Empirical Essays on Health and Human Capital

Fast, Thomas

2013

Empirical Essays on Health and Human Capital

Thomas Eriksson

Lund Economic Studies Number 172

Distribution by the Department of Economics
Lund University
P.O. Box 7082
S-220 07 LUND
Sweden
Telephone: +46 (0)46 222 00 00
www.nek.lu.se

ISBN 978-91-7473-591-8
ISSN 0460-0029

Printed in Sweden by Media-Tryck, Lund University
Lund 2013

To my family

Acknowledgement

Before becoming a PhD candidate, I had the privilege of working at the Lund University Centre for Health Economics (LUCHE). I am grateful for that, Professors Kristian Bolin and Björn Lindgren introduced me to the world of academia and motivated me to apply for the PhD program.

This PhD thesis would not have been what it is today without the help of my supervisors, Professors Kristian Bolin and Petter Lundborg. I appreciate Kristian´s help and comments, which vastly improved this thesis. Furthermore, Petter has guided me and been a good discussion partner regarding econometric techniques.

During my time as a PhD candidate, I have attended various conferences and courses outside the department. These conferences and courses have been important in my progression as a PhD candidate. I gratefully acknowledge financial support from: *LUCHE*, *Stiftelsen för främjande av ekonomisk forskning vid Lunds universitet*, *Health Economics Program* at the Department of Economics, and the Department of Economics. In addition, the seminars with the labor and health group at the Department of Economics and the brown-bag seminars at Centre for Economic Demography have also been important in my academic education.

I would like to thank Professors Sonja Opper and Christer Gunnarsson for letting me work as a coordinator for the Comparative Institutional Analysis research group. Furthermore, I am grateful for the thorough work Håkan Selin from Uppsala University did as the discussant during my final seminar.

The time in the PhD program would not have been the same without the companion of other PhD students. Lina Maria Ellegård, thanks for being my roommate and giving me invaluable help during the years. In addition, I have always enjoyed the evenings where we played board games. Who could know that board games could contain so much economics? Furthermore, I am grateful for having the opportunity to co-author a paper with Sofie Gustafsson. Other, current and former PhD students have helped me with practical issues. In addition, they have been really good colleagues by providing a good working environment. Elvira Andersson Karin Bergman, Daniel Ekeblom, Albin Erlanson, Jens Gudmundsson, Wolfgang Hess, Bujar Huskaj, Erik Jonasson, Ida Johansson, Gustav Kjellson, Åsa Ljungvall, Liu Lu, Pernilla Johansson, Therese Nilsson, Maria Persson, Hilda Ralsmark, Alexander Reffgen, Eric Rehn and Emma Svensson all deserves a special thanks for making the best out of my time as a PhD candidate.

Other colleges at the department should also be mentioned. Peter Jochumzen and David Edgerton, thanks for always having the door open to discuss econometric issues. Birger Nilson and Carl Hampus Lyttkens, thanks for organizing tournaments to ensure the physical standard at the department. Tommy Andersson thanks for taking the nutritious responsibility in the lunch room and introducing the “Kebab-Wednesday”. Several other colleagues should also be mentioned: Mariana Anton, Ulf Gerdtham, Pontus Hansson, Håkan J Holm, Agneta Kruse, Paul Linge, Anna Marciniak, Inga Persson, Peter Schüller, Agnes Bridge Walton and Karin Wandér.

Last, but not the least, I would like to thank my family: Lena, my parents Anna-Lisa and Gunnar, my sister Mikaela and my brother-in-law Per. Thanks for your unconditional support and help during the years. I love you all very much.

Thomas Eriksson
Lund, August 2013

Contents

Chapter 1 Introduction

Besides this introductory chapter, this thesis consists of three empirical papers. The three papers analyze different aspects regarding individual choice in the realm of health and labor economics. The remainder of the introductory chapter is organized as follows: section 1.1 introduces the concept of individual choice in regards to health and labor economics. Section 1.2 will discuss some of the influential theoretical work used in this thesis. Section 1.3 concludes with a summary of the three papers in this thesis.

1.1 Individual choice

An individual is faced with many choices through life. For instance, the individual must choose whether or not to continue to university and when to retire. Both theoretical and empirical work has tried to deepen the understanding of individual choice. Among the theoretical works, Gary Becker's work (1964) on human capital is among the most influential. Regarding empirical work, numerous aspects has been analyzed. For instance, Freeman (1986) and Heckman et al (2001) have analyzed the demand for education while Lazear (1986) and Scholz et al (2006) have analyzed individuals´ retirement decision.

The scarcity of resources (e.g. time and money) is a fundamental notion in individual choice. The scarcity of resources can be illustrated with educational attainment. To pass a course, the individual must allocate time for attending lectures and revision. Given that the individual only have 24 hours to allocate per day, he or she must decide which course to take and not to take. Hence, the cost of passing a course can be thought of in terms of not being able to attend some other course. This is referred to the opportunity cost. For instance, the cost of attending an economics course can be thought of in terms of not being able to attend an economics course[1].

The choices an individual makes, depends on his or hers preferences and the preferences can be illustrated with a utility function. A simple version of a utility function can be illustrated by:

$$U = f(C, L) \qquad (1)$$

where the individual´s utility is a function of consumption of goods (*C*) and leisure (L). It is usually assumed that: (1) the marginal utility of both *C* and L are non-negative, and (2) there is a trade-off between consumption and leisure. The trade-off between consumption and leisure is present because more time spent in leisure leads to less time devoted to work, leading to lower income, and hence, less consumption. Hence, by choosing the optimal bundle of consumption and leisure, the individual maximizes his or her utility.

The trade-off between consumption and leisure may be realized in different time periods. By returning to the example with educational attainment and assuming that education leads to a higher paid job, schooling today implies less leisure today but more consumption in the future. Given that the trade-off between consumption and leisure is not realized in the same time period, discounting may be employed to calculate the present value. The present value calculates future costs and benefits into today´s value. To give

[1] It is certainly possible that gifted students can attend more than one course at a time. However, it is not possible that they can attend all courses at once.

an example, suppose that an individual will receive $100, 5 years from now and that the discount rate is 4 percent on an annual basis. The present value of the $100, 5 years from now, is approximately $82[2]. The discount rate reflects the individual´s time preferences, which may vary between individuals. Several studies have estimated individuals´ discount rate (see Luhmann, 2013 for a recent example).

1.2 Influential theoretical work

Becker´s human-capital model and Grossman´s extension of Becker´s human-capital model has been influential for theoretical and empirical research regarding individual choice. This section will introduce these models and discuss some of the critique towards the models.

In Becker's human-capital model (1964), the individual combines non-working time and goods to produce commodities denoted Z. For instance, the individual can combine own time and goods bought in a grocery store to produce the commodity lunch. The production function for producing commodity Z_i can be written as:

$$Z_i = f_i(x_i, t_i | E_i) \qquad (2)$$

where x_i and t_i denotes the possible combinations of goods and own time to produce a commodity. For instance, by increasing the amount of pre-cooked meals, the amount of own time needed to produce a lunch decreases. E_i denotes human capital (e.g. educational attainment and on-the-job training),

[2] The present value of x is calculated by: $\frac{x}{(1+d)^t}$ where *d* is the discount rate and *t* is the time in years. Hence, the present value in the example is: $\frac{100}{(1+0.04)^5}$

ability and other environmental factors influencing the production of commodity Z_i. For instance, the size of the individual's household can be an environmental factor.

Compared with the utility function in equation (1), where the individual received utility directly by consuming goods and leisure, the individual in Becker´s human-capital model receives utility by consuming commodities (e.g. consuming a lunch). Hence, the utility function can be written as:

$$U = f(Z_1, Z_2, \dots Z_N) \qquad (3)$$

To maximize his or hers utility, the individual choose the optimal bundle of commodities, subject to the budget constraint and a time constraint.

In Becker's original formulation of the human-capital, the individual demands human-capital for two reasons. First, given that human-capital influences the production function of commodities, the individual demands human-capital to increase the production of commodities. For instance, the individual can attend a cooking-class to improve his or her efficiency in preparing lunches. Second, given that human-capital is positively correlated with wage rate, the individual can invest in human-capital to increase his or hers labor-income, which, in turn, increases the individual's budget constraint. Hence, by investing in human-capital, the individual can buy more inputs needed to produce commodities.

1.2.1 Introducing health

In Becker's human-capital model, health is incorporated in the stock of human-capital. Hence, the individual is assumed to demand health to both increase the production in the labor market and the efficiency at which the individual produces commodities and performance on the labor market.

Grossman (1972a; 1972b)[3] notes that the individual could also demand health because good health enables the individual to work for a longer period of time as less time is spent being sick. This is referred to as the investment aspect of health in Grossman's framework. In addition, Grossman notes that the individual is likely to demand health because good health is enjoyable. This is usually referred to as the consumption aspect of health in the Grossman's framework.

The stock of health-capital can be written as:

$$HEC_j = HEC_{j-1}(1-\mu) + I_{j-1}^{HEC} \qquad (4)$$

where HEC_{j-1} is the stock of health-capital at time $j-1$, μ is the depreciation rate of health-capital and I_{j-1}^{HEC} is the health-capital investment (e.g. physical activity) at time $j-1$. The depreciation rate is assumed to increase by age, and hence, the demand for health-capital is likely to increase with age.

