Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

ECONOMICS

Ciclo 32°

Settore Concorsuale: 13/A1

Settore Scientifico Disciplinare: SECS-P/01

Essays in Applied Economics

Presentata da: Giacomo Danda

Coordinatore Dottorato

Prof.ssa Maria Bigoni

Supervisore

Prof. Davide Dragone

Esame finale anno 2020

Abstract

The present thesis is made up by three separates chapters in applied microeconomics touching the realms of labor, health and family economics. The first one considers individual genetic information to explore the interplay between genes and environmental factors in shaping individual labor outcomes. The second one looks at old age health and provide an estimate of the causal effect of retirement on a syndrome of health deficit accumulation called frailty. The third one investigates and describes the role of preferences in the screening and matching process of child adoption with the use of a novel dataset.

Contents

List of Figures

List of Tables

Chapter 1

The gene-environment interplay in late working age labor market outcomes

Abstract

We investigate the interaction between negative labour market shocks in the United States and genetic predisposition to educational attainment as measured by polygenic score in determining unemployment in late working age. We exploit variation at State and industry level in the peak time of mass lay-off events and initial unemployment benefit claims in the period following the great recession to identify exogenous labour market shocks to the probability of unemployment. We provide evidence for polygenic score for educational attainment being a protective factor against unemployment status in late working life even after controlling for education and college completion.

1.1 Introduction

Late working life has been extensively investigated as a crucial transitional period with significant potential consequences not only with respect to individual health trajectories in later life (Moon et al. (2012);Coe and Zamarro (2011); Behncke (2012); Hessel et al. (2018)) but also with respect to pension systems and public finance (Hicks (2011),Brugiavini and Peracchi (2005)). With this study, by exploiting recent findings in the genomic and bio-informatics literature we explore the interplay between environmental factors and individual genetic endowment in late working life labor market outcomes. More precisely, we study the interaction between negative labour market shocks in the United States and individual genetic predisposition for educational attainment in determining unemployment status.

Before going to the reasons why we think genetic predisposition for educational attainment is relevant in the context of late working life it is useful to think of the general implications of unemployment in this life period. Of course, experiencing adverse labor market conditions can be associated with less employment, but interestingly, in late working life this is also related to decreased health insurance coverage and longevity (Coile et al. (2012)). Moreover, older workers experiencing unemployment have significantly lower re-employment rates because of reasons spanning from skills mismatch to age discrimination (Lahey (2005); Malul (2009); Axelrad et al. (2018)). In this context, the possibility to identify individuals at higher risk of negative labor market outcomes could be leveraged to better target preventive measures or even active labour market interventions.

To justify the link between genetic predisposition to educational attainment and labor market outcomes we build on recent findings by Papageorge and Thom (2018) who suggest genetic predisposition for educational attainment being also indicative of a general predisposition to keep acquiring skills related to analytical non-routine tasks beyond formal education. This suggests that genetic predisposition for education could indeed be telling something more, for example, predisposition to learn or acquire new competences and skills.

If this is indeed the case, one could expect individuals with higher genetic predisposition to educational attainment, to be also more able to expand their human capital beyond education itself, and so to be on average more productive in the workplace, all the rest being equal. As an example one could think of learning to use a new software, a novel programming language or generally facing a new working context. Most interestingly, it would be in late working age where differences in genetic endowment (beyond education itself) would reflect more the underlying differences in the latent human capital and, as a consequence, in productivity. Then, to the extent to which predisposition to educational attainment also contribute to general human capital accumulation beyond education and so late working age productivity, we would expect individuals with higher predisposition for educational attainment to be less likely to be laid off or lose their job once a negative labour market shock hits. In other words this mean expecting workers with lower genetic predisposition to educational attainment to be less productive and so more likely to represent the so called marginal worker. Alternatively, but along the same reasoning, one could think of individuals with higher predisposition for education to be more likely to find a new job after having lost one. It is fair to say that our framework of analysis does not allows to perfectly distinguish between these two cases as in our data individuals are observed only every two years.

We test this interpretation of a defined measure of genetic predisposition to educational attainment by identifying a negative labor market shock and assessing heterogeneous effects to probability of unemployment along this genetic dimension. Before going into the details of the study, we provide a sort of first landing yet not exhaustive guide to genetics in social science.[1]

1.2 From Genetics and Economics to Geno-economics

Since the completion of the human genome project in 2001 (et al. (2001),Craig Venter et al. (2001), the role of genetic information have largely and increasingly been questioned and explicitly investigated. While at first, also because of significant knowledge entrance barriers, most of the research was exclusively

[1]For further details readers can refer to Conley (2009), Beauchamp et al. (2011) and Benjamin et al. (2012).

held within the realm of biology and molecular medicine, as sequencing and imputation technologies as well as analytical methodologies became more and more accessible, more and less contiguous research fields became aware of the relevance of genetics and, as a consequence, interested in incorporating it into more traditional frameworks.

This happened with economics as well, with theorists working in insurance debating on the relevance of such information for an industry deeply dissected with regard to issues of adverse selection and moral hazard (Hoy and Polborn (2000), Barigozzi and Henriet (2011). However, aside empirical studies on twins (Ashenfelter and Krueger (2019); Conley and Strully (2012)), little has been done so far to explicitly incorporate genetic information into economics.

Recent developments and continuous decreasing costs of genotyping promise extremely easily available individual genetic information leading some scholars to focus on the explicit role of such information for social science (Beauchamp et al. (2011); Benjamin et al. (2012)) and related policy relevant issues (Lehrer and Ding (2017)). In principle, opportunities for social scientists, and economists in particular, are related to two main approaches with regard to the use of genetic information: as instrumental variable in a so called Mendelian Randomization Study (Conley and Zhang (2018); Van Kippersluis and Rietveld (2018))[2], and finally, gene-environment interaction studies or "GXE" (Conley (2009)). In the current study we focus on the latter.

1.2.1 Polygenic Scores

In practical terms, applied researchers can exploit genetic information in the form of polygenic scores (PGSs). A PGS is a synthetic measure of genetic predisposition for a given phenotype (i.e. an observable outcome or condition determined both by genetic and environmental factors) built on the basis of the presence, for a given individual, of a certain set on genetic mutations called Single Nucleotide

[2]While the use of individual level genetic information as instrumental variable (i.e. Mendelian Randomizazion) is not entirely new to fields like epidemiology, research to date have almost entirely overlooked the potential violation of the key exclusion restriction assumption due to highly likely pleyotropic effects of genetic endowment.

Polymorphisms (SNPs). Such genetic mutation is the single most commonly occurring mutation that happens at the level of single building block of our DNA sequence: the nucleotide. Nucleotides are molecules made by a nitrogenous base, sugar and a phosphate group which generally defines the nucleotide as A, T, C, or G. Whenever there is a mutation of a given nucleotide in the DNA sequence, in a relevant share of the whole population (i.e. more than 1%), we call it a SNP. As already mentioned, SNPs are extremely common forms of mutation in our DNA and their position with respect to a gene can either have no effect at all on its functioning or influence its expression (modulation with environmental factors), heritability or favour the development of some disease (Beauchamp et al. (2011); Benjamin et al. (2012)).

This being said, one can define a PGS for a give phenotype as a weighted sum of the SNPs' effect on such phenotype where weights are defined on the basis of genome-wide association studies (GWAS). These association studies consist in the estimation through separate regressions of the effect of each genotyped SNP on the outcome of interest[3]. Once the estimates are computed, their significance is tested at a p-value threshold adjusted for multiple testing hypothesis by Bonferroni correction. The so called "genome-wide significance" threshold level is normally set at $5.5\ 10^{-8}$.

For individual i and a given phenotype, the PGS is defined as:

$$PRS_i = \sum_{j=1}^{n} w_j x_{ij} \tag{1.1}$$

where x_{ij} is the number of reference allele (zero, one or two) for individual i and SNP j and w_j is the SNP's coefficient from the reference GWAS. By reference allele we mean the presence of a mutation, with respect to the most common genetic variant, in either one or two copies of an individual's DNA. The reference allele is zero when there is no mutation at the precise locus in the DNA sequence where a

[3]Note that at this stage control covariates such as gender, age polynomials, age and gender interactions and the first ten genetic principal components are included. Most importantly, the first ten genetic principal components are included to control for possible population stratification in the genetic endowment of the individuals. By this we refer to the systematic difference in allele frequencies between different groups of a population (see Ware et al. (2018), Beauchamp et al. (2011) and Benjamin et al. (2012) for further details).

SNP could be observed. Whenever there is a SNP only in one copy of the DNA sequence instead, this mean that the individual inherited the mutation from only one of the two parents (heterozygosity) and the allele reference is therefore one while, if the mutation is present in both copies of the DNA, it means that both parents transmitted the mutation (homozygosity). In this latter case the reference allele is two (Ware et al. (2018); Beauchamp et al. (2011)). The polygenic score is then standardized to have mean zero and variance equal to one.

It is important to note that unless researchers are capable of controlling for parental genotype, the nature of the PGS itself does not allow a causal interpretation with respect to the outcome of interest. In other words, it is difficult, if not impossible, to completely disentangle the "true genetic effect" from environmental channels taking place in the downstream of it. As an example, taking into account a PGS for BMI (Locke et al. (2015)) we could in principle expect particular familiar environment to play a role in addition to genetic endowment in determining individual's BMI. This is to say that as a newborn inherits his/her DNA from the parents, he/she will also experience a certain environment which is correlated with his/her parents' DNA. Again, as suggested in Schmitz and Conley (2016)), also upstream mediating channels such as in-utero environment affected by mothers eating or smoking behaviour are theoretically possible, making the relation between polygenic score potentially spurious. Given this limitations, it is possible to claim SNPs to be as good as randomly assigned at conception only conditionally on parental genotype (Schmitz and Conley (2017)). Whenever it is not possible to account for parental genotypes, as it is in most of the cases due to data limitation, SNPs and consequently polygenic scores should be interpreted as predetermined.

The lack of causal interpretation, does not make the analysis useless though. While it is true that it is mostly impossible to pin down the true genetic effect reflecting solely biological pathways from the DNA sequence to given phenotypes, SNPs weights and polygenic scores do capture it. The genetic effect is likely inflated by environmental factors which are not strictly genetic but which are inherited alongside the genetic structure. This includes any sort of discrimination happening in the society which can not be accounted for by population stratification or ancestry group.

1.3 Literature

First and foremost we broadly contribute to the already cited literature on the use of genetic information in economics and social science (Beauchamp et al. (2011); Benjamin et al. (2012); Lehrer and Ding (2017); Conley and Zhang (2018); Van Kippersluis and Rietveld (2018)).

More precisely, we contribute to a recent but rapidly growing literature on the interaction between genetic and environmental factors in determining observable outcomes. By explicitly including genetic information in more traditional frameworks, research shifted heavily from the so called nature versus nurture debate to investigating the actual interplay between the two. While GXE studies are not new to medicine or realms like epidemiology and biology (Ottman (1996)), recent discoveries in genome-wide significant SNPs for traits such as educational attainment (Okbay et al. (2016); Lee et al. (2018)) but also BMI (Locke et al. (2015)) and risk tolerance (Karlsson Linnér et al. (2019)) which are traditionally highly relevant in economic research are driving a surge of interest also among economists with most of the studies referring to health and education economics.

Starting from the health domain, Biroli (2015) proposes a theoretic framework for BMI and human capital formation that explicitly account for the role of genetics and molecular biology in which GxE interaction play a "pivotal role" in the evolution of BMI. More specifically, the author presents a model where genes influence both the health production function (*"genetic productivity effect"*: how productively inputs are converted into outputs) as well as individuals' preferences, thus affecting the implicit cost of investment in health capital (*"genetic cost effect"*). Schmitz and Conley (2016) look at the interaction between unemployment status and a PGS for BMI in the Health and Retirement Study showing heterogeneity in the effect of job loss on BMI, with high risk individuals being more likely to gain weight when losing jobs. With respect to mental health instead and again within the Health and Retirement Study, Domingue et al. (2017) focus on the interaction between a score for genetic predisposition to subjective well-being and a stressor event like the death of a spouse in determining depressive symptoms.

Results show a significant protective effect of genetic predisposition for subjective well-being with respect a spouse's death.

Bridging the health and the education realms, Amin et al. (2017) investigate gene-environment interactions between education and BMI in a sample from UK and Finland. The study evidences a statistically significant negative association between education and BMI as well as a statistically positive association between the genetic endowment and BMI but no significant interaction effect.

Shifting instead completely to education, Schmitz and Conley (2017) investigate the effect on education of veteran status as instrumented by the Vietnam lottery draft and interacted with a PGS for educational attainment. The authors report veterans with below average PGS being more likely to collect less years of schooling as compared to non veterans with similar polygenic scores.

More interestingly for our framework, Papageorge and Thom (2018) investigates the interaction between a PGS for educational attainment[4] and childhood socio-economic status in the Health and Retirement Study finding a significant main genetic effect with respect to college graduation as well as a significant interaction with childhood socio economic status suggesting possible concerns of wasted potential when growing up in lower socio-economic status. More relevant to our setting, Authors also find that the PGS for educational attainment predicts labour market outcomes such as earning and employment even after controlling for education and college completion. By analysing the time variation in earnings and the relationship with the PGS they argue the score to be indicative of something more than just predisposition for educational attainment and rather, predisposition to accommodate ongoing skill biased technological changes.

Table 1.1 replicate some of the results in Papageorge and Thom (2018) with few modifications. First and foremost, not having access to external data other than the HRS, we rely on self reported earnings for the outcome variable. On the other hand we use a more updated version of the polygenic score (Lee et al. (2018)) and we also able to control for State and industry of occupation. Results are qualitatively comparable to those of the original study. Even after controlling for

[4]The authors use a polygenic score based on the weights as estimated in Okbay et al. (2016).

Table 1.1: OLS for labor earnings over education and polygenic score

	Log(Earnings)				
	(a)	(b)	(c)	(d)	(e)
PGS_EA	0.112***	0.055***	0.031***	0.029**	0.022**
	(0.013)	(0.015)	(0.011)	(0.011)	(0.011)
Educational Attainment		0.011**	0.016***	0.016***	0.014***
		(0.005)	(0.004)	(0.005)	(0.004)
College		0.399***	0.345***	0.351***	0.330***
		(0.037)	(0.029)	(0.030)	(0.029)
Constant	10.148***	8.857***	9.843***	10.127***	10.084***
	(0.080)	(0.789)	(0.269)	(0.140)	(0.217)
PC 1-10	Y	Y	Y	Y	Y
PC 1-10 * Educ. Att.	Y	Y	Y	Y	Y
PC 1-10 * College	Y	Y	Y	Y	Y
Gender	Y	Y	Y	Y	Y
Age	Y	Y	Y	Y	Y
Birth year	Y	Y	Y	Y	Y
Interview year	Y	Y	Y	Y	Y
State & Ind. dummies					Y
Age Range	25-65	25-65	40-65	50-65	50-65
Employment earnings	>0	>10,000	>10,000	>10,000	>10,000
N	13755	9923	8879	8427	8427
R2	0.08	0.09	0.14	0.14	0.18

Note: Columns (a) to (d) replicate the results from Papageorge and Thom (2018) with our working sample and few consequential modifications. The dependent variable is self reported labor earnings. PGS_EA is Lee et al. (2018). Column (e) also controls for State and industry dummies.

formal education, its interaction with the first ten genetic principal components as well as State and industry fixed effects (not to mention other usual controls), the polygenic score remains predictive of the $Log(earning)$.

Beyond the GXE setting, other studies focused on the polygenic score for educational attainment and its interpretation. Along the same line of Papageorge and Thom (2018), investigating wealth inequality, Barth et al. (2019) provide evidence of polygenic scores for educational attainment being associated with wealth not only through education and earnings but also via financial decision making. To the best of our knowledge, Rustichini et al. (2018) and citeWilloughby2019, provide to date the only educational attainment polygenic score analysis capable of controlling for parental genotypes and family environment. Rustichini et al. (2018) propose a model of intergenerational mobility and subsequently investigate the channels between PGS and educational outcomes disentangling cognitive skills from personality traits. Results points to both spheres being relevant, with intelligence accounting for a larger proportion. Most interestingly, by controlling for mothers' and fathers' PGSs, the study provide evidence of the presence of parental environmental effect. Nevertheless, the coefficients of the parental PGS for educational attainment becomes very close to zero and non significant once controls like parental education and family income are taken into account in regressions for both GPA and years of education. This can give an idea of the extent to which estimates of main genetic effect could be inflated by environmental channels if one can not control for the appropriate familiar variables. In a model for GPA the polygenic score's effect remains highly significant decreasing by roughly 30% in point size estimate. On the other hand, in a model for educational attainment the coefficient drops by only 8% still remaining highly significant. This is important evidence of the fact that while parent environmental channels are non negligible, they are not the only responsible for educational outcomes of off-springs with the larger proportion of the effect captured by polygenic score being indeed genetic.

1.4 Data

The study uses two separate sources of data. On the one hand, information regarding individuals' working status, demographic controls and polygenic scores are from the Health and Retirement Study (HRS) while, negative labor market shocks are built using data on mass lay-off and initial unemployment benefit claims from the Bureau of Labor Statistics (BLS).

1.4.1 Health and Retirement Study and polygenic score

The Health and Retirement Study is a representative panel data survey collecting information every two years on domains like health, labor and socio-economic status every two years among individuals of 50 or more years old, together with their spouses and partners. Since 2006, by mean of a so called "enhanced face to face interview" the study started collecting saliva samples for genotyping together with other general biomarkers. In 2006, half of the interviewees were randomly selected for such enhanced interview of which, roughly 85% agreed. The other half was instead selected in the following wave in 2008. The same was done in the two subsequent waves in 2010 and 2012.

Polygenic scores are provided for a rich set of phenotypes ranging from BMI to neuroticism, Alzheimer disease and of course educational attainment. The HRS provides polygenic scores computed using all available SNPs overlapping with those of the original genome-wide meta-analysis without any p-value thresholding. To obtain externally valid SNP weights, whenever the original discovery sample included HRS individuals, the GWAS analysis was repeated excluding HRS data. We use the latest PGS for educational attainment as developed by Lee et al. (2018) with an original discovery sample of 1,1 million individuals and 1,271 genome-wide significant SNPs found.

For our analysis we restrict the sample to individual between 50 and 65 years of age, interviewed between 2006 and 2014 who were either self-declared employed or self-employed in 2006 which is the last wave before the Great Recession. We do so because the nature of the labour market data we exploit does not allows us to capture positive fluctuations of the labor market but only negative shocks. As a matter of fact we are interested in the transition from employment or self-employment to unemployment and/or exiting the labor force via (early)retirement or disability.

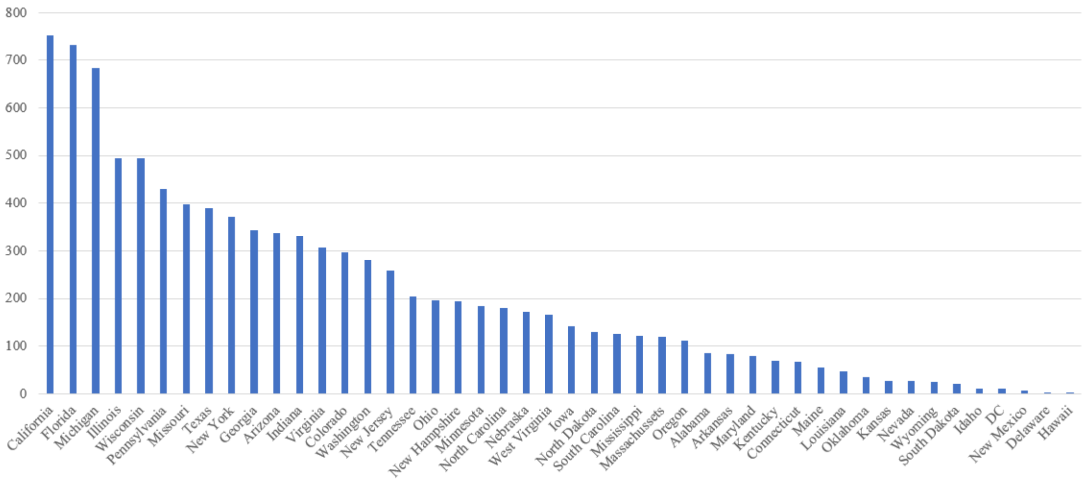
Table 1.2: HRS pooled sample.

	Total HRS	Genotyped	European Ancestry
	(1)	(2)	(3)
age	59.131	59.240	59.252
number of children	0.440	0.384	0.354
years of education	13.642	13.918	14.036
unemployed	0.027	0.024	0.024
income	33,154	34,346	35,720
diabetes	0.166	0.158	0.146
Hispanic	0.098	0.000	0.000
Afro-american	0.131	0.144	0.000
Caucasian	0.804	0.850	0.995
N	15350	10893	8499

In order to control for confounders at ancestry group level we follow usual recommendations from the literature and the Health and Retirement Study itself and further restrict to individual with European ancestry excluding individuals with Afro-American ancestry. This restriction also reflects a general limitation of the genomic literature to date which is the lack of predictive power of estimated SNPs' effect outside of the European ancestry group. Such lack of predictive power is due to under-representation of non European in the data suitable GWAS analysis. As a result of this under-representation, SNPs weight for Afro-American ancestry groups are considerably less predictive of which ever phenotype of interest is under investigation. It is worth noting that for the same reasoning the Health and Retirement Study do not provide polygenic score for Hispanic individuals.

This lead us to a pooled working sample of 8,499 observations consisting in roughly 3,500 individuals observed throughout the considered waves. As summarized in Table 1.2, it is fair to say that both the 85% take-up rate of the enhanced face-to-face interview and the restriction to European ancestry contribute to a non negligible selection of the sample with respect to certain dimensions which

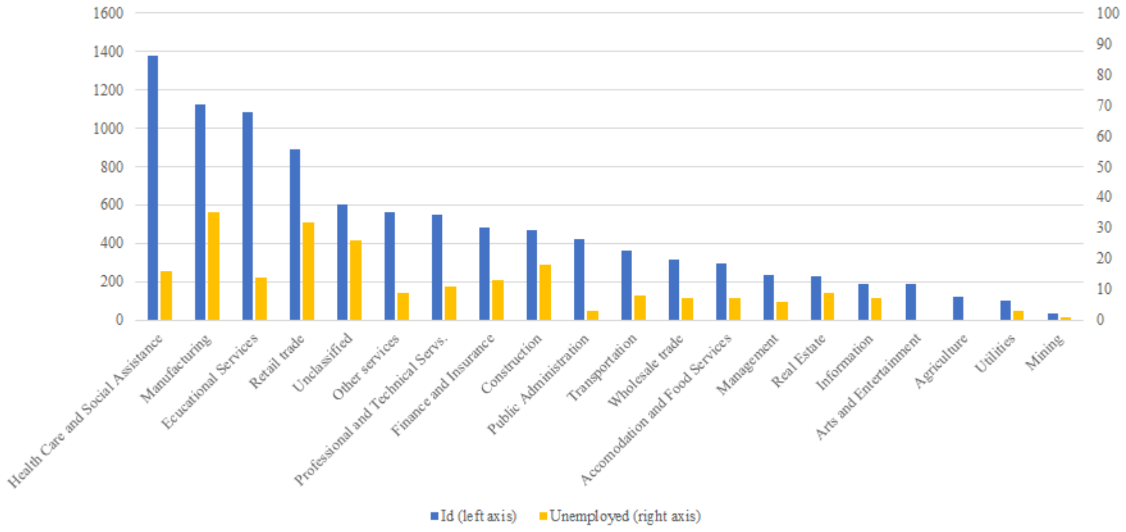
Figure 1.1: Sample distribution across States. HRS, waves 2006 - 2014.



are indeed interesting for the outcomes under investigations. With respect to the general population of the Health and Retirement Study, column 3 of Table 1.2 describes a sample which is on average slightly more educated, with less kids, with higher income less chances of self reporting unemployment status as well as lower risk of being diagnosed with diabetes. As expected, no self reported Afro-American as well as Hispanic individual is present in the final sample. Overall, despite representing a limitation for the external validity of our findings this forced selection is also making our results more salient. In fact, focusing within a sub-population with higher average PGS and which is for whatever reason less likely to experience the environmental shock at the basis of the gene-environment interaction under investigation would suggest to interpret point estimates as a lower bound.

