or parent-in-law residing within 100 km, as a distance above 100 km renders the provision of daily care almost impossible. While smaller distances may lead to more caregiving, caring for a parent or parent-in-law living farther away adds to the caregiving costs in terms of transportation time. With retirement time becomes less costly, also time spent on the road, thereby potentially affecting the caregiving intensity for a parent or parent-in-law living at a longer distance from the home of the adult child.

4.1 Informal Care

Since the caregiving intensity generally seems to matter for results in other studies (Heitmueller, 2007), I focus on caregivers who provided daily care for parents and parents-in-law inside or outside the household (or both). I construct the main dependent variable of the analysis as a binary variable for informal care, taking the value 1, if the respondent gave daily care to a parent or parent-in-law, and 0 otherwise.

In the survey, respondents are asked about informal caregiving both inside and outside the home, and to be consistent with this, I first create two variables. For 'daily caregiving inside the household' respondents are asked whether they in the last twelve months daily (or almost daily) helped someone living in their household with personal care[3], and the individuals then specify the relationship (mother, father, mother-in-law, or father-in-law). All respondents are asked this question if their household included more than just one person. If the respondent provided daily help with personal care to at least

[3]In all five waves, the respondents are asked "Is there someone living in this household whom you have helped regularly during the last twelve months with personal care, such as washing, getting out of bed, or dressing? By regularly we mean daily or almost daily during at least three months."

one parent or parent-in-law inside the household, the dummy variable 'daily care inside the household' is equal to 1, and 0 otherwise.

How the respondents in the survey are asked about informal caregiving for individuals outside the household has changed over time. In wave 1, 2, and 6 help outside the household is explicitly specified for either personal care, practical household help, or help with paperwork.[4] In waves 4 and 5 respondents are simply asked whether they have given personal care or practical household help to someone outside the household.[5] To be consistent across waves, I define informal caregiving outside the household as having given either personal care or practical household help (or both). In wave 1, 2, and 6, all respondents are asked about care outside the household, while for wave 4 and 5, questions about informal care outside the household are only asked to family respondents (73.3 % of the individuals in the sample are family respondents). As with caregiving inside the household, the respondents specify the recipients of their informal care (mother, father, mother-in-law, or father-in-law). Moreover, the respondents are asked about the caregiving intensity outside the household, from daily (or almost daily) to less often than monthly. I construct the dummy variable 'daily care outside the household', taking the value 1, if the respondent provided daily help with personal care or practical household help (or both) to at least one parent or parent-in-law living outside the household, and 0 otherwise.

Table 1 reports summary statistics on informal caregiving by men and women across countries. Column (2), (3), (6), and (7) show the fractions of men and women by country, who provided daily care inside *and* outside the household. The main dependent variable of this study is constructed by the combination of the two caregiving variables; see column (4) and (8), taking the value 1, if the respondent gave

[4] In waves 1, 2, and 6, the respondents are asked: "In the last twelve months, have you personally given any kind of help listed on this card to a family member from outside the household, a friend or neighbor?" If the answer is yes, the respondents are asked: "Which types of help have you given to this person in the last twelve months? (1) personal care, (2) practical household help, (3) help with paperwork."

[5] In waves 4 and 5, the respondents are asked: "In the last twelve months, have you personally given personal care or practical household help to a family member living outside your household, a friend or neighbor?"

'daily care' to a least one parent or parent-in-law inside or outside the household (or both), and 0 otherwise. Column (1) and (5) report the total number of observations by country sorted by European regions for men and women in the sample.

Table 1 Descriptive statistics on informal caregiving variables

	Daily caregiving - **Males**				Daily caregiving - **Females**			
		Inside household	Outside household	Total*		Inside household	Outside household	Total*
	Obs. (1)	Mean (2)	Mean (3)	Mean (4)	Obs. (5)	Mean (6)	Mean (7)	Mean (8)
Northern Europe								
Denmark	1,357	0.002	0.023	0.025	1,328	0.003	0.035	0.038
Sweden	1,083	0.005	0.026	0.030	1,150	0.004	0.029	0.033
Western Europe								
Austria	1,065	0.013	0.038	0.050	1,097	0.041	0.088	0.122
Belgium	2,010	0.004	0.051	0.055	1,777	0.011	0.088	0.098
France	1,555	0.014	0.040	0.051	1,633	0.016	0.078	0.091
Germany	1,376	0.017	0.048	0.060	1,457	0.024	0.109	0.130
Luxembourg	250	0.004	0.016	0.020	237	0.013	0.076	0.089
Netherlands	1,068	0.004	0.021	0.023	1,037	0.006	0.060	0.065
Switzerland	861	0.005	0.017	0.022	897	0.008	0.033	0.039
Southern Europe								
Greece	894	0.011	0.028	0.039	542	0.028	0.111	0.131
Italy	1,345	0.047	0.061	0.103	1,197	0.057	0.122	0.169
Portugal	238	0.050	0.029	0.080	196	0.097	0.071	0.158
Spain	1,029	0.044	0.037	0.077	846	0.073	0.099	0.164
Eastern Europe								
Croatia	254	0.047	0.031	0.075	239	0.075	0.071	0.138
Czech Republic	1,214	0.024	0.061	0.078	1,528	0.033	0.110	0.135
Estonia	669	0.040	0.034	0.073	878	0.067	0.074	0.125
Poland	307	0.029	0.029	0.052	377	0.058	0.069	0.119
Slovenia	596	0.020	0.062	0.074	690	0.043	0.119	0.145
All	17,171	0.018	0.039	0.055	17,106	0.029	0.081	0.105

Source: SHARE 2004 to 2015, ages 50 to 70. *Total is the combination of daily care inside *and* outside the household, and it constructs the main dependent variable of this study; the dummy variable *daily caregiving*, taking the value 1 for respondents that gave daily (or almost daily) care to a least one parent or parent-in-law inside or outside the household (or both), and 0 otherwise.

Across countries, 2,733 of the 34,277 individuals in the sample provided daily care (8.0 %). The average participation rates for daily caregiving were 5.5 % for men and 10.5 % for women. Table 1 shows a great deal of heterogeneity across the European regions. Individuals in the Scandinavian countries, the Netherlands, and Switzerland provided less informal care to parents and parents-in-law compared to individuals in Southern and Eastern European countries, such as Italy, Portugal, Spain, Croatia, and Slovenia, that showed much higher average rates of daily informal care. Looking at all countries individually, women gave daily care to parents and parents-in-law at higher average rates than men, without exception. In some countries, the gender gap in daily caregiving is large, for example, Greece (9.2 %), Spain (8.8 %), and Portugal (7.8 %), as opposed to less than 2 % for individuals in Sweden, Denmark, and the Netherlands[6].

Most of the caregivers (75.5 %) provided daily care outside their household, while 29.2 % gave daily care to someone within their household, and a small percentage (4.7 %) of the individuals were daily caregivers both inside and outside their household. Other studies distinguish between the two types of care; extra-residential and co-residential care, arguing that caregivers living with the recipient of care are likely to provide different types of care compared to those providing care to parents or parents-in-law living outside the household (Michaud et al., 2010). *Daily care* being the combination of inside and/or outside daily caregiving, is important because of the heterogeneity in the type of care across countries, and an analysis only looking at one or the other would be limiting. Respondents in Northern and Western Europe provide informal care mostly outside of the household, while individuals in Southern and Eastern Europe more equally provide caregiving inside and outside the household, and for caregivers in Croatia and Portugal the higher rates of daily care are inside the household.

[6]Bear in mind that the actual number of observations is small.

4.2 Retirement

I construct the key variable 'retired' based on survey questions about the retirement status, filtering out individuals, who despite self-declaring to be retired, did paid work during the last four weeks before the interview. Retirement is considered an absorbing state with no transitions from retirement back into employment. I restrict my sample to individuals with previous work experience and exclude individuals who were unemployed, permanently sick or disabled, or homemakers at the time of the interview. One could argue that the same constraints are not removed through the transition into retirement, if the person is unemployed prior to retirement, and under such other circumstances I would not expect the same impact of retirement on the supply of informal care. Likewise, disabled individuals may face different choice sets with respect to retirement and care provision. As Van Houtven et al. (2013), I use the respondents' self-reported retirement statuses, without checking whether the individuals received a public pension or had reached an institutionally relevant pension age at the time of retirement. I also exclude individuals that retired from work before age 40. 26.9 % of the individuals in the total sample are in retirement. Appendix 2 includes more details and the full list of information sources on country-specific pension age eligibility criteria.

Table 2 below shows summary statistics on the retirement variables by country, reporting the mean age of the respondent, and the fraction of retired individuals in the sample, along with the mean pension eligibility ages for early and statutory retirement for men and women, respectively. The pension eligibility ages are reported as intervals here, as the sample includes respondents from five survey waves covering the period 2004 to 2015. In figures 1 and 2 in Appendix 1, I present histograms of the respondents' reported ages at actual retirement by country, respectively for the retired men and women with the institutional changes in early and statutory pension eligibility ages.

Table 2 Summary statistics on retirement variables by country

	Males	Age	Retired	Early ret. age	Statutory ret. age	***Females***	Age	Retired	Early ret. age	Statutory ret. age
	Obs.	Mean	%			Obs.	Mean	%		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Northern Europe										
Denmark	1,357	58	0.136	60-62	65-67	1,328	57	0.130	60-62	65-67
Sweden	1,083	59	0.153	60-61	65-67	1,150	58	0.160	60-61	65-67
Western Europe										
Austria	1,065	59	0.456	60-64	65	1,097	58	0.480	55-59	60
Belgium	2,010	57	0.283	60-61	65	1,777	57	0.239	55-61	60-65
France	1,555	59	0.391	55	60-62	1,633	58	0.325	55	60-62
Germany	1376	58	0.216	60-63	65	1,457	57	0.158	60-63	65
Luxembourg	250	58	0.420	57	65	237	57	0.274	57	65
Netherlands	1068	58	0.162	60	65	1037	57	0.106	60	65
Switzerland	861	59	0.106	63-64	65	897	57	0.094	62	62-64
Southern Europe										
Greece	894	58	0.265	60-62	62-67	542	56	0.284	55-62	57-67
Italy	1,345	59	0.354	52-63	60-66	1,197	58	0.308	52-63	56-66
Portugal	238	60	0.508	55	65-66	196	58	0.327	55, 60	62-66
Spain	1,029	59	0.276	60-63	65	846	57	0.144	60-63	65
Eastern Europe										
Croatia	254	59	0.472	57-60	60-65	239	57	0.444	52-56	55-60
Czech Republic	1,214	59	0.329	57-59	60-62	1,528	58	0.396	53-58	56-61
Estonia	669	58	0.184	60	63	878	57	0.133	55-58	58-61
Poland	307	58	0.270	60	65	377	57	0.432	55	60
Slovenia	596	60	0.497	55-58	58-65	690	59	0.548	50-58	53-63
All	17,171	58	0.281			17,106	57	0.257		

Source: SHARE 2004 to 2015, ages 50 to 70. Gruber and Wise (2002-2018): Information on pension eligibility ages for early and statutory retirement. See Appendix 2 for full list of sources.