Based on the consumption aspect and investment aspect of health-capital, the demand for health-capital is likely to differ between individuals. Starting with the consumption aspect, an individual enjoying an active lifestyle is likely to demand more health-capital, compared with an individual who enjoys a passive lifestyle. Continuing to the investment aspect, the demand for health-capital depends on the possibility in the labor market. To give an example, consider two individuals (A and B) who are similar in every aspect, except their possibilities in the labor market. Individual A works in a sector,

---

[3] An overview regarding theoretical extensions and empirical applications can be found in Grossman (2000)

which is currently booming and individual B works in a sector, which is currently in a bust. Given that the returns on investment in heath-capital are larger for individual A than for individual B, individual A is likely to invest in more health capital than individual B.

Grossman's framework has received some critique. A recent exchange of views can be found in Zweifel (2012; 2013) and Kaestner (2013). The main critique against Grossman's framework is that it is completely deterministic, implying that the individual has full control of its stock of health-capital. Zweifel et al (2009) offers an alternative framework where the individual faces a stochastic process in which the individual can either become sick or healthy. If the individual becomes sick, the individual will invest in health-capital (e.g. visit a physician). If the individual becomes healthy, the individual will not invest in health-capital. Therefore, compared with Grossman's framework where health is considered to be continuous, the individual in Zweifel's framework do not engage in preventive health-capital investments, which is a strong assumption. Hence, the position in this thesis is that even if Grossman's framework has received some critique, Grossman's framework offers a good starting point for applied research regarding individual's health investments.

1.3 Summary of the three papers

The first paper adds to the broad literature regarding the effect of alcohol consumption on educational attainment. Given that no consensus exists in the literature regarding this issue, this paper adds to the literature by analyzing the effect of alcohol consumption on educational attainment, using a life cycle approach. In addition, whereas many studies have analyzed a US context, this study used a dataset covering Swedish adolescents.

The results in this paper show that alcohol consumption above the equivalence of two bottles of wine per week, during high school (tenth to twelfth grade), reduces the probability of continuing to university. Furthermore, among university students, the results indicate that those who abstain from alcohol finish university courses in a slower rate compared with those who do not abstain from alcohol.

The second paper analyzed the effect of health on the probability of retiring. This paper used a dataset covering individuals who are 50 years or older from eleven European countries to analyze if the effect of health on the probability of retiring, differs between the countries. To ensure cross-country comparability, this paper used overnight hospitalization as a proxy for a negative health shock. Given that individuals might not be randomized to either experience an overnight hospitalization or not, the paper used propensity score matching to avoid selection bias.

When all eleven European countries were analyzed together, overnight hospitalization has a positive effect on the retiring probability. When analyzed separately, the results are mixed; overnight hospitalization has a positive effect on the retiring probability in some countries and no effect in some countries. The disparities in results can reflect differences in the institutional settings across countries. For instance, the access to rehabilitory care may differ between the countries.

The third paper (co-authored with Sofie Gustafsson) analyzed the pharmaceutical-based health investments differences between immigrants and natives in Sweden. This paper used an interview survey combined with a registered database containing all prescribed pharmaceuticals. Immigrants were divided into the groups according to their region of origin: Nordic, Western and non-Western. The results show that there are differences in the utilization of prescribed pharmaceuticals between immigrants and natives. For instance, when looking at all pharmaceuticals analyzed together, males with a non-Western origin are less likely to access prescribed pharmaceuticals, compared with native males. Turning to specific pharmaceutical groups, the results show that immigrants, compared with natives, are less likely to have access to first-line treatments for high blood pressure, heart failure and kidney diseases.

References

Becker, Gary S. 1964. Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education. Chicago, University of Chicago Press.

Cameron, Stephen V., and James J. Heckman. 2001. "The Dynamics of Educational Attainment for Black, Hispanic, and White Males." The Journal of Political Economy (3): 455.

Freeman, Richard B. 1986. "Demand for Education." In Handbook of Labor Economics, edited by Orley C. Ashenfelter and Richard Layard, Volume 1:357–386. Elsevier.

Grossman, Michael. 1972a. "On the Concept of Health Capital and the Demand for Health." Journal of Political Economy 80 (2): 223–256.

———. 1972b. The Demand for Health : a Theoretical and Empirical Investigation / Michael Grossman. National Bureau of Economic Research Occasional Paper: 119. New York, 1972.

———. 2000. "The Human Capital Model." In Handbook of Health Economics. Vol. 1A, edited by Anthony J. Culyer and Joseph P. Newhouse. Amsterdam : Elsevier, cop. 2000.

Kaestner, Robert. 2013. "The Grossman Model after 40 Years: A Reply to Peter Zweifel." European Journal of Health Economics 14 (2) (April): 357–360.

Lazear, Edward P. 1986. "Retirement from the Labor Force." In Handbook of Labor Economics, edited by Orley C. Ashenfelter and Richard Layard, Volume 1:305–355. Elsevier. http://www.sciencedirect.com/science/article/pii/S1573446386010088.

Luhmann, Christian C. 2013. "Discounting of Delayed Rewards Is Not Hyperbolic." Journal of Experimental Psychology: Learning, Memory, and Cognition (January 28).

Scholz, J.K. (1, A. (1) Seshadri, and S. (3) Khitatrakun. 2006. "Are Americans Saving 'Optimally' for Retirement?" Journal of Political Economy 114 (4): 607–643.

Zweifel, P., F. Breyer, and M. Kifmann. 2009. Health Economics. Springer Science & Business

Zweifel, Peter. 2012. "The Grossman Model after 40 Years." European Journal of Health Economics 13 (6) (December): 677–682.

———. 2013. “The Grossman Model after 40 Years: Response to Robert Kaestner.” The European Journal of Health Economics 14 (2) (April 1): 361–362.

2. The effect of alcohol consumption on university initiation – the case of Sweden

Abstract

This paper adds to the broad literature regarding the effect of alcohol consumption on educational attainment. Given that no consensus exists in the literature regarding this issue, this paper adds to the literature by analyzing the effect of alcohol consumption on educational attainment, using a life cycle approach. In addition, whereas many studies have analyzed a US context, this study used a dataset[4] covering Swedish adolescents.

The results in this paper show that alcohol consumption above the equivalence of two bottles of wine per week, during high school (tenth to twelfth grade), reduces the probability of continuing to university. Furthermore, among university students, the results indicate that those who abstain from alcohol finish university courses in a slower rate compared with those who do not abstain from alcohol.

[4] Financial support from the Swedish Council for Working Life and Social Research FAS (dnr 2007–0318) is gratefully acknowledged.

2.1 Introduction

Adolescents´ alcohol consumption has attracted much attention from researchers. The negative effect of alcohol on health is well established (e.g. Corrao et al. 1999; Rehm et al. 2010) and adolescents´ alcohol consumption has, for instance, been discussed in relation to traffic fatalities (e.g. Dee 1999; Lovenheim and Slemrod 2010). In addition, much research has been directed towards how alcohol consumption affects educational attainment. However, there is no consensus regarding the effect of alcohol consumption on educational attainment.

A large number of studies have found that alcohol consumption has a negative effect on educational attainment. Yamada et al (1996) found that alcohol consumption during high school had a negative effect on the probability of high school graduation. In addition, alcohol initiation before the age of 14 and binge drinking during high school increases the probability of late high school graduation (Koch and McGeary 2005; Renna 2007). Furthermore, binge drinking in the final year of high school increases the probability of graduating with a General Education Development (GED), instead of a high school diploma (Renna 2008). Even though a GED is supposed to be a substitute for a high school diploma, individuals with a GED have a labor market position equivalent to high school dropouts, and the years of schooling completed after high school are fewer compared to individuals with a high school diploma (Cameron & Heckman, 1993).

Grade point average (GPA) has also been used to measure educational attainment. GPA is important for school admission and future earnings (e.g. Jones and Jackson 1990). Balsa, Giuliano, & French (2011) found a statistically significant negative effect of alcohol consumption on GPA for males, but no statistically significant effect for women. Some of this GPA reduction can be explained via a reduction in hours spent studying due to alcohol consumption (Wolaver 2002; Williams, Powell, and Wechsler 2003). Alcohol consumption has also been shown to reduce the academic performance of 21 year olds in the US (Scott E. Carrell, Mark Hoekstra, and James E. West 2003). Furthermore, heavy drinking during high school has a negative effect on the years of completed schooling after high school (Cook and Moore 1993). A similar study found that heavy alcohol use at the age of 16 has a negative effect on educational attainment at the age of 42 for men, but not for women (Staff et al. 2008).

The evidence that alcohol consumption adversely affects educational attainment has been questioned. Chatterji (2006) used a bivariate model approach to jointly estimate alcohol use and educational attainment. She concludes that the results are sensitive to the choice of statistical specification and, hence, that there is little evidence that the association between alcohol consumption and educational quality is causal. Similarly, Koch & Ribar (2001) found that results were sensitive to specification, and concluded that the effect of youthful drinking on educational attainment was likely to be small. It has also been shown that instead of a causal effect, it is more likely an association as some of the negative effect of alcohol consumption on GPA disappears when controlling for factors such as discounting, risk aversion and the use of other drugs (DeSimone 2010).