For each individual at every point in time we know his/her residence location at county level as well as the industry of occupation at two digits NAICS level. We use information on the geographical location and industry of occupation to merge individual level data from the Health and Retirment Study with labor market data from the Bureau of Labor Statistics. California, Florida and Michigan are the three most populous States in the HRS give the sample restriction adopted (see Figure 1.1). In terms of industries, most of the working sample is employed in healthcare, manufacturing and educational services sector. Interestingly, with the exception of art and entertainment and agriculture, which nevertheless represent a small

Figure 1.2: Workers and unemployed distribution across industries. HRS, waves 2006 - 2014.



proportion of the sample) in every sector we find an appreciable proportion of the pooled sample is unemployed going from the 0.7% for public administration to 3.9% of real estate.

1.4.2 Labor Market

It is important to note that while providing evidence of negative labor market shocks impacting individual employment status could be seen as effort to prove the obvious, we should still pay attention at the way we estimate the effect of these shocks. As a matter of fact, a proper identification of the environmental effect is key to providing a thoughtful interpretation of the interaction with genetic endowment.

In order to identify negative labor market shocks we rely on data from the Bureau of Labor Statistics in the US. We use data at state, industry and month level about mass layoff events and initial unemployment benefit claims in alternative specifications for robustness purposes. To make sure that the effect we identify is not artificially driven by the way we define peaks in either mass lay-off or benefit claims, we also provide an alternative set of results based on unemployment rate at county level.

The Bureau of Labor Statistics defines a mass layoff event whenever a firm lays off at least 50 employees within a 5 weeks period. On the other hand, initial unemployment benefit claims are the number of claims filed in order to obtain unemployment benefits. The two measures are obviously correlated but yet capture slightly different margins of the same phenomenon. On the one hand, mass lay-off events capture the intensive margin of negative labour market shocks while unemployment benefit claims do capture the extensive one. In the main set of results we define a negative labor market shock as a peak in mass lay-off, devoting benefit claims data to robustness. As a matter of fact, initial benefit claims data are more likely to capture seasonal turnover in the labor force, making peaks harder to identify due to noisier trends.

We define a peak in mass lay-offs ($PeakML_{dst}$) for industry d, State s at time t as a dichotomous variable taking value one for that industry state and time cell whenever mass lay-offs are greater or equal than k times the mass-layoffs' average along the considered period. In the main set of results the factor k takes a value of three.[5]

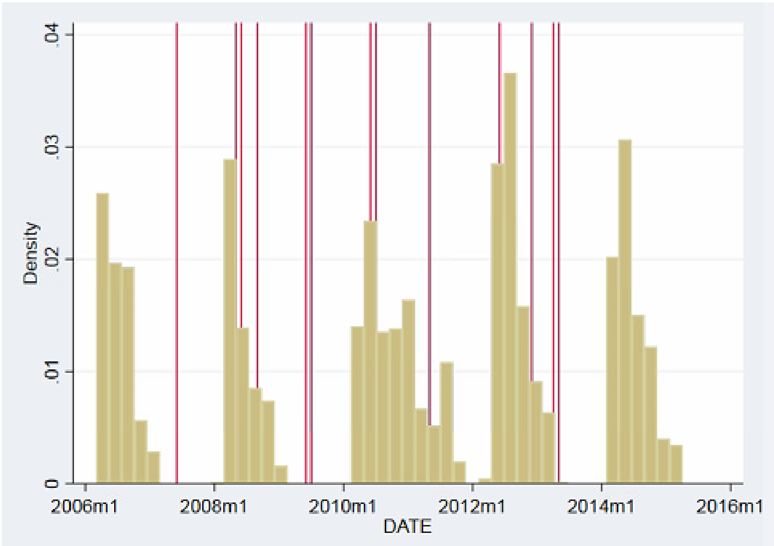
$$PeakML_{dst} = 1 \quad if \quad ml_{dst} \geqslant k * \overline{ml}_{ds} \tag{1.2}$$

As we will describe in more detail in the following section, we rely on the timing of the peaks among state-industry pairs to identify the effect of negative labor market shocks on one's employment status. The reader will have noticed that $PeakML_{dst}$ allows for multiple peaks within state-industry pairs but, exploiting the panel dimension of our dataset, we will focus on the first peak along the considered time frame as measure of negative labor market shock. This reconciles the need to capture shocks throughout the entirety of the considered waves as well as to emphasize the consequences of the Great Recession for those industries that are typically more affected by the economic cycle.

It comes with no surprise that industries such as manufacturing, retail trade or construction can suffer more the economic cycle with respect to industries such as healthcare or education for which lay-offs, if present, are more likely to be due to

[5] For robustness purposes, we provide results using alternative values for the factor k.

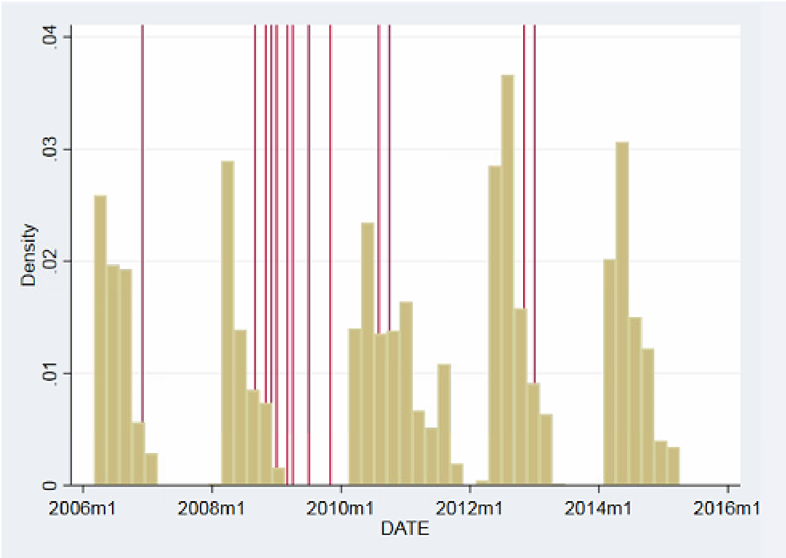
Figure 1.3: Interviews distribution over time and mass lay-offs peaking time for the Health Care and Social Assistance sector.



regular turnover or seasonality than macro economic conditions. In the first case we do expect peaks to appear really close to recessions, no matter how they are actually measured. On the other hand, as the industry is not immediately affected by the economic cycle, mass lay-offs could either manifest later on, not happen at all or show some kind of repeated pattern just above the zero without any clear peak. Defining a peak as in (2) with large enough k allows to exclude these minor movements and focus on more salient cases.

Figure 1.3 and Figure 1.4 describe interviews distribution and highlight the timing of peak in mass-layoff as described in equation (2) across different States for the two most popular industries in the sample: health care and manufacturing. As anticipated, for manufacturing most of the States exhibit a peak in mass lay-off around the Great Recession, between late 2008 and 2009. For the health care sector instead, peaks are more evenly distributed. Another interesting point to emphasize is the interviews distribution. For each wave to be collected, interviews take no less than a year, giving us the opportunity to exploit the combination of heterogeneity in the timing of the peaks and timing of interviews for identification. As a consequence, our identification of the effect of labor market shocks do not solely exploit peaks happening at different times between waves but also during waves.

Figure 1.4: Interviews distribution over time and mass lay-offs peaking time for the Manufacturing sector.



1.5 Methodology

Rather then starting directly with the aim of a specification for the interaction between the polygenic score and the environmental shock, for the sake of clarity and interpretative purposes with respect to the final results, we firstly focus our attention to the proper identification of the labor market shocks' effect on unemployment status. In this context the main issues for a credible identification are simultaneity, omitted variable bias and selection. The first one implies a bi-directional causal link between treatment and outcome which in our case are peaks in mass lay-off and individual unemployment status. To cope with this threat, first of all we rely on the nature of our measure for negative labor market shocks. By considering peaks of mass lay-off we use information at State-industry level over a 9 years time period to identify an event which predicts individual level probability of unemployment. While at first glance the simultaneous link between the two seems obvious, it is indeed hard to claim that a single worker displacement is not simply a cause of a single mass lay-off but a peak in mass lay-offs for a given industry in a given State. It is worth recalling that a mass lay-off is registered whenever a single firm displaces at least 50 employee within a five weeks period. A further but by no mean less important factor which makes us confident in the exogenous nature of our shock variable (at least with respect to simultaneity concerns) is the particular time frame under consideration which highlight the role of the Great Recession (e.g. Figure 1.4).

Before tackling the other two issues it is useful to introduce a first tentative approach in modelling probability of unemployment. Given the panel structure of our data, the obvious starting point is a difference in difference approach (DD).

$$Y_{it} = \beta_0 + \beta_1 T_i + \beta_2 Post_t + \beta_3 (Post * T)_{it} + \beta_8 X_{it} + \epsilon_{it} \tag{1.3}$$

Y_{it} is a dichotomous variable being equal to one if individual i is unemployed at the time of the interview t and zero otherwise. T_i is a dichotomous variable equal to one if individual i has been exposed to a shock in the labor market.[6] $Post_t$ is another dichotomous variable being activated once individuals in the treated group have received the treatment, $(Post * T)_{it}$ is the interaction term X_{it} represents a vector of time varying controls for individual i and ϵ_i an error term.

The model in equation (3) suffers from two important problems. To begin with, it does not consider the structure of labor market shocks data. As treated workers are affected by labor market shocks at different times depending on which State-industry pair they belong to, it is not possible to identify $Post_t$ for non treated individuals. In other words, equation (3) neglects the main source of variation we can exploit to identify the environmental shock, that is time variation in the peak of mass lay-off across States and industries. A second major problem is omitted variable bias which may arise from T_i being correlated with the error term. Suffering a negative labor market shock in a given State and industry could in fact be correlated with unobserved individual characteristics spanning from preferences over the choice of industry to work in, the choice of the State or even to skills as well as familiar socio-economic status. To some extent the same problem could be seen as a selection issue but for completeness we will specifically address this when considering the genetic endowment. Of course, the choice of

[6]Note that a this stage this simply means belonging to a State-industry pair for which we identify a peak in mass-lay-off at some point between 2006 and 2015.

the industry and State can not solely be due to preferences or skills as educational attainment and its genetic predisposition are likely to play a role.

To face these problem a second step is to exploit the panel dimension of the data by estimating a generalized difference in difference with treatment at different times by group with a fixed effect (FE) model as in equation (4). This allows to avoid omitted variable bias as long as the omitted information correlated with the shock variable is time constant.

$$
\begin{aligned}
Y_{it} = \beta_0 + \beta_1 Post.PeakML_{idst} + \beta_2 X_{it} \\
+ a_i + \lambda_t + \gamma_{ds} + \nu_{dst} + \epsilon_{it}
\end{aligned}
\tag{1.4}
$$

$Post.PeakML_{idst}$ is a dummy equal to one if individual i in industry d and State s has suffered a peak of mass lay-off in between time t and $t-1$. The variable is zero for every individual in the sample at first observation in 2006, it gets a value of one once a peak happens in individual i's State-industry group and it remains equal to one throughout the available waves. As a consequence, $Post.PeakML_{idst}$ captures the effect of being exposed to a negative labor market shock on the probability of unemployment. Following Hansen (2007) and Bertrand et al. (2004), equation (4) includes an individual specific term a_i, time effects λ_t, group effects γ_{ds} as well as a set of time/group interactions ν_{dst} and an error term ϵ_{it}.

In our case groups are indeed the product of States and industries pairs. Considering that in the United States there are 50 States and according to the the two digits NAICS industry classification we can distinguish 20 industries, on paper this translates into a thousand groups. If this was not enough of a concern, considering the interaction between groups and time effects we would theoretically have four thousand control.[7]

For feasibility constraints as well as interpretative concerns we follow Imbens and Wooldridge (2008) and Bertrand et al. (2004) and disregard ν_{dst} assuming individual level observation as independent. Again, it is not feasible to estimate a model including all of our group fixed effects γ_{ds} and, even if possible, that would most likely be a huge source over-fitting considering the dimension of our panel.

[7] One out of the five waves would be considered as reference category.

As a second best we control for industry effects ρ_d and State effects η_s which, nevertheless are the main components of our groups. This leads to equation (5).

$$\begin{aligned} Y_{it} &= \beta_0 + \beta_1 Post.PeakML_{idst} + \beta_2 X_{it} \\ &+ a_i + \lambda_t + \rho_d + \eta_s + \epsilon_{it} \end{aligned} \tag{1.5}$$

Notice that both in equation (4) and (5), $Post.PeakML_{idst}$ is equivalent to $(Post{*}T)_{it}$ of equation (3). The main difference between (3) and the subsequent models (4) and (5) lays on the fact that in the latter models individual time constant characteristics are accounted for by construction thanks to the within estimator. Interestingly enough, this implies that not only that T_i is wiped out while time trend are accounted for by λ_t but also that we would not be able to explicitly include any individual time constant variable such as a polygenic score.

To cope with this we dichotomize the poligenic score at one standard deviation creating $HighPGS_i$: a dummy variable equal to one for individuals with polygenic score beyond one and zero otherwise and interact it with $Post.PeakML_{idst}$. The resulting model (6) resembles a generalized triple difference in difference (DDD) with treatment occurring at different times.

$$\begin{aligned} Y_{it} = \beta_0 &+ \beta_1 Post.PeakML_{idst} + \beta_2 X_{it} \\ &+ \beta_3 HighPGS_i * Post.PeakML_{idst} \\ &+ \beta_4 College_i * Post.PeakML_{idst} \\ &+ a_i + \lambda_t + \rho_d + \eta_s + \epsilon_{it} \end{aligned} \tag{1.6}$$

Notice that in order to assess the additional informative power of the polygenic score we would like to control for formal education but, as for any other individual specific time constant variables, the within estimator cancels it out. Consequently, as for the polygenic score we create a dummy variable for having obtained a college degree and we interact it with the shock variable. In terms of interpretation β_1 has positive expected sign and represents the increase in the probability of unemployment due to having suffered a negative labor market shock as defined by a peak in mass lay-off in a given state-industry pair for individuals without a college degree and a polygenic score for educational attainment below one standard deviation. Conversely, β_3 would have negative expected sign and represent the change in probability of unemployment when suffering a shock for individuals with a polygenic score above one standard deviation. β_4 would also have negative expected sign and represent the change in probability of unemployment when suffering a shock for individuals with a college degree.

While simultaneity and omitted variable bias have been tackled with regard to our identification of labor market shocks, few consideration should be made with regard to the interpretation of the polygenic score introduced from equation (6). Notice that since we are not able to control for parental genotype for individual in our sample, this should not be intended as being in the pursuit of a causal interpretation of the genetic endowment but rather we hope to precisely highlight which aspect represents the more salient threat to assessing the protective effect of high polygenic score for educational attainment.

It is quite straightforward that being the sequence of SNPs determined at conception, simultaneity is not an issue for the interpretation of β_3. As a matter of fact, there is no way labor market outcomes in late working age could be a causal factor for a polygenic score which is in fact fixed at conception. On the other hand, omitted variable may be an issue as long as the omitted information is both correlated with the polygenic score and probability of unemployment and it is time variant. Of course, a major candidate would be industry or state specific labor market fluctuations which we already control for. Formal education would also be a candidate but as mentioned above we already control for it both directly including an interaction term of a dummy for college degree and labor market shock and indirectly using the within estimator.

The most problematic aspect for interpreting the β_3 in equation (6) as protective effect of genetic predisposition to education with respect to negative labor market shocks is given by selection. As the exposure to shocks (as well as its timing) is a function of the combination of State and industry to which the worker belongs to, one may argue that β_3 reflects the selection into specific State-industry groups rather than a some protective effect rooted into skills or human capital. With this regard it is worth noticing that in model (6) the poligenic score dummy is itself interacted with $Post.PeakML_{idst}$. This means that β_3 describes the change in probability of unemployment for those who actually experienced a shock having a high polygenic score. In other words it captures the effect of heterogeneity in the polygenic score within those experiencing a peak in mass lay-off as described in (2). However, while our shock measure identifies spikes beyond three times the average of mass lay-off, it does not distinguish differences across industries and States that could in principle be present above this threshold. This is to say that shocks as identified via peaks in mass layoff might indeed differ across states and industries in intensity and actual labor market consequences in ways we cannot control for.

Furthermore, while selection as a function of polygenic score might happen in a non mutually exclusive way both for geographical location and industry, in our framework it seems reasonable to say that the most salient channel would be selection into industry. In a time frame surrounding the Great Recession, the propagation of negative macro economic conditions, and so mass lay-off, would first happen at industry level rather than geography. Interestingly, as we will see in the following of the study we do find evidence of selection into industry which nevertheless, does not seem to be a driver of our result.

Another significant aspect that we should take into account for the interpretation of the results is the definition of unemployment. As mentioned above, the outcome of the models from (3) to (6) is self-reported unemployment at time of the interview. This should rise two concerns: on the one hand it neglects other important possible outcomes of late working age such as early retirement or disability; on the other hand it overlooks what happens in between two waves. Since individuals are interviewed approximately every two years, using self-declared unemployment status does not tell a complete story of what happened in between two subsequent interviews. In fact, a worker might well go unemployed for few months and then find a new job before the next interview, or use unemployment as a corridor for retirement.

For the sake of the first point we not only look at unemployment but rather we turn also to broader changes in working status, considering as outcomes for alternative specifications having experienced any change in working status and going out of the labor force, with the latter being defined as either self-declared retired or self-declared disabled. In general one would expect both education and polygenic score to be protective against being affected by negative labor market shocks but being affected by such shock could mean different things for different workers. While one might expect genetic predisposition to educational attainment (being associated with predisposition to accumulate human capital beyond formal education) to be protective with respect to unemployment, then same might not hold true, for example, with early retirement. For example, factors not necessarily related to polygenic scores such as contribution years could in fact play a role in the decision of taking the chance to retire when a recession or a mass lay-off knocks at the door. Along this line, to the extent to which contribution years are related to, for example, to formal education this would turn out to be the only protective characteristics in this case. As we will see this seems to be the case as it is expectable that individuals with less educational achievements entered earlier on in the labor force keeping age constant. This would mean that as a labor market shock kicks in these individuals could be more likely to accept early retirement agreement to accommodate firm's contingent needs.

With regard to the second point mentioned just above we propose robustness tests looking at a more encompassing definition of unemployment. Rather than looking at self reported unemployment status, we take into account the receipt of unemployment benefit claims in the last calendar year as a proxy for unemployment. Yet, since the question is only limited to the last calendar year we still are not able to get a complete picture of the entire two years before the interview. Nevertheless this allows us to capture a significant larger amount job losses. Note that we do not adopt such a definition of unemployment for the main set of results because it would not be consistent with the other set of job market outcomes we investigate.

In case of significant protective effect of the polygenic score in the main set of results, another crucial problem to be addressed would be the channel through which this relation is actually working. As our interviews are carried out every two years one could think of individuals with higher polygenic score to be either less likely to loose job when a shock hits the State-industry pair or, alternative, even if displaced with the same likelihood, they might be more likely to find re-employment before the next interview. while we are aware that our setting is not the perfect one to disentangle these two cases we propose a specification that goes in the direction of providing at least some suggestive evidence.

$$\begin{aligned} W_{it} = \beta_0 + \beta_1 JL_{it} + \beta_2 X_{it} \\ + \beta_3 HighPGS_i * JL_{it} \\ + \beta_4 College_i * JL_{it} \\ + a_i + \lambda_t + \rho_d + \eta_s + \epsilon_{it} \end{aligned} \tag{1.7}$$

Equation (7) is very similar to equation (6) except that for two aspects. First of all, rather than modelling the probability of unemployment as self- reported unemployment status we build a W_{it}: a dichotomous variable taking value one if in between two subsequent waves an individual has received some unemployment benefit but, by the time of the following interview he/she reported to be either employed of self-employed. This, rather than adopting a broader definition of unemployment means to focus on those cases in which workers did in fact experienced some unemployment (which is proxied by receiving unemployment benefit) but where able to find another job before the next wave. The second difference is that rather than regressing the outcome on a measure of labor market shock we regress it on a dummy for having lost the job in between two subsequent waves and interact it with the usual dummy for a polygenic score beyond one standard deviation and one for college degree. Notice that in equation (7), β_1 is by construction heavily correlated with the outcome W_{it} which is in fact build as an intersection of individuals who have lost their job in between interviews (according toJL_{it}) but who have a job at the time of the interview. So, while in (7) the significance of β_1 would not be an interesting result, what we aim to test is in fact the significance of $beta_3$. In fact, the significance of the coefficient of the interaction between $HighPGS_i$ and JL_{it} would provide at least suggestive evidence of the polygenic score being protective against unemployment not only by making job loss less probable but also by increasing the probability of re-employment after a job loss.

To test against the possibility that our results are a product of the particular way we identify the labor market shock measures or the point in which we set the threshold of the polygenic score dummy, we propose a robustness based on county level unemployment rate and non-dichotomized polygenic score for educational attainment.

1.6 Results

Table 1.3 presents estimates as from model (6) for changing working status, going out of the labor force, and going unemployed. Starting from the whole sample, we see that experiencing a peak of mass lay-off in between two subsequent waves increases the probability of experiencing a general change in the working status by 3.6%. The interaction between the shock and the dummy for a high polygenic score has the negative expected sign but it is not statistically significant. On the other hand, the interaction with the college degree has negative expected sign and it is statistically significant describing a decrease in the probability of changing working status of about 7%. To see what is hidden behind this change of working status we distinguish between going out of the labor force (i.e. retirement or disability[8]) and going unemployed. Looking at the second and third column of the top panel in Table 1.3 we see an interesting pattern. The coefficient for experiencing a peak in mass lay-off is, if anything, only barely significant in the model for unemployment; its interaction with the polygenic score dummy is significant with negative expected sign only for unemployment while the interaction with the college degree dummmy is only significant in the model for exiting the labor force.

While the limited significance of the non interacted shock measure is not per se a concern, the alternative significance of the polygenic score and college degree in the last two specifications needs further interpretation. Moreover, one would expect formal education to be negatively associated with unemployment when hit by a negative labor market shock. The estimates suggest that workers with a college degree are less likely to exit the labor force when exposed to a labor market shock. In fact, as a worker become older his/her contribution years (and/or savings) increase and the sooner a worker entered the labor force, the more likely

[8]While we are aware that suffering a certain degree of disability per se does not preclude employment, we use the self-reported disability status as alternative to the employed self-employed one to assume the exit from the labor market due to disability. Moreover, we grouped it with retirement to acknowledge the known channel of early retirement through disability benefit.

he/she will be to retire before 65 years of age. It is straightforward that since the workers with a college degree entered later on into the labor force with respect to those without college degree, the former will be less likely to be ready to take the chance to retire when exposed to negative labor market conditions in the 50 to 65 age range. If this is true then, it comes with no surprise that the interaction with formal education does not have any effect on the probability of unemployment while, the only significant interaction for individuals remaining in the labor market is that with the polygenic score. In other words, as workers experience a negative labor market shock, those with lower educational attainment, being entered erlier on in the labor market, will be more likely to exit the labor market and retire. On the other hand, those with higher educational attainment will more likely seek to remain in the labor force as they entered it later on in their life. If in this setting of older workers educational attainment is mainly associated with lower chances of retirement, having a polygenic score beyond one is associated with lower chances of unemployment. Finally, it should be noted that while the R^2 in the models for change in working status and exit from the labor force are respectively 26% and 28%, in the model for unemployment the goodness of fit remains at 6%.

Following Papageorge and Thom (2019) we further split the sample by gender in order to acknowledge possible different selection paths into education and/or labor market which are likely to be present for the cohorts under consideration. It is worth noticing that in contrast to the above mentioned authors, for the sake of completeness we decide to include women in the analysis sample. The resulting estimates are reported in the middle and bottom panel of Table 3. It emerges quite clearly that the pattern observed in the whole sample is heavily driven by men. In the middle panel we see the lack of significance and almost zero point estimates for the mass lay-off shock along every scenario of analysis. The interaction with the polygenic score dummy remain insignificant while the one with college degree is significant and with expected sign only in the model for exiting the labor force. In the case of male workers instead, experiencing a labor market shock is associated with a 7.1% increase in the probability of general change in the working status. When the shock is suffered by an individual with a polygenic score greater than one, this probability decreases by 8% while the change in probability for an individual with college degree decreases by 8.5%. With regard to exiting the labor force, as for the previous two cases, only the interaction with college degree is significant with a point estimate of -7.6%. More interestingly, looking at the estimates for unemployment, the statistical significance is stronger if compared with the panels above and experiencing a shock in the labor market in between two subsequent waves is associated with an increased probability of unemployment of about 3.9% while, this probability decreases by 6.2% if the worker has a high polygenic score for educational attainment. As for the other cases, no significant association is detected for the interaction with college degree. Notice that the point estimate for the interaction with the polygenic score more than offset the estimate of the coefficient for the shock alone.