Column (2) and (7) in Table 2 report that the mean age in the sample is 58 for men and 57 for women, with little variation across countries. The proportion of retired individuals in the overall sample is 28.1 % for men and 25.7 % for women (who on average are one year younger); see column (3) and (8), and it greatly varies by both country and gender. Column (4), (5), (9), and (10) show the intervals of age thresholds for early and statutory retirement which were and are applicable for the individuals in the sample. The share of retired individuals is significantly higher among caregivers (38.1 %) than non-caregivers (25.9 %).

4.3 Control variables

Standard demographic variables such as age (at the time of the interview), age^2, marital status (here 'having a partner'), children, level of education, household size, and being the family respondent are used as controls in the standard, baseline model. The binary variable 'having a partner' takes the value 1, if the respondent is married or in a registered partnership; and 0 otherwise, which is true for 81.0 % of the individuals in the sample. There is a significant difference in the marital status between caregivers (74.3 %) and individuals who did not provide care (81.6 %) (MWW, $p<0.001$). 4.9 % of men who have a partner as opposed to 8.7 % of the men without a partner provided daily care to a parent or parent-in-law, while the participation rates are 10.0 % for women with a partner and 12.2 % for women without a partner.

The measure 'children' includes all natural children, fostered, adopted, stepchildren, and the partner's children; 91.2 % of the respondents have children. I control for the number of household members with 'household size'; the mean being 2.5, and in the regression analysis, I also add a dummy variable for being the family respondent. The level of education is included with education level dummies (1-3; with 3 being the highest level) based on the 1997 International Standard Classification of Education (ISCED), which is a statistical standard coding of education for international and national comparisons.[7]

In the extended models, I add measures for physical and mental health along with siblings and parental health measures. The respondent's measure of mental health is the EURO-D caseness variable (eurodcat), which is generated from questions in the mental health module of the SHARE questionnaire (Börsch-Supan et al., 2005). 19.3 % of the respondents in the sample (12.8 % of the men and 25.8 % of

[7] Education level 1 comprises early childhood education, primary education, and lower secondary education. Education level 2 includes upper secondary and post-secondary education, while education level 3 signifies short-cycle tertiary education and bachelor or the equivalent level of education.

the women) self-identified as having some depressive symptoms. Measures of physical health are based on self-reports on limitations with 'activities of daily living' (ADL; such as dressing, walking, getting in and out of bed, and eating) and 'instrumental activities of daily living' (IADL; such as preparing a meal, shopping for groceries, and making telephone calls). Difficulties with one or more of these activities indicate a lower mobility and physical functioning of the respondent, which is the case for 3.7 % (ADL) and 5.5 % (IADL) of the individuals in the sample. The health statuses of parents or parents-in-law seem to be highly correlated with the caring decision, and they are therefore included in the analysis as dummy variables, being equal to 1, if the respondent has a parent or parent-in-law in poor health; and 0 otherwise. I also include a binary variable for having siblings. All regressions include wave and country dummies to capture time trends and variation across countries.

Comparing the sample characteristics of caregivers versus non-caregivers and testing simple mean comparisons, I find that on average caregivers in the sample are less likely to be 'having a partner,' and they are less likely to have children. The respondents, who provided daily care, are older (on average almost one year). Though I find no significant difference between the fraction of caregivers and non-caregivers in poor physical health, the respondents providing daily care were much more likely to have depressive symptoms (27.1 % vs. 18.6 %; MWW, $p<0.001$). Caregivers were less likely to have siblings (MWW, $p<0.001$), and naturally much more likely to have a parent or parent-in-law in poor health (MWW, $p<0.001$).

5 Results

The aim of this study is to examine the impact of retirement on informal caregiving, and the regression models in Table 3 below show the effect of retirement on daily caregiving for parents and parents-in-law. Through the standard ordinary least squares (OLS) regression models (1)-(3) and 2SLS regression models (4)-(6), I gradually add extra control variables.[8] The standard errors are clustered on a combination of wave, gender, and country, and the number of clusters is equal to 142.

Table 3 OLS and 2SLS regressions, the effect of retirement on daily caregiving

	(1) OLS	(2) OLS	(3) OLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
	Dependent variable: Daily care					
Retired	0.025***	0.025***	0.024***	0.065***	0.066***	0.065***
	(0.005)	(0.005)	(0.005)	(0.018)	(0.019)	(0.019)
Female	0.054***	0.052***	0.052***	0.053***	0.051***	0.051***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Age	0.015**	0.015**	0.013*	0.026***	0.027***	0.025***
	(0.008)	(0.008)	(0.007)	(0.010)	(0.010)	(0.010)
Age²	-0.010	-0.010	-0.009	-0.022**	-0.022**	-0.021**
	(0.006)	(0.006)	(0.006)	(0.009)	(0.009)	(0.009)
Having a partner	-0.033***	-0.032***	-0.033***	-0.034***	-0.033***	-0.034***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Children	-0.043***	-0.043***	-0.042***	-0.043***	-0.043***	-0.043***
	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)
Education dummies: Education level 3 (highest level) is baseline						
Education level 1	0.006	0.006	0.008**	0.003	0.003	0.005
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Education level 2	0.006	0.006	0.008**	0.004	0.004	0.006
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Household size	0.017***	0.017***	0.017***	0.018***	0.018***	0.018***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Family respondent	0.032***	0.032***	0.032***	0.031***	0.031***	0.032***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Depressed		0.026***	0.021***		0.025***	0.020***
		(0.004)	(0.004)		(0.004)	(0.004)
ADL limitations		-0.003	-0.005		-0.005	-0.007
		(0.008)	(0.008)		(0.008)	(0.008)

[8]The 2SLS and LIML regression estimates are practically almost of identical size, and the LIML estimates are therefore not reported here in Table 3.

IADL limitations		-0.022***	-0.025***		-0.026***	-0.029***
		(0.008)	(0.008)		(0.008)	(0.008)
Siblings			-0.038***			-0.038***
			(0.006)			(0.005)
Parent in poor health			0.073***			0.073***
			(0.006)			(0.006)
Wave dummies: Wave 6 is baseline						
Wave 1	0.016**	0.016***	0.027***	0.010	0.011	0.021***
	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)
Wave 2	0.003	0.004	0.004	0.000	0.001	0.001
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Wave 4	-0.007	-0.007	-0.008	-0.009*	-0.009*	-0.010*
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Wave 5	0.004	0.004	0.003	0.003	0.003	0.002
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Country dummies: Denmark is baseline						
Sweden	-0.008	-0.009	-0.009	-0.005	-0.005	-0.005
	(0.010)	(0.010)	(0.010)	(0.011)	(0.011)	(0.010)
Austria	0.039***	0.039***	0.040***	0.029***	0.029***	0.030***
	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)
Belgium	0.034***	0.033***	0.038***	0.030***	0.028***	0.033***
	(0.007)	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)
France	0.029***	0.025***	0.022***	0.024***	0.020**	0.017**
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Germany	0.057***	0.056***	0.053***	0.056***	0.055***	0.052***
	(0.009)	(0.008)	(0.008)	(0.009)	(0.009)	(0.008)
Luxemburg	-0.001	-0.004	-0.005	-0.010	-0.012	-0.013
	(0.008)	(0.008)	(0.008)	(0.010)	(0.010)	(0.010)
Netherlands	0.006	0.006	0.012*	0.007	0.007	0.012*
	(0.007)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)
Switzerland	-0.012	-0.013	-0.008	-0.009	-0.010	-0.005
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Greece	0.032**	0.030**	0.034***	0.027**	0.025**	0.029**
	(0.013)	(0.013)	(0.012)	(0.013)	(0.012)	(0.012)
Italy	0.084***	0.081***	0.079***	0.079***	0.076***	0.074***
	(0.009)	(0.008)	(0.008)	(0.009)	(0.008)	(0.008)
Portugal	0.076***	0.072***	0.061***	0.068***	0.064***	0.054***
	(0.010)	(0.010)	(0.009)	(0.011)	(0.011)	(0.010)
Spain	0.070***	0.070***	0.069***	0.070***	0.069***	0.068***
	(0.010)	(0.010)	(0.010)	(0.011)	(0.010)	(0.011)
Croatia	0.063***	0.061***	0.060***	0.051***	0.048***	0.048***
	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)
Czech Republic	0.064***	0.063***	0.060***	0.058***	0.057***	0.054***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Estonia	0.064***	0.062***	0.055***	0.064***	0.062***	0.055***

	(0.010)	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)
Poland	0.040**	0.036**	0.035**	0.032**	0.028*	0.027
	(0.016)	(0.016)	(0.016)	(0.016)	(0.017)	(0.017)
Slovenia	0.060***	0.060***	0.062***	0.048***	0.047***	0.049***
	(0.011)	(0.011)	(0.011)	(0.012)	(0.012)	(0.012)
Observations	34,277	34,277	34,277	34,277	34,277	34,277
R-squared	0.037	0.038	0.052	0.035	0.036	0.050
First stage - Dependent variable: Retired						
EligibleER				0.141***	0.141***	0.141***
				(0.022)	(0.022)	(0.022)
EligibleSR				0.281***	0.279***	0.279***
				(0.033)	(0.033)	(0.006)
F-test statistic[♂]				77.22	76.89	76.95
Sargan-Hansen J-test[♀]				0.405	0.279	0.090
				(0.524)	(0.597)	(0.764)
Hausman test[☿]				4.771**	4.843**	76.95**
				(0.029)	(0.027)	(0.026)

Source: SHARE 2004 to 2015, ages 50 to 70. Model (1)-(3) include OLS regression estimates, while model (4)-(6) include IV 2SLS regression estimates. Standard errors in parentheses are clustered on wave, gender, and country. ♂ The F-test of excluded instruments tests the strength of the joint instruments. ♀ The Sargan-Hansen J-statistic reports the overidentification test of all instruments. ☿ The Hausman statistic tests the endogeneity of the endogenous regressor, retired. *** p<0.01, ** p<0.05, * p<0.1.

The linear regression model of interest is estimated both by OLS, maintaining the exogeneity assumption that retirement is exogenous with respect to the caregiving decision, and by 2SLS, allowing for the potential endogeneity of retirement (simultaneity and omitted variable bias). In the OLS regression model in column (1), retirement increases the likelihood of daily caregiving by 2.5 %-points (p<0.001). For women, the probability of providing informal care is 5.4 %-points higher than for men. Both 'having a partner' (-3.3 %-points) and having children (-4.3 %-points) decrease the likelihood of giving daily care, while a larger household size (1.7 %-points) and being the family respondent (3.2 %-points) have a positive impact on daily caregiving. There is no significant effect for the education level dummies. Model (2) adds control variables of the individual's mental and physical health. Being depressed (2.6 %-points) increases the likelihood of informal caregiving, while the physical health measure, IADL (limitations with instrumental activities of daily living), decreases the probability of daily caregiving by 2.2

%-points (the effect of ADL is not significant). In model (3), having a parent in poor health (7.3 %-points) makes it more likely that the individual engages in daily care, while having siblings decreases the probability of being a daily caregiver by 3.8 %-points.