Thus, the empirical evidence regarding the relationship between adolescent alcohol consumption and human-capital investment later in life is inconclusive. Improved evidence is needed, for instance, to formulate and construct cost-effective public policies.

This paper has two objectives. The first objective is to estimate the effect of alcohol consumption among high school students5 on the probability of continuing to university. The second objective is to estimate the effect of

[5] This study focuses on students in the Swedish “Gymnasium”, which corresponds to years 10 to 12 in the Swedish school system. Henceforth will I only use the term “high school”

alcohol consumption among university students on the probability of increasing their university education. Both objectives use information on drinking behavior among Swedish adolescents. To put drinking habits in Sweden in relation to other countries, a comparison of 15 year olds shows that Swedes drink below the mean in the OECD countries, but more than their peers in the USA (OECD, 2009).

The reminder of this paper is organized as follows. Section 2.2 presents the framework. Section 2.3 describes the data and Section 2.4 the method. The results are presented in Section 2.5 and Section 2.6 concludes the paper.

2.2 Theoretical framework

An extensive literature, initiated by Gary Becker´s development of the theory of human capital, examines both theoretically and empirically, individual human-capital related behaviors (Becker, 1964). In Becker´s original formulation, the individual invests in human capital (e.g. education and on-the-job training) as this enhances future labor market productivity which, in turn, enhances labor market earning. The human capital investment (henceforth only educational investments) depends on the marginal cost and marginal benefit of educational investments. Given that the marginal cost of educational investment is assumed to increase and the marginal benefit is assumed to decrease, an individual will choose the optimal level of educational investment.

Individual characteristics can influence the marginal benefit and marginal cost. For instance, an individual's ability increases the marginal benefit of educational attainment as an individual with high ability, compared with an individual with low ability, is more able to produce human capital from education. In addition, family conditions, such as the number of siblings, may affect the marginal cost of educational investment. On the one hand, the presence of siblings may lower the marginal cost of educational investment if siblings help each other with homework. On the other hand, having siblings may increase the marginal cost of educational investment if siblings compete for their parents' resources (Becker 1964). For instance, the parents may not be able to help with the schoolwork of all their children.

Alcohol consumption may affect educational attainment in several ways. First, alcohol consumption leads to less time devoted to studies (Williams, Powell, and Wechsler 2003; Balsa, Giuliano, and French 2011). Second, human capital investment (e.g. buying books and tuition fees) may be crowded out by the purchase of alcoholic beverages. Third, alcohol consumption may influence educational attainment via health. As already mentioned, the negative effect of alcohol on health is well established (e.g. Corrao et al., 1999; Rehm et al., 2010). As the total payoff of an investment is determined by the length of time over which benefits may be utilized, a decrease in the expected length of life decreases the expected return on human capital investments (Becker 2007). Hence, the negative effect of alcohol consumption on health leads to a lower payoff for human capital investments and, hence, educational attainment.

Further, alcohol could be a mediator of social conditions such as family background (Fitzgerald and Zucker 1995; Fischer et al. 2007) or social capital (Winstanley et al., 2008). That is, bad social conditions can lead to harmful alcohol consumption, which in turn leads to lower educational attainment. Finally, even though the possible negative effect of alcohol may not be anticipated by the individual (e.g. Cook & Moore, 1993), it is possible that a difference in drinking behavior reflects individuals' different time preferences – a more forward looking individual will drink less compared with a less forward looking individual.

2.3 Data

This paper used the dataset called the Swedish Survey of Living conditions (ULF), complemented with the Longitudinal Integration Database for Health Insurance and Labor Market Studies (LISA), both originating from Statistics Sweden. The ULF survey uses a representative sample of the Swedish population aged 16 and older. The interviews are conducted either via telephone or via home visits. Every wave lasts for two years. Besides the bi-annual waves, the ULF survey also consists of a health focusing panel, which includes topics such as family background and alcohol consumption and educational attainment. Four health waves (1980/1981, 1988/1989, 1996/1997 and 2004/2005) have been conducted so far. Unfortunately, the first wave (1980/1981) did not contain any questions regarding alcohol consumption and, hence, it is not included in this study. The response rate was 80 percent in 1988/1989, 78 percent in 1996/1997 and 75 percent in 2004/2005 (Statistics Sweden 2013).

The LISA database is a register database containing all individuals aged 16 and older in Sweden. The information in the LISA database is not self-reported, but originates from different Swedish registers containing information on, for instance, educational attainment and income. For this study, early data on educational attainment from 1994 through 2006 was available from the LISA database.

Dependent variable

Educational attainment is coded according to the Swedish Educational Terminology 2000 (SUN 2000) classification in both the ULF survey and the LISA database. The SUN 2000 classification, which is comparable with the International Standard Classification of Education 1997 (ISCED 97), includes information on whether the individual has graduated from high school, has at least two years of university education, the number of years of university education exceeding two years or has a PhD. A thorough description of the SUN 2000 classification can be found in Statistics Sweden (2004). Unfortunately, neither the LISA database nor the ULF survey contains any information regarding grades.

There are two advantages of using the LISA database over the ULF survey. First, self-reported educational attainment may be biased (Balsa, Giuliano, and French 2011), and as the LISA database is not self-reported, the self-reported bias is eliminated. To give an example of the bias; when looking at those interviewed in 1996, 30 percent reported having an educational attainment in the ULF survey different to the attainment reported in the LISA database. Second, given that the LISA database is available on a yearly basis between 1994 and 2006, it is possible to measure educational attainment at the same age across individuals, and compare measures of educational attainment between individuals.

Alcohol measure

The ULF survey contains a detailed measure of alcohol consumption. The respondents were asked how often and how much beer, wine and hard liquor they drank per week. There was a slight change in wording in the questions between waves. In 1988/1989, the respondents were asked how much they drank during a week, on average, but in 1996/1997 and 2004/2005 the respondents were asked how much they had drunk the week before. To obtain a large enough sample for the econometric analysis, the waves are analyzed together.

Given that beers and wines differ in alcohol content, the respondents were asked to differentiate, for instance, between beers with different alcohol contents. All alcohol measures are recoded into one single measure of alcohol consumption, according to their alcohol content. The variable containing overall alcohol consumption is coded into categorical variables to control for non-linearity. Several thresholds of alcohol consumption have been suggested in the literature (e.g. Holman, English, Milne & Winter, 1996). However, these levels are defined out of medical considerations and not educational attainment per se. Hence, to create easily comprehensible levels, alcohol consumption is coded in the equivalence of wine bottles (70 cl)[6].

Control variables

Based on the Becker human capital model, variables that may influence educational attainment are included in the analysis. For instance, the number of siblings and health are controlled for. In addition, the individual's ability is likely to influence the cost of educational investments. Unfortunately, the dataset does not include any information regarding individual ability. Instead, parents' working situation is included in the analysis as proxy variables for the individual´s ability (see Björklund & Salvanes (2010), Pronzato (2012) and Anger & Heineck (2010) for discussions regarding intergenerational mobility). Table 1 presents the descriptive statistics for the dependent and independent variables by the age groups 16-19 and 24-27.

28 percent in the age group 24-27 have begun their university studies whereas, as expected, none in the age group 16-19 have started at university. It can be noted that "1≤x<2" means that the individual drank between one and two bottles of wine each week. An extra threshold, 0.5 bottles per week, is included as proportionally many individuals answered between zero and up to one bottle per week. It can also be noted the alcohol consumption increases by age, which is expected, as the legal age of buying alcohol is 18 in Sweden.

[6] For instance, 33 cl of strong beer (defined in the ULF survey as containing 4.5 volume percent) is equivalent to 14.85 cl of wine given that the ULF survey assumes that wine contains 10 volume percent.