As mentioned in the methodological section of the study, so far we have been considering only job status at time of interview which happens every two year. For this reason we might be losing track change of working status that happens and resolve within the same waves. This is to say that while we show that at least for male workers having a high polygenic score is associated with lower chances of being re-interviewed while unemployed, it says little of weather this is due to lower chances actually experiencing unemployment in the first place or also higher chances of finding an employment once again after being displaced. To dig deeper in this direction we exploit a proxy to detect the experiencing of unemployment status disregarding the actual job status at the time of the interview that is having received any unemployment benefit during last calendar year.

The first three columns of Table 1.4 reports estimates results for linear probability models for having experienced some unemployment as proxied by unemployment benefit receipt. Here we see a very similar pattern as in Table 1.3 with the interaction between the shock and the dummy for high polygenic score being significant with comparable size effect as in Table 1.3. The last three columns of Table 1.4 instead report estimate for the model described by equation (7) where the dependent variable is a dummy for having experienced unemployment and having found another job in between the same two waves. In these columns we focus on experiencing unemployment as main variable of interest to build

Table 1.3: Fixed effect linear probability model for labor market outcomes and mass lay-off events

	Changed working status	Out of the labour force	Unemployment
	Whole Sample		
Post.PeakML	0.036**	0.019	0.015*
	(0.017)	(0.015)	(0.009)
Post.PeakML * HighPGS	-0.013	-0.006	-0.033**
	(0.030)	(0.027)	(0.012)
Post.PeakML * College	-0.069***	-0.061***	0.005
	(0.023)	(0.022)	(0.011)
constant	11.823***	13.254***	-1.045**
	(1.149)	(1.060)	(0.531)
N	8499	8499	8499
R2	0.26	0.28	0.06
	Women		
Post.PeakML	0.008	-0.003	0.002
	(0.022)	(0.019)	(0.011)
Post.PeakML * HighPGS	-0.016	0.030	-0.022
	(0.040)	(0.036)	(0.018)
Post.PeakML * College	-0.073**	-0.060**	0.016
	(0.035)	(0.033)	(0.017)
constant	10.137***	11.829***	-1.398*
	(1.495)	(1.371)	(0.643)
N	4850	4850	4850
R2	0.27	0.30	0.08
	Men		
Post.PeakML	0.071***	0.038	0.039***
	(0.027)	(0.024)	(0.013)
Post.PeakML * HighPGS	-0.080*	-0.036	-0.062***
	(0.048)	(0.043)	(0.016)
Post.PeakML * College	-0.085***	-0.076**	-0.008
	(0.034)	(0.032)	(0.015)
constant	15.011***	15.254***	-0.011
	(1.877)	(1.783)	(1.004)
N	3649	3649	3649
R2	0.29	0.31	0.09

Note: all specifications include the following controls: time dummies, marital status, age, age squared, number of children, a dummy for ever being diagnosed with diabetes, industry dummies, state dummies, income (all source) and tenure with last employer or job.

Table 1.4: Fixed effect linear probability model for experiencing unemployment and re-employment in between subsequent waves.

	Experiencing Unemployment (receiving unemployment benefit)			Re-employment after experiencing unemployment		
	W & M	W	M	W & M	W	M
Post.PeakML	0.025*	0.021	0.032			
	(0.014)	(0.017)	(0.023)			
Post.PeakML * HighPGS	-0.038*	-0.021	-0.079***			
	(0.021)	(0.029)	(0.030)			
Post.PeakML * College	-0.010	-0.018	0.000			
	(0.018)	(0.026)	(0.025)			
Experiencing unemployment				0.497***	0.510***	0.480***
				(0.030)	0.041)	(0.045)
Experiencing unemployment * HighPGS				0.068	-0.046	0.253**
				(0.071)	(0.089)	(0.104)
experiencing unemployment * College				-0.033	0.010	-0.118
				(0.058)	(0.075)	(0.089)
constant	-1.167	-0.924	-1.216	0.308	0.766	-0.800
	(0.736)	(0.924)	(1.408)	(0,396)	(0.469)	(0.727)
N	8499	4850	3649	8499	4850	3649
R2	0.05	0.06	0.07	0.48	0.50	0.50

Note: all specifications include the following controls: time dummies, marital status, age, age squared, number of children, a dummy for ever being diagnosed with diabetes, industry dummies, state dummies, income (all source) and tenure with last employer or job.

the interaction. As mentioned in the methodological section, the non-interacted variable for experiencing unemployment is significant by construction throughout the different sample split as it is a necessary but not sufficient condition for the dependent variable to be equal to one. On the other hand, what we are really interested in is, one again, its interaction with the polygenic score dummy. As the reader can see at least for male individual, such interaction has sizeable and significant coefficient with positive sign, suggesting that even when experiencing unemployment, individuals with above one polygenic scoreare statistically more likely to find another job by the time they are interview in the following wave. On the other hand estimates remain difficult to interpret for women.

1.6.1 Robustness

As first set of robustness, in Table 1.5 we repeat the main analysis using data on the unemployment benefit claims as a measure of labor market shock. As already mentioned there is no reason to expect these data to be less than highly correlated with mass lay-off but nevertheless they capture dynamics of the labor market at level of single workers rather than firms, as it is in the case of statistics on mass lay-off events. Overall, in terms of point estimates results seem be in line with those obtained with mass lay-offs. The biggest difference is given by the interaction between the shock and the polygenic score dummy showing some limited statistical significance also in the the model for men exiting the labor force. With regard to unemployment robustness confirm the significance of the coefficient for the interaction term with similar point estimates, even though in this case the non interacted shock measure shows a non significance coefficient also for the men sample split.

As a second set of robustness we address possible concerns towards the study design in relation to the splitting of the sample on the basis of some discretionary value of the polygenic score as well as the identification of shocks in the labor market. To be more explicit, one may argue that results are purely a function of the way the working sample is split with regard to the polygenic score. While splitting the sample at one standard deviation is quite a common practice in many gene environment interaction studies which are mostly characterized by statistical power issues and tiny effects (e.g. Domingue et al. (2017) it is nevertheless fair to address such a concern. Moreover, another possible concern relates to the environmental shock as some may look at the way we identify the labor market shocks as too discretionary as well. In Table 1.6 we address both these concerns by estimating linear probability models for changing working status, exiting the labor force and self-declared unemployment at time of interview as a function county level unemployment rate and its interaction with the polygenic score for educational attainment as continuous measure and an interaction with years of schooling. In this way we are actually ruling out the above mentioned concerns all at once while still exploiting the panel dimension of our dataset. Notice that the Bueau of Labor Statistics only provide access to county level unemployment

Table 1.5: Fixed effect linear probability model for labor market outcomes and initial unemployment benefit claims events

	Changed working status	Out of the labour force	Unemployment
		Women	
Post.PeakUBC	0.029	0.006	0.013
	(0.022)	(0.019)	(0.010)
Post.PeakUBC * HighPGS	-0.002	0.015	-0.013
	(0.042)	(0.035)	(0.023)
Post.PeakUBC * College	-0.058*	-0.060**	0.023
	(0.034)	(0.033)	(0.017)
constant	10.146***	11.884***	-1.365**
	(1.494)	(1.368)	(0.643)
N	4850	4850	4850
R2	0.27	0.30	0.08
		Men	
Post.PeakUBC	0.059**	0.045**	0.014
	(0.027)	(0.024)	(0.013)
Post.PeakUBC *HighPGS	-0.094**	-0.066*	-0.044***
	(0.042)	(0.075)	(0.015)
Post.PeakUBC * College	-0.078**	-0.075***	-0.004
	(0.034)	(0.032)	(0.015)
constant	14.992***	15.294***	-0.063
	(1.883)	(1.790)	(1.001)
N	3649	3649	3649
R2	0.29	0.31	0.08

Note: all specifications include the following controls: time dummies, marital status, age, age squared, number of children, a dummy for ever being diagnosed with diabetes, industry dummies, state dummies, income (all source) and tenure with last employer or job.

rate as yearly data. In this sense, with respect to the main set of results, we are on the one hand gaining geographical variation while on the other hand we are losing time variability.

Starting from column (a) of Table 1.6 we have a coefficient for county level unemployment of 0.812 which is, not surprisingly, highly significant as we are using the change in unemployment rate in the county of residence between two wave to predict the entering into unemployment of a single individual living in that very same area. As county level unemployment is in percentage points, one would expect its coefficient to be very close to one while, in our case the coefficient is appreciably below it. This is most likely due to the fact that while county level unemployment refers to the whole labor market, in this study we restrict to individuals with 50 to 65 years of age. A below average unemployment rate for such cohort would then explain such a deviation from the expected point estimate of the coefficient.

Introducing in the specification the interaction terms we in columns (b) to (d), we obtain a significance pattern which is closely resembles that of Table 1.3 and Table 1.5. The interaction terms in the model for changing working statu are both significant and with expected negative sign. With regard to the modelfor exiting the labor force the interaction with the polygenic score loses significance while the significant ne is the one for educational attainment. Lastly and most interestingly for us, in the model for unemployment (column (c)) the interaction with the polygenic score is significant with sizeable point estimates if compared the coefficient of county level unemployment itself. Such remarkable similarity with the main set of results is extremely promising in coping with the possible concerns highlighted just above.

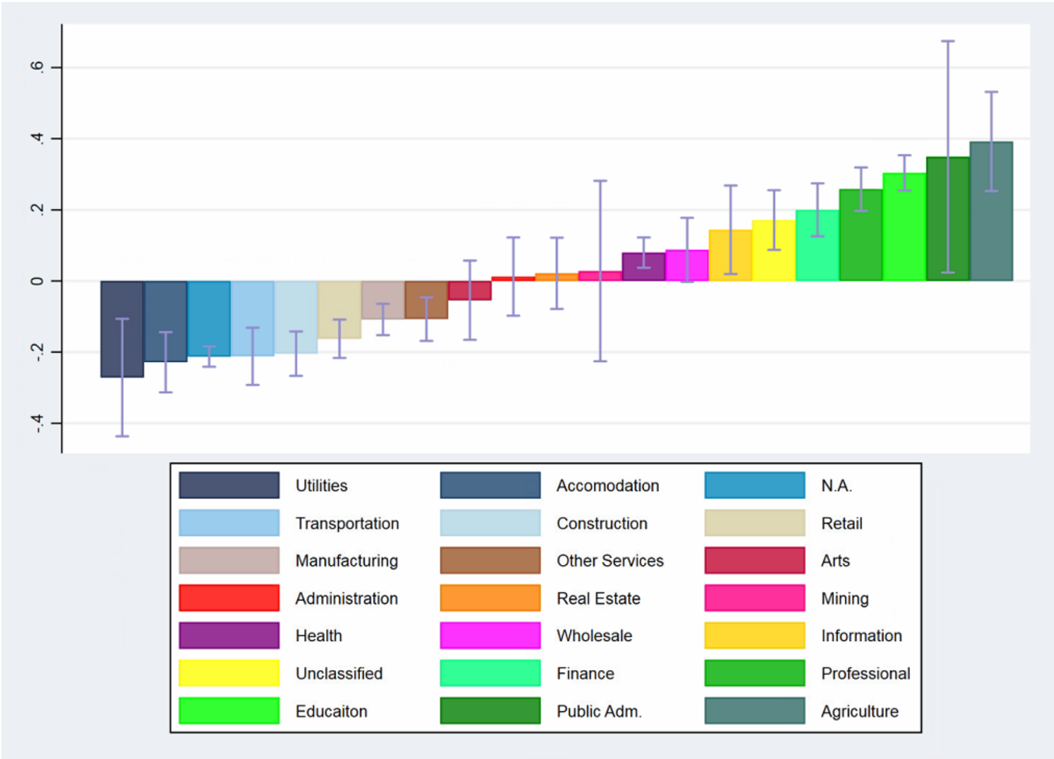
As anticipated in Section 5 an important aspect to consider with regard to the interpretation of the results is possible selection into exposure to labor market shocks driven by the polygenic score itself. To asses whether this could be is indeed an issue we plot in Figure 1.5 the average polygenic score by industry with a 95% confidence interval. We document a surprisingly evident degree of selection into industry. In our working sample the polygenic score is on average below zero for workers in industries like utilities, accommodation, construction,

Table 1.6: Fixed effect linear probability model for labor market outcomes over county level unemployment rates

	Unemployment	Changed Working Status	Out of the Labour Force	Unemployment
	(a)	(b)	(c)	(d)
County UR	0.812***	5.582***	3.677**	1.263*
	(0.093)	(1.352)	(1.236)	(0.683)
County UR * PGS_EA		-0.401**	-0.245	-0.224**
		(0.199)	(0.182)	(0.100)
County UR * EA		-0.374***	-0.242**	-0.075
		(0.091)	(0.082)	(0.046)
Constant	-0,031***	12.280***	13.219***	-0.836*
	(0,006)	(1.157)	(1.064)	(0.469)
N	8499	8499	8499	8499
R2	0.02	0.28	0.29	0.08

Note: County level unemployment are yearly average. All specifications include the following controls: time dummies, marital status, age, age squared, number of children, a dummy for ever being diagnosed with diabetes, industry dummies, state dummies, income (all source) and tenure with last employer or job.

Figure 1.5: Polygenic score for educational attainment by industry of occupation. HRS waves 2006 - 2014.



retail and manufacturing. On the other hand, industries like health services, finance, professional services, education, public administration and agriculture have average polygenic score above zero. With the exception of the latter one, the resulting ranking is somehow close to what one could expect, with workers having below average score employed in sectors where the labor market is typically more volatile or where temporary or low skill jobs are more present. On the other hand it is worth noting that such pattern might reflect, at least to some degree, intergenerational persistence of occupational choices within families.

To the extent to which the belonging to an industry as function of the polygenic score determines in the State-industry pairs a systematic pattern of exposure to labor market shocks, our estimates could in fact reflect such a selection rather than any protective effect. While there is no perfect way to test against this, in Table 1.7 we report estimated results of OLS models where the polygenic score and a dummy for a score above 1 (as used in Table 1.3 to 1.5) are alternatively regressed on the shock measure and a dummy for being exposed to a shock at some point. This, controlling education and all the usual controls used throughout the paper. Estimates show that conditionally on educational attainment or college degree and all the other controls (including States and industry dummies) our labor market shock measure do not significantly correlate with the polygenic score or a dichotomous transformation of it. This is sound evidence that, despite the documented non negligible degree of selection in industries, our results can not anyhow driven by correlation between our measures of labor market shocks and the polygenic score.

Table 1.7: OLS for polygenic score over labor market shock measures.

	PGS for educational attainment			
Post.PeakML	0.016	0.009		
	(0.040)	(0.040)		
Ever exposed to a peak in ML			0.037	0.024
			(0.050)	(0.050)
College	0.485***		0.484***	
	(0.047)		(0.047)	
Educational attainment		0.119***		0.118****
		(0.011)		(0.011)
Constant	3.184	2.329	3.443	2.495
	(2.147)	(2.119)	(2.142)	(2.115)
N	9216	9216	9216	9216
R2	0.12	0.13	0.12	0.13
	HighPGS			
Post.PeakML	-0.006	-0.008		
	(0.015)	(0.015)		
Ever exposed to a peak in ML			0.009	0.006
			(0.019)	(0.019)
College	0.160***		0.159***	
	(0.019)		(0.019)	
Educational attainment		0.034***		0.034***
		(0.004)		(0.004)
Constant	0.168	-0.141	0.239	-0.094
	(0.792)	(0.790)	(0.793)	(0.792)
N	9216	9216	9216	9216
R2	0.08	0.08	0.08	0.08

Note: all specifications include the following controls: time dummies, marital status, age, age squared, number of children, a dummy for ever being diagnosed with diabetes, industry dummies, state dummies, income (all source) and tenure with last employer or job.

1.7 Discussion

Results highlight a lower probability of unemployment after being exposed to a negative labor market shock for individuals with polygenic score above one standard deviation. For male workers the decrease in probability is significant and between 4.4% and 6.2% depending on whether labor market shocks are measured with unemployment benefit claims or mass lay-offs. For women the decrease is estimated between 1.3% and 2.2% but it does not reach statistical significance. The lack of a significant estimated coefficient for the interaction term among female workers is consistent with previous studies from the same literature disregarding such group because of systematic selection patterns into education and labor force for specific cohorts.

Results are consistent with an interpretation of the polygenic score for educational attainment that goes beyond simple predisposition for education. In fact also in our study, the polygenic score remain significant and predictive of outcome observed

On the other hand, it is worth remembering that mainly for reasons of constraint in data availability rooted in the current developmental stage of genomic science we are using a considerably selected sample that limits the external validity of our results as well as possible policy implication.

1.8 Conclusion

In this study we provided evidence of interaction between individual level genetic information on predisposition to educational attainment and labor market data in predicting labor market outcomes in later working life. After controlling for education, a polygenic score above 1 implies a 4.4 to 6.2% decrease in the probability of unemployment when exposed to a negative labor market shock as measured via unemployment benefit claims or mass lay-off respectively. On the other hand, no significant interaction is found for women. Education plays a role only with respect to transitioning out of the labor force when exposed to a labor market shock. Results are robust to alternative definitions of environmental shocks and unemployment. Furthermore, we present suggestive evidence of high polygenic score decreasing the probability of unemployment both via decreasing the probability of losing the job in the first place and via increasing probability of re-employment after a job loss. We document a non negligible degree of selection into industries as function of the polygenic score with lower average polygenic scores concentrated in more volatile sectors or sectors in which is traditionally more likely to observe temporary job contracts. Nevertheless, such selection is unable to explain our results as after considering all the available controls, the polygenic score is not significantly correlated with the exposure to our measure of negative labor market shock.

Chapter 2

Estimating the Causal Effect of Retirement on Frailty

Abstract

In this study we estimate the causal effect of retirement on frailty, a syndrome defined as a multi-systemic health deficit accumulation with good predictive properties with respect negative health outcomes like further health decline, dependency and death. We use data from the Survey on Health, Ageing and Retirement in Europe (SHARE), instrumenting retirement with the reaching of early and statutory retirement eligibility criteria in nine countries. Exploiting the panel structure of the dataset we are able to exploit both cross and within country variation in such criteria. Results highlight a significant protective effect of retirement among male retirees. Results are robust to alternative definitions of retirement and to partial definitions of the frailty index.

2.1 Introduction

Ageing represents one of the many challenges that both developed and developing countries will face in the next decades. To give a sense of the magnitude of the phenomenon, for European countries, old-age dependency ratio (the share of people 65 and older over the population in the 18-65 age range) is expected to increase from less than 30 percent in 2015 to more than 50 percent by 2050 while countries such as Greece, Portugal and Italy are expected to reach respectively 70, 65 and 62 percent (Eurostat 2015). This, in addition to increased life expectancy, poses a number of question both regarding quality of life and sustainability of the healthcare sector.

In such a context, frailty represents a central aspect in describing the health decline of an elder individual, with considerable documented consequences for the healthcare system (Comans et al. (2016), Bock et al. (2016)). Frailty is a clinical syndrome entailing a vulnerable health status resulting from a multisystem reduction in older people's health capacity or physiological reserve (Fried et al. (2001), Staudinger et al. (1995),Lally and Crome (2007),Theou et al. (2015), Harttgen et al. (2013)). A frail individual is affected by a pathological multisystemic syndrome different from aging itself (meant as "chronological" ageing), stemming from both genetic (Viña et al. (2016), Inglés et al. (2019)) and environmental stressor factors (Fried et al. (2001), Staudinger et al. (1995) and Roiland et al. (2015)).

A common definition of such a concept in the applied literature is the one by Fried et al. (2001). As pointed out by Sirven (2012), it is mainly because of its distinction from disability and comorbidity, the easily implemntable operationalization and, finally, its parsimony. In Fried et al. (2001), the definition of frailty relies on only five dymensions: weakness, exhaustion, slowness, low physical activity levels and shrinking. A frailty index going from zero to five synthetize into a single measure the realized status.[1] The aim of the present study is to identify the causal effect of

[1] A valuable alternative to Fried et al. (2001) is the definition of frailty given by Rockwood et al. (2007) which although being extremely encompassing (it is measured over about 70 variables covering also cognitive changes, attitudes and behavioural risks in addition to physiologic status) it is hardly implementable without a professional health assessment.

retirement on frailty as an important outcome for the health trajectory in old age population.

2.2 Literature

From a theoretical standpoint frailty has already been introduced by Strulik (2015) in the economic literature as a model of ageing and longevity which highlight the role of deficit accomulation as opposed to that of health capital accumulation as in Grossman (1972) and most of the following literature in health economics.

The relevance of frailty in both geronthological medicine and social sciences is due to its predictive power with respect to disability and health outcomes in general such as dependence, falls and finally death (Bergman et al. (2007)) as well as to the opportunity for early detection and reversibility (Fried et al. (2004)). In addition, as anticipated above, letting aside the epidemilogical and medical realm of interest, recent studies found positive association between frailty and healthcare costs (Bock et al. (2016), Comans et al. (2016), Sirven and Rapp (2016)). Interestingly, the latter seem to be driven by a combination of disabilities, chronic conditions and frailty. This contrast the belief that age is the main driver of healthcare cost among elderly population. Bock et al. (2016) provide evidence of how frailty is associated with increased health care costs and highlight frailty as one of the main factor for healthcare costs independently from pure age. Overall thi indicates how the overlapping concepts of multimorbidity and frailty are necessary to explain health care use and corresponding costs among older adults. Comans et al. (2016) analyze the cost of frailty by comparing, on the basis of resource use data, patient cohorts respectively entering a community-based post-acute program and entering residential care. Findings confirm association between pre-frailty and frailty statuses and increase in healthcare costs. Finally, Sirven and Rapp (2016) investigate the incremental cost of frailty with respect to ambulatory health care expenditures in the 65 and older French population in 2012. Findings suggest frailty's significant additional explanatory power toward expenditures whatever other health covariates are considered, meaning that frailty can indeed represent a source of omitted variable bias whenever it is not accounted for.

In addition to purely financially driven concerns, recent studies in gerontology and aging such as Woods et al. (2013) or Ruan et al. (2015) highlighted the link between physiological reserve and cognitive decline. By this, the authors explicitly claim the existence of a relationship between what they call physical frailty and the so called cognitive frailty: "an heterogeneous clinical syndrome of cognitive impairment that develops in elderly individuals, is caused by physical factors (e.g., physical frailty and pre-physical frailty) and is excluded from dementia resulting from Alzheimer's disease or other conditions" (Ruan et al. (2015)).

Although being still far from an easily implementable and operationally friendly definition for cognitive frailty, this proposed link is suggestive of the central role of physiological functioning in triggering impaired cognitive states. This seems especially interesting considering the lack of any systematic consideration for the concept of frailty in the literature on health effects of retirement which nevertheless investigated some aspects of cognition.

On the other hand, although a not clear causal direction, also cognitive capabilities characterizing the so called executive functions (EF)[2] are expected to play a role in early stages of pre-frail conditions. In fact, looking for a mediation channel between EF and frailty, Roiland et al. (2015) propose (cross-sectional) evidence of stress exposure and regulation function emerging as a significant predictors of pre-frail condition.

Retirement is recognized as a delicate transition period and many studies already attempted to assess its effect on a number of health outcomes. For studies regarding non physically related outcomes of retirement, the main variable of interest have so far been depression (Charles (2002), Belloni et al. (2016)) and cognitive functioning (Coe and Zamarro (2011), Coe et al. (2012), Mosca (2017)) highlighting mixed results.