The OLS estimates of the coefficient on retired yield unbiased results, if the orthogonality condition is satisfied; i.e. the retirement status must be uncorrelated with the error term. However, this seems unlikely to hold as pointed out by the literature above, and concerns about endogeneity must be considered. In the 2SLS model (4), the instruments, EligibleER and EligibleSR, are strong to predict retirement in the first-stage equation (The F-statistic of excluded instruments is 77.22, $p<0.001$). The coefficients of the instrumental variables are positively correlated with the retirement status, being eligible for statutory retirement (28.1 %) more so than being eligible for early retirement (14.1 %), as could be expected. With one explanatory variable allowed to be endogenous and two exclusion restrictions, the model is overidentified. The Sargan-Hansen J-statistic reported in Table 3 tests the overidentification of all instruments. As the null hypothesis cannot be rejected (p-value=0.524), the overidentification restriction is valid. The Hausman test checks whether the endogenous regressor is in fact exogenous, and as the test statistic is significant (p-value=0.029), the variable (retired) being tested must be treated as endogenous.

The second stage estimate for 'retired' reported in column (4) implies that retirement significantly increases the probability of daily caregiving to parents and parents-in-law by 6.5 %-points. For women compared to men, the likelihood of caregiving is 5.3 %-points higher. 'Having a partner' decreases the probability of engaging in daily care by 3.4 %-points, while having children decreases the likelihood by 4.3 %-points. Being older makes caregiving more likely (+2.6 %-points). The coefficient estimate on household size is positive and significant but small (+1.8 %-points), meaning that a larger household increases the probability of having to care for a parent or parent-in-law. Being a family respondent increa-

ses the probability of being a caregiver by 3.1 %-points. The effect of the education level dummies on daily caregiving is small and insignificant, but in the first stage when determining the retirement status, the coefficients on education level 1 (+8.2 %-points) and education level 2 (+5.0 %-points) are positive and significant at a 1 %-points-level, meaning that a lower education level increases the likelihood of being retired.

Model (5) adds measures of mental and physical health to the specified regression model. Being depressed has a significantly positive effect on the likelihood of informal caregiving (+2.5 %-points). In contrast, IADL (limitations with instrumental activities of daily living) which is the physical health measure, decreases the probability of daily caregiving by 2.6 %-points. In model (6), adding having a parent in poor health, increases the likelihood of engaging in daily care by 7.3 %-points. The dummy variable of ever having had siblings decreases the likelihood of daily caregiving by 3.8 %-points.

Rates of informal care provision differ greatly across the European regions, and this perhaps helps explain the mixed effect of retirement on daily caregiving across regions. Using Denmark as baseline for comparisons among countries, I find in model (1) that individuals in Austria (+3.9 %-points), Belgium (+3.4 %-points), France (+2.9 %-points), and Germany (+5.7 %-points) are more likely to provide daily care for parents and parents-in-law. Individuals from Southern Europe, especially Italy (+8.4 %-points), Portugal (+7.6 %-points), and Spain (+7.0 %-points), are even more likely to give daily care to parents or parents-in-law, and so are respondents in Eastern Europe; for example, Croatia (+6.3 %-points), Czech Republic (+6.4 %-points), Estonia (+6.4 %-points), and Slovenia (+6.0 %-points). As more control variables are added, for example, measures on parental health, this heterogeneity diminishes, and for some, the country dummy effect decreases. Looking more broadly at the European regions, using the Scandinavian countries as baseline[9], I find that respondents in the country group Western Europe are 3.0

[9] Here Denmark and Sweden without The Netherlands.

%-points more likely to be caregivers, while for individuals in Southern Europe it is 6.2 %-points, and for Eastern Europe it is 5.9 %-points.[10]

In Table 4, I add the measure of having a parent-in-law in poor health to the above models, and since this information is more limited, it lowers the total the number of observations from 34,277 to 28,803. In the OLS model (1) and 2SLS regression model (4), I start with the fully specified models from Table 3 but with fewer observations.[11] Model (2) and (5) include the control 'parent-in-law in poor health' instead of 'parent in poor health', while model (4) and (6) use a combination of both measures of parental health.

Table 4 OLS and 2SLS regressions, the effect of retirement on daily caregiving with parent-in-law in poor health

	(1) OLS	(2) OLS	(3) OLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
	Dependent variable: Daily care					
Retired	0.021***	0.021***	0.021***	0.065***	0.064***	0.065***
	(0.005)	(0.005)	(0.005)	(0.017)	(0.017)	(0.017)
Female	0.052***	0.052***	0.052***	0.050***	0.050***	0.050***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Age	0.013*	0.013*	0.013*	0.025***	0.025***	0.025***
	(0.007)	(0.007)	(0.007)	(0.009)	(0.009)	(0.009)
Age²	-0.009	-0.009	-0.009	-0.021***	-0.021***	-0.022***
	(0.006)	(0.006)	(0.006)	(0.008)	(0.008)	(0.008)
Having a partner	0.010	0.010	0.010	0.009	0.009	0.009
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Children	-0.015*	-0.014*	-0.015*	-0.015**	-0.015*	-0.015**
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Education dummies: Education level 3 (highest level) is baseline						
Education level 1	-0.000	0.000	-0.000	-0.003	-0.003	-0.004
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Education level 2	0.006	0.006	0.006	0.004	0.004	0.004
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)

[10] Looking only at the women in the sample and keeping respondents in the Scandinavian countries as baseline, I find that women in Western Europe are 5.0 %-points more likely to take on the caregiver role for a parent or parent-in-law, for respondents in Southern Europe it is 9.6 %-points, and 8.2 %-points for respondents in Eastern European countries.

[11] In the models in Table 4, I include wave dummies and country dummies as in Table 3, but as the coefficient estimates are near identical in both size and significance, I do not report them in Table 4.

Household size	0.016***	0.016***	0.016***	0.016***	0.017***	0.016***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Family respondent	0.029***	0.030***	0.028***	0.028***	0.030***	0.027***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Depressed	0.021***	0.024***	0.021***	0.020***	0.023***	0.020***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
ADL limitations	-0.006	-0.004	-0.006	-0.008	-0.007	-0.009
	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.008)
IADL limitations	-0.020**	-0.019**	-0.021**	-0.024***	-0.023**	-0.024***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Siblings	-0.032***	-0.032***	-0.032***	-0.032***	-0.032***	-0.031***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Parent in poor health	0.066***	-	0.084***	0.066***	-	0.084***
	(0.006)		(0.010)	(0.006)		(0.010)
Parent-in-law in poor health		0.052***	-0.024***		0.052***	-0.024***
		(0.006)	(0.009)		(0.006)	(0.009)
Wave dummies	x	x	x	x	x	x
Country dummies	x	x	x	x	x	x
Observations	28,803	28,803	28,803	28,803	28,803	28,803
R-squared	0.045	0.040	0.046	0.043	0.037	0.043
	First stage - Dependent variable: Retired					
EligibleER				0.149***	0.149***	0.149***
				(0.022)	(0.022)	(0.022)
EligibleSR				0.280***	0.280***	0.280***
				(0.034)	(0.034)	(0.034)
F-test statistic[σ]				77.30	77.32	77.30
Sargan-Hansen J-test[φ]				0.353	0.503	0.335
				(0.553)	(0.478)	(0.563)
Hausman test[ψ]				6.967***	6.774***	7.031***
				(0.008)	(0.009)	(0.008)

Source: SHARE 2004 to 2015, ages 50 to 70. Model (1)-(3) include OLS regression estimates, while model (4)-(6) include IV 2SLS regression estimates. Standard errors in parentheses are clustered on wave, gender, and country. For the first stage equation, I only report estimates for the two instrumental variables used. [σ] The F-test of excluded instruments tests the strength of the joint instruments. [φ] The Sargan-Hansen J-statistic reports the overidentification test of all instruments. [ψ] The Hausman statistic tests the endogeneity of the endogenous regressor, retired. *** p<0.01, ** p<0.05, * p<0.1.

In model (1)-(3) retirement increases the likelihood of daily caregiving by 2.1 %-points. With the lower number of observations, the estimated coefficients are in general a little bit smaller and less significant. In the way the data sample is constructed with information about the health status of a parents-in-law, the respondent in the sample needs to have a partner (also in the sample) or be in good contact with

the parent of an ex-partner, to able to have an observation here. 94.5 % of the individuals in this smaller sample 'have a partner', while in the original sample used in Table 3 it is 81.0 %. This can help explain why the effect of 'having a partner' on daily care becomes lesser in both size and significance. Compared to before, the percentage of respondents who have children is larger, and similarly the effect of children on daily caregiving is smaller and less significant. Some average characteristics of the individuals in the sample have simply changed, and this is my motivation for reporting the regression models that include the control "parent-in-law in poor health" in a separate table.

In model (4) having a parent in poor health increases daily caregiving by 6.6 %-points, while in model (5) having a parent-in-law in poor health increases caregiving by 5.2 %-points. In model (6) with both measures of parental health, the effect of retirement of daily caregiving to 6.5 %-points. The likelihood of being a caregiver increases with having a parent in poor health by 8.4 %-points, but decreases with a having a parent-in-law in poor health by 2.4 %-points. Are the care needs of a parent and a parent-in-law in conflict with each other? Is it the partner who takes care of the parent-in-law, if the respondent's own parent also needs informal care? 7,788 individuals of the 34,277 individuals in Table 3 have a parent in poor health. The individuals with a parent-in-law in poor health all also have a parent in poor health. Of the 28,803 respondents with information about the health of a parent-in-law, 6,590 individuals have a parent in poor health, and among them 4,762 also have a parent-in-law in poor health. Overall, 2,361 individuals (86.4 % of the caregivers) care daily for a parent, while 541 individuals (19.8 %) take care of a parent-in-law, and 169 of the individuals (6.2 %) provide care for both a parent and a parent-in-law.

6. Robustness checks

Among the 18 countries in the analysis, there are great differences both in terms of the participation rates for informal care, the availability of formal care provided by state, and the percentage of retired individuals in the sample. To test that the estimation results are geographically robust, I examine whether the regression coefficients estimated for one country region group of the data are equal to those of the other groups. With a structural stability test (Chow test), I can check whether the data can be pooled together, supposing that all groups have the same slope and intercept. For the OLS regression, I assume that $y = X\beta + \varepsilon$ for all observations, and I split the sample in three groups, fitting separate regressions for each country region group.[12]

$y_1 = X_1\beta_1 + \varepsilon_1$ for group 1 (Central European countries)

$y_2 = X_2\beta_2 + \varepsilon_2$ for group 2 (Denmark, Sweden, and the Netherlands)

$y_3 = X_3\beta_3 + \varepsilon_3$ for group 3 (Southern European countries)

To test $\beta_1 = \beta_2 = \beta_3$, I convert the multiple equations into one big regression equation, omitting the overall intercept as the separate models each have their own intercept,

$$y = d_1 * (X_1\beta_1 + \varepsilon_1) + d_2 * (X_2\beta_2 + \varepsilon_2) + d_3 * (X_3\beta_3 + \varepsilon_3)$$

$$= d_1 * X_1\beta_1 + d_2 * X_2\beta_2 + d_3 * X_3\beta_3 + d_1 * \varepsilon_1 + d_2 * \varepsilon_2 + d_3 * \varepsilon_3$$

and I test whether the coefficients are the same for the three country groups with $H_0: \beta_1 = \beta_2 = \beta_3$. d_1, d_2, and d_3 are the country dummies for the three country groups.