Table 1. Sample means, by age group

Variables	Description	Age group 16-19	Age group 24-27
		Fractions	
Dependent variables			
HIGHEDUC	1 if having started higher education	0%	28.5%
Respondents weekly alcohol consumption in wine bottles (70 cl)			
Abstent	No alcohol consumption	20.7%	6.8%
0<wine<0.5	Between 0 and 0.5 bottle	60.0%	36.6%
0.5≤wine<1	Between 0.5 and 1 bottle	5.9%	17.3%
1≤wine<2	Between 1 and 2 bottles	7.8%	19.7%
2≤wine	2 or more bottles	5.4%	19.3%
Control variables			
MALE	Male	49.2%	49.2%
AGE	Age at interview	17.52	25.52
no_sibling	Have no sibling	8.1%	6.4%
one_sibling	Have one sibling	35.9%	38.3%
two_sibling	Have two siblings	33.4%	33.4%
three+_sibling	Have three or more siblings	22.5%	21.9%
M_BCOLLOR	Mother have a blue color job	50.7%	50.7%
M_WCOLLOR	Mother have a white color job	44.1%	44.1%
M_ENTREPREN	Mother is an entrepreneur	4.2%	4.2%
M_FARMER	Mother is a farmer	1.0%	1.0%
F_BCOLLOR	Father have a blue color job	45.9%	45.9%
F_WCOLLOR	Father have a white color job	38.3%	38.3%
F_ENTREPREN	Father is an entrepreneur	16.3%	16.3%
F_FARMER	Father is a farmer	3.2%	3.2%
LIVEHOME	Live with parent or parents	91.5%	3.4%
CLOSFRIEND	Feel to have close friends	92.5%	91.5%
LARGE_CITY	Lives in any of the three biggest cities in Sweden	26.3%	34.2%
SWEDISH_PARENTS	Both parents Swedish	85.6%	85.6%
2GENIM	Second generation immigrant	11.0%	11.0%
BORNABROAD	Born abroad, now swedish citizen	3.4%	3.4%
ECONCONSTRAIN	Economically constrained*	37.8%	18.5%
BADHEALTH	Feel to have bad health	2.4%	5.4%
ANIXIETY	Feel to suffer from anxiety	7.1%	14.6%
UNDERW	BMI<18.5	9.0%	3.1%
NORMALW	18.5≤BMI<25	82.9%	70.0%
OVERW	25≤BMI	8.1%	26.9%
SMOKES	Currently smokes	12.7%	14.7%

Note: *Defined as not being able to within a week provide a certain amount. 9,000 SEK in 88/89, 14,000 SEK in 96/97 and 15,000 SEK in 04/05 (1 USD≈7 SEK).

2.4 Empirical model

Logit models are used to estimate the effect of alcohol consumption on educational attainment[7]. Based on the theoretical framework in section 2 and in line with previous studies (e.g. Renna 2007; DeSimone 2010; Duarte and Escario 2006), it is possible to specify the level of educational attainment as a function of individual drinking habits, in addition to a set of variables (X) commonly believed to affect educational attainment:

$$EDUCATIONAl\ ATTAINMENT_i = \alpha + \beta\ \mathrm{ALCOHOL}_i + \gamma X_i + \varepsilon_i$$

where $EDUCATIONAl\ ATTAINMENT_i$ measures the individual educational attainment. α is the intercept and $\mathrm{ALCOHOL}_i$ is a set of dummy variables measuring weekly alcohol consumption measured in units of wine bottles8. ε_i is the error term and is assumed to be uncorrelated with the other variables. X_i is a set of control variables.

Estimations of equation (1) can be biased if alcohol consumption is endogenous. For instance, educational attainment and alcohol consumption are determined simultaneously and educational attainment may affect alcohol consumption (i.e. reverse causality).

Endogeneity can be controlled for via an instrumental variable approach. Regional differences, such as state tax levels on alcohol or different state minimum legal drinking ages, have been used to instrument alcohol consumption in a US context. However, Dee (1999) and Koch & McGeary (2005) argue that state-level variables yield imprecise estimations of

[7] Probit models yield similar results

[8] Estimations using linear, quadratic, cubic and log-linear specifications of alcohol consumption can be obtained from the author upon request.

adolescence drinking. In addition, Chatterji (2006) notes that a particular state policy can be associated with other state policies that underlie both alcohol consumption and educational attainment. In a Swedish context, the tax level on alcohol and the minimum legal drinking age are uniform across the country[9]. In addition, alcohol above 3.5 volume percent is only sold in the state-owned monopoly (Systembolaget), which ensures uniform pricing in Sweden. Instead, two reforms have been used in the Swedish context to instrument increased alcohol availability. First, a working paper (Nilsson 2008) used a policy experiment between November 1967 and July 1968, which temporarily allowed sales of beer up to 5.6 volume percent in ordinary grocery shops to instrument increased alcohol availability. Second, another working paper (Priks et al. 2011) used the openings of new Systembolaget outlets from 1978 onwards to instrument increased alcohol availability. Unfortunately, neither of these reforms can be used in this study. First, even though these reforms increased availability for those of eligible age, it is not certain that it increased the alcohol use of high school students. Second, given that these reforms took place some time ago, its relevance for today's high school students can be questioned.

Due to these issues with an instrumental variables approach, a life cycle approach can be used to overcome reverse causality. Koch & Ribar (2001) used a life cycle approach[10] and found that the age at onset of drinking alcohol had only a small effect on years of completed schooling at age 25. A similar study (Burgess and Propper 1998) concluded that alcohol and soft

---

[9] Individuals in Sweden above the age of 18 can buy beer with an alcohol content of up to 3.5 volume percent in a grocery store and stronger alcohol on-premises in bars and restaurants. Individuals above the age of 20 can only buy alcohol off-premises from Systembolaget, the state-owned monopoly.

[10] A life cycle approach can intuitively be motivated by the fact that your decision today can affect your decisions tomorrow. However, the opposite cannot be true.

drug consumption amongst young men has no harmful effect on economic prospects later in life. Koch & McGeary (2005) argue that the time spans used by Koch & Ribar (2001) (8 years on average) and Burges & Propper (1998) (10 years) are too long as individuals with problematic alcohol consumption may be able to return to a more productive lifestyle. To avoid reverse causality, the paper used a life cycle approach when analyzing the effect of alcohol consumption on educational attainment

Given that the objective of this study is two-fold, I used two different specifications of the outcome variable, educational attainment. The first objective, to analyze the effect of alcohol consumption among high school students on the probability of continuing to university, uses the following specification:

$$\text{CONTINUING TO UNIVERSITY}_i = \alpha + \beta\ \text{ALCOHOL}_i + \gamma X_i + \varepsilon_i$$

where $\text{CONTINUING TO UNIVERSITY}_i$ can take either the value one if the individual has started a university education[11], or zero otherwise. Given that high school students in Sweden are usually between 16 and 19 years old, individuals older than 19 are excluded (the youngest individual in the ULF survey is 16 years old). Regarding continuing to university, the median age is just above 22 (Swedish National Agency for Higher Education, 2010). Hence, to include late starters, educational attainment is measured at age 25 for all respondents. Information on educational attainment originates from the LISA database. Given that information regarding educational attainment in the LISA database was available from 1994 through 2006, the ULF waves of

[11] There is a distinction between universities and university colleges ("högskola" in Swedish) in Sweden; master and graduate degrees can only be acquired at a university. As this paper focuses on continuation to universities or university colleges, no distinction is necessary.

1988/1989 and 1996/1997 was used. This yields a dataset containing 590 individuals.

As a comparison, the estimations were rerun with the same individuals, except that the explanatory variables originate from the subsequent wave. That is, instead of using the ULF waves of 1988/1989 and 1996/1997, the waves of 1996/1997 and 2004/2005 are used and, hence, the individuals are 24 to 27 years old, whereas educational attainment is still measured at the age of 25. Given that educational attainment for some individuals are measured before they are included in the ULF survey, this estimation may be hampered by reverse causality. Still, by comparing these two specifications, lagged and contemporaneous explanatory variables, it is possible to analyze if and when alcohol consumption affects educational attainment.

The second objective, to analyze the effect of alcohol consumption among university students on the probability of increasing their educational attainment, uses the following specification:

$$INCREASED\ EDUCATIONAL\ ATTAINMENT_i = \alpha + \beta\ \mathrm{ALCOHOL}_i + \gamma X_i + \varepsilon_i$$

where $INCREASED\ EDUCATIONAL\ ATTAINMENT_i$ can take either the value one if the individual has increased his or her university educational attainment, or zero if not. Given that information regarding educational attainment is available from 1994 through 2006, the ULF wave of 1996/1997 was used. Only individuals with some university education are included in the analysis, and individuals above 30 years old at the interview were excluded as older individuals are not likely to still be university students. Besides using the same control variables as in equation 2, educational attainment is controlled for at the time of interview, as some individuals may already have finished their education.

2.5 Results

All estimations are done in two steps. First, only alcohol consumption and an intercept are controlled for. Second, other control variables are added. By doing the estimations in two steps, it is possible to get an indication of how sensitive the alcohol variables are to adding control variables.

Tables 2 and 3 report the marginal effects of the logit estimations. Table 2 shows the results of estimating the probability of having begun university studies (equation 2 above) and table 3 shows the results from equation 3 above (the probability of increasing university educational attainment).

Starting with table 2, educational attainment is measured at age 25 in all the columns and all the columns have the same set of respondents. Columns (a) and (b) present the results for the respondents between 16 and 19 years old. The specification in column (a) only controls for alcohol consumption, whereas the specification in column (b) uses the full set of control variables. Both columns (a) and (b) show that alcohol consumption has a negative effect on the probability of starting university studies. The magnitudes seen in columns (a) and (b) are similar in magnitude. For instance, individuals who drink between 1 and 2 bottles of wine per week have a 25 percent lower probability, compared with those with no alcohol consumption, of continuing to university.

Table 2 - Marginal effects from logit estimations on having continued to university at the age of 25. Separate estimations for when the covariates originates from when the individual was 16-19 years old or 24-27 years old.