With regard to physical health, Coe and Zamarro (2011) also looked at general health predicting self reported health status over serveral objective measures of health such as the number of hospitalizations in the last year, obesity, the number

[2]For a comprehensive analysis and definition of Executive Functions please, refer to Diamond (2014)

of chronic diseases and mobility limitations. [3] On the other hand, outcomes such as cholesterol and blood pressure (Behncke (2012)) as well as BMI (Godard (2016)) have also been directly investigated highlighting positive effects of retirement on these outcomes. The most closely related paper to the present one is Bertoni et al. (2018) where the authors assess the causal effect of retirement on the loss of muscle strength, an important component of frailty. Findings point to a short term protective effect of retirement on grip strength but do not consider frailty as a systemic concept in the health of retirees.

Despite a quite extensive literature, most of the empirical literature on retirement and health has so far mainly addressed short term effects while long term ones remain an overlooked question(Avendano and Berkman, 2014). With this regard, investigating the first stages of deficit accumulation in the aging process as described by a frailty score, would allow to shed some light on the long term effects. To the best of our knowledge only few studies investigate the socioeconomic aspects determining frailty (Sirven (2012), Lu et al. (2017) and Kalousova and de Leon (2015)) and even though they seem to point in the direction of the presence of socio-economic gradient Sirven (2012), correlation with employment histories Lu et al. (2017) and psychosocial working condition on the job Kalousova and de Leon (2015), all the three studies fail to give any particular causal interpretation.

Using panel and retrospective data from SHARE accross three waves, Sirven (2012) investigates the determinants of frailty by mean of a Fixed Effect Poisson and Mundlak Random Effect model highlighting the presence of an income gradient having positive effects on frailty and a positive effect given by social capital. On the other hand, the study highlights higher frailty associated to being in the labour force. Interestingly, this seems counter-intuitive with respect to the so called *Healthy Worker Effect* (the selection into retirement of the weaker/ill workers and a consequent healthier remaining labour force) leading the reader to question whether retirement could represent (for some worker) a valuable preventive measure for the considered health outcome.

[3]More on this in Figure 1.1.

Kalousova and de Leon (2015) using wave I and wave IV in a multilevel linear model framework found that working in a position with high effort and low reward predicts the greater increase in frailty while also "effort-to-control" ratio is associated with increased level of frailty. With respect to retirement the authors report a negative effect (decrease in frailty) associated with retiring from working positions with low reward.

Finally, using data from ELSA with multilevel models, Lu et al. (2017) found lower levels of frailty among women who experienced distinct periods of work and family care over the life course. On the other hand, among men, retiring before 65 seems beneficial for slowing down frailty trajectories.

2.3 Data

We use data from the Release 6 of the last three waves (2011, 2013 and 2015) of the Survey of Health, Ageing and Retirement in Europe (SHARE), a multidisciplinary and cross-national panel database with individual information on health, socio-economic status as well as social and family networks. To date, SHARE collected more than 120,000 Computer Assisted Personal Interview (CAPI) covering 27 European countries and Israel.

Our sample includes all individuals aged 50-65 from nine European countries (Austria, Germany, France, Spain, Italy, Slovenia, Belgium, Czech Republic, and Estonia), who were working at baseline and who declared in each wave to be either retired or employed. Thus, we only consider transitions to retirement from employment, neglecting other possible paths such as through unemployment or disability. Moreover, we drop respondents for whom variables needed for the computation of our frailty index are missing as well as respondents with missing values for covariates in our final specification.

2.3.1 Defining Frailty

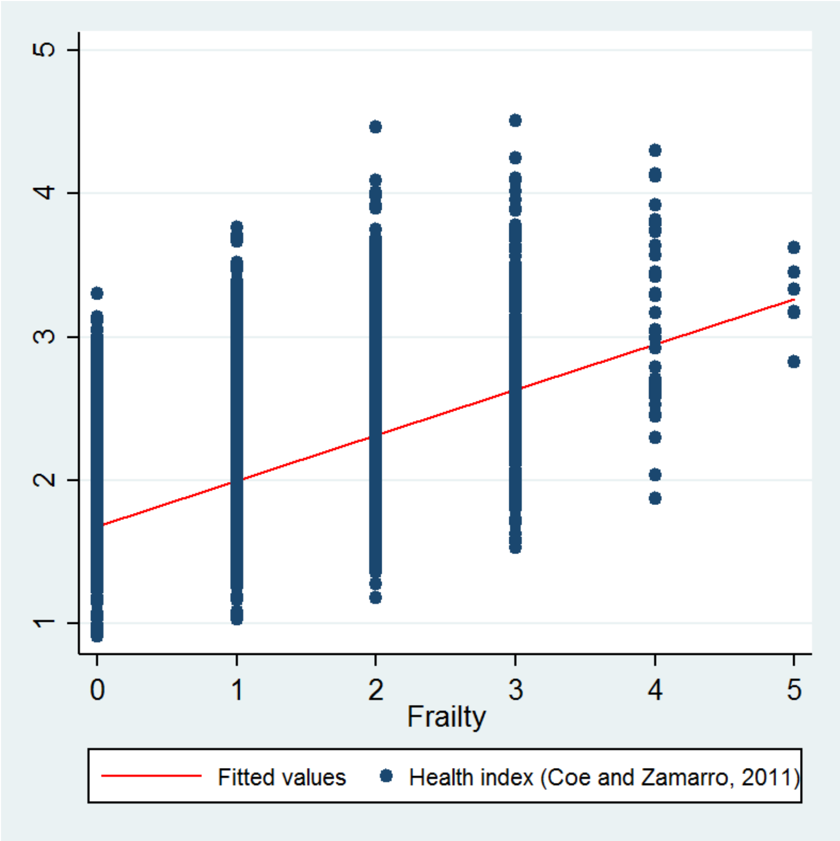
From an operational perspective, we define frailty on the basis of the frail phenotype definition proposed in Fried et al. (2001). As mentioned above, the frail phenotype definition identifies an index going from zero to five according to the number of markers present for the individual. A frailty index equal to zero identifies a fit individual, if the index takes either values of one or two we refer to a pre-frail individual, while for an index equal to three and more we have a frail individual.

The five dymensions of the index are: shrinking, exhaustion, low grip strength (weakness), slowness, and low physical activity. For each of these markers we define a dummy variable whose activation depend on the following measures and answer, as in Santos-Eggimann et al. (2009). The dummy for shrinking identifies individuals reporting a "*diminuition in the desire for food*" to the question "*What has your appetinte been like?*" or "*less*" to the question "*So have you been eatin more or less than usual?*". For exhaustion, the dummy takes value one if a positive reponse is given to the question "*In the last month, have you had too little energy to do things you wanted to do?*". As for grip strength, we attributed a deficit to individuals with a maximum hand grip strength over four trials below some thresholds depending on gender and bmi.[4] The dummy for slowness takes value one if a positive answer was give to "*Because of a health problem, do you have difficulties walking 100 meters*" or "*...climbing one flight of stairs without resting?*". Finally, an individual is defined as carrying out low physical activity if reporting "*hardly ever or never*" or "*one to three times a month*" to the question "*How often do you engage in activities requiring a low or moderate level of energy such as gardening, cleaning the car or going for a walk?*" Overall, this frailty index relies in part on subjective measures of health. Nevertheless, covering different domains of individuals' health, the encompassing and systemic nature of the indicator can rule out, or at the very least attenuate, concerns regarding justification bias.

To further motivate the study we can compare the frailty index computed as just described with the health index estimated in Coe and Zamarro (2011). In their assessment of the effect of retirement on health, the authors predict self-reported health status going from "*excellent*" to "*poor*" with a five point scale over a vector

[4]For men the cutoff are 29 Kg for a bmi below 24; 30 Kg for a bmi between 24 and 28; 32 Kg for bmi above 28. For women, thresholds are 17 Kg for a bmi below 23; 17.3 Kg for a bmi between 23 and 26; 18 Kg for a bmi between 26 and 29 and 21 for a bmi above 29 (Fried et al. (2001)). In this ought to point out that while the measure for grip strength is the only truly objective one that we can retrieve from our data, the BMI measures on which the thresholds are attributed are in fact resulting from self declared height and weight.

Figure 2.1: Frailty vs. Health Index



of objective measure of health including the number of limitations in activities of the daily living (ADLs), BMI, grip strength, EURO-D depression score, number of chronic diseases etc.

Figure 2.1 shows a scatter plot with frailty index on the x-axis and the estimated values of the health index computed as in Coe and Zamarro (2011) on the y-axis. To begin with, despite exploiting partially overlapping information, the two measures are only correlated at about 50%. Moreover, looking at Figure 2.1 we can appreciate how, with respect to the Frailty score, a general health index whose estimates are obtained on the basis of self reported health as an outcome seems to overestimate the erosion of health capacity on the fit end of the spectrum while underestimating it on the other extreme of the distribution.

Given the relatively low age range we consider, the probability of observing a large share of severely frail individuals is ex-ante rather low. In fact, in the cross section of our sample only 1.5% of the individuals exhibit three or more markers contemporaneously while more than 38% (34.4% and 42.5% respectively for men and women) exhibit at least a pre-frail status. Therefore, as anticipated above, considering the systemic nature of the syndrome and so how each marker represents a risk factor for the others in the path toward frailty, we focus on the development of pre-frail status. In order to do so we dichotomize the frailty index between zero and one, identifying the causal effect of retirement on the probability of being at least in a pre-frail status.

As for retirement, we adopt two distinct definitions, in the attempt to assess the robustness of the results. A first one, in line with Godard (2016), is solely based on self-declared status while a second one, further condition on not having done any paid job during the last four weeks. While self-declared status nicely correlates with eligibility criteria and receiving pension benefits, not working for pay in the last four weeks allows to explicitly account for labor supply. A possible problem with the first definition would be given by individuals declaring themselves retired because they left the job that mostly characterized their career while eventually carrying out some other activities.

Table 2.1 describes a sample of 7698 and 7690 pooled observations respectively for men and women. Retirees are not surprisingly on average older and slightly less educated which is suggestive of an earlier entry in the labor market. In terms of socio-economic conditions retiring individuals seem comparable with those remaining in the labor force. Again,not surprisingly, income is appreciably lower among retirees and it is worth noting the significant lower averages for women throughout both groups. In terms of health, the average frailty index seems is 0.46 and 0.56 among retiring men and women respectively against 0.43 and 0.57 among non-retiring male and women. The distribution across fit, pre-frail and frail statuses is very comparable between the two groups both within male and female individuals. Figures in Table 2.1 highlight how, for the given range of age, pre-frail statuses tend to be significantly more common among women then men and finally, as expected given the age we are considering, frailty conditions affect only less than 2% of the working sample. Finally, retiring individuals tend to be more likely to have chronic conditions while the presence of two or more limitations with daily activities only sightly higher for them as compared to non retiring individuals. On the same line, also scores on the EURO-D depression

Table 2.1: Summary Statistics

Variable	Male Whole sample	Male Retired across waves	Male Working all along	Female Whole sample	Female Retired across waves	Female Working all along
Demographics						
Age	60.54	62.43	59.42	60.14	61.90	59.17
Years of education	12.76	12.18	13.11	12.72	12.22	13.00
One child	0.12	0.10	0.13	0.17	0.17	0.18
More than one child	0.55	0.55	0.55	0.58	0.59	0.58
Retirement	0.19	0.51	0.00	0.18	0.51	0.00
Socio-Economics						
MEM very hardly	0.06	0.06	0.06	0.06	0.05	0.07
MEM easily	0.37	0.37	0.37	0.34	0.35	0.34
Income (1000 €)	16.67	12.20	19.33	11.23	8.09	12.96
Health						
Frailty index (0-5)	0.44	0.46	0.43	0.57	0.56	0.57
Fit	0.65	0.64	0.65	0.56	0.57	0.56
Pre-frail	0.34	0.34	0.34	0.42	0.41	0.42
Frail	0.01	0.02	0.01	0.02	0.02	0.02
2+ Chronic cond.	0.32	0.35	0.30	0.34	0.38	0.32
2+ LDA	0.03	0.04	0.03	0.04	0.04	0.04
EURO-D (0-12)	1.59	1.62	1.57	2.33	2.28	2.36
Country						
Austria	0.08	0.09	0.07	0.08	0.13	0.05
Germany	0.12	0.08	0.14	0.12	0.09	0.14
Spain	0.09	0.06	0.11	0.06	0.04	0.08
Italy	0.13	0.12	0.13	0.09	0.06	0.10
France	0.11	0.13	0.10	0.12	0.13	0.11
Belgium	0.14	0.15	0.13	0.12	0.13	0.12
Czech Rep	0.13	0.14	0.13	0.12	0.18	0.10
Slovenia	0.05	0.05	0.04	0.03	0.04	0.03
Estonia	0.17	0.17	0.16	0.25	0.21	0.27

Note: Retirement across waves refers to self-declared retirement status. MEM stands for "Make ends meet". LDA stands for "Limitations in activities of daily living". EURO-D represents a geriatric depression scale going from 0 to 12.

scale are only marginally higher for retiring workers. Overall Table 2.1 describes a situation with retiring workers having on average worse health outcomes as compared to non-retiring individuals. This being said, it does not say anything regarding any possible change occurring around the time of retirement which is indeed the aim of the current study.

2.4 Identification Strategy

Endogeneity is a recurrent threat when estimating any effect of retirement. It may stem from a variety of sources such as omitted variables affecting both retirement and our dependent variable or reverse causality. In the former case we may think to omitted unobservable time preference while in the latter we could expect individuals exhibiting higher frailty indexes being more likely to self select into early retirement.

A first step toward a causal interpretation of the coefficient of retirement would be to exploit the panel dimension of our data considering a fixed effect model. In so doing, we would control for unobserved individual specific constant variables such as time preferences (assuming them to be constant), allowing explanatory variable to be endogenous as far as their endogeneity arises from correlation with unobserved heterogeneity. Still, reverse causality remains a problem. To address reverse causality we estimate a fixed effect instrumental variable model. By instrumenting retirement in a fixed effect framework, we are able to both control for time constant unobserved omitted variables as well as reverse causality.[5]

As for the majority of studies using SHARE data, we rely on instruments based on retirement age thresholds such as in Coe and Zamarro (2011),Mazzonna et al. (2014) andGodard (2016). The main intuition underlying this approach is to exploit discontinuities in the probability to retire around country-specific eligiblity ages. This is to say that we rely on the fact that, as showed by Gruber et al. (1999),

[5]The choice to model individual time constant heterogeneity through a fixed effect rather than a random effect is driven by Hausman test leading to reject the null hypothesis of no systematic difference between the coefficient under the two models with a p-value of 0.044. Using a fixed effect also allows us to exploit functionalities of the *xtivreg2* Stata package like standard errors clustered at individual level and tests for the validity of the instrumental variables.

individuals seem to be willing to retire as soon as possible given their country's retirement rules. Such age thresholds can be easily demonstrated to be relevant in workers' decision to retire and at the same time seem arguably exogenous. We define early retirement age (ERA) as the earliest age - conditional on contribution years - at which individuals are entitled to reduce pension benefits while ordinary retirement eligibility age (ORA) is the earliest age at which workers are entitled to full old-age pension, regardless of contribution history.

Our main model of interest is:

$$Pr(Y_{it} = 1|R_{it}, x_{it}, a_i, \lambda_t) = \beta_0 + \beta_1 R_{it} + \gamma^{'} x_{it} + a_i + \lambda_t + \varepsilon_{it} \tag{2.1}$$

where Y_{it} is bivariate variable for being at least in a pre-frail status for individual i at time t; R_{it} is a dummy identifying the retirement status; x_{it} is a vector of time varying characteristics; a_i is an individual fixed effect while λ_t represents a time dummy and ε_{it} an error term.

We instrument the retirement status R_{it} with two variables defined as:

$$Z_{ict} = 1_{\{age_{it} \geq ERA_{ct}\}} \tag{2.2}$$

$$W_{ict} = 1_{\{age_{it} \geq ORA_{ct}\}} \tag{2.3}$$

Therefore, Z_{ict} and W_{ict} are two dummies describing the eligibility to either early or normal retirement as a function of individuals' age and country-specific eligibility rules at time t. It follows a first stage of the kind:

$$Pr(R_{it} = 1|Z_{ict}, W_{ict}, x_{it}, a_i, \lambda_t) = \alpha_0 + \alpha_1 Z_{ict} + \alpha_2 W_{ict} + \eta^{'} x_{it} + a_i + \lambda_t + \varepsilon_{it} \tag{2.4}$$

where in addition to the instrument just described we have the set of covariates x_{it} as in equation (1) as well as the individual fixed effect a_i, a time dummy λ_t and the error term ε_{it}.

With regard to our instruments, having information on workers from nine countries and three waves, we rely both on cross-country variation in retirement eligibility rules as well as on within country variation. As a matter of fact, in our sample, every country exhibit a shift in at least one eligibility criteria for retirement for either men or women. Interestingly, in contrast to the majority of study on this matter that exploit upward variations in retirement age as seen in many European countries in the last fifteen years, our framework allows to exploit a rather unique downward shift in eligibility criteria for normal retirement in Germany in 2015.

As reported in Table 2.2, each country in the panel faced some sort of change in retirement eligibility rule. For some of them the change is due to reforms meant to ensure long-run financial stability such as for Italy, in other cases instead, we see more subtle shifts reflecting adjustments to life expectancy like for Czech Republic or adjustment to target ages defined by law before 2011.

In the case of Italy for example, the statutory retirement age for men passed from 65 to 66 regardless of the working sector. On the other hand, the statutory retirement age for women changed as function of the working sector getting from 60 to 66 years old for public sector workers, to 63 for self-employed and to 62 for private sector employees. Finally, with respect to early retirement, age requirement passed from 60 to 62 (OECD (2011), OECD (2013b), OECD (2015)).[6]

On the other hand, in Germany statutory retirment age was set to 65 years old in 2011 provided an increasing path reaching 67 in twenty years from 2012. In addition, with an insurance record of at least 35 years[7], early retirement was allowed at 63 years old with a permanent benefit reduction of 3.6% per year of early retirement (OECD (2011) OECD (2013a)). Effectively from July 1st 2014, both male and female German workers with an insured working life of 45 years who were born before 01/01/1953 can retire at 63 years old without any reduction of the pension benefits. Such a threshold is increased by 2 months for each year of birth after 1952 (OECD (2015)). Considering the time span in our analysis this change would affect three cohorts: from 1950 to 1952.

[6]No benefits reduction for early retirement with 42 years and one months and 41 years and one months of contribution respectively for men and women. Early retirement regardless of the contribution period triggers a 1% payments reduction for each year of early retirement up to 60 years old and a 2% reduction per year if retirement occurs before 60 years old. For individuals under the contributive and mixed system with twenty years of contribution the early retirement age is instead 63.

[7]Insured time from employment, child care and child raising periods up to age 10 or periods of short unemployment (OECD (2015))

Table 2.2: Early and normal retirement age

Early retirement age (ERA)						
	Males			Females		
Country	Wave			Wave		
	2011	2013	2015	2011	2013	2015
Austria	62	62	62	57	58.6	59.3
Germany	63	63	63	63	63	63
Spain	60 / 61	60 / 63	60 / 63	60 / 61	60 / 63	60 / 63
Italy	60 / 61	62	62	60 / 61	62	62
France	59	60	60	59	60	60
Belgium	60	60.5	61.5	60	60.5	61.5
Czech Rep	59.3	59.5	59.8	59.6	60.3	61
Slovenia	58	60	60	58	60	60
Estonia	60	60	60	58.5	59	59.5
Ordinary Retirement Age (ORA)						
	Males			Females		
Country	Wave			Wave		
	2011	2013	2015	2011	2013	2015
Austria	65	65	65	60	60	60
Germany	65	65	65 / 63	65	65	65 / 63
Spain	65	65	65	65	65	65
Italy	65	66	66	60	62 / 63 / 66	63 / 64 / 66
France	65	65	66	65	65	66
Belgium	65	65	65	65	65	65
Czech Rep	62.3	62.5	62.8	60.6	61.3	62
Slovenia	58	58	58,6	56	57.3	58.3
Estonia	63	63	63	61.5	62	62.5

Source: OECD (2011-2015) and MISSOC (2011-2015)

For our instruments to be valid, they need not only to be relevant for retirement decisions but also exogenous, that is, related with the outcome of interest only through retirement. In this regard, a threat to our identification could be given by possible systematic discontinuities in workers' health at their countries' early and ordinary retirement ages. However, considering the cross and within country variation for ERA and ORA, this seems highly unlikely. For men, the lowest considered retirement age is 58 years while the highest arrives to 66 years. For women, the range of retirement ages goes instead from 57 to 66. Therefore we are able to estimate the effect of retirement over a considerable range of ages for both genders.[8]

As reported in Table 2.3, looking at the cross-sectional dymension of our sample in column 1, on average 37% of the individuals retired during the considered years. Columns 2-5 exploit instead the panel dymension of our sample to describe respectively the share of workers becoming eligible for early retirement and ordinary retirement (columns 2-3) as well as the share of individuals who become eligible in between two waves and decide to retire in between the same two waves (columns 4-5). Overall, more than 64% of workers who became eligible for ordinary retirement decided to retire in between the same two waves while, only 36% of individuals who became eligible for early retirement decided to retire in between the same waves. This seems suggestive of workers considering less the early retirement age threshold as compared to the normal retirement age in driving their retiring behaviour. While at first this might seem in contrast with Gruber and Köszegi (2001), this could simply reflect a change in behaviour given by heavier disincentives toward early retirement as a result of the many pension system reforms occurring in Europe in the last fifteen years.

Considering the identification strategy and the IVFE model described above, we are able to interpret the estimated coefficient for retirement as a local average

[8] Another possible threat would be given by retirement age thresholds being set as a function of workers' health in each country. While we cannot neglect the fact that position on requirement for retirement could be salient points upon which to set voting behaviour, this seems unlikely given that we are looking at a number of countries where retirement age thresholds have been shifted upward mainly for general financial sustainability constraints. Yet , in any case, such a limitation would apply to every study relying on early and normal retirement age as instrument.

Table 2.3: Eligibility and retirement behaviour

Country	Retired	Become Eligible ERA	Become Eligible ORA	Retired when reaching ERA	Retired when reaching ORA
Austria	0.50	0.17	0.30	0.07	0,18
Germany	0.26	0.06	0.22	0.04	0.12
Spain	0.31	0.15	0.15	0.04	0.14
Italy	0.24	0.11	0.07	0.03	0.06
France	0.41	0.29	0.05	0.16	0.03
Belgium	0.39	0.37	0.13	0.15	0.10
Czech Rep	0,45	0,19	0.33	0.05	0.25
Slovenia	0.44	0.16	0.34	0.05	0.21
Estonia	0.34	0.27	0.30	0.05	0.13
Total	0.37	0.20	0.21	0.07	0.13

treatment effect (LATE) where our treatment is simply retiring between two subsequent waves between which the worker became eligible for either early or normal retirement. In other words, if our compliers are those individuals whose behaviour is shifted by our instrument, becoming eligible and retiring in between the same two waves defines our compliers.

2.5 Results

Table 2.4 summarizes our main set of results, taking into consideration self reported retirement. Columns (1) and (2) report estimated coefficients of the first stage for men and women respectively so, the dependent variable is a dummy for retirement. In both groups our instruments are highly statistically significant. In particular, as anticipated by descriptives in Table 2.3, reaching the ordinary retirement age seem to be a more salient threshold in driving retiring decisions as compared to reaching the eligibility criteria for early retirement. While reaching

Table 2.4: Results: Retirement as self declared status.

Effect of self-reported retirement on the probability of being at least in a pre-frail status. Odd columns refer to male individuals while even ones are for women. Columns (1) and (2) report estimates for the first stage. Columns (2) to (8) refer instead to the second stage.