[12] In regional comparisons, the literature often groups the Netherlands with the Scandinavian countries because of similar public care institutions and the same level of social support (Ciccarelli and Van Soest, 2018). Group 1: Central European countries include Austria, Belgium, Czech Republic, Estonia, France, Germany, Luxembourg, Switzerland, and Poland. Group 2: Scandinavia and the Netherlands include Denmark, the Netherlands, and Sweden. Group 3: Southern European countries include Croatia, Greece, Italy, Portugal, Slovenia, and Spain.

In the Chow test, I also include a test for equality of the individual intercepts. I add country group interactions with the full number of covariates, education, wave, and country dummies, and I run a combined test of structural stability for the models in Table 3. For the fully interacted model (3) with the *F*-test statistic (67, 141) = 65.63, p<0.001, I reject the null hypothesis of equal coefficients. The estimated parameters of the model for the region subsamples are statistically different from each other.

For the instrumental variable model, the same principle applies when comparing the coefficients between the subsamples. As before, combining the separate models into one, I write

$$y = d_1 * X_1\beta_1 + d_2 * X_2\beta_2 + d_3 * X_3\beta_3 + d_1 * \varepsilon_1 + d_2 * \varepsilon_2 + d_3 * \varepsilon_3$$

$X_1 = Z_1\pi_1 + \vartheta_1$ for group 1

$X_2 = Z_2\pi_2 + \vartheta_2$ for group 2

$X_3 = Z_3\pi_3 + \vartheta_3$ for group 3

where the region group dummies are interacted with X and Z, so that

$X = (d_1 * X_1, d_2 * X_2, d_3 * X_3)$ and $Z = (d_1 * Z_1, d_2 * Z_2, d_3 * Z_3)$

which allow me to test the null hypothesis $H_0: \beta_1 = \beta_2 = \beta_3$.

For the 2SLS model (6) with the test statistic, $\chi^2_{(71)} = 4054.51$, p<0.001, the joint null hypothesis of equal coefficients across subgroups is rejected. Running three regressions instead of one may likely provide a better fit. However, when dividing the sample in three (or possibly four), I run into problems with small sample sizes, and because of the small number of individuals, who perform intensive informal caregiving, the effect of retirement on daily care will be difficult to detect in individual regressions. Adding the interactions with the many covariates, wave dummies, and country dummies, I end up with very high values for the Chow test, but the exercise of writing up and running the combined model of the individual

regression equations does provide some understanding of the regional differences and dissimilar coefficient estimates.

Table 5 reports the coefficient estimates for the combined, fully specified model with the three individual equations, both for the OLS models (1)-(3) and the 2SLS models (4)-(6) from Table 3. I only include the most important variables and covariates of interest in Table 5, since the fully specified model is very long, but Table 7 in Appendix 1 shows the combined total, and much more detailed, models (1)-(6).

Table 5 The fully specified OLS models (1)-(3) and 2SLS models (4)-(6) with Chow test

	(1) OLS	(2) OLS	(3) OLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
	Dependent variable: Daily care					
*Retired * Central Europe*	0.015**	0.015**	0.014**	0.055**	0.056**	0.053**
	(0.007)	(0.007)	(0.007)	(0.025)	(0.025)	(0.025)
*Retired * Scandinavia and Netherlands*	0.033***	0.034***	0.034***	0.159**	0.163**	0.164**
	(0.011)	(0.011)	(0.011)	(0.065)	(0.066)	(0.067)
*Retired * Southern Europe*	0.038***	0.037***	0.032***	0.052	0.047	0.052
	(0.010)	(0.010)	(0.010)	(0.039)	(0.039)	(0.039)
*Female * Central Europe*	0.054***	0.052***	0.052***	0.053***	0.051***	0.051***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)
*Female * Scandinavia and Netherlands*	0.020***	0.020***	0.020***	0.017***	0.017***	0.018***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
*Female * Southern Europe*	0.084***	0.078***	0.078***	0.083***	0.078***	0.077***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
*Having a partner * Central Europe*	-0.034***	-0.033***	-0.033***	-0.035***	-0.034***	-0.034***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
*Having a partner * Scandinavia and Netherlands*	-0.001	-0.001	-0.002	-0.003	-0.002	-0.004
	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)
*Having a partner * Southern Europe*	-0.060***	-0.058***	-0.056***	-0.060***	-0.058***	-0.056***
	(0.013)	(0.013)	(0.012)	(0.013)	(0.013)	(0.012)
*Children * Central Europe*	-0.039***	-0.039***	-0.040***	-0.039***	-0.039***	-0.039***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
*Children * Scandinavia and Netherlands*	-0.017*	-0.017	-0.016	-0.015	-0.016	-0.015
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
*Children * Southern Europe*	-0.061***	-0.063***	-0.062***	-0.062***	-0.063***	-0.063***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
*Depressed * Central Europe*		0.022***	0.018***		0.021***	0.017***
		(0.005)	(0.005)		(0.005)	(0.005)
*Depressed * Scandinavia and Netherlands*		0.011	0.010		0.009	0.007
		(0.007)	(0.007)		(0.007)	(0.007)

*Depressed * Southern Europe*		0.043***	0.035***		0.043***	0.034***
		(0.010)	(0.010)		(0.010)	(0.010)
*IADL limitations * Central Europe*		-0.016	-0.018*		-0.020*	-0.022**
		(0.011)	(0.011)		(0.010)	(0.010)
*IADL limitations * Scandinavia and Netherlands*		-0.015	-0.017		-0.031**	-0.033**
		(0.011)	(0.011)		(0.014)	(0.014)
*IADL limitations * Southern Europe*		-0.044*	-0.046*		-0.045*	-0.047**
		(0.024)	(0.024)		(0.024)	(0.024)
*Siblings * Central Europe*			-0.033***			-0.033***
			(0.007)			(0.007)
*Siblings * Scandinavia and Netherlands*			-0.027***			-0.026***
			(0.008)			(0.009)
*Siblings * Southern Europe*			-0.054***			-0.054***
			(0.013)			(0.012)
*Parent in poor health * Central Europe*			0.064***			0.064***
			(0.007)			(0.007)
*Parent in poor health * Scandinavia and Netherlands*			0.035***			0.035***
			(0.009)			(0.009)
*Parent in poor health * Southern Europe*			0.125***			0.124***
			(0.013)			(0.013)
Interactions with education dummies	x	x	x	x	x	x
Interactions with wave dummies	x	x	x	x	x	x
Interactions with country dummies	x	x	x	x	x	x
Interactions with other control variables¤	x	x	x	x	x	x
Observations	34,277	34,277	34,277	34,277	34,277	34,277
R-squared	0.117	0.119	0.133	0.114	0.115	0.130
	First stage - Dependent variable: Retired					
*EligibleER * Central Europe*				0.095***	0.095***	0.095***
				(0.037)	(0.037)	(0.037)
*EligibleER * Scandinavia and Netherlands*				-0.017	-0.016	-0.016
				(0.034)	(0.034)	(0.034)
*EligibleER * Southern Europe*				0.254***	0.254***	0.254***
				(0.058)	(0.058)	(0.058)
*EligibleSR * Central Europe*				0.317***	0.315***	0.315***
				(0.038)	(0.038)	(0.038)
*EligibleSR * Scandinavia and Netherlands*				0.333***	0.327***	0.327***
				(0.060)	(0.060)	(0.060)
*EligibleSR * Southern Europe*				0.199**	0.195**	0.195**
				(0.086)	(0.086)	(0.086)
Chow-test	57.32	63.49	65.63	2778.04	3789.2	4054.51

Source: SHARE 2004 to 2015, ages 50 to 70. Standard errors in parentheses are clustered on wave, gender, and country. *** p<0.01, ** p<0.05, * p<0.1. For the first stage equation, I only report estimates for the two instrumental variables used. ¤Other control variables include *Age, age², household size, family respondent,* and *ADL limitations*. See Table 7 in Appendix 1 for the fully detailed specified models (1)-(6).

Focusing solely on the estimates reported in the instrumental variable models in Table 5, the effect of retirement seems stronger for individuals in Scandinavia and the Netherlands, where in model (6)

retirement increases the likelihood of daily care by 16.4 %-points. This is much larger than for individuals in both Central Europe (+5.3 %-points) or Southern Europe (where the effect is insignificant, but +5.2 %-points). Labor market participation puts time constraints on caregiving activities, and the time conflict between labor supply and informal care indeed seems to be more restricting for individuals in Scandinavia and the Netherlands. An explanation could be that once retired, the moral obligation to provide daily care for a parent or parent-in-law is stronger for individuals in Scandinavia and the Netherlands, while for individuals in Central Europe and especially Southern Europe it is always inherently there, retired or not retired. Being female significantly increases the likelihood of daily caregiving in model (6), but the effect is heterogenous across regions and largest for individuals in Southern Europe (+7.7 %-points), smallest in Scandinavia and the Netherlands (+1.8 %-points), and somewhere in between for individuals in Central Europe (+5.1 %-points).

In model (6), 'having a partner' decreases the likelihood of daily caregiving by 5.6 %-points for respondents in Southern Europe and by 3.4 %-points for respondents in Central Europe. The negative effect of having children on caregiving is largest in Southern Europe (-6.3 %-points) and -3.9 %-points for individuals in Central Europe. For both 'having a partner' (-0.4 %-points) and children (-1.5 %-points), the effect on daily caregiving is insignificant among individuals in Scandinavia and the Netherlands. Same pattern is observed for poor mental health; being depressed increases daily caregiving by 3.4 %-points for individuals in Southern Europe and by 1.7 %-points for individuals in Central Europe. IADL limitations decreases the likelihood of engaging in daily care by 4.7 %-points in Southern Europe, by 3.3 %-points for individuals in Scandinavia and the Netherlands, and by 2.2 %-points for individuals in Central Europe.

Except for retired and IADL, I observe a similar pattern when looking at coefficients for having siblings and a parent in poor health. The impact of these controls seems heightened for individuals in Southern Europe. Ever had siblings has a negative impact on daily caregiving; for respondents in Sou-

thern Europe the effect is -5.4 %-points, in Central Europe it is -3.2 %-points and in Scandinavia and the Netherlands it is 2.6 %-points. Having a parent in poor health increases the likelihood of daily caregiving by 12.4 %-points for individuals in Southern Europe, by 6.4 %-points in Central Europe, and by 3.5 %-points for individuals in Scandinavia and the Netherlands. The rest of the coefficient estimates for the interactions with the country region groups in Table 5 (and in the full models in Table 7 in Appendix 1) largely mirror and emphasize the findings communicated in this paragraph.