	16-19 years old								**24-27 years old**		
	(a)			(b)			(c)		(d)		
	M.E.		Std. Err.	M.E.		Std. Err.	M.E.	Std. Err.	M.E.		Std. Err.
Abstent	ref.			ref.			ref.		ref.		
0<wine<0.5	-0.01		(0.05)	-0.01		(0.05)	-0.07	-0.08	-0.09		(0.08)
0.5≤wine<1	-0.16	**	(0.08)	-0.16	**	(0.08)	0.04	-0.09	-0.02		(0.09)
1≤wine<2	-0.25	***	(0.06)	-0.25	***	(0.06)	-0.07	-0.08	-0.13		(0.09)
2≤wine	-0.27	***	(0.06)	-0.24	***	(0.07)	-0.1	-0.08	-0.15	*	(0.09)
MALE				-0.03		(0.04)			0.01		(0.04)
no_sibling				ref.					ref.		
one_sibling				0.12	*	(0.07)			0.12	*	(0.07)
two_sibling				0.03		(0.06)			0.06		(0.07)
three+_sibling				0.02		(0.07)			-0.01		(0.08)
M_BCOLLOR				ref.					ref.		
M_WCOLLOR				0.06		(0.04)			0.06		(0.04)
M_ENTREPREN				-0.03		(0.08)			-0.04		(0.08)
F_BCOLLOR				ref.					ref.		
F_WCOLLOR				0.08	*	(0.04)			0.10	**	(0.05)
F_ENTREPREN				0.08		(0.05)			0.08		(0.05)
LIVEHOME				0.16	*	(0.09)			-0.09		(0.11)
CLOSFRIEND				-0.08		(0.07)			-0.01		(0.07)
LARGE_CITY				0.04		(0.04)			0.02		(0.04)
SWEDISH_PARENTS				ref.					ref.		
2GENIM				-0.03		(0.06)			-0.01		(0.06)
BORNABROAD				-0.07		(0.11)			-0.01		(0.11)
ECONCONSTRAIN				-0.07	*	(0.04)			-0.02		(0.05)
BADHEALTH				0		(0.15)			-0.05		(0.09)
ANIXIETY				-0.14	*	(0.08)			-0.07		(0.06)
NORMALW				ref.					ref.		
OVERW				-0.04		(0.08)			-0.11	**	(0.05)
SMOKES				-0.08		(0.06)			-0.08		(0.06)
Intercept	YES			YES			YES		YES		
Control for age	NO			YES			NO		YES		
R squared	0.04			0.10			0.01		0.07		
No. obs.	590			590			590		590		

Note: R squared from the logit regression. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3. Conditional on having initiated university at the time of the interview, marginal effects from logit estimations on having increased educational attainment 3 and 5 years after the interview. Individuals above the age of 30 at the interview are excluded.

	3 years after interview						5 years after interview					
	(a)			(b)			(c)			(d)		
	M.E.		Std. Err.	M.E.		Std. Err.	M.E.		Std. Err.	M.E.		Std. Err.
Abstent	ref.			ref.			ref.			ref.		
0<wine<0.5	0.15	**	(0.07)	0.15	**	(0.07)	0.08		(0.09)	0.02		(0.07)
0.5≤wine<1	0.17	**	(0.08)	0.13	*	(0.07)	0.12		(0.10)	0.02		(0.08)
1≤wine<2	0.18	**	(0.08)	0.16	**	(0.08)	0.17	*	(0.09)	0.06		(0.08)
2≤wine	0.12		(0.08)	0.1		(0.08)	0.09		(0.10)	-0.03		(0.08)
MALE				-0.06		(0.04)				-0.07	*	(0.04)
no_sibling				ref.						ref.		
one_sibling				-0.12		(0.07)				-0.08		(0.07)
two_sibling				-0.07		(0.08)				-0.03		(0.07)
three+_sibling				-0.13		(0.09)				-0.09		(0.08)
M_BCOLLOR				ref.						ref.		
M_WCOLLOR				0.04		(0.04)				0.03		(0.04)
M_ENTREPREN				0.07		(0.09)				0.1		(0.07)
F_BCOLLOR				ref.						ref.		
F_WCOLLOR				0.07		(0.05)				0.08	**	(0.04)
F_ENTREPREN				0.04		(0.06)				-0.03		(0.05)
LIVEHOME				0.1		(0.08)				0.00		(0.09)
CLOSFRIEND				0.06		(0.06)				0.00		(0.05)
LARGE_CITY				0.03		(0.04)				0.04		(0.04)
SWEDISH_PARENTS				ref.						ref.		
2GENIM				0.02		(0.06)				-0.09	*	(0.06)
BORNABROAD				0.07		(0.13)				-0.15		(0.13)
ECONCONSTRAIN				0.08		(0.06)				-0.04		(0.06)
BADHEALTH				-0.03		(0.07)				0.00		(0.08)
ANIXIETY				0.07		(0.05)				0.08		(0.06)
NORMALW				ref.						ref.		
OVERW				-0.02		(0.05)				0.04		(0.04)
SMOKES				-0.17	***	(0.06)				0.05		(0.06)
Intercept	YES			YES			YES			YES		
Control for age	NO			YES			NO			YES		
R squared	0.01			0.23			0.01			0.38		
No. obs.	525			525			525			525		

Note: R squared from the logit regression. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Continuing with column (b), having a sibling has a positive effect on educational attainment. Both the coefficients for anxiety and economically constrained (see table 2 for details) have a statistically significant effect on educational attainment.

Columns (c) and (d) present the results for respondents between 24 and 27 years old. The alcohol variables in column (c) are negative but not statistically significant. The only significant statistically significant alcohol variable in column (d) is alcohol consumption above 2 bottles of wine per week. Having one sibling, compared with not having a sibling, increases the probability of having begun university studies at the age of 25 by 12 percent. Having a father who is either white collar or an entrepreneur increases the probability of beginning university studies by 10 percent, compared with having a father who is blue collar.

All columns in table 3 estimate the probability of increasing university educational attainment after the ULF survey. In columns a and b, 3 years after the ULF survey, and in columns c and d, 5 years after the ULF survey. Looking at 3 years after the ULF survey, when only controlling for the level of alcohol consumption at the time of the interview (column a), drinking up to the equivalence of 2 bottles of wine per week increases the probability of increasing university educational attainment by 14 to 18 percent compared with individuals who do not drink any alcohol. For alcohol consumption equivalent to drinking more than two bottles per week, no statistical effect is found.

Turning to column c, alcohol consumption equivalent to 1-2 bottles of wine per week increases the probability of a higher university educational attainment five years after the interview, compared with those who do not consume alcohol. When control variables are included (column d), no alcohol coefficients are statistically significant. Having a mother who is white collar, compared with having a mother who is blue collar, increases the likelihood of increasing university attainment by 8 percent.

2.6 Discussion

This study analyzed the influence of adolescent alcohol consumption on the probability of beginning a university education. Even though the relationship between alcohol consumption and educational attainment has been analyzed in numerous previous studies, the results are not conclusive. In addition, little attention has been paid to the question of alcohol consumption during high school and the likelihood of continuing to university.

The main finding in this paper is that alcohol consumption (above the equivalence of a half bottle of wine per week) during high school (significantly) decreases the probability of continuing to university. Moreover, contemporaneous alcohol consumption is found to have no effect (except above the equivalence of drinking two bottles of wine per week) on continuing to university. This suggests that the consumption history is more important than contemporary consumption. Nonetheless, one should keep in mind the fact that the potential negative effect of alcohol consumption is offset by university students drinking habits, which sometimes include heavy drinking

The results in this study are in line with the results obtained by Koch & McGeary (2005), who found that alcohol initiation before age 14 reduces the probability of completing high school by between 7 percent and 22 percent. However, the results obtained here are not consistent with the results reported in the life cycle studies by Koch & Ribar (2001) and Burgess & Propper (1998). Compared with this study, Koch & Ribar (2001) and Burgess & Propper (1998) use a longer timespan between alcohol consumption and educational attainment. As mentioned in section 4, a too long timespan can reduce the observed effect of alcohol consumption on later educational attainment, as heavy drinkers may return to a more productive lifestyle (Koch and McGeary 2005). Hence, given that this study measured if the individual had continued to university at age 25, those who do not go straight to university after high school, but perhaps travel or work for a couple of years before continuing to university, should be included.

The negative effect of overweight on educational attainment is as expected. It may seem surprising that none of the other health indicators appear to affect educational attainment. However, as noted by Renna (2008), the correlation between health indicators makes it difficult to distinguish the separate effects. The effect of family background (working situation for parents) on educational attainment is consistent with previous empirical work (e.g Fischer, Fortun, Pidcock & Dowd, 2007; Renna, 2008).

Turning to the effect of alcohol consumption on the probability of increasing university attainment among university students, the results differ depending on the time horizon. 3 years after the interview, alcohol consumption up to the equivalence of 2 bottles of wine per week has a statistically significant positive effect on increased university attainment. 5 years after the interview, alcohol consumption does not (when control variables were included) have a statistically significant effect on university attainment. The difference in results could indicate that alcohol consumption does not affect overall educational attainment in the long run, but speeds up educational attainment in the short run. It may seem counterintuitive that alcohol consumption has a positive effect on educational attainment in the short run, but one should remember that alcohol consumption is common in many of the students' social activities. For instance, it is possible that students involved in many student activities, compared with students not involved in so many student activities, meet more students and, as a result, have a larger network from which they can receive help. This notion is in line with research which has found a wage bonus for moderate alcohol consumption, compared with abstainers and heavy drinkers (see Lye and Hirschberg 2010 for an overview).