	(1) FE	(2) FE	(3) IV pooled	(4) IV pooled	(5) IV pooled	(6) IV pooled	(7) IVFE	(8) IVFE
retirement			-0.204**	-0.073	-0.237**	-0.081	-0.284**	-0.051
			(-2.28)	(-1.15)	(-2.50)	(-0.97)	(-2.26)	(-0.48)
Z	0.062***	0.062***						
	(3.52)	(3.86)						
W	0.239***	0.210***						
	(9.27)	(10.22)						
Education #yy	-0.009	0.006	-0.003**	-0.001	-0.003*	-0.0009	-0.005	-0.007
	(-0.81)	(0.46)	(-2.03)	(-0.50)	(-1.80)	(-0.58)	(-0.24)	(-0.43)
One child	-0.025	0.055**	-0.006	0.021	-0.003	0.025	0.034	0.039
	(-1.05)	(2.17)	(-0.37)	(1.28)	(-0.14)	(1.52)	(1.05)	(1.15)
More than one child	0.014	0.037***	-0.006	-0.001	-0.004	0.005	0.005	0.010
	(1.14)	(2.59)	(-0.49)	(-0.01)	(-0.33)	(0.43)	(0.27)	(0.47)
Difficulties in MEM	-0.017**	-0.028***	-0.029***	-0.024***	-0.023***	-0.012*	-0.004	0.007
	(-2.45)	(-4.14)	(-5.13)	(-4.02)	(-3.69)	(-1.86)	(-0.34)	(0.77)
2+ Chronic cond.	0.018	0.017	0.064***	0.065***	0.067***	0.067***	0.032*	-0.0122
	(1.43)	(1.38)	(5.50)	(5.72)	(5.68)	(5.93)	(1.83)	(-0.73)
2+ LDA	-0.012	0.008	0.137***	0.113***	0.139***	0.115***	0.053	0.072**
	(-0.41)	(0.32)	(4.46)	(4.41)	(4.55)	(4.52)	(1.34)	(2.09)
EUROD	-0.005	-0.002	0.125***	0.120***	0.125***	0.120***	0.122***	0.122***
	(-1.61)	(-0.84)	(41.02)	(50.72)	(40.26)	(50.59)	(24.06)	(30.76)
Constant	-0.146	-0.336	0.519***	0.220***	-2.152**	-2.745**		
	(-0.69)	(-1.46)	(6.55)	(2.81)	(-1.98)	(-2.01)		
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country dummies					Yes	Yes		
N	7698	7690	7698	7690	7698	7690	7678	7676
R^2	0.457	0.434	0.220	0.275	0.219	0.283	0.117	0.184
JPval			0.469	0.035	0.460	0.487	0.492	0.437

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

the latter increases probability of retirement by only 6.2%, the former is associated with an increased probability of retirement of about 24% and 21% for male and females respectively.

As we move from column (3) to column (6) we have estimated coefficients for IV models with pooled observations both with and without taking into account country dummies. We do so to account for possible country specific health trends that might be correlated with retirement decisions. As already mentioned, the dependent variable is a dichotomous variable taking value one for individuals being at least in a pre-frail status. The coefficients for retirement entail a decrease between 20.4% and 23.7% of the probability of being in a pre-frail status or beyond form men. For women the coefficient would translate in a decrease between 7.3% and 8.1% but, for this group there is no statistical significance.

Columns (7) and (8) display estimated coefficients for our IV fixed effect model. Similarly to previous estimates the coefficient of retirement is significant for males but not for females. For the formers, according to IVFE estimates, retirement as instrumented by reaching either early or statutory retirement age, causes a decrease in the probability of being in (at least a) pre-frail status by 28%. The coefficient for female is of the expected sign with a point estimate of -0.051.

A possible interpretation of the lack of significance for women draws on the different timing in the onset of the first markers of frailty between the two gender and its relation with retirement patterns. As a matter of fact, as confirmed in Table 2.1, women have a higher average frailty index and a prevalence of pre-frail statuses higher by 8% with respect to men. Such difference is true for the age range considered in our study but in the literature (Theou et al. (2015), Harttgen et al. (2013)) it is a well known fact holding at any age. To the extend to which, as compared to men, women's pre-frail status arises earlier with respect to possible retirement channels (and taking into account the cumulative nature of such syndrome in which every marker represent a risk factor for the others leading to deficits accumulation), women's exit from the labor force might occur on average at a relative later stage of development of frailty. This could translate in lower chances of reversibility from retirement.

Another possible explanation refers to the persistence of gender norms within the family. As women retire, they could substitute (at least more, as compared to men) working for pay with work at home. This would entail a difference in the persistence of stressor factor after retirement that could translate into a limited change in the path of defict accumulation for women versus a decrease among men.

To conclude the analysis of the first set of results, despite the almost zero point estimates, it is interesting to note how the fixed effect model do not drop controls like education years or any of the dummies for children, indicating at least a tiny within variation in the sample. Lastly, going back to the vality of the instruments, throughout the whole set of results, looking at the p-value for the Hansen's J statistic we fail to reject the joint null of valid instrument.

In Table 2.5 we replicate the analysis of Table 2.4 with a more restrictive definition of retirement. Rather than simply relying on self-reported job status, we further condition on not having done any paid job during the last four weeks. In so doing we cope with possible downward biases in the estimates arising from sampled individuals framing retirement as retiring from the main job characterizing the career of an individual rather than a real exit from the labor force. As individuals who retire from their "career job" self-declare as retired while carrying out some other paid activity, the coefficient for retirement would not really capture the effect of exiting the labor force. As a consequence, if this is indeed an issue one would expect to obtain much larger point estimates (in absolute value) when this further condition is applied.

Looking a Table 2.5 we see that in term of significance, estimated coefficients are are overall following the same patterns as in Table 2.5. With particular regard to the point estimates of the coefficient for retirement, we see larger point estimates (again in absolute value) in the pooled IV models where the coefficient passes from -0.204 and -0.237 (-0.073 and -0.081) to -0.356 and -0.412 (-0.092 and -0.090) for men (women). Interestingly, in the IV model with fixed effects the coefficient is larger for male individuals as compared to Table 2.4 but not for female. As a male worker reaches either early or statutory retirement eligibility criteria in between two subsequent waves and decide to retire in between the same two waves, its probability of being in at least a pre-frail status is lower by almost 52%.

It is worth remembering that given the LATE interpretation of the coefficient of retirement, we can only attribute the estimated effect to those individuals whose behaviour has been shifted by our instruments, which in our case are having reached retirement eligibility criteria. According to the definition of retirement of Table 2.5 these individuals are roughly 18% of the male sample and 21% female one.[9]

[9]Percentages are taken by summing the coefficient of the instrumental variables in columns (1) and (2) and should be referred to the pooled sample.

Table 2.5: Results: Retirement as not working for pay.

Effect of retirement defined self-reported status conditioned on not having done any paid job during the last 4 weeks on the probability of being at least in a pre-frail status. Odd columns refer to male individuals while even ones are for women.Columns (1) and (2) report estimates for the first stage. Columns (2) to (8) refer instead to the second stage

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	FE	FE	IV pooled	IV pooled	IV pooled	IV pooled	IVFE	IVFE
Retirement v2			-0.356**	-0.092	-0.412**	-0.090	-0.517**	-0.036
			(-2.14)	(-0.94)	(-2.54)	(-0.77)	(-2.28)	(-0.24)
Z	0.055***	0.080***						
	(3.18)	(5.12)						
W	0.126***	0.132***						
	(4.92)	(6.67)						
Education #yy	0.008	-0.005	-0.004**	-0.001	-0.004**	-0.001	0.002	-0.007
	(0.76)	(-0.50)	(-2.24)	(-0.53)	(-2.12)	(-0.57)	(0.12)	(-0.46)
One child	-0.019	0.036	-0.007	0.019	-0.003	0.023	0.031	0.038
	(-0.86)	(1.57)	(-0.38)	(1.17)	(-0.18)	(1.41)	(0.90)	(1.11)
More than one child	0.007	0.027**	-0.007	-0.001	-0.005	0.005	0.004	0.009
	(0.58)	(2.07)	(-0.60)	(-0.09)	(-0.45)	(0.36)	(0.23)	(0.44)
Difficulties in MEM	-0.019***	-0.019***	-0.033***	-0.024***	-0.028***	-0.012 *	-0.009	0.008
	(-2.93)	(-2.87)	(-5.36)	(-4.05)	(-4.00)	(-1.84)	(-0.81)	(0.84)
2+ Chronic cond.	0.016	0.021	0.066***	0.066***	0.069***	0.068***	0.036 *	-0.013
	(1.33)	(1.73)	(5.44)	(5.74)	(5.61)	(5.89)	(1.92)	(-0.73)
2+ LDA	-0.032	-0.018	0.138***	0.111***	0.140***	0.112***	0.039	0.071**
	(-1.12)	(-0.74)	(4.44)	(4.32)	(4.53)	(4.42)	(0.95)	(2.05)
EUROD	-0.002	-0.004	0.125***	0.120***	0.125***	0.120***	0.122***	0.122***
	(-0.72)	(-1.61)	(40.52)	(50.56)	(39.72)	(50.47)	(23.56)	(30.35)
Constant	-0.172	-0.136	0.583***	0.212***	-1.749	-2.857**		
	(-0.85)	(-0.67)	(5.55)	(2.74)	(-1.45)	(-2.03)		
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country dummies					Yes	Yes		
N	7698	7690	7698	7690	7698	7690	7678	7676
R^2	0.344	0.307	0.188	0.27	0.176	0.283	0.050	0.186
JPval			0.342	0.026	0.773	0.354	0.847	0.385

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

2.5.1 Robustness

A significant threat to the relevance of this study could lay on the effect of retirement working only through one of the five markers of frailty taken into consideration to build the index. To address this concern, we replicate the analysis of Table 2.4 excluding one marker at the time by the frailty index and still dichotomizing the resulting score distinguishing between individuals at least in a pre-frail status and fit ones. Results are displayed in Table 2.6.

The estimated coefficients remain significant at least at 10% significance level for each model among male individuals. In absolute terms, the minimum is reached excluding *exhaustion* (-0.163) while the highest estimated coefficient is obtained disregarding *slowness* (-0.319). Coefficient remain insignificant among women for every model under consideration with even a positive coefficient for the model excluding *limited physical activity*. Overall, these results seem indicative of an actual protective role of retirement with respect to frailty at least among men and it does not appear to be mainly driven by one of the distinct markers under consideration.

Table 2.6: Robustness: partial definition of frailty

IVFE estimates of the effect of self-declared retirement status on the probability of being at least in a pre-frail status excluding one marker at the time from the definition of the frailty index. Model and specification as in columns (7) and (8) of Table 2.4 and Table 2.5.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	noLPA	noLPA	noSLO	noSLO	noWEAK	noWEAK	noSHK	noSHK	noEXH	noEXH
retirement	-0.204*	0.0543	-0.319***	-0.111	-0.238*	-0.0612	-0.285**	-0.0704	-0.163*	-0.008
	(-2.29)	(0.62)	(-3.32)	(-1.18)	(-2.56)	(-0.65)	(-2.97)	(-0.76)	(-2.06)	(-0.10)
N	7676	7673	7676	7673	7676	7673	7676	7673	7676	7673

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

2.6 Conclusion

In this study we investigated the causal effect of retirement on frailty among older Europeans exploiting cross and within country variation in retirement eligibility criteria as instrumental variables. Results highlight a significant and sizeable protective role of retirement among men but no significant results for women. As a male worker become eligible for either eraly or statutory retirement consequently decide to retire in between the same two waves, the probability of being at least in a pre-frail status decreases between 28.4% and 51.7% depending on whether retirement is defined as self-declared status or further conditioning on not having done any work for pay for the past four weeks. The larger point estimates for the model using a more restrictive definition of retirement are indicative of some individuals considering retirement as withdrawal from their main occupation rather than actual exit from the labor force.

Possible non mutually exclusive explanations with regard the asymmetric significance of the results between gender are briefly described. On the one hand the lack of significant results for women could be due to earlier onset of the pre-frail statuses which could make retirement less effective in protecting against the development of deficit accumulation. On the other hand, another explanation could refer to the possible different meaning that retirement have between the two gender. While in general we think of retirement as a moment of abstention from stressor working activities, to the extent to which women simply substitute working for pay on the the job with working at home, retirement might not represent the unbend period that could instead represent for men.

Results are robust to partial definition of the index used to operationalize the frail phenotype definition. Overall, at least for men, results indicate a significant protective role of retirement with respect to frailty and in particular to pre-frail conditions. Due to the high predictive power of frailty with respect to deficit accumulation, health decline and extreme outcomes such as dependency and death, this study can at least partially shed some light on the long term effect of retirement on health. This being said it is also true that a conclusive assessment of the health effect of retirement should also consider all the dimensions of health that are disregarded by the present study, such as cognitive abilities and mental health.

Finally we want to briefly mention possible further steps of the research on frailty in economics. In particular, given recent developments in the investigation of the genetic roots of frailty (Inglés et al. (2019)) it would be possible to estimate individual level polygenic scores of genetic predisposition to frailty which could be leveraged to further investigate the heterogeneous effect of retirement along the genetic dimension.[10]

[10] By polygenic score we refer a measure of genetic predisposition to a given phenotype that is computed starting from the estimation of the effect on the phenotype of interest (in this case frailty) of genetic mutation hapening at the level of the single building block of our DNA. With this regard,more details can be found in the second chapter.

Chapter 3

Child Adoption: the Role of Couples' Preferences in the Screening and Matching Process

joint with

Davide Dragone - University of Bologna

Nektaria Glynia - University of Bologna

Abstract In this paper we investigate prospective adoptive parents' preferences thanks to a completely novel dataset from the Emilia-Romagna Region in Italy. Thanks to detailed information on the screening process from 2007 to 2010, we are able to understand how couples self select into the national vs international adoption process and how their preferences over adoptive children affect their chances of being matched to a child with specific characteristics.

3.1 Introduction

In Italy, roughly one out of ten individuals is interested by the phenomenon of adoption either as adoptive parent, adopted child or other degree of kinship.[1] Despite the magnitude, there is a significant lack of available data on the matter, especially outside the US (e.g. Baccara (2014) and Skidmore et al. (2013)).

The strength of this novel dataset is that we not only have data on the matching between couples and children but we also have extensive information on the the application process on the universe of applicant couples, regardless of their success in the adoption process. This allows to explicitly investigate the degree of selection into adoption. More precisely, in this paper we answer the following questions: what are parental preferences over adoptive children? Do preferences reflect quality throughout the screening process? Do preferences affect the probability of adoption? Do they affect the waiting time? Do preferences really map into the adopted child characteristics?

3.2 Literature

Despite the size of the phenomenon, as well as the particular interesting matching mechanisms that occur in the adoption market between the demand side (prospective adoptive parents) and the supply (relinquished children), adoption has not received much attention by economists. Early works on adoption includes studies that focus on the demand side of adoption. Jones (2009) provides a descriptive overview of the adoption market in the US using survey data. Declining fertility rates has been identified as an important factor driving the demand for children through the adoption channel. Gumus and Lee (2012) study the substitutability of children adoption with reproductive technologies as alternative ways to parenthood. Their results indicate a substitutability of child adoption with assisted reproductive technology (ART), i.e. a 10% increase in adoption will result in a 1.5% decrease in ART. Doyle and Peters (2007) study the relationship between state subsidies paid to foster families with the quality of foster care services provided. Their starting point is the shortage of foster care homes and excess

[1] Italia Adozioni Survey Report, 2019.

demand identified in the 1980s and 1990s. They show that state governments with high demand for foster care can use, up to a point, subsidies to attract foster families More recently, Oldani (2017) reports that the demand for adoptive children is empirically investigated in lights of the relation between parental satisfaction and the characteristics of both the procedure itself and those of the adoptive children. Findings show that adoptive parents' satisfaction is inversely correlated with the child's age and the duration of the adoption procedure.

This study also touches the literature related to the effect early of childhood conditions on later children outcomes. Aizer and Doyle (2014) review methods and studies regarding child adoption and support, family planning, education and health policy to analyze causal effects of child welfare interventions on child outcomes. Hansen (2008b) proves that child adoptions provide important long-term benefits to children. Her analysis considers a wide range of post-adoption outcomes including the child's behavior, health, education, criminal activity and eventually employment. Recent papers focus on the prospective parents' preferences over children's relinquished for adoption characteristics. For instance, Skidmore et al. (2013) study which factors affect adoptive parents' preferences over adoptive child characteristics and how these are expressed and translated to differences in the costs of adoption. The analysis shows the source of variation in adoption costs is child characteristics. These costs are significantly lower for older children, children of African descent, and children with special needs. Baccara et al. (2014) study the preferences of prospective adoptive parents over (US-born or unborn) children's attributes like gender and race. Using a novel dataset of the US child-adoption market they show that prospective adoptive parents show significant preference in favor of girls and unborn children while at the same time they are much less in favor for African American children. The results are robust when they control for differences with respect to prospective parents' characteristics such as same-sex, heterosexual or single mothers. How they show that racial preferences are 3 times higher in magnitude for straight couples while same-sex couples show higher persistence in submitting applications for prospective adoption.

3.3 Institutional Framework

Child adoption in Italy is ruled under the law 184/1983. In Italy adoption is allowed only to heterosexual couples married since at least three years or with at least three years of demonstrable partnership with cohabitation. Adoptive parent shall not have less than 18 years of age difference with the adopted child and more than 45, at least for one adoptive parent, and 55 for the the other.

Adoption can be generally declined in three big categories: national adoption, international adoption and special cases. The first one is organized under the actives of the juvenile court together with social assistants for screening, matching and assessment. The second one is dependent on an official declaration of eligibility to adoption by the juvenile court and then follows different paths depending on the couple's preference over country and the private entity assisting the couple in the proceedings. The last one mainly refers to cases of stepchild adoption. In the present chapter, we will focus on the first one, referring to international adoption only as general outside option.

The journey to adoption of a prospective adoptive couple in Italy starts with seminars and workshops held by social assistants units, where the couple is introduced to the procedures and expectations on both the proceedings and adoptive parenthood. Once couples decide to start the procedures they are required to undergo to an inquiry by the social assistants covering their background, current status, motivations and expectation. The completion of the inquiry generally takes between two and six months depending on social assistance units. Once the inquiry is complete, couples can deposit their application to the juvenile court. The application can refer to the national adoption process, the international one or even both. The court examines the inquiry sent directly from the social assistants and set an interview with the couple. After taking vision of the inquiry, an honorary member of the court interviews the couple with the aim of the assessment of the required capabilities of the prospective parents as well as their availabilities with respect to kid's health conditions, legal condition, age, gender and even their number.[2] In the Juvenile Court of Emilia Romagna the overall quality of a couple is summarized in a color scheme which can be translated into a score going from one to four where four is the maximum quality. [3] From the time of the submitting of the application the the couple stays in the pool of potential pick for up to three years. As new cases of abandoned kids emerge, the court examines the list of couples available at a given time. If a couple do not finalize the adoption within such a period it can repeat the procedure and re-file the application.

3.4 Data

Overall, the dataset contains information on the screening and matching activities of the Emilia Romagna Region Juvenile Court from 2005 to 2010. For years 2005 to 2007 we have detailed demographic information (e.g. from age at application to date of marriage and occupation to whether parents are alive and available to assist the couple in raising a kid), detailed information on their preferences (e.g. self declared preferences at beginning of the screening procedure and final preferences over age, gender, health, legal risk, religion and ethnicity) as well as information on the timing of the application and its proceedings (e.g. time of the beginning of the inquiry with social assistant units, time of presentation of the application, time of the meeting with honorary judges, evaluators' id as well as their evaluation etc...).

For 2005 and 2006 we only observe unmatched couples: those whose application has remained outstanding for the entire three years consideration period. For 2007 we observed both unmatched and matched couples. For the latter group we therefore also observe the kid with which the couple has been matched. Such information refers to the demographics of the child (e.g. date of birth, gender and for a sub-sample of the records also the nationality of the biological mother.)

[2]By legal condition we refer to the presence of a so called legal risk. A kid is considered under legal risk if he/she has been withdrawn from his/her biological family who is (or could) take action against the decision of the court.

[3]Formally, no application for national adoption is rejected as filing the application legally simply means declaring the availability to adopt. On the other hand, for international adoption, after the screening interview the court decides whether to grant to the couple a declaration of adaptability which is necessary condition to continue the adoption proceedings abroad.

and information on the timing of the matching (e.g. child age at declaration of adaptability, date at first contact with adoptive parents). For 2007, we also tract some information on the international adoption such as the country from which the child comes from.

For years 2008 to 2010 we focus on national adoptions tracking both matched and unmatched couples with slightly less detailed information. For these years we observe general demographics like age at application, education and occupation; final preferences (excluding ethnicity and religion) and couples' evaluation. For those who are matched we also observe the date of birth of the kid, the date of declaration of adoptability, the date of the kid entering the adoptive family. Despite not tracking any application specific for the international adoption proceeding, we observe whether or not a couple who applied for both national and international adoption finally adopt a child from the international channel.

Table 3.1: Descriptive

Demographics and adoptive preferences by application outcomes.

	(1) Avg. whole sample	(2) Avg. Unmatched	(3) Avg. Matched Int.	(4) Avg. Matched Nat.
Demographics				
age male partner	41.56	42.01	42.14	39.17
age female partner	40.10	40.59	40.65	37.63
years since married	8.37	8.89	7.71	8.08
years of education male partner	13.33	12.92	13.55	14.03
years of education female partner	14.14	13.85	14.15	14.97
biological kid(s)	0.13	0.19	0.07	0.04
at least another adopted kid	0.14	0.15	0.15	0.06
Preferences				
maximum age	5.29	5.21	5.45	5.22
age range	4.99	4.82	5.16	5.13
any gender preference	0.04	0.05	0.03	0.01
available for gender male	0.96	0.94	0.97	1
available for gender female	0.99	0.99	0.99	0.99
number	1.19	1.17	1.19	1.23
Avlbl. legal risk	0.55	0.53	0.5	0.73
Avlbl. health risk	0.91	0.87	0.92	0.96
Avlbl. handicap	0.14	0.13	0.11	0.21
Score (1 to 4)	2.78	2.43	2.95	3.52
N.	1530	777	489	264

Table 3.1 summarises demographics informations and preferences across different subsets of applicants couples from 2007 as a function of their application outcomes. The average applicant couple is married since 8 years with partners being 40 and 42 years old for female and male partner respectively. Age is generally comparable in the case of unmatched couples and couples going in the international market while couples matched in the national adoption process are significantly younger by roughly two years. On average, the female partner has a higher educational attainment than the male one, and this holds true throughout the different samples. In the whole sample 13% of the applicant couples have a biological kid while 14% have already adopted. Interestingly, among couples who successfully pursue international adoption the percentages of those with a biological kid is roughly half that of the entire sample while 15% have already adopted. On the other hand, couples matched in national adoption process have significantly lower probabilities of both having a biological kid (4%) and having already adopted (6%). This could be both due to preferences over the international market for couples who have already adopted or preference of the court over couples without either a biological or adopted kids with respect to matching. Both explanations have anecdotal support. From the analysis of the application packages and the inquiries of the social assistance emerges clearly how couples with adopted kids tends to prefer the international market to ensure some homogeneity of origins and traits between kids adopted sequentially. At the same time, judges in charge of the evaluation tend to consider carefully the potential negative externalities on the existent household other kids.

Preferences in the unrestricted sample define an average desired kid of a maximum of five years of age. Preference over a given gender is present in 4% of the population of applicant couples and when it is present it is practically always for a female kid. 55% of the applicants are available to bear legal risk and more than 91% is willing to face some sort of health risk while only 14% would accept an handicapped kid. The profile of the desired kid changes only slightly for matched couples with lower prevalence of any preference for gender. On the other hand prospective adoptive parents who manage to be matched in the national adoption process are significantly more likely to be willing to bear legal risk (73%) as well as general health risk (86%) and handicap conditions (21%).

As expected, the distribution of the evaluation score reflects the outcome taken under considerations. For couples presenting their application in 2007, 17% managed to get a kid under the national adoption process (average evaluation score of 3.52 out of 4) while a little under 32% completed international adoption (average evaluation score of 2.95 out of 4).

3.5 Methodology

In line with the descriptive aim of the paper, we adopt simple regression analysis as well as multinomial logit to assess to investigate the association between preference and screening and matching outcomes. In the same spirit, with respect to waiting times we adopt survival analysis and Cox regressions. With this regard we model the duration of the time remaining in the pool waiting to be matched. Thus, longer survival time actually translates in waiting longer to be matched, up to a maximum of three years of elapsed time. In this survival framework, matching represent the "death" event. We proxy the time of the matching with the day since when the adopted kid starts living with the adoptive family, regardless of the actual officialization of the adoption.

In order to avoid unbalanced samples given by the structure of the information collected for the different years we only consider data from 2007 onward. The main dependent variables taken into account are the outcome of the screening process, the probability of the matching and the duration of the waiting time to matching.