7. Conclusion

A flurry of pension reforms across Europe have led to a remarkable increase in the labor force participation among men and women at the end of their working lives. This has boosted the sustainability of public pension systems, but its consequences for the lives of this group of individuals and their relatives need careful investigation. In particular, increasing the labor supply with pension reforms that delay retirement may negatively influence the future supply of informal care for elderly parents. As the share of caregivers in the population, who provide care for parents and parents-in-law, is highest in the years around their retirement, it is important to ask whether a transition from employment to retirement causes an increase in the provision of informal care.

In this chapter, I have analyzed the causal effect of retirement on daily caregiving for parents and parents-in-law, using data from five waves of the SHARE on individuals aged 50 to 70 in 18 European countries, covering the period 2004 to 2015. Exploiting information on country-specific pension eligibility ages, I have constructed an instrument for the retirement status with having reached country-specific pension eligible ages for early or statutory retirement. Although previous studies have examined the effect of intensive caregiving on labor force participation, and only more recently the impact of caregiving on retirement (Jacobs et al., 2017), to my knowledge this is the first study to explore the causal effect of

retirement on informal caregiving for parents and parents-in-law, applying an instrumental variables approach on a cross-country sample. I find that retirement increases the probability of engaging in daily caregiving by 6.5 %-points. After including extra control variables, such as measures for own health and parental health, the effect size remains substantial (6.5 %-points), and this may have implications for policy, even if the share of individuals who provide daily care is limited, with average participation rates of 5.5 % and 10.5 % for men and women, respectively, across all individuals in the sample. Checking the robustness of the results with a structural stability test, I find that the effect of retirement seems stronger for individuals in Scandinavia and the Netherlands, where retirement increases the likelihood of daily caregiving by 16.4 %-points, which is much larger than for the other country region groups. In countries where older individuals, both men and women, on average have high employment rates, the time conflict between care provision and the labor supply may be stronger.

Due to population ageing and rising costs of formal care provisions, governments are expected to shift some of the responsibility for the long-term care of the elderly onto family members, most likely the adult children, and thereby increasing the demand for informal caregiving. The rising pension ages seen across Europe may on the other hand reduce the supply of informal care. Negative effects of later retirement should be taken into consideration, if government are to rely on informal care provision to meet the increased demand for long-term care of the elderly. To cope with this potential trade-off between the provision of informal care and the labor supply of retirement-aged individuals, governments may wish to provide suitable incentives to ensure the supply of informal care meets the demand, for example, by offering more financial support to caregivers.

The policy ambition to extend the working lives of retirement-aged individuals and increase women's labor force participation, by pushing pension eligibility ages to be more equal to that of men's and tightening the generosity of pension benefits, may diminish the supply of informal care as individuals

seem to significantly increase their provision of informal through retirement. For future research it could be interesting to shift the perspective and study how the care recipients (the parents and parents-in-law) benefit from the informal care of this particular age group of caregivers in their fifties and early sixties and estimate how current and future pension reforms potentially will diminish this supply of care and their related consequences for the health of the elderly .

References

Angelini, V., Brugiavini, A., and Weber, G. (2009). Ageing and unused capacity in Europe: is there an early retirement trap? *Economic Policy, 24*(59), 463-508.

Bolin, K., Lindgren, B., and Lundborg, P. (2008). 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.

Börsch-Supan, A., Brugiavini, A., Jürges, H., Kapteyn, A., Mackenbach, J., Siegrist, J., and Weber, G. (2005). *Health, ageing and retirement in Europe: first results from the Survey of Health, Ageing and Retirement in Europe*. Mannheim Research Institute for the Economics of Aging (MEA).

Carmichael, F., and Charles, S. (2003). The opportunity costs of informal care: does gender matter? *Journal of Health Economics, 22*(5), 781-803.

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

Celidoni, M., and Rebba, V. (2017). Healthier lifestyles after retirement in Europe? Evidence from SHARE. *The European Journal of Health Economics, 18*(7), 805-830.

Ciccarelli, N., and Van Soest, A. (2018). Informal caregiving, employment status and work hours of the 50+ population in Europe. *De Economist, 166*(3), 363-396.

Crespo, L. (2006, November). Caring for parents and employment status of European mid-life women. In *Second Workshop Economics of the Family at the University of Zaragoza* (pp. 9-10).

Crespo, L., and Mira, P. (2014). Caregiving to elderly parents and employment status of European mature women. *Review of Economics and Statistics, 96*(4), 693-709.

Ciani, E. (2012). Informal adult care and caregivers' employment in Europe. *Labour Economics, 19*(2), 155-164.

Dobrescu, L. I. (2015). To love or to pay savings and health care in older age. *Journal of Human Resources, 50*(1), 254-299.

Fischer, B., and Müller, K. U. (2019). *Time to Care? The Effects of Retirement on Informal Care Provision* (No. 1809). DIW Berlin, German Institute for Economic Research.

Gans, D., Silverstein, M., and Lowenstein, A. (2009). Do religious children care more and provide more care for older parents? A study of filial norms and behaviors across five nations. *Journal of Comparative Family Studies*, *40*(2), 187-201.

Gruber, J., and Wise, D. A. (Eds.). (2002). *Social security and retirement around the world*. University of Chicago Press.

Gruber, J. and Wise, D. A. (Eds.). (2008). *Social security and retirement around the world.* University of Chicago Press.

Gruber, J. and Wise, D. A. (Eds.). (2010). *Social security and retirement around the world.* University of Chicago Press.

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 Michaud, P. (2006). Informal Care and Employment in England: Evidence from the British Household Panel Survey. IZA Discussion Papers 2010. *Institute for the Study of Labor (IZA).*

Jacobs, J. C., Laporte, A., Van Houtven, C. H., and Coyte, P. C. (2014). Caregiving intensity and retirement status in Canada. *Social Science and Medicine*, *102,* 74-82.

Jacobs, J. C., Van Houtven, C. H., Laporte, A., and Coyte, P. C. (2017). The impact of informal caregiving intensity on women's retirement in the United States. *Journal of Population Ageing*, *10*(2), 159-180.

Lilly, M. B., Laporte, A., and Coyte, P. C. (2010). Do they care too much to work? The influence of caregiving intensity on the labour force participation of unpaid caregivers in Canada. *Journal of Health Economics, 29*(6), 895-903.

Lilly, M. B., Jacobs, J. C., Ng, C., and Coyte, P. C. (2011). Exploring the influence of caregiving intensity on the labor supply of Family/Friend caregivers in Canada. *Ottawa, Canada: Human Resources and Skills Development Canada.*

Mazzonna, F., and Peracchi, F. (2017). Unhealthy retirement? *Journal of Human Resources, 52*(1), 128-151.

Michaud, P. C., Heitmueller, A., and Nazarov, Z. (2010). A dynamic analysis of informal care and employment in England. *Labour Economics*, *17*(3), 455-465.

Nizalova, O. (2012). The wage elasticity of informal care supply: Evidence from the health and retirement study. *Southern Economic Journal*, *79*(2), 350-366.

OECD. (2017). *Pensions at a Glance 2017,* Organisation for Economic Co-operation and Development.

Van Houtven, C. H., and Norton, E. C. (2004). Informal care and health care use of older adults. *Journal of Health Economics*, *23*(6), 1159-1180.

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

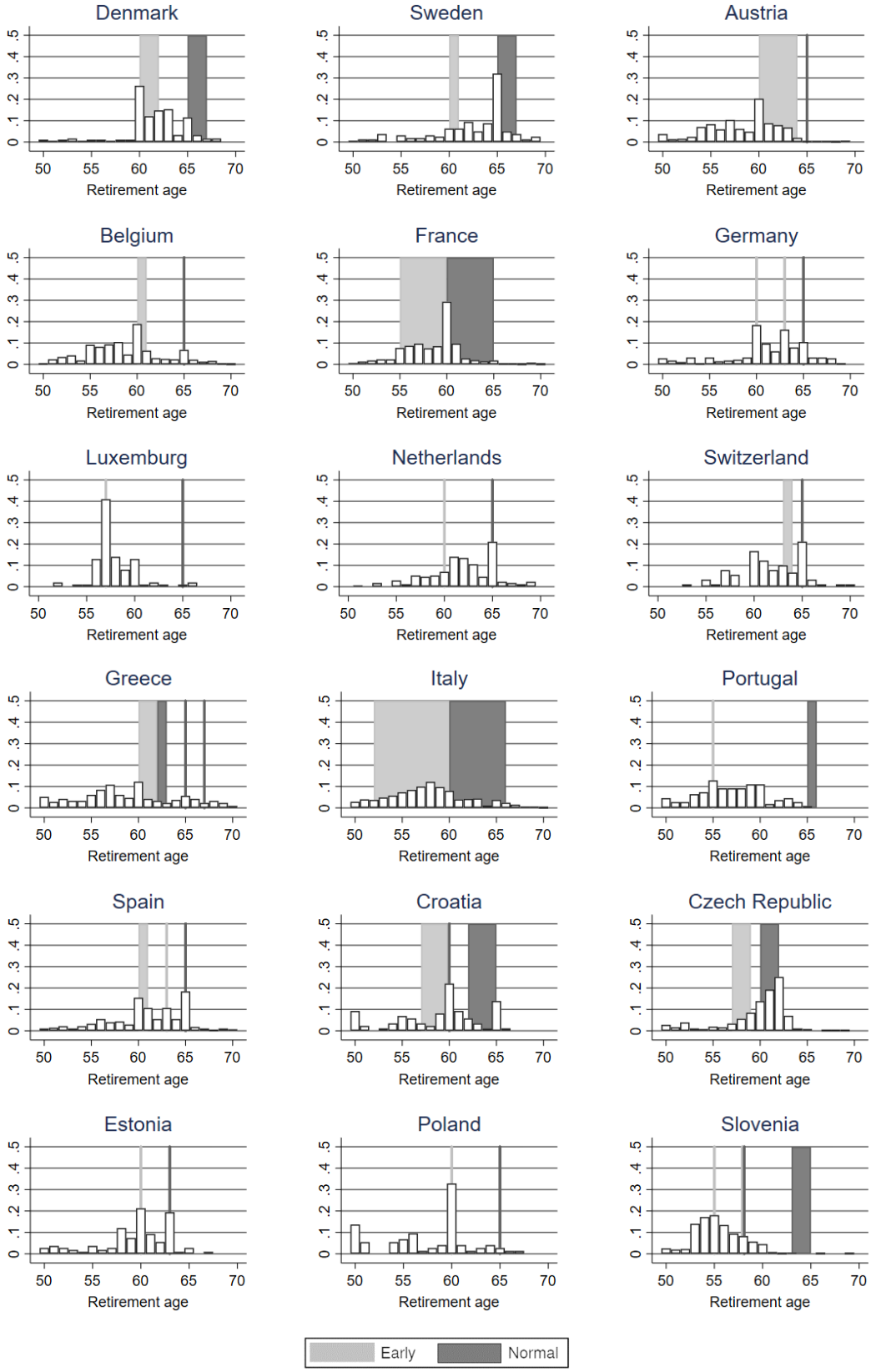
Wolf, D. A. and Soldo, B. J. (1994). Married women's allocation of time to employment and care of elderly parents. *Journal of Human Resources,* 1259-1276.