Even though this paper used a life cycle approach to avoid reverse causality, the results in this study can be biased if omitted variables affect both alcohol consumption and educational attainment. However, by including a rich set of control variables that are in line with previous studies (e.g. Renna 2007; DeSimone 2010; Duarte and Escario 2006), the risk of omitted variable bias is minimized in this paper.

The finding that high school drinking is more important than contemporaneous consumption for the likelihood of entering university suggests that policies aiming at increasing the share of young people entering higher education should focus on curbing alcohol consumption during high school, rather than among those of university-eligible ages. For instance, as many in high school are legally ineligible to buy alcohol, one should concentrate on efforts to decrease peddling to minors. In addition, help should be provided for adolescents who already are, or run the risk of becoming, heavy drinkers to minimize the negative effects of alcohol consumption on educational attainment.

References

Anger, Silke, and Guido Heineck. 2010. "Do Smart Parents Raise Smart Children? The Intergenerational Transmission of Cognitive Abilities." *Journal of Population Economics* 23 (3): 1105–1132.

Balsa, Ana I., Laura M. Giuliano, and Michael T. French. 2011. "The Effects of Alcohol Use on Academic Achievement in High School." *Economics of Education Review* 30 (1): 1–15. doi:doi: 10.1016/j.econedurev.2010.06.015.

Becker, Gary S. 2007. "Health as Human Capital: Synthesis and Extensions." *Oxford Economic Papers* 59 (3): 379.

Becker, Gary S. 1964. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. Chicago, University of Chicago Press.

Björklund, Anders, and Kjell G. Salvanes. 2010. "Education and Family Background: Mechanisms and Policies." In *Handbook of the Economics of Education*, edited by Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, 3:201–247. Amsterdam: Elsevier. http://www.elsevier.com/wps/find/bookdescription.cws_home/724027/description#description.

Burgess, Simon M., and Carol Propper. 1998. "Early Health-related Behaviours and Their Impact on Later Life Chances: Evidence from the US." *Health Economics* 7 (5): 381–399.

Cameron, Stephen V., and James J. Heckman. 1993. "The Nonequivalence of High School Equivalents." *Journal of Labor Economics* 11 (1) (January 1): 1–47.

Chatterji, P. 2006. "Does Alcohol Use During High School Affect Educational Attainment?: Evidence from the National Education Longitudinal Study." *Economics of Education Review* 25 (5): 482–497.

Cook, P.J., and M.J. Moore. 1993. "Drinking and Schooling." Edited by Axel Börsch-Supan, Agar Grugiavini, Hendrik Jürges, Arie Kapteyn, Johan Mackenbach, Johannes Siegrist, and Gugliemo Weber. *Journal of Health Economics* 12 (4): 411–429.

Corrao, Giovanni, Vincenzo Bagnardi, Antonella Zambon, and Sarino Arico. 1999. "Exploring the Dose-response Relationship Between Alcohol Consumption and the Risk of Several Alcohol-related Conditions: a Meta-analysis." *Addiction* 94 (10): 1551–1573.

Dee, Thomas S. 1999. "State Alcohol Policies, Teen Drinking and Traffic Fatalities." *Journal of Public Economics* 72 (2) (May 1): 289–315. doi:doi: DOI: 10.1016/S0047-2727(98)00093-0.

DeSimone, Jeff. 2010. "Drinking and Academic Performance in High School." *Applied Economics* 42 (12): 1481.

Duarte, R., and J.J. Escario. 2006. "Alcohol Abuse and Truancy Among Spanish Adolescents: A Count-data Approach." *Economics of Education Review* 25 (2): 179–187. doi:doi: DOI: 10.1016/j.econedurev.2005.01.007.

Fischer, Judith, Larry Forthun, Boyd Pidcock, and Duane Dowd. 2007. "Parent Relationships, Emotion Regulation, Psychosocial Maturity and College Student Alcohol Use Problems." *Journal of Youth and Adolescence* 36 (7) (October 7): 912–926.

Fitzgerald, Hiram E., and Robert A. Zucker. 1995. "Socioeconomic Status and Alcoholism: The Contextual Structure of Developmental Pathways to Addiction." In *Children of Poverty: Research, Health, and Policy Issues*, edited by Hiram E Fitzgerald, Barry M. Lester, and Barry Zuckerman, 23:125–145. Reference Books on Family Issues. http://www.researchgate.net/publication/232527942_Socioeconomic_status_and_alcoholism_The_contextual_structure_of_developmental_pathways_to_addiction.

Holman, C D'Arcy, Dallas R English, Elizabeth Milne, and Miachael G Winter. 1996. "Meta-Analysis of Alcohol and All-cause Mortality: a Validation of NHMRC Recommendations." *The Medical Journal of Australia.* 164: 141–145.

Jones, Ethel B., and John D. Jackson. 1990. “College Grades and Labor Market Rewards.” *The Journal of Human Resources* (2): 253. edsjad.

Koch, Steven F, and Kerry Anne McGeary. 2005. “The Effect of Yought Alcohol Initiation on High School Completion.” *Economic Inquiry* 43 (4): 750–765.

Koch, Steven F, and David C Ribar. 2001. “A Siblings Analysis of the Effects of Alcohol Consumption Onset on Educational Attainment.” *Contemporary Economic Policy* 19 (2): 162–175.

Lovenheim, Michael F., and Joel Slemrod. 2010. “The Fatal Toll of Driving to Drink: The Effect of Minimum Legal Drinking Age Evasion on Traffic Fatalities.” *Journal of Health Economics* 29 (1) (January): 62–77. doi:10.1016/j.jhealeco.2009.10.001.

Lye, Jenny, and Joe Hirschberg. 2010. “Alcohol Consumption and Human Capital: a Retrospective Study of the Literature.” *Journal of Economic Surveys* 24 (2): 309–338.

Nilsson, Peter. 2008. “Does a Pint a Day Affect Your Child’s Pay? The Effect of Prenatal Alcohol Exposure on Adult Outcomes”. Working paper. cemmap.ifs.org.uk/wps/cwp2208.pdf.

OECD. 2009. *Health at a Glance 2009 - OECD Indicators*. OECD Publishing.

Priks, Mikael, Anders Andersson, Mårten Palme, and Per Pettersson-Lidbom. 2011. “The Effect of Alcohol Availability on Crime and Health.”

Pronzato, Chiara. 2012. “An Examination of Paternal and Maternal Intergenerational Transmission of Schooling.” *Journal of Population Economics* 25 (2): 591–608.

Rehm, Juergen, Dolly Baliunas, Guilherme L. G. Borges, Kathryn Graham, Hyacinth Irving, Tara Kehoe, Charles D. Parry, et al. 2010. “The Relation Between Different Dimensions of Alcohol Consumption and Burden of Disease: An Overview.” *ADDICTION* 105 (5) (May): 817–843. doi:10.1111/j.1360-0443.2010.02899.x.

Renna, Francesco. 2007. “The Economic Cost of Teen Drinking: Late Graduation and Lowered Earnings.” *Health Economics* 16 (4): 407–419.

———. 2008. “Teens’ Alcohol Consumption and Schooling.” *Economics of Education Review* 27 (1): 69–78.

Scott E. Carrell, Mark Hoekstra, and James E. West. 2003. “Does Drinking Impair College Performance? Evidence from a Regression Discontinuity Approach.” *Journal of Public Economics* 95: 54–62. edselp.

Staff, Jeremy, Megan E. Patrick, Eric Loken, and Jennifer L. Maggs. 2008. “Teenage Alcohol Use and Educational Attainment.” *Journal of Studies on Alcohol and Drugs* 69 (6) (October 17): 848.

Statistics Sweden. 2004. “SUN2000 (English Version).” http://www.scb.se/statistik/UF/UF0506/SUN2000English%20version.doc.

———. 2013. “Statistics Sweden.” http://www.scb.se/Pages/SSD/SSD_TablePresentation____340486.aspx?rxid=6b565665-f0f6-4050-abe7-fbcd5f84d8a2.

Swedish National Agency for Higher Education. 2010. “Nyckeltal - Högskoleverket.” http://www.hsv.se/statistik/statistikomhogskolan/nyckeltal.

Williams, Jenny, Lisa M Powell, and Henry Wechsler. 2003. “Does Alcohol Consumption Reduce Human Capital Accumulation? Evidence from the College Alcohol Study.” *Applied Economics* 35 (10): 1227.

Winstanley, Erin L., Donald M. Steinwachs, Margaret E. Ensminger, Carl A. Latkin, Maxine L. Stitzer, and Yngvild Olsen. 2008. “The Association of Self-reported Neighborhood Disorganization and Social Capital with Adolescent Alcohol and Drug Use, Dependence, and Access to Treatment.” *Drug and Alcohol Dependence* 92 (1-3) (January 1): 173–182. doi:doi: DOI: 10.1016/j.drugalcdep.2007.07.012.