3.6 Results

The first measurable outcome in the adoption process is the screening after application is filed. As mentioned above, in the Juvenile Court of Emilia Romagna, such screening is performed by honorary judges evaluating the couple's motivation and capabilities from the inquiry by the social assistants and the interview with the couple. The score goes from 1 to 4 but there is no official cut-off to reach to be eligible for adoption.

For both regressions (1) and (2) in Table 3.2 we control for a set of dummies capturing male and female partner occupational role. Adding the set of covariates describing couple's preference over adoption, the point estimates and the significance of the coefficient vary only slightly. The age of the female is significant and negatively related to the score. Estimates suggests the same for the years since when the couple is married and having any gender preferences. Higher scores for younger couples could probably be explained by stronger motivations. In terms of what has a significant and positive impact on the evaluation of the couple, the estimates highlight educational attainment of both the male and female partner (with the second one having a higher point estimate), having already adopted, age range being available for legal risk, health risk and handicapped minors.

In Table 3.3 columns (1) and (2) we take the same exact specification as in Table 3.2 with the only addition of the score. As a dependent variable we adopt a dichotomous variable for being matched in a national adoption. With no surprises the score is positive and significant and the age of both partners is significant and negatively related with probability of adoption. Educational attainment is significant only for the female partner while having a biological kid and having already adopted is significantly negatively related with probability of matching. With respect to preferences over adoption the only significant coefficient is that of availability to legal risk.[4]

These two set of estimates suggest how screening and matching are indeed separate processes taking into account different dimensions. The most interesting results are the ambivalent role of having already adopted (and having biological kids) with regard to the different dependent variable. While on the side of the score specification such variable capture a higher level of familiar capabilities gained

[4]The careful reader could highlight the possibility of omitted variable bias with regard to dimensions such as preferences of the judges or information conveyed in the inquiry are not included in the specification. This could indeed be the case and for this reason, in an extended version of the current paper we are actually taking into account also these measurements. Including dummies for evaluating judges reduces the magnitude of the point estimates but not the significance of the coefficients. We are currently also working on the summarizing the textual information of the inquiries by means of principal component analysis to further account for possible omitted variable bias.

through experience, on the side of the matching this represent a possible added risk to the stability of the future adoptive family. Similarly, while availability to legal risk health risk and handicapped conditions are positive predictors of a high score, only the first one remain significant in the second stage of the adoptive process.

Column (3) and (4) of Table 3.3 focus being matched to an infant conditionally on being matched to some kid. We define an infant as a kid who has been abandoned at the hospital right after being born. These kids are generally declared adoptable within few weeks from their birth and with regard to the preferences over adoption taken into account by the process they could be seen as kids with highest degree of desirability because of their age and the lack ant legal risk. The significant variable for these models are the age of the male partner, with negative sign and the preference profile over the age of the kid which sums to zero for couple with preferred age range equal to the maximum desired age (which imply negative effect overall for couples with preference for a smaller age range with respect to their maximum age). Despite the lack of significance, it is interesting to note the negative sign of the coefficient for availability to legal risk, suggesting the attempt to optimal matching taking into account the characteristics of the kids and the preferences of the couples.

Table 3.2: Evaluation score regression on demographics and adoption preferences.

	(1) score	(2) score
age male partner	-0.00856	-0.00985
	(-1.06)	(-1.27)
age female partner	-0.0257***	-0.0201**
	(-2.99)	(-2.36)
year since married	-0.0164**	-0.0143**
	(-2.50)	(-2.27)
educational attainment male partner	0.0268**	0.0285***
	(2.50)	(2.79)
educational attainment female partner	0.0403***	0.0350***
	(3.51)	(3.24)
any biological kids	0.0531	0.0598
	(0.51)	(0.58)
at least another adopted kid	0.436***	0.535***
	(4.98)	(6.13)
maximum age		-0.0452
		(-1.55)
age range		0.107***
		(3.54)
any gender preference		-0.412**
		(-2.24)
gender preference for female		-0.439
		(-0.92)
number		0.0270
		(0.36)
Avlbl. legal risk		0.139**
		(2.38)
Avlbl. Health risk		0.507***
		(4.23)
Avlbl. Handicap		0.247***
		(2.90)
constant	3.882***	5.486***
	(7.31)	(8.37)
male partner occupation	Y	Y
female partner occupation	Y	Y
N	1044	1007
R-sq	10.6	15.4

t statistics in parentheses

* p<0.10, **p<0.05, *** p<0.010

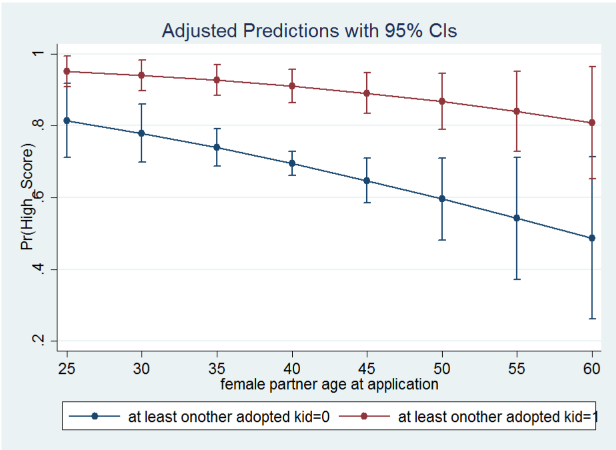
Table 3.3: Logistic regressions for adoption matching on demographics and adoption preferences.

	(1)	(2)	(3)	(4)
	Matched	Matched	Matched	Matched
	in national adoption	in national adoption	with infant	with infant
score	1.117***	1.025***	0.333	0.329
	(8.63)	(7.40)	(1.06)	(1.04)
age male partner	-0.127***	-0.130***	-0.210***	-0.207***
	(-3.91)	(-3.95)	(-3.62)	(-3.54)
age female partner	-0.0735***	-0.0785***	0.0300	0.0414
	(-2.60)	(-2.66)	(0.49)	(0.67)
year since married	0.0345	0.0377	-0.0841*	-0.0950*
	(1.20)	(1.29)	(-1.70)	(-1.77)
educational attainment male partner	0.0700*	0.0673*	0.0207	0.0354
	(1.79)	(1.65)	(0.28)	(0.46)
educational attainment female partner	0.121***	0.123***	-0.0864	-0.103
	(2.60)	(2.59)	(-1.19)	(-1.41)
any biological kids	-1.844***	-1.792***	1.242	1.232
	(-3.42)	(-3.22)	(1.29)	(1.25)
at least another adopted kid	-2.139***	-1.815***	-0.427	-0.577
	(-4.13)	(-3.34)	(-0.37)	(-0.49)
maximum age		0.0248		-4.169***
		(0.13)		(-13.56)
age range		0.0958		4.196***
		(0.44)		(13.77)
any gender preference		1.272		0.001
		(1.38)		(0.02)
gender preference for female		0.002		0.003
		(0.01)		(0.01)
number		-0.0873		-0.162
		(-0.33)		(-0.36)
Avlbl. legal risk		0.582***		-0.310
		(2.71)		(-0.69)
Avlbl. Health risk		0.328		-0.247
		(0.66)		(-0.32)
Avlbl. Handicap		0.403		-0.189
		(1.21)		(-0.35)
constant	-128	-0.817	6.937***	7.150***
	(-0.11)	(-0.62)	(2.80)	(2.59)
male partner occupation	Y	Y	Y	Y
female partner occupation	Y	Y	Y	Y
N	1038	970	171	168
pseudo R-sq	0.274	0.271	0.171	0.183

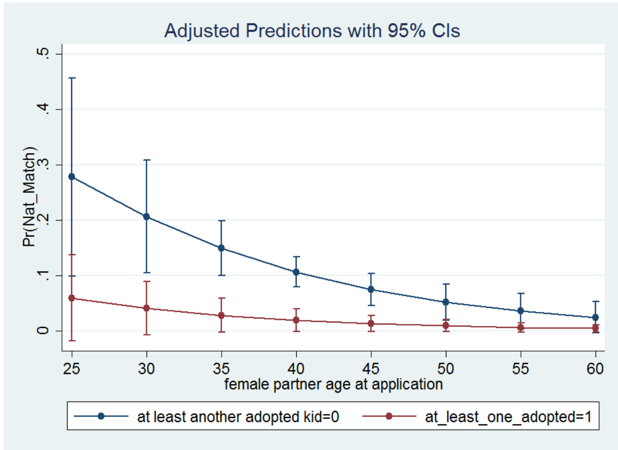
t statistics in parentheses

* p<0.10, **p<0.05, *** p<0.010

Figure 3.1: Marginal effects of female partner age at application on (a) the probability of having a score higher or equal than three and (b) probability of being matched



(a)



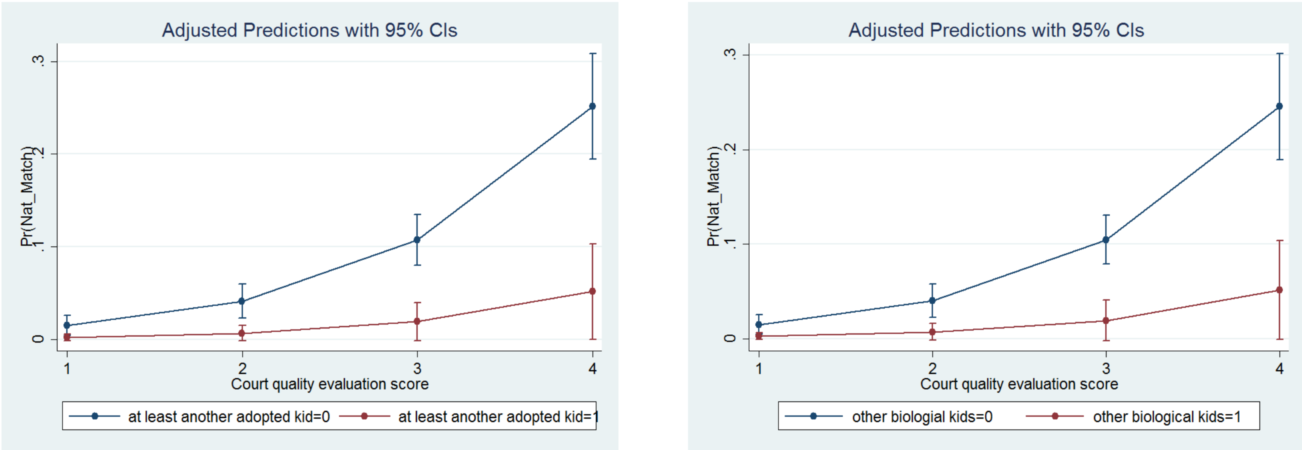
(b)

To better grasp the role in terms of probability, of the most important and interesting independent variables from the model above we compute and plot marginal effect. Figure 3.1 shows the marginal effect of age at application with both respect to score (a) and matching (b) by groups determined by having already adopted or not. In addition to the clear negative impact of age with respect to both outcomes, the graphs highlights the dual role of preavious experience with adopted kid.

Figure 3.2 shows how low the probability of matching in the national adoption process really is for couples who either have a biological kid or either have already adopted. For applicants without any previous experience the probability of successful national adoption are between 20 and 30% in case of a score of 4 out of 4. The probability shrinks to around 10% for a score of 3 out of 4 going practically to zero for a score of 1 out of 4.

In the last part of this section we focus on waiting times. Figure 3.4 shows survival curves of the applicant couples by evaluation score. Average survival time in the pool of couple available to adoptions is 3 and 2.97 years respectively for couples with evaluation score respectively of 1 and 2. Couples with a score of 3 wait on average 2.46 years while applicants with the highest evaluation wait on average 2 years. Assuming a constant inflow of kids, this averages and graphs

Figure 3.2: Marginal effects of the evaluation score on matching probability by (a) having already adopted and (b) having other biological kids.



(a) (b)

are consistent with a first in first out processing of the application of prospective adoptive parents.

In Figure 3.5 we condition on having a score higher or equal to three and plot the survival curves by availability to legal risk. As expected couples with availability to legal risk exit the pool at a higher rate but timing at which they start exiting does not seem to be statistically different from one group to another.

Finally, Figure 3.6 summarizes hazard ratios from Cox regressions for being matched. As it was for the estimates from the logistic probability models, the most relevant predictors are having a high score (dichotomous variable for score greater or equal to three), availability to legal risk and college degree of the female partner.

3.7 Conclusion

In this last chapter we provided quantitative descriptive evidence from a novel dataset on the phenomenon of national adoption taking a close look at the screening and matching process and the role of preferences of prospective adoptive parents. In the Emilia Romagna region in Italy, screening and matching processes are distinct phases. Applicants couple have on average preferences for kids in pre-scholar age. If they have any preference over the gender of the kid it is for

Figure 3.3: Survival curve by evaluation score

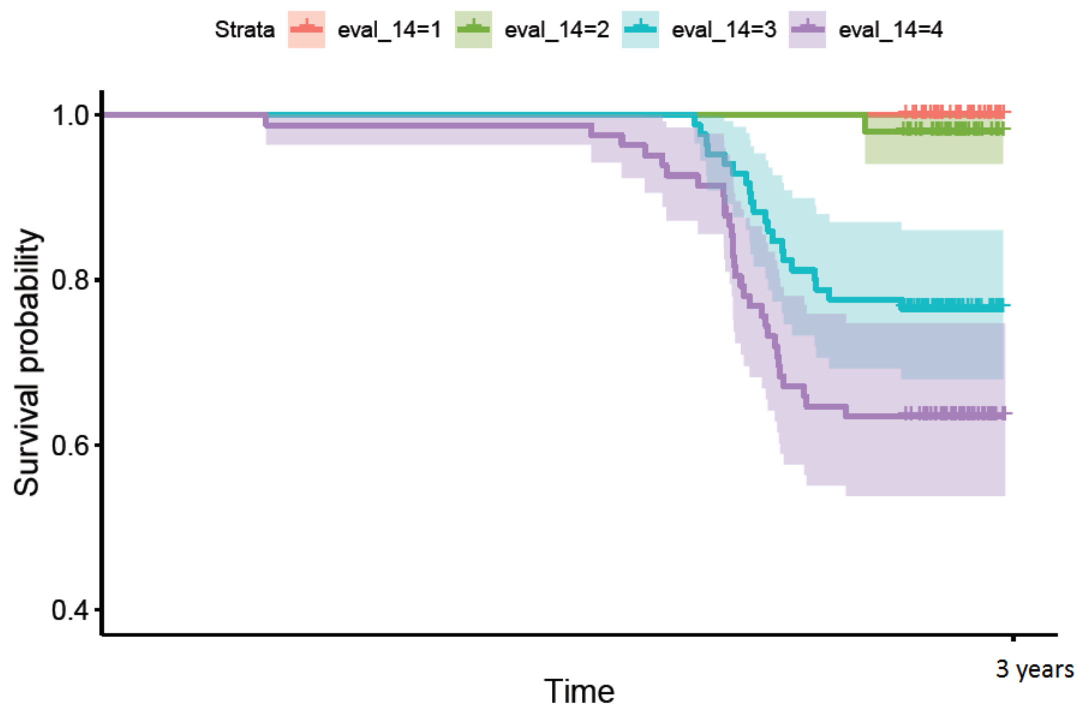


Figure 3.4: Survival curve by availability to legal risk

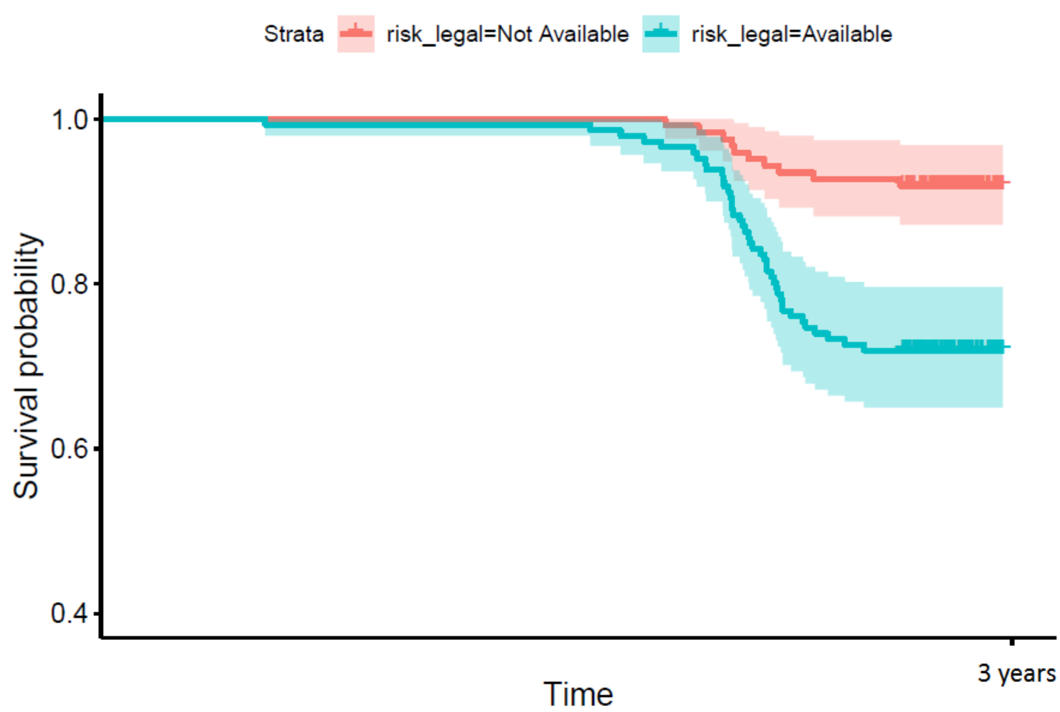
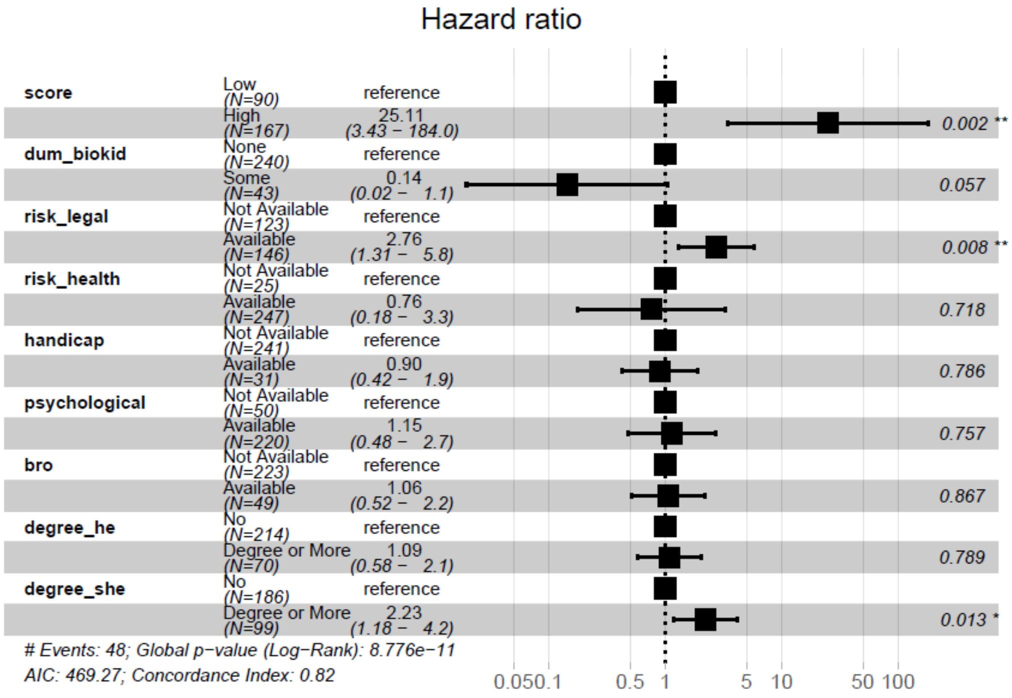


Figure 3.5: Hazard ratios of being matched



females. Most of the applicant couples declare themselves willing to bear some sort of health risk while only 50% is available to face legal risk, which is the possibility appeal from the biological family of the minor declared adoptable. The availability to be matched with an handicapped minor is stated only by 14% of the applicants. Only couples with a sufficient screening evaluation have a real chance of being matched to a kid. On top of it, the most important aspect to consider in order to increase the chances of a successful national adoption application is the availability to bear legal risk. The most significant predictor for a successful screening process are the age of the couple and the duration of their marriage at time of the application, the educational attainment of the female partner, the lack of any gender preference and the availability for health risk, legal risk and handicapped condition. Interestingly, having already adopted a kid in the past has a positive and significant impact on the screening phase but is negatively associated with matching probability. Waiting times are heterogeneous across different evaluation scores and reflects the quality of the couple summarized by the score. Couples with a score of one wait in list "forever" as they are never chosen for matching.

Bibliography

V. Amin, P. Böckerman, J. Viinikainen, M. C. Smart, Y. Bao, M. Kumari, N. Pitkänen, T. Lehtimäki, O. Raitakari, and J. Pehkonen. Gene-environment interactions between education and body mass: Evidence from the UK and Finland. *Social Science and Medicine*, 195(June):12–16, 2017. ISSN 18735347. doi: 10.1016/j.socscimed.2017.10.027. URL https://doi.org/10.1016/j.socscimed.2017.10.027.

O. Ashenfelter and A. Krueger. Estimates of the Economic Return toSchooling from New Sample of Twins. *American Economic Association*, 84(5):1157–1173, 2019.

M. Avendano and L. F. Berkman. Labor Markets, Employment Policies, and Health. In *Social Epidemiology*, chapter 6. Oxford University Press, 2014.

H. Axelrad, M. Malul, and I. Luski. Unemployment among younger and older individuals : does conventional data about unemployment tell us the whole story ? *Journal for Labour Market Research*, 2018. ISSN 2510-5027. doi: 10.1186/s12651-018-0237-9. URL https://doi.org/10.1186/s12651-018-0237-9.

F. Barigozzi and D. Henriet. GENETIC INFORMATION:COMPARING ALTERNATIVE REGULATORY APPROACHES WHEN PREVENTION MATTERS. *Journal of Public Economic Theory*, 13(1):23–46, 2011.

D. Barth, N. W. Papageorge, and K. Thom. Genetic Endowments and Wealth Inequality. *Journal of Political Economy*, page 705415, 2019. ISSN 0022-3808. doi: 10.1086/705415. URL https://www.journals.uchicago.edu/doi/10.1086/705415.

J. P. Beauchamp, D. Cesarini, M. Johannesson, M. J. H. M. van der Loos, P. D. Koellinger, P. J. F. Groenen, J. H. Fowler, J. N. Rosenquist, A. R. Thurik, and N. A. Christakis. Molecular Genetics and Economics. *Journal of Economic Perspectives*, 25(4):57–82, 2011. ISSN 0895-3309. doi: 10.1257/jep.25.4.57. URL http://pubs.aeaweb.org/doi/10.1257/jep.25.4.57.

S. Behncke. DOES RETIREMENT TRIGGER ILL HEALTH ? y. *Health Economics*, 300(February 2011):282–300, 2012. doi: 10.1002/hec.

M. Belloni, E. Meschi, and G. Pasini. The Effect on Mental Health of Retiring During the Economic Crisis. *Health economics*, 25(2):126–140, 2016.

D. J. Benjamin, D. Cesarini, C. F. Chabris, E. L. Glaeser, I. David, V. Gu, T. B. Harris, L. J. Launer, S. Purcell, A. V. Smith, M. Johannesson, P. K. E. Magnusson, P. Jonathan, N. A. Christakis, C. S. Atwood, and B. Hebert. The Promises and Pitfalls of Genoeconomics. *Annu Rev Econom*, 4:627–662, 2012. doi: 10.1146/annurev-economics-080511-110939.The.

H. Bergman, L. Ferrucci, J. Guralnik, D. Hogan, S. Hummel, S. Karunananthan, and C. Wolfson. Frailty: an emerging research and clinical paradigm - issues and controversies. *J Gerontol A Biol Sci Med Sci*, 62(7):731–7, 2007. ISSN 15378276. doi: 10.1016/j.mib.2008.09.004.Bacteriophage.

M. Bertoni, S. Maggi, and G. Weber. Work, retirement, and muscle strength loss in old age. *Health Economics (United Kingdom)*, 27(1):115–128, 2018. ISSN 10991050. doi: 10.1002/hec.3517.