Appendix 1

Table 6 Countries and waves included in dataset

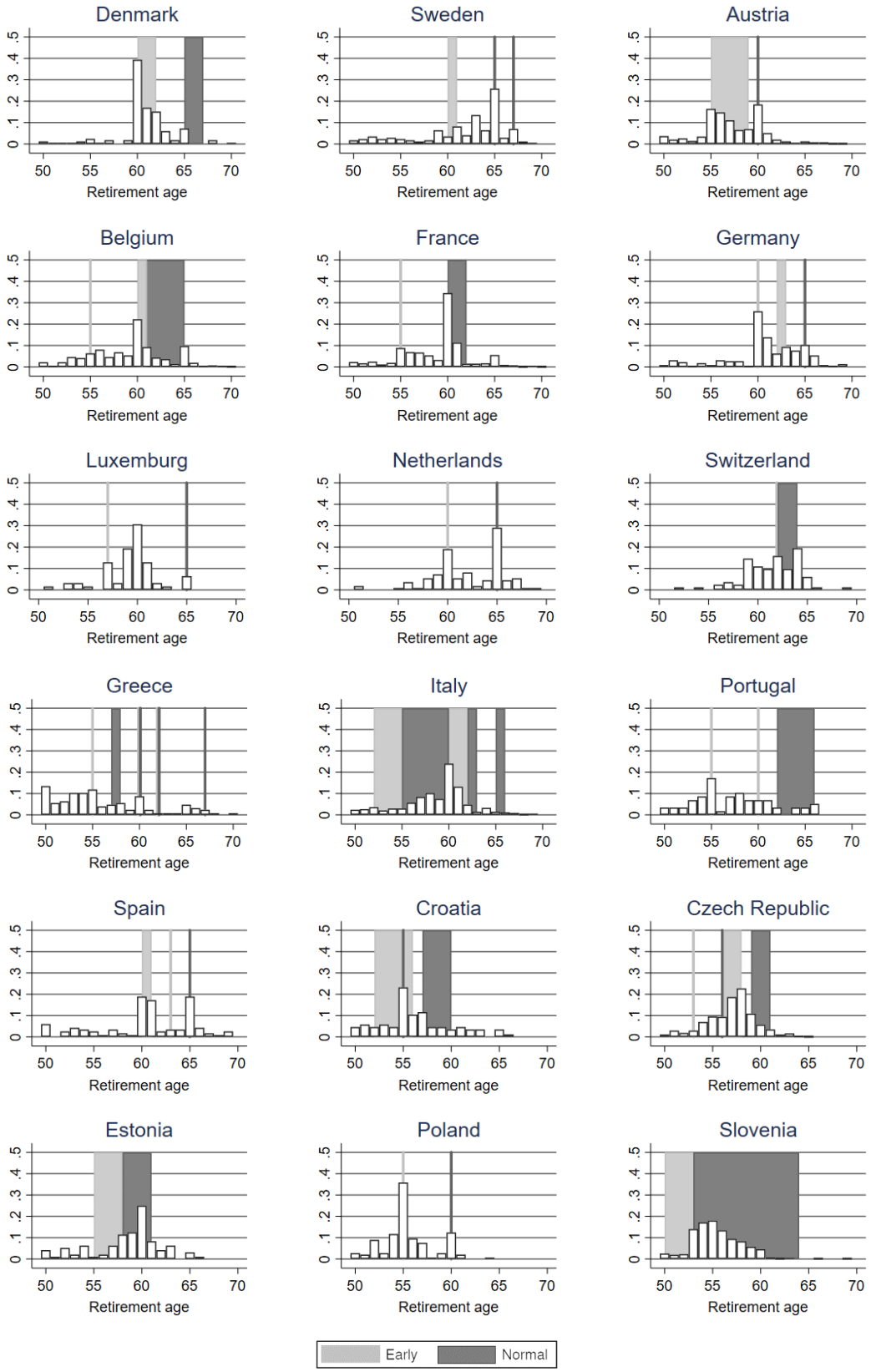
	Wave 1	Wave 2	Wave 4	Wave 5	Wave 6
Northern Europe					
Denmark	x	x	x	x	x
Sweden	x	x	x	x	x
Western Europe					
Austria	x	x	x	x	x
Belgium	x	x	x	x	x
France	x	x	x	x	x
Germany	x	x	x	x	x
Luxemburg				x	x
Netherlands	x	x	x	x	x
Switzerland	x	x	x	x	x
Southern Europe					
Greece	x	x			x
Italy	x	x	x	x	x
Portugal			x		x
Spain	x	x	x	x	x
Eastern Europe					
Croatia					x
Czech Republic		x	x	x	x
Estonia			x	x	x
Poland		x	x		x
Slovenia			x	x	x

Figure 1 Early and statutory eligibility ages for pension benefits, by country for men.



Note: In the individual figures, pension eligibility ages for early retirement are indicated in light grey, pension eligibility ages for statutory (normal) retirement are in dark grey, and the distribution of the actual ages at retirement is shown in white.

Figure 2 Early and statutory eligibility ages for pension benefits, by country for women.



Note: In the individual figures, pension eligibility ages for early retirement are indicated in light grey, pension eligibility ages for statutory (normal) retirement are in dark grey, and the distribution of the actual ages at retirement is shown in white.

Table 7 The fully specified model, OLS and 2SLS with Chow test

	(1) OLS	(2) OLS	(3) OLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
	Dependent variable: Daily care					
*Retired * Central Europe*	0.015**	0.015**	0.014**	0.055**	0.056**	0.053**
	(0.007)	(0.007)	(0.007)	(0.025)	(0.025)	(0.025)
*Retired * Scandinavia and Netherlands*	0.033***	0.034***	0.034***	0.159**	0.163**	0.164**
	(0.011)	(0.011)	(0.011)	(0.065)	(0.066)	(0.067)
*Retired * Southern Europe*	0.038***	0.037***	0.032***	0.052	0.047	0.052
	(0.010)	(0.010)	(0.010)	(0.039)	(0.039)	(0.039)
*Female * Central Europe*	0.054***	0.052***	0.052***	0.053***	0.051***	0.051***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)
*Female * Scandinavia and Netherlands*	0.020***	0.020***	0.020***	0.017***	0.017***	0.018***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
*Female * Southern Europe*	0.084***	0.078***	0.078***	0.083***	0.078***	0.077***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
*Age * Central Europe*	0.015	0.016	0.015	0.025*	0.026*	0.024*
	(0.011)	(0.011)	(0.011)	(0.013)	(0.013)	(0.013)
*Age * Scandinavia and Netherlands*	0.010	0.011	0.010	0.077**	0.078**	0.078**
	(0.014)	(0.014)	(0.014)	(0.036)	(0.037)	(0.037)
*Age * Southern Europe*	0.011	0.009	0.006	0.013	0.011	0.009
	(0.017)	(0.017)	(0.016)	(0.017)	(0.018)	(0.017)
*Age² * Central Europe*	-0.010	-0.010	-0.009	-0.020*	-0.021*	-0.019
	(0.009)	(0.009)	(0.009)	(0.012)	(0.012)	(0.012)
*Age² * Scandinavia and Netherlands*	-0.009	-0.010	-0.009	-0.070**	-0.072**	-0.072**
	(0.012)	(0.012)	(0.012)	(0.033)	(0.033)	(0.033)
*Age² * Southern Europe*	-0.006	-0.004	-0.002	-0.008	-0.006	-0.005
	(0.014)	(0.014)	(0.014)	(0.015)	(0.015)	(0.015)
*Married * Central Europe*	-0.034***	-0.033***	-0.033***	-0.035***	-0.034***	-0.034***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
*Married * Scandinavia and Netherlands*	-0.001	-0.001	-0.002	-0.003	-0.002	-0.004
	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)
*Married * Southern Europe*	-0.060***	-0.058***	-0.056***	-0.060***	-0.058***	-0.056***
	(0.013)	(0.013)	(0.012)	(0.013)	(0.013)	(0.012)
*Children * Central Europe*	-0.039***	-0.039***	-0.040***	-0.039***	-0.039***	-0.039***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
*Children * Scandinavia and Netherlands*	-0.017*	-0.017	-0.016	-0.015	-0.016	-0.015
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
*Children * Southern Europe*	-0.061***	-0.063***	-0.062***	-0.062***	-0.063***	-0.063***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Education dummies: Education level 3 is baseline						
*Education level 1 * Central Europe*	0.001	0.000	0.004	-0.003	-0.004	-0.000
	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.007)
*Education level 1 * Scandinavia and Netherlands*	0.003	0.003	0.005	-0.003	-0.002	-0.000
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
*Education level 1 * Southern Europe*	0.019**	0.018*	0.015	0.017*	0.017*	0.013
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
*Education level 2 * Central Europe*	0.001	0.001	0.003	-0.002	-0.002	0.001
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
*Education level 2 * Scandinavia and Netherlands*	0.011**	0.011**	0.012**	0.008	0.009*	0.010*
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
*Education level 2 * Southern Europe*	0.015	0.014	0.011	0.014	0.014	0.010
	(0.010)	(0.010)	(0.010)	(0.009)	(0.009)	(0.010)