Wolaver, A.M. 2002. “Effects of Heavy Drinking in College on Study Effort, Grade Point Average, and Major Choice.” *Contemporary Economic Policy* 20 (4): 415–428. edselc.

Yamada, Tetsuji, Michael Kendix, and Tadashi Yamada. 1996. “The Impact of Alcohol Consumption and Marijuana Use on High School Graduation.” *Health Economics* 5 (1): 77–92.

3 Does the effect of a negative health shock on the retirement decision differ among European countries?

Abstract

This paper analyzed the effect of health on the probability of retiring. The paper used a dataset covering individuals who are 50 years or older from eleven European countries to analyze if the effect of health on the probability of retiring, differs between the countries. To ensure cross-country comparability, this paper used overnight hospitalization as a proxy for a negative health shock. Given that individuals might not be randomized to either experience an overnight hospitalization or not, the paper used propensity score matching to avoid selection bias.

When all eleven European countries were analyzed together, overnight hospitalization has a positive effect on the retiring probability. When analyzed separately, the results are mixed; overnight hospitalization has a positive effect on the retiring probability in some countries and no effect in some countries. The disparities in results can reflect differences in the institutional settings across countries. For instance, the access to rehabilitory care may differ between the countries.

3.1 Introduction

Current demographical changes in many Western countries - decreasing fertility and increasing life expectancy - will lead to ageing populations with a larger share of retirees and, as a consequence, increased financial strain on pension systems and tax payers.

Health and socioeconomic factors have been shown to influence the retirement decision. For instance, (1) schooling tends to postpone retirement (Alavinia & Burdorf, 2008), (2) the spouse´s retirement decision – spouses tend to retire simultaneously (Bingley & Lanot, 2007; Denaeghel, Mortelmans, & Borghgraef, 2011), (3) unemployment – unemployed individuals are more likely to retire (Christensen & Kallestrup-Lamb, 2012), (4) the generosity of the benefit system (Gruber & Wise, 2000; Börsch-Supan, Hank, Jürges, & Schröder, 2009; Engelhardt, 2012), and (5) poor work quality have been shown to increase the probability of early retirement (Siegrist, Wahrendorf, von dem Knesebeck, Jürges, & Börsch-Supan, 2006). The effect of health on the retirement decision has been studied in several institutional settings. For instance, empirical studies have found that poor health is associated with a higher probability of retirement in Britan (Disney, Emmerson, & Wakefield, 2006), in France (Vaillant & Wolff, 2012), in Denmark (Christensen & Kallestrup-Lamb, 2012; Gupta & Larsen, 2010; Gortz, 2012), and in the US (Dwyer & Mitchell, 1999; McGarry, 2004; McGeary, 2009; Bound, Stinebrickner, & Waidmann, 2010; Renna & Thakur, 2010). In addition, studies analyzing European countries jointly (Alavinia & Burdorf, 2008; García-Gómez, 2011) have also found that poor health is associated with an increase in the probability of retiring.

Different kinds of negative health shocks have been used to estimate a causal health effect on retirement decision. For instance, transitions to lower self-assessed health have been shown to have a positive effect on the probability of retiring (Bound et al., 2010; Disney et al., 2006; García Gómez & López Nicolás, 2006; García-Gómez, 2011). Moreover, health shocks such as accidents and chronic illness also increase the probability of retirement (McClellan, 1998; Gupta & Larsen, 2010).

The objective of this study was to analyze the effect of a negative health shock (hospitalization) on the probability of retiring among individuals aged 50 and above in several European countries. From a policy perspective, it could be interesting to analyze individuals below the age of 50 as younger individuals are likely to work for a longer time period. However, given the demographical changes with increased longevity and an increased share of the elderly, individuals above the age of 50 are interesting to analyze as the retirement age is likely to increase and, hence, understanding the retirement behavior of the population above the age of 50 will increase in importance.

Whereas most previous studies analyzing the effect of health on the retirement decision employ single-country data (e.g. Dwyer & Mitchell, 1999; Vaillant & Wolff, 2012), this study, in contrast, utilizes a cross-national survey (the Survey of Health Ageing and Retirement in Europe, see section 3). Thus, this study was able to compare the effects of a negative health shock between countries, and between different institutional structures. For instance, the institutional structure for early retirement age, net replacement rates, physician density and access to specialist care differ among the European countries (Bolin, Lindgren, Lindgren, & Lundborg, 2009; Engelhardt, 2012). The influence of health on the likelihood of retirement has been analyzed in several studies (e.g. Siegrist et al., 2006; Alavinia & Burdorf, 2008) using self-assessed health, which is the empirical measure of health in the SHARE survey. To the best of the author's knowledge, no previous study using the SHARE survey has used hospitalization as a proxy for a negative health shock. This study used propensity score matching to control for selection bias, as individuals who have experienced hospitalization are likely to differ from individuals who have not.

The paper proceeds as follows. Section 3.2 discusses the theoretical framework and the empirical model is explained in Section 3.3. The data is described in Section 3.4 and the results are presented in Section 3.5. Section 3.6 concludes the paper.

3.2 Theoretical framework

Following Burkhauser (1979) and Hamermesh (1984), this paper assumed that a worker will retire if the utility as a retiree is greater than the utility as a worker:

$$U(retiree) > U(worker) \qquad (1)$$

Further, the individual enjoys utility from consumption and leisure and, hence, the utility function can be written as:

$$U(C, L) \qquad (2)$$

where consumption (C) and leisure (L) are assumed to have a positive but diminishing effect on utility. Consumption, in turn, depends on the individual’s preferences and the relative prices of consumption and leisure. More productive individuals will earn higher wages and, hence, more combinations of consumption and leisure are feasible via a larger budget set. Productivity is assumed to depend on the individual’s human capital which, in turn, depends heavily on educational attainment. If the individual´s productivity decreases, for instance due to an unexpected negative health shock (e.g. sickness), the wage may fall below the wage rate at which the individual is willing to work (i.e. reservation wage), in which case it is rational for the individual to retire. Further, whether or not it is possible to return to work after a severe negative health shock depends on the structure of the particular institutional setting, for instance the availability of rehabilitation care. In addition, the generosity of the benefit system may also influence the probability of returning to work after a negative health shock. For instance, it is possible that individuals living in a country with a generous benefit system return to work after a negative health shock to a lesser extent than those in a country with a less generous benefit system.

Besides health being a determinant of labor productivity, the relationship between a negative health shock and the labor supply may be strictly mechanical; the individual can work, earn income and consume goods or have leisure. In the case of a negative health shock, which could result in hospitalization, the individual might not be able to work while he or she is in hospital.

Studies (e.g. Bingley & Lanot, 2007; Denaeghel et al., 2011) have found that spouses tend to retire at the same time, suggesting that consumption and *own* leisure may not be the only goods providing utility – spousal and own leisure may be complements.

Given this framework, several factors are likely to influence the retirement decision and, hence, the probability of retiring can be written as:

$$\Pr(retired) = f(health, spouse\ retired, education, X) \quad (3)$$

where *health* represents the health stock, *spouse retired* is a binary variable indicating if the spouse is working or retired, *education* represents the stock of educational attainment and *X* represents a set of variables also determining the probability of retiring (e.g. institutional setting).

It is possible that the effect of health on the retirement decision depends on education. For instance, an individual with a large stock of education is likely to access a larger labor market than an individual with a small stock of educational attainment. The possibility of changing occupation may be important after a negative health shock and thus individuals with a large stock of educational attainment, compared with individuals with a small stock of educational attainment, may be better prepared for a negative health shock.

3.3 Data

This paper used the Survey of Health Ageing and Retirement in Europe (SHARE)[12], which is a multidisciplinary and cross-national panel database covering individuals 50 years of age and older. The survey interviews a representative sample of non-institutionalized individuals and their spouses, and other individuals aged 50 or older in the household. The first wave, conducted in 2004 included: Austria, Germany, Sweden, The Netherlands, Spain, Italy, France, Denmark, Greece, Switzerland and Belgium. The Czech Republic, Poland and Ireland joined in the second wave, which was conducted in 2006-2007. The third wave, called SHARELIFE (2008-2009), differed from waves one and two by being retrospective and collecting information about the respondent´s early years. Estonia, Hungary and Slovenia joined in the fourth wave, which was conducted in 2010-2012. The overall household response rate was 61.6 percent in the first wave, ranging from 38.8 percent in Switzerland to 81.0 percent in France. In the first wave, 31,115 were interviewed, of whom 18,742 were re-interviewed in the second wave.

The sample used in the analyses was restricted in the following way. First, incomplete interviews were excluded. Second, spouses that are below 50 years old were excluded as they are very few in number. Third, respondents older than 70 years of age were excluded as they are not likely to still be working. Fourth, homemakers and those who are permanently sick or disabled were excluded. Fifth, the empirical strategy used here requires two consecutive waves to avoid reversed causality when estimating the propensity score. Given that the third wave (SHARELIFE) is a retrospective survey and asks about early life conditions, either the two first waves or

[12] See Mazzonna & Peracchi (2012) for a thorough description of the SHARE data.

waves two and four may be used. Because a shorter time span elapses between the two first waves, compared to the second and fourth waves, the two first waves were used in the analysis. More specifically, the propensity score for hospitalization in the second wave was estimated using covariates from the first wave, whereas retirement status in the second wave was regressed using covariates from the second wave. After these exclusion restrictions, this paper's working sample consists of 8,538 individuals.