M. Bertrand, E. Duflo, and S. Mullainathan. HOW MUCH SHOULD WE TRUST DIFFERENCES-IN-DIFFERENCES ESTIMATES ?* Differences-in-Differences (DD) estimation has become an increasingly popular way to estimate causal relationships . DD estimation consists of identifying a speci c intervention or treat-. *The Quarterly Journal of Economics*, (February):1–27, 2004.

P. Biroli. Genetic and Economic Interaction in Health Formation: The Case of Obesity. *WP*, 2015. URL http://www.cdc.gov/obesity/data/childhood.html.

J.-O. Bock, H.-H. König, H. Brenner, W. E. Haefeli, R. Quinzler, H. Matschinger, K.-U. Saum, B. Schöttker, and D. Heider. Associations of frailty with health care costs - results of the ESTHER cohort study. *BMC health services research*, 16(1):128, 2016. ISSN 1472-6963. doi: 10.1186/s12913-016-1360-3. URL http://www.ncbi.nlm.nih.gov/pubmed/27074800%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4831082.

A. Brugiavini and F. Peracchi. Fiscal implications of Pension Reforms in Italy. 2 (6), 2005.

K. K. Charles. Is Retirement Depressing? Labor Force Inactivity and Psychological Well-being in Later Life. 2002.

N. B. Coe and G. Zamarro. Retirement effects on health in Europe. *J Health Econ.*, 30((1)):77–86, 2011.

N. B. Coe, H.-M. Von Gaudecker, M. Lindeboom, and J. Maurer. The effect of retirement on cognitive functioning. *Health economics*, 21:913–927, 2012.

C. C. Coile, P. B. Levine, and R. McKnight. RECESSIONS, OLDER WORKERS, AND LONGEVITY: HOW LONG ARE RECESSIONS GOOD FOR YOUR HEALTH? *NBER Working Paper 18361*, 2012. URL http://www.nber.org/papers/w18361%0ANATIONAL.

T. A. Comans, N. M. Peel, R. E. Hubbard, A. D. Mulligan, L. C. Gray, and P. A. Scuffham. The increase in healthcare costs associated with frailty in older people discharged to a post-acute transition care program. *Age and Ageing*, 45 (2):317–320, 2016. ISSN 14682834. doi: 10.1093/ageing/afv196.

D. Conley. The promise and challenges of incorporating genetic data into longitudinal social science surveys and research. *Biodemography and Social Biology*, 55(2):238–251, 2009. doi: 10.1080/19485560903415807. URL https://doi.org/10.1080/19485560903415807. PMID: 20183907.

D. Conley and K. W. Strully. Birth weight , infant mortality , and race : Twin comparisons and genetic / environmental inputs. *Social Science & Medicine*, 75 (12):2446–2454, 2012. ISSN 0277-9536. doi: 10.1016/j.socscimed.2012.09.024. URL http://dx.doi.org/10.1016/j.socscimed.2012.09.024.

D. Conley and S. Zhang. The promise of genes for understanding cause and effect. *Proceedings of the National Academy of Sciences of the United States of America*, 115(22):5626–5628, 2018. ISSN 10916490. doi: 10.1073/pnas.1805585115.

J. Craig Venter, M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural, G. G. Sutton, H. O. Smith, M. Yandell, C. A. Evans, R. A. Holt, J. D. Gocayne, P. Amanatides, R. M. Ballew, D. H. Huson, J. R. Wortman, Q. Zhang, C. D. Kodira, X. H. Zheng, L. Chen, M. Skupski, G. Subramanian, P. D. Thomas, J. Zhang, G. L. Gabor Miklos, C. Nelson, S. Broder, A. G. Clark, J. Nadeau, V. A. McKusick, N. Zinder, A. J. Levine, R. J. Roberts, M. Simon, C. Slayman, M. Hunkapiller, R. Bolanos, A. Delcher, I. Dew, D. Fasulo, M. Flanigan, L. Florea, A. Halpern, S. Hannenhalli, S. Kravitz, S. Levy, C. Mobarry, K. Reinert, K. Remington, J. Abu-Threideh, E. Beasley, K. Biddick, V. Bonazzi, R. Brandon, M. Cargill, I. Chandramouliswaran, R. Charlab, K. Chaturvedi, Z. Deng, V. di Francesco, P. Dunn, K. Eilbeck, C. Evangelista, A. E. Gabrielian, W. Gan, W. Ge, F. Gong, Z. Gu, P. Guan, T. J. Heiman, M. E. Higgins, R. R. Ji, Z. Ke, K. A. Ketchum, Z. Lai, Y. Lei, Z. Li, J. Li, Y. Liang, X. Lin, F. Lu, G. V. Merkulov, N. Milshina, H. M. Moore, A. K. Naik, V. A. Narayan, B. Neelam, D. Nusskern, D. B. Rusch, S. Salzberg, W. Shao, B. Shue, J. Sun, Z. Yuan Wang, A. Wang, X. Wang, J. Wang, M. H. Wei, R. Wides, C. Xiao, C. Yan, A. Yao, J. Ye, M. Zhan, W. Zhang, H. Zhang, Q. Zhao, L. Zheng, F. Zhong, W. Zhong, S. C. Zhu, S. Zhao, D. Gilbert, S. Baumhueter, G. Spier, C. Carter, A. Cravchik, T. Woodage, F. Ali, H. An, A. Awe, D. Baldwin, H. Baden, M. Barnstead, I. Barrow, K. Beeson, D. Busam, A. Carver, A. Center, M. Lai Cheng, L. Curry, S. Danaher, L. Davenport, R. Desilets, S. Dietz, K. Dodson, L. Doup, S. Ferriera, N. Garg, A. Gluecksmann, B. Hart, J. Haynes, C. Haynes, C. Heiner, S. Hladun, D. Hostin, J. Houck, T. Howland, C. Ibegwam, J. Johnson, F. Kalush, L. Kline, S. Koduru, A. Love, F. Mann, D. May, S. McCawley, T. McIntosh, I. McMullen, M. Moy, L. Moy, B. Murphy, K. Nelson, C. Pfannkoch, E. Pratts, V. Puri, H. Qureshi, M. Reardon, R. Rodriguez, Y. H. Rogers, D. Romblad, B. Ruhfel, R. Scott, C. Sitter, M. Smallwood, E. Stewart, R. Strong, E. Suh, R. Thomas, N. Ni Tint, S. Tse, C. Vech, G. Wang, J. Wetter, S. Williams, M. Williams, S. Windsor, E. Winn-Deen, K. Wolfe, J. Zaveri, K. Zaveri, J. F. Abril, R. Guigo, M. J. Campbell, K. V. Sjolander, B. Karlak,

A. Kejariwal, H. Mi, B. Lazareva, T. Hatton, A. Narechania, K. Diemer, A. Muruganujan, N. Guo, S. Sato, V. Bafna, S. Istrail, R. Lippert, R. Schwartz, B. Walenz, S. Yooseph, D. Allen, A. Basu, J. Baxendale, L. Blick, M. Caminha, J. Carnes-Stine, P. Caulk, Y. H. Chiang, M. Coyne, C. Dahlke, A. Deslattes Mays, M. Dombroski, M. Donnelly, D. Ely, S. Esparham, C. Fosler, H. Gire, S. Glanowski, K. Glasser, A. Glodek, M. Gorokhov, K. Graham, B. Gropman, M. Harris, J. Heil, S. Henderson, J. Hoover, D. Jennings, C. Jordan, J. Jordan, J. Kasha, L. Kagan, C. Kraft, A. Levitsky, M. Lewis, X. Liu, J. Lopez, D. Ma, W. Majoros, J. McDaniel, S. Murphy, M. Newman, T. Nguyen, N. Nguyen, M. Nodell, S. Pan, J. Peck, M. Peterson, W. Rowe, R. Sanders, J. Scott, M. Simpson, T. Smith, A. Sprague, T. Stockwell, R. Turner, E. Venter, M. Wang, M. Wen, D. Wu, M. Wu, A. Xia, A. Zandieh, and X. Zhu. The sequence of the human genome. *Science*, 291 (5507):1304–1351, 2001. ISSN 00368075. doi: 10.1126/science.1058040.

A. Diamond. Executive Functions. *Annual review of clinical psychologyPsychol.*, 64: 135–168, 2014. ISSN 15378276. doi: 10.1146/annurev-psych-113011-143750. Executive.

B. W. Domingue, H. Liu, A. Okbay, and D. W. Belsky. Genetic heterogeneity in depressive symptoms following the death of a spouse: Polygenic score analysis of the U.S. Health and retirement study. *American Journal of Psychiatry*, 174(10): 963–970, 2017. ISSN 15357228. doi: 10.1176/appi.ajp.2017.16111209.

L. et al. Initial sequencing and analysis of the human genome. 409(February), 2001.

L. Fried, C. Tangen, J. Walston, A. Newman, C. Hirsch, J. Gottdiener, T. Seeman, R. Tracy, W. Kop, B. G, and M. MA. Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci.*, 56(3):M146–56, 2001.

L. P. Fried, L. Ferrucci, J. Darer, J. D. Williamson, and G. Anderson. Untangling the Concepts of Disability, Frailty, and Comorbidity: Implications for Improved Targeting and Care. 59(3):255–263, 2004. ISSN 10795006. doi: 10.1093/gerona/59.3.M255.

M. Godard. Gaining weight through retirement? Results from the SHARE survey. *Journal of Health Economics*, 45:27–46, 2016. ISSN 18791646. doi: 10.1016/j.jhealeco.2015.11.002. URL http://dx.doi.org/10.1016/j.jhealeco.2015.11.002.

M. Grossman. On the concept of health capital and the demand for health. *Journal of Political Economy*, 80(2):223–55, 1972. URL https://EconPapers.repec.org/RePEc:ucp:jpolec:v:80:y:1972:i:2:p:223-55.

J. Gruber and B. Köszegi. Is Addiction "Rational"? Theory and Evidence. *The Quarterly Journal of Economics*, 116(4):1261–1303, 2001. ISSN 00335533. doi: http://www.mitpressjournals.org/loi/qjec. URL http://qje.oxfordjournals.org/cgi/doi/10.1162/003355301753265570%5Cnhttp://www.jstor.org/stable/2696459%5Cnhttp://www.jstor.org/page/.

J. Gruber, J. Kim, and D. Mayzlin. Physician fees and procedure intensity: The case of cesarean delivery. *Journal of Health Economics*, 18(4):473–490, 1999. ISSN 01676296. doi: 10.1016/S0167-6296(99)00009-0.

C. B. Hansen. Generalized least squares inference in panel and multilevel models with serial correlation and fixed effects. *Journal of Econometrics*, 140(2):670–694, 2007. ISSN 03044076. doi: 10.1016/j.jeconom.2006.07.011.

K. Harttgen, P. Kowal, H. Strulik, S. Chatterji, and S. Vollmer. Patterns of frailty in older adults: Comparing results from higher and lower income countries using the survey of health, ageing and retirement in europe (share) and the study on global ageing and adult health (sage). *PLOS ONE*, 8(10):1–11, 10 2013. doi: 10.1371/journal.pone.0075847. URL https://doi.org/10.1371/journal.pone.0075847.

P. Hessel, C. J. Riumallo-Herl, A. K. Leist, L. F. Berkman, and M. Avendano. Economic Downturns, Retirement and Long-Term Cognitive Function among Older Americans. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, 73(4):744–754, 2018. ISSN 10795014. doi: 10.1093/geronb/gbx035.

P. Hicks. The surprisingly large policy implications of changing retirement durations. (284), 2011. URL http://login.ezproxy.lib.umn.edu/?url=http://search.proquest.com/docview/896015348?accountid=14586%5Cnhttp://tc.liblink.umn.edu/sfx_local?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:journal&genre=unknown&sid=ProQ:ProQ:econlitshell&atitle=The+surpri.

M. Hoy and M. Polborn. The value of genetic information in the life insurance market. *Journal of Public Economics*, 78:235–237, 2000.

M. Inglés, C. Mas-Bargues, L. Gimeno-Mallench, R. Cruz-Guerrero, F. J. García-García, J. Gambini, C. Borrás, L. Rodríguez-Mañas, and J. Viña. Relation between genetic factors and frailty in older adults. *Journal of the American Medical Directors Association*, 2019. ISSN 1525-8610. doi: https://doi.org/10.1016/j.jamda.2019.03.011. URL http://www.sciencedirect.com/science/article/pii/S1525861019302981.

L. Kalousova and C. M. de Leon. Increase in frailty of older workers and retirees predicted by negative psychosocial working conditionson the job. *Social Science and Medicine*, 124:275–283, 2015. ISSN 18735347. doi: 10.1016/j.socscimed.2014.11.055. URL http://dx.doi.org/10.1016/j.socscimed.2014.11.055.

R. Karlsson Linnér, P. Biroli, E. Kong, S. F. W. Meddens, R. Wedow, M. A. Fontana, M. Lebreton, S. P. Tino, A. Abdellaoui, A. R. Hammerschlag, M. G. Nivard, A. Okbay, C. A. Rietveld, P. N. Timshel, M. Trzaskowski, R. de Vlaming, C. L. Zünd, Y. Bao, L. Buzdugan, A. H. Caplin, C. Y. Chen, P. Eibich, P. Fontanillas, J. R. Gonzalez, P. K. Joshi, V. Karhunen, A. Kleinman, R. Z. Levin, C. M. Lill, G. A. Meddens, G. Muntané, S. Sanchez-Roige, F. J. Rooij, E. Taskesen, Y. Wu, F. Zhang, M. Agee, B. Alipanahi, R. K. Bell, K. Bryc, S. L. Elson, N. A. Furlotte, K. E. Huber, N. K. Litterman, J. C. McCreight, M. H. McIntyre, J. L. Mountain, C. A. Northover, S. J. Pitts, J. F. Sathirapongsasuti, O. V. Sazonova, J. F. Shelton, S. Shringarpure, C. Tian, J. Y. Tung, V. Vacic, C. H. Wilson, M. Agbessi, H. Ahsan, I. Alves, A. Andiappan, P. Awadalla, A. Battle, F. Beutner, M. Jan Bonder, D. I. Boomsma, M. Christiansen, A. Claringbould, P. Deelen, T. Esko, M. J. Favé, L. Franke, T. Frayling, S. A. Gharib, G. Gibson, B. Heijmans,

G. Hemani, R. Jansen, M. Kähönen, A. Kalnapenkis, S. Kasela, J. Kettunen, Y. Kim, H. Kirsten, P. Kovacs, K. Krohn, J. Kronberg-Guzman, V. Kukushkina, Z. Kutalik, B. Lee, T. Lehtimäki, M. Loeffler, U. M. Marigorta, A. Metspalu, L. Milani, G. W. Montgomery, M. Müller-Nurasyid, M. Nauck, M. G. Nivard, B. Penninx, M. Perola, N. Pervjakova, B. Pierce, J. Powell, H. Prokisch, B. M. Psaty, O. Raitakari, S. Ring, S. Ripatti, O. Rotzchke, S. Rüeger, A. Saha, M. Scholz, K. Schramm, I. Seppälä, M. Stumvoll, P. Sullivan, P. B. t. Hoen, A. Teumer, J. Thiery, L. Tong, A. Tönjes, J. van Dongen, J. van Meurs, J. Verlouw, P. M. Visscher, U. Völker, U. Võsa, H. J. Westra, H. Yaghootkar, J. Yang, B. Zeng, F. Zhang, A. Okbay, J. P. Beauchamp, M. A. Fontana, J. J. Lee, T. H. Pers, C. A. Rietveld, P. Turley, G. B. Chen, V. Emilsson, S. F. W. Meddens, S. Oskarsson, J. K. Pickrell, K. Thom, P. Timshel, R. de Vlaming, A. Abdellaoui, T. S. Ahluwalia, J. Bacelis, C. Baumbach, G. Bjornsdottir, J. H. Brandsma, M. P. Concas, J. Derringer, N. A. Furlotte, T. E. Galesloot, G. Girotto, R. Gupta, L. M. Hall, S. E. Harris, E. Hofer, M. Horikoshi, J. E. Huffman, K. Kaasik, I. P. Kalafati, R. Karlsson, A. Kong, J. Lahti, S. J. der Lee, C. de Leeuw, P. A. Lind, K. O. Lindgren, T. Liu, M. Mangino, J. Marten, E. Mihailov, M. B. Miller, P. J. der Most, C. Oldmeadow, A. Payton, N. Pervjakova, W. J. Peyrot, Y. Qian, O. Raitakari, R. Rueedi, E. Salvi, B. Schmidt, K. E. Schraut, J. Shi, A. V. Smith, R. A. Poot, B. S. Pourcain, A. Teumer, G. Thorleifsson, N. Verweij, D. Vuckovic, J. Wellmann, H. J. Westra, J. Yang, W. Zhao, Z. Zhu, B. Z. Alizadeh, N. Amin, A. Bakshi, S. E. Baumeister, G. Biino, K. Bønnelykke, P. A. Boyle, H. Campbell, F. P. Cappuccio, G. Davies, J. E. De Neve, P. Deloukas, I. Demuth, J. Ding, P. Eibich, L. Eisele, N. Eklund, D. M. Evans, J. D. Faul, M. F. Feitosa, A. J. Forstner, I. Gandin, B. Gunnarsson, B. V. Halldórsson, T. B. Harris, A. C. Heath, L. J. Hocking, E. G. Holliday, G. Homuth, M. A. Horan, J. J. Hottenga, P. L. de Jager, P. K. Joshi, A. Jugessur, M. A. Kaakinen, M. Kähönen, S. Kanoni, L. Keltigangas-Järvinen, L. A. Kiemeney, I. Kolcic, S. Koskinen, A. T. Kraja, M. Kroh, Z. Kutalik, A. Latvala, L. J. Launer, M. P. Lebreton, D. F. Levinson, P. Lichtenstein, P. Lichtner, D. C. Liewald, A. Loukola, P. A. Madden, R. Mägi, T. Mäki-Opas, R. E. Marioni, P. Marques-Vidal, G. A. Meddens, G. McMahon, C. Meisinger, T. Meitinger, Y. Milaneschi, L. Milani, G. W. Montgomery, R. Myhre, C. P. Nelson, D. R. Nyholt, W. E. Ollier, A. Palotie, L. Paternoster, N. L. Pedersen, K. E. Petrovic, D. J. Porteous, K. Räikkönen,

S. M. Ring, A. Robino, O. Rostapshova, I. Rudan, A. Rustichini, V. Salomaa, A. R. Sanders, A. P. Sarin, H. Schmidt, R. J. Scott, B. H. Smith, J. A. Smith, J. A. Staessen, E. Steinhagen-Thiessen, K. Strauch, A. Terracciano, M. D. Tobin, S. Ulivi, S. Vaccargiu, L. Quaye, F. J. Rooij, C. Venturini, A. A. Vinkhuyzen, U. Völker, H. Völzke, J. M. Vonk, D. Vozzi, J. Waage, E. B. Ware, G. Willemsen, J. R. Attia, D. A. Bennett, K. Berger, L. Bertram, H. Bisgaard, D. I. Boomsma, I. B. Borecki, U. Bültmann, C. F. Chabris, F. Cucca, D. Cusi, I. J. Deary, G. V. Dedoussis, C. M. Duijn, J. G. Eriksson, B. Franke, L. Franke, P. Gasparini, P. V. Gejman, C. Gieger, H. J. Grabe, J. Gratten, P. J. Groenen, V. Gudnason, P. van der Harst, C. Hayward, D. A. Hinds, W. Hoffmann, E. Hyppönen, W. G. Iacono, B. Jacobsson, M. R. Järvelin, K. H. Jöckel, J. Kaprio, S. L. Kardia, T. Lehtimäki, S. F. Lehrer, P. K. Magnusson, N. G. Martin, M. McGue, A. Metspalu, N. Pendleton, B. Penninx, M. Perola, N. Pirastu, M. Pirastu, O. Polasek, D. Posthuma, C. Power, M. A. Province, N. J. Samani, D. Schlessinger, R. Schmidt, T. I. Sørensen, T. D. Spector, K. Stefansson, U. Thorsteinsd. Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. *Nature Genetics*, 51(2):245–257, 2019. ISSN 15461718. doi: 10.1038/s41588-018-0309-3.

J. N. Lahey. DO OLDER WORKERS FACE DISCRIMINATION ? *An Issue in Brief , Center For Retiremennt Research as Boston College*, (33), 2005.

F. Lally and P. Crome. Understanding frailty. *Postgraduate medical journal*, 83(975):16–20, 2007. ISSN 1469-0756. doi: 10.1136/pgmj.2006.048587. URL `http://www.ncbi.nlm.nih.gov/pubmed/17267673%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2599957`.

J. J. Lee, R. Wedow, A. Okbay, E. Kong, O. Maghzian, M. Zacher, T. A. Nguyen-Viet, P. Bowers, J. Sidorenko, R. Karlsson Linnér, M. A. Fontana, T. Kundu, C. Lee, H. Li, R. Li, R. Royer, P. N. Timshel, R. K. Walters, E. A. Willoughby, L. Yengo, M. Agee, B. Alipanahi, A. Auton, R. K. Bell, K. Bryc, S. L. Elson, P. Fontanillas, D. A. Hinds, J. C. McCreight, K. E. Huber, N. K. Litterman, M. H. McIntyre, J. L. Mountain, E. S. Noblin, C. A. Northover, S. J. Pitts, J. F. Sathirapongsasuti, O. V. Sazonova, J. F. Shelton, S. Shringarpure, C. Tian, V. Vacic, C. H. Wilson, A. Okbay, J. P. Beauchamp, M. A. Fontana, J. J. Lee, T. H. Pers, C. A. Rietveld,

P. Turley, G. B. Chen, V. Emilsson, S. F. W. Meddens, S. Oskarsson, J. K. Pickrell, K. Thom, P. Timshel, R. de Vlaming, A. Abdellaoui, T. S. Ahluwalia, J. Bacelis, C. Baumbach, G. Bjornsdottir, J. H. Brandsma, M. P. Concas, J. Derringer, N. A. Furlotte, T. E. Galesloot, G. Girotto, R. Gupta, L. M. Hall, S. E. Harris, E. Hofer, M. Horikoshi, J. E. Huffman, K. Kaasik, I. P. Kalafati, R. Karlsson, A. Kong, J. Lahti, S. J. van der Lee, C. de Leeuw, P. A. Lind, K. O. Lindgren, T. Liu, M. Mangino, J. Marten, E. Mihailov, M. B. Miller, P. J. van der Most, C. Oldmeadow, A. Payton, N. Pervjakova, W. J. Peyrot, Y. Qian, O. Raitakari, R. Rueedi, E. Salvi, B. Schmidt, K. E. Schraut, J. Shi, A. V. Smith, R. A. Poot, B. St Pourcain, A. Teumer, G. Thorleifsson, N. Verweij, D. Vuckovic, J. Wellmann, H. J. Westra, J. Yang, W. Zhao, Z. Zhu, B. Z. Alizadeh, N. Amin, A. Bakshi, S. E. Baumeister, G. Biino, K. Bønnelykke, P. A. Boyle, H. Campbell, F. P. Cappuccio, G. Davies, J. E. De Neve, P. Deloukas, I. Demuth, J. Ding, P. Eibich, L. Eisele, N. Eklund, D. M. Evans, J. D. Faul, M. F. Feitosa, A. J. Forstner, I. Gandin, B. Gunnarsson, B. V. Halldórsson, T. B. Harris, A. C. Heath, L. J. Hocking, E. G. Holliday, G. Homuth, M. A. Horan, J. J. Hottenga, P. L. de Jager, P. K. Joshi, A. Jugessur, M. A. Kaakinen, M. Kähönen, S. Kanoni, L. Keltigangas-Järvinen, L. A. Kiemeney, I. Kolcic, S. Koskinen, A. T. Kraja, M. Kroh, Z. Kutalik, A. Latvala, L. J. Launer, M. P. Lebreton, D. F. Levinson, P. Lichtenstein, P. Lichtner, D. C. Liewald, A. Loukola, Life Lines Cohort Study, P. A. Madden, R. Mägi, T. Mäki-Opas, R. E. Marioni, P. Marques-Vidal, G. A. Meddens, G. McMahon, C. Meisinger, T. Meitinger, Y. Milaneschi, L. Milani, G. W. Montgomery, R. Myhre, C. P. Nelson, D. R. Nyholt, W. E. Ollier, A. Palotie, L. Paternoster, N. L. Pedersen, K. E. Petrovic, D. J. Porteous, K. Räikkönen, S. M. Ring, A. Robino, O. Rostapshova, I. Rudan, A. Rustichini, V. Salomaa, A. R. Sanders, A. P. Sarin, H. Schmidt, R. J. Scott, B. H. Smith, J. A. Smith, J. A. Staessen, E. Steinhagen-Thiessen, K. Strauch, A. Terracciano, M. D. Tobin, S. Ulivi, S. Vaccargiu, L. Quaye, F. J. van Rooij, C. Venturini, A. A. Vinkhuyzen, U. Völker, H. Völzke, J. M. Vonk, D. Vozzi, J. Waage, E. B. Ware, G. Willemsen, J. R. Attia, D. A. Bennett, K. Berger, L. Bertram, H. Bisgaard, D. I. Boomsma, I. B. Borecki, U. Bültmann, C. F. Chabris, F. Cucca, D. Cusi, I. J. Deary, G. V. Dedoussis, C. M. van Duijn, J. G. Eriksson, B. Franke, L. Franke, P. Gasparini, P. V. Gejman, C. Gieger, H. J. Grabe, J. Gratten, P. J. Groenen, V. Gudnason, P. van der Harst, C. Hayward, D. A.