*Household size * Central Europe*	0.016***	0.016***	0.016***	0.016***	0.016***	0.017***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
*Household size * Scandinavia and Netherlands*	-0.000	-0.000	0.000	0.001	0.001	0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
*Household size * Southern Europe*	0.030***	0.030***	0.029***	0.030***	0.030***	0.029***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
*Family respondent * Central Europe*	0.033***	0.034***	0.034***	0.033***	0.033***	0.033***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
*Family respondent * Scandinavia and Netherlands*	0.008	0.008	0.008	0.007	0.007	0.007
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
*Family respondent * Southern Europe*	0.046***	0.045***	0.046***	0.046***	0.045***	0.046***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
*Depressed * Central Europe*		0.022***	0.018***		0.021***	0.017***
		(0.005)	(0.005)		(0.005)	(0.005)
*Depressed * Scandinavia and Netherlands*		0.011	0.010		0.009	0.007
		(0.007)	(0.007)		(0.007)	(0.007)
*Depressed * Southern Europe*		0.043***	0.035***		0.043***	0.034***
		(0.010)	(0.010)		(0.010)	(0.010)
ADL limitations Central Europe*		-0.005	-0.006		-0.007	-0.008
		(0.011)	(0.010)		(0.010)	(0.010)
ADL limitations Scandinavia and Netherlands*		-0.008	-0.007		-0.015	-0.015
		(0.012)	(0.012)		(0.013)	(0.013)
ADL limitations Southern Europe*		-0.002	-0.009		-0.003	-0.011
		(0.021)	(0.020)		(0.022)	(0.021)
*IADL limitations * Central Europe*		-0.016	-0.018*		-0.020*	-0.022**
		(0.011)	(0.011)		(0.010)	(0.010)
*IADL limitations * Scandinavia and Netherlands*		-0.015	-0.017		-0.031**	-0.033**
		(0.011)	(0.011)		(0.014)	(0.014)
*IADL limitations * Southern Europe*		-0.044*	-0.046*		-0.045*	-0.047**
		(0.024)	(0.024)		(0.024)	(0.024)
*Siblings * Central Europe*			-0.033***			-0.033***
			(0.007)			(0.007)
*Siblings * Scandinavia and Netherlands*			-0.027***			-0.026***
			(0.008)			(0.009)
*Siblings * Southern Europe*			-0.054***			-0.054***
			(0.013)			(0.012)
*Parent in poor health * Central Europe*			0.064***			0.064***
			(0.007)			(0.007)
*Parent in poor health * Scandinavia and Netherlands*			0.035***			0.035***
			(0.009)			(0.009)
*Parent in poor health * Southern Europe*			0.125***			0.124***
			(0.013)			(0.013)
Wave dummies: Wave 6 is baseline						
*Wave 1 * Central Europe*	0.016*	0.016*	0.026***	0.010	0.010	0.020**
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
*Wave 1 * Scandinavia and Netherlands*	0.018**	0.018**	0.023***	0.006	0.006	0.010
	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)	(0.009)
*Wave 1 * Southern Europe*	0.006	0.007	0.026***	0.003	0.005	0.023*
	(0.010)	(0.010)	(0.010)	(0.012)	(0.012)	(0.012)
*Wave 2 * Central Europe*	0.003	0.003	0.004	-0.000	0.000	0.001
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
*Wave 2 * Scandinavia and Netherlands*	0.007	0.007	0.006	-0.000	-0.000	-0.002
	(0.006)	(0.006)	(0.005)	(0.008)	(0.008)	(0.008)
*Wave 2 * Southern Europe*	-0.002	0.000	0.004	-0.003	-0.001	0.002

	(0.012)	(0.012)	(0.010)	(0.013)	(0.012)	(0.010)
*Wave 4 * Central Europe*	-0.011*	-0.011*	-0.011*	-0.013**	-0.014**	-0.013**
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
*Wave 4 * Scandinavia and Netherlands*	-0.001	-0.001	-0.003	-0.000	-0.001	-0.002
	(0.008)	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)
*Wave 4 * Southern Europe*	0.014	0.014	0.008	0.012	0.013	0.006
	(0.011)	(0.011)	(0.011)	(0.012)	(0.011)	(0.012)
*Wave 5 * Central Europe*	0.007	0.007	0.006	0.005	0.005	0.005
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
*Wave 5 * Scandinavia and Netherlands*	0.008	0.008	0.007	0.004	0.004	0.003
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
*Wave 5 * Southern Europe*	-0.007	-0.007	-0.009	-0.007	-0.007	-0.009
	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.008)
Country dummies (Netherlands, Poland, and Croatia cancel out)						
*Denmark * Scandinavia and Netherlands*	-0.011**	-0.010**	-0.013***	-0.012**	-0.012**	-0.015**
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
*Sweden * Scandinavia and Netherlands*	-0.014**	-0.014**	-0.018***	-0.009	-0.009	-0.012*
	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)
*Austria * Central Europe*	-0.003	0.000	0.002	-0.005	-0.001	0.001
	(0.016)	(0.016)	(0.017)	(0.017)	(0.017)	(0.018)
*Germany * Central Europe*	0.012	0.015	0.014	0.019	0.021	0.020
	(0.015)	(0.016)	(0.016)	(0.016)	(0.017)	(0.017)
*France * Central Europe*	-0.014	-0.014	-0.015	-0.011	-0.011	-0.012
	(0.015)	(0.015)	(0.016)	(0.016)	(0.016)	(0.017)
*Switzerland * Central Europe*	-0.057***	-0.054***	-0.048***	-0.046***	-0.043**	-0.038**
	(0.015)	(0.016)	(0.016)	(0.017)	(0.017)	(0.018)
*Belgium * Central Europe*	-0.010	-0.008	-0.002	-0.007	-0.005	0.001
	(0.015)	(0.015)	(0.016)	(0.016)	(0.016)	(0.016)
*Czech Republic * Central Europe*	0.021	0.024	0.023	0.024	0.026	0.025
	(0.015)	(0.015)	(0.016)	(0.016)	(0.016)	(0.017)
*Luxembourg * Central Europe*	-0.045***	-0.044***	-0.043***	-0.046***	-0.044***	-0.044**
	(0.015)	(0.015)	(0.016)	(0.016)	(0.016)	(0.017)
*Estonia * Central Europe*	0.019	0.020	0.015	0.026	0.028*	0.023
	(0.016)	(0.016)	(0.017)	(0.017)	(0.017)	(0.018)
Spain Southern Europe*	0.010	0.012*	0.011	0.014	0.015	0.017
	(0.007)	(0.007)	(0.007)	(0.014)	(0.013)	(0.013)
*Italy * Southern Europe*	0.020**	0.019**	0.016*	0.022*	0.021	0.020
	(0.009)	(0.009)	(0.008)	(0.013)	(0.013)	(0.012)
*Greece * Southern Europe*	-0.025**	-0.025**	-0.019**	-0.022*	-0.023*	-0.016
	(0.010)	(0.010)	(0.009)	(0.012)	(0.012)	(0.011)
*Portugal * Southern Europe*	0.002	-0.001	-0.016*	0.004	0.001	-0.013
	(0.008)	(0.008)	(0.009)	(0.010)	(0.010)	(0.011)
*Slovenia * Southern Europe*	-0.003	0.001	0.004	-0.003	0.001	0.004
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Group 1, Central Europe	-0.510	-0.528*	-0.477	-0.738*	-0.758**	-0.695*
	(0.313)	(0.311)	(0.312)	(0.378)	(0.377)	(0.377)
Group 2, Scandinavia and Netherlands	-0.248	-0.280	-0.232	-2.049**	-2.098**	-2.074**
	(0.400)	(0.408)	(0.403)	(1.002)	(1.009)	(1.011)
Group 3, Southern Europe	-0.409	-0.366	-0.249	-0.453	-0.397	-0.307
	(0.507)	(0.510)	(0.488)	(0.516)	(0.518)	(0.502)
Observations	34,277	34,277	34,277	34,277	34,277	34,277
R-squared	0.117	0.119	0.133	0.114	0.115	0.130
	First stage - Dependent variable: Retired					

*EligibleER * Central Europe*				0.095***	0.095***	0.095***
				(0.037)	(0.037)	(0.037)
*EligibleER * Scandinavia and Netherlands*				-0.017	-0.016	-0.016
				(0.034)	(0.034)	(0.034)
*EligibleER * Southern Europe*				0.254***	0.254***	0.254***
				(0.058)	(0.058)	(0.058)
*EligibleSR * Central Europe*				0.317***	0.315***	0.315***
				(0.038)	(0.038)	(0.038)
*EligibleSR * Scandinavia and Netherlands*				0.333***	0.327***	0.327***
				(0.060)	(0.060)	(0.060)
*EligibleSR * Southern Europe*				0.199**	0.195**	0.195**
				(0.086)	(0.086)	(0.086)
Chow-test	57.32	63.49	65.63	2778.04	3789.2	4054.51

Source: SHARE 2004 to 2015, ages 50 to 70. Standard errors in parentheses are clustered on wave, gender, and country. *** p<0.01, ** p<0.05, * p<0.1. For the first stage equation, I only report estimates for the two instrumental variables used.

Appendix 2

The information on pension age eligibility criteria is primarily based on Gruber and Wise (2010, 2016), and the books' country-specific chapters on early retirement and the capacity to work at older ages. Other important data sources are the reports "Social Security Programs Throughout the World: Europe," that include country summaries with information on qualifying conditions for early and statutory retirement pensions, and the biennial reports "Pensions at a Glance" that presents profiles on the parameters and rules of the national pension systems across OECD and G20 countries. Moreover, I crosscheck with Celidoni and Rebba (2017), Mazzonna and Peracchi (2014), and Angelini et al. (2009). Additional sources are specified below for each individual country. ER stands for early retirement, while SR indicates statutory retirement.

Denmark (see Bingley et al., 2010; Bingley et al., 2016)

ER: 60 for both men and women. ER was opened for long-term unemployed individuals aged 50 to 54 in 1992, extended for 55 to 59 in 1994, and closed again in 1996.

SR: 67 until 2004, and then reduced by six-month steps to 65 by 2006, for both men and women.

From 2014, both ER and SR increase continuously by six-month steps for individuals born after 1953, so that ER is 61 and SR is 66 for individuals born in 1954, and 62 and 67 for individuals born in 1955.

Sweden (see Johansson et al., 2016; Mazzonna and Peracchi, 2017)

ER: No early retirement until 1962, 60 from 1963 to 1997, 61 from 1998 onwards, for both men and women.

SR: 67 for both men and women until 1994, 65 from 1995 onwards, for both men and women.

Austria (see Staubli and Zweimüller, 2011)

ER: 60 for men and 55 for women until 2000. From 2001 to 2004, early retirement age depends on the year of birth (with a step-wise increase of two months each quarter): For men, it is 61 for those born in

1942, and 62 for those born in 1943 or after. For women, it is 56 for those born in 1947, 57 for those born in the period 1948 to 1951, and 58 from 1952 onwards. Early retirement was gradually faced out; for men it increased to 63 in 2012, 64 in 2014, while for women it was raised to 59 in 2014.

SR: 65 for men and 60 for women.

Belgium (see Jousten et al., 2010)

ER: No early retirement until 1966. From 1967, 60 for men and for women 55, before increasing to 60 from 1992. From 2014, 61 for both men and women.

SR: 65 for men. 60 for women until 1996, 61 from 1997, 62 from 2000, 63 from 2003, 64 from 2006 to 2008, and 65 from 2009.

France (see Mazzonna and Peracchi, 2014; Behaghel et al., 2014)

ER: No early retirement until 1962. For both men and women, 60 from 1963, 55 from 1981 onwards. Early retirement is possible from 55 under certain circumstances related to working conditions and years of service.

SR: For both men and women, 65 until 1982, and hereafter 60 from 1983 to 2010. From 2011, 60 for those born up until 1952, 61 for those born in 1953 to1954, and 62 for those born in 1955 and after.

Germany (see Mazzonna and Peracchi, 2014; Jürges et al., 2016)

ER: No early retirement for men until 1972, then 60 from 1973, and 63 from 2004 onwards. For women, no early retirement until 1961, then 60 from 1962, 62 from 2004, 63 from 2006 onwards.

SR: 65 for all.

Luxembourg (see Duval, 2004)

ER: 57 for both men and women with the early retirement scheme ("pré-retraite"), making early retirement possible under certain circumstances related to working conditions, disability, and years of service.

SR: 65 for all.

Netherlands (see Euwals et al., 2010; Kapteyn et al., 2010; Kalwij et al., 2016)

ER: No early retirement until 1974. 60 from 1975 onwards, for both men and women.

SR: 65 for both men and women.

Switzerland (see Dorn and Sousa-Poza, 2003;)

ER: No early retirement for men until 1996 and for women until 2000. 64 for men from 1997, and 63 from 2001 onwards. For women 62 from 2001 onwards.

SR: 65 for men. For women 63 until 1963, 62 in the period 1964 to 2000, 63 from 2001, and 64 from 2006.

Greece

ER: 60 for men until 2012, and 62 from 2013 onwards. 55 for women until 2010, 60 from 2011 to 2012, and 62 from 2013 onwards.

SR: 65 for men until 2012, and 67 from 2013 onwards. 60 for women until 2010, 62 from 2011, and 67 from 2013 and onwards.

Until 2010, workers with at least 10,000 days (33 years) of contributions could retire at 62 (men) or 57 (women). From 2011 to 2012, for men could retire at 63 with at least 10,000 days (33 years) of contributions, or 58 for women with at least 10,800 days (36 years) of contributions. From 2013, 62 for men and women with 12,000 days (40 years) of contributions.

Years of contributions and pension age are lowered for individuals with arduous or unhealthy work and for women with dependent or disabled children.

Italy (see Angelini et al., 2009; Mazzonna and Peracchi, 2014)

ER: from 1965 to 1995, early retirement was possible at any age with 35 years of contributions (25 in the public sector) for both men and women. From 1996, early pension age increased in steps up to 63 for both men and women.

SR: The statutory retirement age was 60 (65 in the public sector) for men and 55 (60 in the public sector) for women from 1961 to 1993. Reforms stepwise increased the statutory retirement age to 66 for both men and women.

Portugal (see Albuquerque et al., 2009)

ER: 55 for both men and women until 2011. ER was suspended from 2012 to 2014. Instead long-term unemployed people aged 55 or over, could retire at age 60 with full pension, and at age 55 with 20 years of contributions for individuals unemployed at age 50 or above. From 2014 again 55 for both.

SR: 65 for men until 2013. 62 for women until 1993, and thereafter gradually rising to 63 from 1994, 64 from 1996, and 65 from 1999. From 2014, 66 for both men and women onwards.

Spain (see Blanco, 2000)

ER: 64 until 1982, 60 from 1983, 61 from 1994 (60 for government employees with at least 30 years of work). 63 from 2013 onwards, for both men and women.

SR: 65 for all.

Croatia (see Holzmann and Guven, 2009)

ER: From 2002, 57 for men and 52 for women (increasing by 6 months each year until 2008). 58 for men and 53 for women in 2004, 59 for men and 54 for women in 2006, and 60 for men and 55 for women in 2008. From 2014, 56 for women.

SR: 60 for men and 55 for women until 2001. From 2002, 62 for men and 57 for women (increasing by 6 months each year until 2008). 63 for men and 58 for women in 2004, 64 for men and 59 for women in 2006, and 65 for men and 60 for women in 2008 and onwards.

Czech Republic

ER: ER is possible up to three years before SR. 57 for men and 53 for women until 2001. From 2002 it is 58 for men and 56 for women, in 2008, 57 for women. From 2010, 59 for men. From 2012, 58 for women. From 2016, 60 for men and 59 for women.

SR: 60 for men and 56 for women until 2001.From 2002, 61 for men and 59 for women. From 2008, 60 for women. From 2010, 62 for men. From 2012, 61 for women.

The pension age is reduced by up to four years for women, depending on the number of children they have raised. This reduction is gradually being phased out and will be eliminated for insured women born after 1976.

Estonia (see Schmähl and Horstmann, 2002)

ER: 60 for men. 55 for women from 2002, 56 from 2004, 57 from 2008, and 58 from 2010 to 2015. Early retirement is possible up to three years before statutory retirement.

SR: 63 for men. The pension retirement age for women has gradually been raised. 58 for women from 2002, 59 from 2004, 60 from 2008, and 61 from 2010 to 2015.

Poland (see Holzmann and Guven, 2009)

ER: 60 for men and 55 for women.

SR: 65 for men and 60 for women.

Slovenia (see Polanec et al., 2013; Guardiancich, 2010; Holzmann and Guven, 2009)

ER: From 1992, 55 for men and 50 for women. For men, 58 from 2002 onwards. For women, 54 from 2000, 55 from 2005, 56 from 2010, 57 from 2012, and 58 from 2014.

SR: From 1992 to 2013 for men, 58 with 40 years of contributions, 63 with 20 years of contributions, or 65 with 15 years of contributions. From 2014 for men, 58 with 40 years of contributions, 64 with 20 years of contributions, or 65 with 15 years of contributions.

From 1992 to 1999 for women, 53 with 35 years of contributions, 58 with 20 years of contributions, or 63 with 15 years of contributions. From 2000 to 2004 for women, 54 with 35 years of contributions, 59 with 20 years of contributions, or 61 with 15 years of contributions. From 2005 to 2009 for women, 55 with 36 years of contributions, 60 with 20 years of contributions, or 62 with 15 years of contributions. In 2010 to 2011, 56 with at least 37 years of contributions, 61 with at least 20 years of contributions, or 63 with at least 15 years of contributions. In 2012 to 2013, 57 with at least 37 years contributions, 61 with at least 20 years of contributions, or 63 with at least 15 years of contributions.

From 2014, 58 with at least 38 years contributions, 62 with at least 20 years of contributions, or 64 with at least 15 years of contributions. In 2015, 58 with at least 38 years contributions, 62 with at least 20 years of contributions, or 64 with at least 15 years of contributions.

Additional references on pension ages

Albuquerque, P., Arcanjo, M., and Escária, V. (2009). Early retirement in Portugal.

Angelini, V., Brugiavini, A., and Weber, G. (2009). Ageing and unused capacity in Europe: is there an early retirement trap? *Economic Policy, 24*(59), 463-508.

Behaghel, L., Blanchet, D., and Roger, M. (2014). *Retirement, early retirement and disability: Explaining labor force participation after 55 in France* (No. w20030). National Bureau of Economic Research.

Bingley, P., Gupta, N. D., and Pedersen, P. J. (2010). Social Security, Retirement and Employment of the Young in Denmark. In *Social Security Programs and Retirement around the World: The Relationship to Youth Employment* (pp. 99-117). University of Chicago Press.

Bingley, P., Gupta, N. D., and Pedersen, P. (2016). Health capacity to work at older ages in Denmark. In *Social Security Programs and Retirement around the World: The Capacity to Work at Older Ages* (pp. 85-110). University of Chicago Press.

Blanco, A. (2000). *The decision of early retirement in Spain.* FEDEA Working Paper no. 76.

Celidoni, M., and Rebba, V. (2017). Healthier lifestyles after retirement in Europe? Evidence from SHARE. *The European Journal of Health Economics, 18*(7), 805-830.

Dorn, D., and Sousa-Poza, A. (2003). Why is the employment rate of older Swiss so high? An analysis of the social security system. *The Geneva Papers on Risk and Insurance-Issues and Practice, 28*(4), 652-672.

Duval, R. (2004). Retirement behaviour in OECD countries. *OECD economic studies*, *2003*(2), 7-50.

Euwals, R., Van Vuuren, D., and Wolthoff, R. (2010). Early retirement behavior in the Netherlands: Evidence from a policy reform. *De Economist, 158*(3), 209-236.

Gruber, J. and Wise, D. A. (Eds.). (2010). *Social Security Programs and Retirement Around the World.*. University of Chicago Press.

Gruber, J. and Wise, D. A. (Eds.). (2016). *Social Security Programs and Retirement Around the World.* University of Chicago Press.

Guardiancich, I. (2010). Current pension system: first assessment of reform outcomes and output. *European Social Observatory, Country Report, Italy.*

Holzmann, R., and Guven, U. (2009). *Adequacy of Retirement Income After Pension Reforms in Central, Eastern and Southern Europe: Nine Country Studies.* The World Bank.

Kalwij, A., Kapteyn, A., and De Vos, K. (2016). Work Capacity at Older Ages in the Netherlands. In *Social Security Programs and Retirement Around the World: The Capacity to Work at Older Ages* (pp. 243-268). University of Chicago Press.

Kapteyn, A., de Vos, K., and Kalwij, A. (2010). Early retirement and employment of the young in the Netherlands. In *Social Security Programs and Retirement Around the World: The Relationship to Youth Employment* (pp. 243-259). University of Chicago Press.

Johansson, P., Laun, L., and Palme, M. (2016). Health, work capacity and retirement in Sweden. In *Social Security Programs and Retirement Around the World: The Capacity to Work at Older Ages* (pp. 301-328). University of Chicago Press.

Jousten, A., Lefèbvre, M., Perelman, S., and Pestieau, P. (2010). The effects of early retirement on youth unemployment: The case of Belgium. In *Social Security Programs and Retirement Around the World: The Relationship to Youth Employment* (pp. 47-76). University of Chicago Press.

Jürges, H., Thiel, L., and Börsch-Supan, A. (2016). Healthy, happy and idle: Estimating the health capacity to work at older ages in Germany. In *Social Security Programs and Retirement Around the World: The Capacity to Work at Older Ages* (pp. 149-180). University of Chicago Press.

Mazzonna, F., and Peracchi, F. (2014). *Unhealthy Retirement?* (No. 1409). Einaudi Institute for Economics and Finance (EIEF).

OECD. (2017). *Pensions at a Glance 2017: OECD and G20 Indicators*, OECD Publishing, Paris.

OECD. (2009). Country Profiles. In *Pensions at a Glance 2009: Retirement-Income Systems in OECD Countries*, OECD Publishing.

Pestieau, P., and Stijns, J. P. (1999). Social security and retirement in Belgium. In *Social Security and Retirement around the World* (pp. 37-71). University of Chicago Press.

Polanec, S., Ahčan, A., and Verbič, M. (2013). Retirement decisions in transition: microeconometric evidence from Slovenia. *Post-Communist Economies, 25*(1), 99-118.

Salem, M. B., Blanchet, D., Bozio, A., and Roger, M. (2010). Labor force participation by the elderly and employment of the young: The case of France. In *Social Security Programs and Retirement Around the World: The Relationship to Youth Employment* (pp. 119-146). University of Chicago Press.

Schmähl, W., and Horstmann, S. (Eds.). (2002). *Transformation of pension systems in central and eastern Europe.* Edward Elgar Publishing.

Staubli, S., and Zweimüller, J. (2011). Does raising the retirement age increase employment of older workers? IZA Discussion Papers 5863. *Institute for the Study of Labor (IZA).*

Verbič, M. (2008). The ageing population and the associated challenges of the Slovenian pension system. *Financial Theory and Practice, 32*(3), 321-338.