Hinds, W. Hoffmann, E. Hyppönen, W. G. Iacono, B. Jacobsson, M. R. Järvelin, K. H. Jöckel, J. Kaprio, S. L. Kardia, T. Lehtimäki, S. F. Lehrer, P. K. Magnusson, N. G. Martin, M. McGue, A. Metspalu, N. Pendleton, B. W. Penninx, M. Perola, N. Pirastu, M. Pirastu, O. Polasek, D. Posthuma, C. Power, M. A. Province, N. J. Samani, D. Schlessinger, R. Schmidt, T. I. Sørensen, T. D. Spector, K. Stefansson, U. Thorsteinsdottir, A. R. Thurik, N. J. Timpson, H. Tiemeier, J. Y. Tung, A. G. Uitterlinden, V. Vitart, P. Vollenweider, D. R. Weir, J. F. Wilson, A. F. Wright, D. C. Conley, R. F. Krueger, G. D. Smith, A. Hofman, D. I. Laibson, S. E. Medland, M. N. Meyer, J. Yang, M. Johannesson, P. M. Visscher, T. Esko, P. D. Koellinger, D. Cesarini, D. J. Benjamin, M. Alver, Y. Bao, D. W. Clark, F. R. Day, N. A. Furlotte, P. K. Joshi, K. E. Kemper, A. Kleinman, C. Langenberg, R. Mägi, J. W. Trampush, S. S. Verma, Y. Wu, M. Lam, J. H. Zhao, Z. Zheng, J. D. Boardman, H. Campbell, J. Freese, K. M. Harris, C. Hayward, P. Herd, M. Kumari, T. Lencz, J. Luan, A. K. Malhotra, A. Metspalu, L. Milani, K. K. Ong, J. R. Perry, D. J. Porteous, M. D. Ritchie, M. C. Smart, B. H. Smith, J. Y. Tung, N. J. Wareham, J. F. Wilson, J. P. Beauchamp, D. C. Conley, T. Esko, S. F. Lehrer, P. K. Magnusson, S. Oskarsson, T. H. Pers, M. R. Robinson, K. Thom, C. Watson, C. F. Chabris, M. N. Meyer, D. I. Laibson, J. Yang, M. Johannesson, P. D. Koellinger, P. Turley, P. M. Visscher, D. J. Benjamin, and D. Cesarini. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics*, 50(8):1112–1121, 2018. ISSN 15461718. doi: 10.1038/s41588-018-0147-3.

S. F. Lehrer and W. Ding. Are genetic markers of interest for economic research? *IZA Journal of Labor Policy*, 6(1), 2017. ISSN 21939004. doi: 10.1186/s40173-017-0080-6.

A. E. Locke, B. Kahali, S. I. Berndt, A. E. Justice, T. H. Pers, F. R. Day, C. Powell, S. Vedantam, M. L. Buchkovich, J. Yang, D. C. Croteau-Chonka, T. Esko, T. Fall, T. Ferreira, S. Gustafsson, Z. Kutalik, J. Luan, R. Mägi, J. C. Randall, T. W. Winkler, A. R. Wood, T. Workalemahu, J. D. Faul, J. A. Smith, J. H. Zhao, W. Zhao, J. Chen, R. Fehrmann, K. Hedman, J. Karjalainen, E. M. Schmidt, D. Absher, N. Amin, D. Anderson, M. Beekman, J. L. Bolton, J. L. Bragg-Gresham, S. Buyske, A. Demirkan, G. Deng, G. Ehret, B. Feenstra, M. Feitosa, K. Fisc-

her, A. Goel, J. Gong, A. U. Jackson, S. Kanoni, M. E. Kleber, K. Kristiansson, U. Lim, V. Lotay, M. Mangino, I. M. Leach, C. Medina-Gomez, S. E. Medland, M. A. Nalls, C. D. Palmer, D. Pasko, S. Pechlivanis, M. J. Peters, I. Prokopenko, D. Shungin, A. Stančáková, R. J. Strawbridge, Y. J. Sung, T. Tanaka, A. Teumer, S. Trompet, S. W. van der Laan, J. van Setten, J. V. Van Vliet-Ostaptchouk, Z. Wang, L. Yengo, W. Zhang, A. Isaacs, E. Albrecht, J. Ärnlöv, G. M. Arscott, A. P. Attwood, S. Bandinelli, A. Barrett, I. N. Bas, C. Bellis, A. J. Bennett, C. Berne, R. Blagieva, M. Blüher, S. Böhringer, L. L. Bonnycastle, Y. Böttcher, H. A. Boyd, M. Bruinenberg, I. H. Caspersen, Y. I. Chen, R. Clarke, E. W. Daw, A. J. de Craen, G. Delgado, M. Dimitriou, A. S. Doney, N. Eklund, K. Estrada, E. Eury, L. Folkersen, R. M. Fraser, M. Garcia, F. Geller, V. Giedraitis, B. Gigante, A. S. Go, A. Golay, A. Goodall, S. D. Gordon, M. Gorski, H. J. Grabe, H. Grallert, T. B. Grammer, J. Gräßler, H. Grönberg, C. J. Groves, G. Gusto, J. Haessler, P. Hall, T. Haller, G. Hallmans, C. A. Hartman, M. Hassinen, C. Hayward, N. L. Heard-Costa, Q. Helmer, C. Hengstenberg, O. Holmen, J. J. Hottenga, A. L. James, J. Jeff, Johansson, J. Jolley, T. Juliusdottir, L. Kinnunen, W. Koenig, M. Koskenvuo, W. Kratzer, J. Laitinen, C. Lamina, K. Leander, N. R. Lee, P. Lichtner, L. Lind, J. Lindström, K. S. Lo, S. Lobbens, R. Lorbeer, Y. Lu, F. Mach, P. K. Magnusson, A. Mahajan, W. L. McArdle, S. McLachlan, C. Menni, S. Merger, E. Mihailov, L. Milani, A. Moayyeri, K. L. Monda, M. A. Morken, A. Mulas, G. Müller, M. Müller-Nurasyid, A. W. Musk, R. Nagaraja, M. M. Nöthen, I. M. Nolte, S. Pilz, N. W. Rayner, F. Renstrom, R. Rettig, J. S. Ried, S. Ripke, N. Robertson, L. M. Rose, S. Sanna, H. Scharnagl, S. Scholtens, F. Schumacher, W. R. Scott, T. Seufferlein, J. Shi, A. V. Smith, J. Smolonska, A. V. Stanton, V. Steinthorsdottir, K. Stirrups, H. M. Stringham, J. Sundström, M. A. Swertz, A. J. Swift, A. C. Syvänen, S. T. Tan, B. Tayo, B. Thorand, G. Thorleifsson, J. Tyrer, H. W. Uh, L. Vandenput, F. C. Verhulst, S. H. Vermeulen, N. Verweij, J. M. Vonk, L. L. Waite, H. R. Warren, D. M. Waterworth, M. N. Weedon, L. Wilkens, C. Willenborg, T. Wilsgaard, M. K. Wojczynski, A. Wong, A. F. Wright, Q. Zhang, E. P. Brennan, M. Choi, Z. Dastani, A. W. Drong, P. Eriksson, A. Franco-Cereceda, J. R. Gådin, A. G. Gharavi, M. E. Goddard, R. E. Handsaker, J. Huang, F. Karpe, S. Kathiresan, S. Keildson, K. Kiryluk, M. Kubo, J. Y. Lee, L. Liang, R. P. Lifton, B. Ma, S. A. McCarroll, A. J. McKnight, J. L. Min, M. F. Moffatt, G. W. Montgomery, J. M.

Murabito, G. Nicholson, D. R. Nyholt, Y. Okada, J. R. Perry, R. Dorajoo, E. Reinmaa, R. M. Salem, N. Sandholm, R. A. Scott, L. Stolk, A. Takahashi, F. M. Van't Hooft, A. A. Vinkhuyzen, H. J. Westra, W. Zheng, K. T. Zondervan, A. C. Heath, D. Arveiler, S. J. Bakker, J. P. Beilby, R. N. Bergman, J. Blangero, P. Bovet, H. Campbell, M. Caulfield, G. Cesana, A. Chakravarti, D. Chasman, P. S. Chines, F. S. Collins, D. Crawford, L. Cupples, D. Cusi, J. Danesh, U. de Faire, H. M. Den Ruijter, A. F. Dominiczak, R. Erbel, J. Erdmann, J. G. Eriksson, M. Farrall, S. B. Felix, E. Ferrannini, J. Ferrières, I. Ford, N. G. Forouhi, T. Forrester, O. H. Franco, R. T. Gansevoort, P. V. Gejman, C. Gieger, O. Gottesman, V. Gudnason, U. B. Gyllensten, A. S. Hall, T. B. Harris, A. T. Hattersley, A. A. Hicks, L. Hindorff, A. Hingorani, A. Hofman, G. Homuth, G. Hovingh, S. E. Humphries, S. C. Hunt, E. Hyppönen, T. Illig, K. B. Jacobs, M. R. Jarvelin, K. H. Jöckel, B. Johansen, P. Jousilahti, J. Jukema, A. Jula, J. Kaprio, J. J. Kastelein, S. M. Keinanen-Kiukaanniemi, L. A. Kiemeney, P. Knekt, J. S. Kooner, C. Kooperberg, P. Kovacs, A. T. Kraja, M. Kumari, J. Kuusisto, T. Lakka, C. Langenberg, L. L. Marchand, T. Lehtimäki, V. Lyssenko, S. Männistö, A. Marette, T. Matise, C. A. McKenzie, B. McKnight, F. L. Moll, A. D. Morris, A. P. Morris, J. C. Murray, M. Nelis, C. Ohlsson, A. J. Oldehinkel, K. K. Ong, P. A. Madden, G. Pasterkamp, J. F. Peden, A. Peters, D. S. Postma, P. P. Pramstaller, J. F. Price, L. Qi, O. Raitakari, T. Rankinen, D. C. Rao, T. K. Rice, P. Ridker, J. D. Rioux, M. Ritchie, I. Rudan, V. Salomaa, N. Samani, J. Saramies, M. A. Sarzynski, H. Schunkert, P. E. Sch. Genetic studies of body mass index yield new insights for obesity biology. *Nature*, 518(7538):197–206, 2015. ISSN 14764687. doi: 10.1038/nature14177.

W. Lu, R. Benson, K. Glaser, and E. Al. Relationship between employment histories and frailty trajectories in later life: evidence from the English Longitudinal Study of Ageing. Technical report, 2017.

M. Malul. Older workers' employment in dynamic technology changes. *Journal of Socio-Economics*, 38(5):809–813, 2009. ISSN 10535357. doi: 10.1016/j.socec.2009.05.005.

F. Mazzonna, F. Peracchi, and Others. Unhealthy Retirement? 2014. ISSN 0022-166X. doi: 10.3368/jhr.52.1.0914-6627R1.

J. R. Moon, M. M. Glymour, S. V. Subramanian, M. Avendaño, and I. Kawachi. Transition to retirement and risk of cardiovascular disease: Prospective analysis of the US health and retirement study. *Social Science and Medicine*, 75(3): 526–530, 2012. ISSN 02779536. doi: 10.1016/j.socscimed.2012.04.004. URL http://dx.doi.org/10.1016/j.socscimed.2012.04.004.

I. Mosca. The Effect of Retirement on Cognitive Decline: Evidence from the Abolition of the Marriage Bar in Ireland. 2017.

OECD. Pensions at a Glance 2011. Technical report, 2011. URL http://www.oecd-ilibrary.org/social-issues-migration-health/pensions-at-a-glance-2015_pension_glance-2015-en.

OECD. *Pensions at a Glance 2013: OECD and G20 Indicators*. 2013a. ISBN 9789264203921. doi: 10.1787/pension_glance-2013-en. URL http://www.oecd-ilibrary.org/finance-and-investment/pensions-at-a-glance-2013/gross-pension-replacement-rates_pension_glance-2013-12-en.

OECD. Pensions at a Glance 2013 Italia. Technical report, 2013b.

OECD. Pensions at a Glance 2015. Technical report, 2015.

A. Okbay, J. P. Beauchamp, M. A. Fontana, J. J. Lee, T. H. Pers, C. A. Rietveld, P. Turley, G. B. Chen, V. Emilsson, S. F. W. Meddens, S. Oskarsson, J. K. Pickrell, K. Thom, P. Timshel, R. De Vlaming, A. Abdellaoui, T. S. Ahluwalia, J. Bacelis, C. Baumbach, G. Bjornsdottir, J. H. Brandsma, M. Pina Concas, J. Derringer, N. A. Furlotte, T. E. Galesloot, G. Girotto, R. Gupta, L. M. Hall, S. E. Harris, E. Hofer, M. Horikoshi, J. E. Huffman, K. Kaasik, I. P. Kalafati, R. Karlsson, A. Kong, J. Lahti, S. J. Van Der Lee, C. Deleeuw, P. A. Lind, K. O. Lindgren, T. Liu, M. Mangino, J. Marten, E. Mihailov, M. B. Miller, P. J. Van Der Most, C. Oldmeadow, A. Payton, N. Pervjakova, W. J. Peyrot, Y. Qian, O. Raitakari, R. Rueedi, E. Salvi, B. Schmidt, K. E. Schraut, J. Shi, A. V. Smith, R. A. Poot, B. St Pourcain, A. Teumer, G. Thorleifsson, N. Verweij, D. Vuckovic, J. Wellmann, H. J. Westra, J. Yang, W. Zhao, Z. Zhu, B. Z. Alizadeh, N. Amin, A. Bakshi, S. E. Baumeister, G. Biino, K. Bønnelykke, P. A. Boyle, H. Campbell, F. P. Cappuccio,

G. Davies, J. E. De Neve, P. Deloukas, I. Demuth, J. Ding, P. Eibich, L. Eisele, N. Eklund, D. M. Evans, J. D. Faul, M. F. Feitosa, A. J. Forstner, I. Gandin, B. Gunnarsson, B. V. Halldórsson, T. B. Harris, E. G. Holliday, A. C. Heath, L. J. Hocking, G. Homuth, M. A. Horan, J. J. Hottenga, P. L. De Jager, P. K. Joshi, A. Jugessur, M. A. Kaakinen, M. Kähönen, S. Kanoni, L. Keltigangas-Järvinen, L. A. Kiemeney, I. Kolcic, S. Koskinen, A. T. Kraja, M. Kroh, Z. Kutalik, A. Latvala, L. J. Launer, M. P. Lebreton, D. F. Levinson, P. Lichtenstein, P. Lichtner, D. C. Liewald, A. Loukola, P. A. Madden, R. Mägi, T. Mäki-Opas, R. E. Marioni, P. Marques-Vidal, G. A. Meddens, G. Mcmahon, C. Meisinger, T. Meitinger, Y. Milaneschi, L. Milani, G. W. Montgomery, R. Myhre, C. P. Nelson, D. R. Nyholt, W. E. Ollier, A. Palotie, L. Paternoster, N. L. Pedersen, K. E. Petrovic, D. J. Porteous, K. Raïkkönen, S. M. Ring, A. Robino, O. Rostapshova, I. Rudan, A. Rustichini, V. Salomaa, A. R. Sanders, A. P. Sarin, H. Schmidt, R. J. Scott, B. H. Smith, J. A. Smith, J. A. Staessen, E. Steinhagen-Thiessen, K. Strauch, A. Terracciano, M. D. Tobin, S. Ulivi, S. Vaccargiu, L. Quaye, F. J. Van Rooij, C. Venturini, A. A. Vinkhuyzen, U. Völker, H. Völzke, J. M. Vonk, D. Vozzi, J. Waage, E. B. Ware, G. Willemsen, J. R. Attia, D. A. Bennett, K. Berger, L. Bertram, H. Bisgaard, D. I. Boomsma, I. B. Borecki, U. Bültmann, C. F. Chabris, F. Cucca, D. Cusi, I. J. Deary, G. V. Dedoussis, C. M. Van Duijn, J. G. Eriksson, B. Franke, L. Franke, P. Gasparini, P. V. Gejman, C. Gieger, H. J. Grabe, J. Gratten, P. J. Groenen, V. Gudnason, P. Van Der Harst, C. Hayward, D. A. Hinds, W. Hoffmann, E. Hyppönen, W. G. Iacono, B. Jacobsson, M. R. Järvelin, K. H. Jöckel, J. Kaprio, S. L. Kardia, T. Lehtimäki, S. F. Lehrer, P. K. Magnusson, N. G. Martin, M. Mcgue, A. Metspalu, N. Pendleton, B. W. Penninx, M. Perola, N. Pirastu, M. Pirastu, O. Polasek, D. Posthuma, C. Power, M. A. Province, N. J. Samani, D. Schlessinger, R. Schmidt, T. I. Sørensen, T. D. Spector, K. Stefansson, U. Thorsteinsdottir, A. R. Thurik, N. J. Timpson, H. Tiemeier, J. Y. Tung, A. G. Uitterlinden, V. Vitart, P. Vollenweider, D. R. Weir, J. F. Wilson, A. F. Wright, D. C. Conley, R. F. Krueger, G. Davey Smith, A. Hofman, D. I. Laibson, S. E. Medland, M. N. Meyer, J. Yang, M. Johannesson, P. M. Visscher, T. Esko, P. D. Koellinger, D. Cesarini, and D. J. Benjamin. Genome-wide association study identifies 74 loci associated with educational attainment. *Nature*, 533(7604):539–542, 2016. ISSN 14764687. doi: 10.1038/nature17671.

R. Ottman. Gene – Environment Interaction : Definitions and Study Designs. 25: 764–770, 1996.

N. W. Papageorge and K. Thom. GENES, EDUCATION, AND LABOR MARKET OUTCOMES: EVIDENCE FROM THE HEALTH AND RETIREMENT STUDY. *NBER Working Papers*, 2018. ISSN 0717-6163. doi: 10.1007/s13398-014-0173-7. 2. URL http://www.nber.org/papers/w25114.

K. Rockwood, M. Andrew, and A. Mitnitski. A comparison of two approaches to measuring frailty in elderly people. *J Gerontol A Biol Sci Med Sci.*, 62((7)): 738–43, 2007.

R. A. Roiland, F. Lin, C. Phelan, and B. P. Chapman. Stress regulation as a link between executive function and pre-frailty in older adults. *Journal of Nutrition, Health and Aging*, 19(8):828–838, 2015. ISSN 17604788. doi: 10.1007/s12603-015-0476-1.

Q. Ruan, Z. Yu, M. Chen, Z. Bao, J. Li, and W. He. Cognitive frailty, a novel target for the prevention of elderly dependency. *Ageing Research Reviews*, 20:1–10, 2015. ISSN 18729649. doi: 10.1016/j.arr.2014.12.004. URL http://dx.doi.org/10.1016/j.arr.2014.12.004.

A. Rustichini, W. G. Iacono, J. J. Lee, and M. McGue. Polygenic score analysis of educational achievement and intergenerational mobility. *Working Paper*, pages 1–57, 2018. URL http://users.econ.umn.edu/$\sim$rusti001/Research/Genetics/Polygenic_Analysis.pdf.

B. Santos-Eggimann, P. Cuénoud, J. Spagnoli, and J. Junod. Prevalence of frailty in middle-aged and older community-dwelling Europeans living in 10 countries. *Journals of Gerontology - Series A Biological Sciences and Medical Sciences*, 64(6): 675–681, 2009. ISSN 10795006. doi: 10.1093/gerona/glp012.

L. Schmitz and D. Conley. Modeling Gene-Environment Interactions With Quasi-Natural Experiments. *Journal of Personality*, 85(1):10–21, 2017. ISSN 14676494. doi: 10.1111/jopy.12227.

L. L. Schmitz and D. Conley. The Impact of Late-Career Job Loss and Genotype on Body Mass Index. *NBER Working Paper Series*, 22348:43, 2016.

N. Sirven. On the socio-economic determinants of frailty: findings from panel and retrospective data from share. 2012. URL http://www.irdes.fr/EspaceAnglais/Publications/WorkingPapers/DT52SocioEconomicDeterminantsFrailty.pdf.

N. Sirven and T. Rapp. The cost of frailty in France. *European Journal of Health Economics*, 18(2):1–11, 2016. ISSN 16187601. doi: 10.1007/s10198-016-0772-7.

U. Staudinger, M. Marsiske, and P. Baltes. Resilience and reserve capacity in later adulthood: Potentials and limits of development across the life span. In *Developmental psychopathology*, pages 801–847. Wiley, New York, 1995.

H. Strulik. Frailty, mortality, and the demand for medical care. *Journal of the Economics of Ageing*, 6:5–12, 2015. ISSN 2212828X. doi: 10.1016/j.jeoa.2015.09.001. URL http://dx.doi.org/10.1016/j.jeoa.2015.09.001.

O. Theou, L. Cann, J. Blodgett, L. M. K. Wallace, T. D. Brothers, and K. Rockwood. Modifications to the frailty phenotype criteria: Systematic review of the current literature and investigation of 262 frailty phenotypes in the survey of health, ageing, and retirement in Europe. *Ageing Research Reviews*, 21:78–94, 2015. ISSN 18729649. doi: 10.1016/j.arr.2015.04.001. URL http://dx.doi.org/10.1016/j.arr.2015.04.001.

H. Van Kippersluis and C. A. Rietveld. Pleiotropy-robust Mendelian randomization. *International Journal of Epidemiology*, 47(4):1279–1288, 2018. ISSN 14643685. doi: 10.1093/ije/dyx002.

J. Viña, F. J. Tarazona-Santabalbina, P. Pérez-Ros, F. M. Martínez-Arnau, C. Borras, G. Olaso-Gonzalez, A. Salvador-Pascual, and M. C. Gomez-Cabrera. Biology of frailty: Modulation of ageing genes and its importance to prevent age-associated loss of function. *Molecular Aspects of Medicine*, 50:88–108, 2016. ISSN 18729452. doi: 10.1016/j.mam.2016.04.005. URL http://dx.doi.org/10.1016/j.mam.2016.04.005.

E. Ware, L. Schmitz, A. Gard, and J. Faul. HRS Documentation Report: HRS Polygenic Scores – Release 2; 2006-2012 Genetic Data. (October):44, 2018.

A. J. Woods, R. A. Cohen, and M. Pahor. Cognitive frailty: Frontiers and challenges. *Journal of Nutrition, Health and Aging*, 17(9):741–743, 2013. ISSN 12797707. doi: 10.1007/s12603-013-0398-8.

Academic Activities Report

The doctoral path was structured in three years with a first year of joint courses and research and two more completely devoted to research. In addition to the courses offered by the University of Bologna, I took part to the Winter School on Advanced Method for Causal Inference and Policy Evaluation held in Trento at the IRVAAP Bruno Kessler Foundation and the Summers School in Genome-Wide Data Analysis held in Amsterdam at the Tinbergen Institute. To develop the first chapter of the present thesis, I spent three months at the Department of Applied Economics of the Erasmus School of Economics in Rotterdam where I presented my research, hosted by Professor C.A. Rietveld. During the three years, the quality of the research carried out was evaluated by the Faculty during three distinct PhD Forums held in July 2017, March 2018 and July 2019.