Dependent variable

The respondents are asked the following question: "In general, how would you describe your current employment situation" and given five alternatives; (1) retired, (2) employed or self-employed, (3) unemployed and looking for work, (4) permanently sick or disabled, or (5) homemaker. The SHARE survey also includes a question on whether the individual receives pension benefits. The two variables - "In general, how would you describe your current employment situation" and if the individual receives pension benefits are highly correlated (0.8) and are equal in 90 percent of the interviews. Given that individuals can work at the same time as they receive pension benefits, the question regarding the current employment situation to measure retirement status is used. The distribution of the answers from waves one and two are shown in table 1 for the working sample.

Table 1 - Current job situation for wave 1 and 2

Current job situation in wave 1	Current job situation, wave 2					
	Retired	Employed or self-employed	Unemployed	Permanently sick or disabled	Homemaker	Total
Retired	31.6%	0.5%	0.1%	0.8%	1.5%	34.5%
Employed or self-employed	5.5%	32.6%	1.1%	0.8%	1.1%	41.1%
Unemployed	1.2%	1.0%	1.7%	0.2%	0.5%	4.5%
Permanently sick or disabled	1.2%	0.3%	0.1%	1.9%	0.2%	3.6%
Homemaker	2.3%	0.9%	0.3%	0.4%	12.5%	16.3%
Total	41.7%	35.3%	3.2%	4.1%	15.7%	100.0%

31.6 percent were retired in both waves whereas 32.6 percent are either employed or self-employed in both waves; 5.5 percent go from either employed or self-employed to retired, whereas 0.5 percent goes from retired to either employed or self-employed. Hence, even though the share of individuals going from either employed or self-employed to retirement is larger than the share going in the opposite direction, retirement is not an absorbing state. If retirement were an absorbing state, it would not be interesting to include those who are retired in the first wave. Given that retirement is not an absorbing state; those who are retired in the first wave are included in the analysis.

Independent variables

The variable is coded 1 if the individual has spent at least one night in hospital during the last year, and zero otherwise. By only looking at overnight hospital admission, vaccinations, regular check-ups and minor injuries (e.g. sprained ankle) are excluded. Hospitalizations due to giving birth are likely to be rare in the data, given that respondents are at least 50 years of age.

Other independent variables

Based on the discussion in the theoretical section, variables with a potential effect on retirement decision are included as control variables. The descriptive statistics are presented in table 2. Education is coded into three dummy categories: (1) primary school or less, (2) secondary school, and (3) tertiary education. A dummy variable indicates if the individual is married. Household size is controlled for by including one dummy variable indicating whether or not the household consists of three members, and another dummy variable indicating whether or not the household comprises four or more members. Further, dummy variables are used to capture whether or not respondents have a retired spouse; whether or not the respondent was born in the country in which he or she currently resides; the respondent resides in a metropolitan area. Moreover, dummy variables are used to control for age (five-year intervals), and gender.

Descriptive statistics are shown in table 2 and is reported separately for respondents reported to have been admitted to hospital overnight within one year before wave 2 (treated) or not (controls). In addition, table 2 reports if the means are statistically significantly different between the treated and controls. 973 respondents reported being treated and 7,565 respondents reported being controls.

Table 2 shows that there are statistically significant differences between the means of the treated and the controls. For instance, 63.5 percent of the respondents in the treatment group are retired in wave 2, whereas 46.8 percent of the respondents in the control group are retired in wave 2. Further, 25.8 percent of the respondents in the control group and 30.9 percent of the respondents in the treatment group have third level education in wave 2. Hence, regarding the descriptive statistics, the treatment group appears to differ from the control group

Table 2 - Descriptive statistics for those who participated in wave 1 and 2. Data is presented separately for respondents having an overnight hospital admission within one year before wave 2.

		Wave 2		
		Hospital admission		Diff.
		Yes	No	
Number of observation		973	7565	
Variable	Description			
Dependent variable				
Retired	1 if retired	63.5%	46.8%	***
Independent variable				
Married	1 if married	76.2%	77.7%	
Educational attainment				
Primary school	1 if primary school or less	38.3%	37.5%	
Secondary school	1 if secondary school	35.8%	31.7%	**
Tertiary education	1 if tertiary education	25.8%	30.9%	***
Household size				
HHsize3	1 if three in household	12.9%	14.0%	
HHsize4+	1 if at least four in household	7.3%	11.9%	***
Native	1 if born in residential country	92.5%	93.0%	
Spouse retired	1 if spouse retired	61.6%	47.7%	***
Age groups				
Age50-54	1 if 50 to 54 years old	10.9%	15.4%	***
Age55-59	1 if 55 to 59 years old	21.6%	29.4%	***
Age60-64	1 if 60 to 64 years old	26.1%	25.8%	
Age65-70	1 if 65 to 70 years old	41.3%	29.2%	***
Male	1 if male	57.6%	53.8%	**
Large town	1 if living in a large town or larger	47.6%	49.5%	
		Wave 1		
Married	1 if married	76.9%	78.4%	
Educational attainment				
Primary school	1 if primary school or less	38.3%	37.5%	
Secondary school	1 if secondary school	35.8%	30.9%	**
Tertiary education	1 if tertiary education	25.9%	30.9%	***
Household size				
HHsize3	1 if three in household	15.1%	16.2%	
HHsize4+	1 if at least four in household	10.5%	15.1%	***
Native	1 if born in residential country	92.5%	93.0%	
Age groups				
Age50-54	1 if 50 to 54 years old	19.3%	29.3%	***
Age55-59	1 if 55 to 59 years old	23.0%	27.2%	***
Age60-64	1 if 60 to 64 years old	30.8%	24.6%	***
Age65-70	1 if 65 to 70 years old	23.8%	16.3%	***
Male	1 if male	57.6%	53.8%	**
Large town	1 if living in a large town or larger	48.1%	51.2%	*

3.4 Empirical strategy

There are several empirical measures of health – self-assessed health, different constructed health indices and functional measures of health etcetera – and there is no consensus on which is the most appropriate measure. Self-assessed health is a candidate and is often used in empirical work. However, self-assessed health can be biased if retired respondents report health levels below their actual levels, to justify their retirement. This is also known as justification bias and puts an upward bias on the effect of health on the retirement decision (Bound, 1991; Dwyer & Mitchell, 1999; Kerkhofs, Lindeboom, & Theeuwes, 1999) In addition, it is possible that individuals adapt to their current health status or change their subjective scale of reference by age (Wim, 2000; Au, Crossley, & Schellhorn, 2005). Therefore, self-assessed health may not be a good measure of health in terms of the retirement decision.

Other, more objective measures besides self-assessed health have been suggested to overcome the justification bias. For instance, McGarry (2004) used specific conditions such as hypertension, diabetes and stroke. While specific conditions may be warranted for use in specific research, they do not capture the overall health status. Instead of specific conditions, different kinds of health shocks have been used to analyze a causal effect of health on the retirement decision. For instance, accidents and hospitalization stays have been utilized as health shocks (McClellan, 1998; Gupta & Larsen, 2007; Gomez et al., 2013).

The use of hospitalization, an easily comparable measure of health, is further warranted when using a cross-national survey, as differences in wording can bias the result. The SHARE questionnaire is first constructed in a generic, English, version and then translated into the respective languages. Jürges (2007) has shown that the SHARE survey wording for self-assessed health differs in the participating countries. Thus, an easily comparable measure of health (i.e. hospitalization) may be better than self-assessed health when analyzing the SHARE survey.

Following the discussion in section 2 and the usage of hospitalization as a measure of a negative health shock, the following probability model of retirement can be estimated:

$$Retirement_i = \alpha_i + \beta Hospitalization_i + \gamma X_i + \varepsilon_i \qquad (4)$$

where α_i is the intercept, X_i is a vector of control variables (e.g. if the spouse is retired and educational attainment) and ε_i is the error term.

The estimation of β may be biased if $Hospitalization_i$ is correlated with ε_i (i.e. conditional independence assumption). $Hospitalization_i$ may be correlated with ε_i if the individuals are not randomly assigned to experience hospitalization (treated) or not (control).

For instance, it is possible that individuals, depending on socioeconomic status, differ in their propensity to seek health care and ultimately get hospitalized.

The potential bias can be corrected for by using Propensity Score Matching (PSM), developed by Rosenbaum and Rubin (1983). PSM has been demonstrated to produce results similar to those obtained when individual are randomized to treatment and control groups (Dehejia & Wahba, 1999). In addition, PSM reduces the potential problem of omitted variable bias by removing the correlation between the omitted variables, and reducing the correlation between the omitted and included variables (Imbens & Wooldridge, 2009). PSM has, for instance, been used when analyzing the effect of private education on academic achievements (Anand, Mizala, & Repetto, 2008) and the impact of health care on health (Wang, Yip, Zhang, & Hsiao, 2009). An overview of studies using PSM can be found in Thoemmes & Kim (2011).

PSM is analogues to weighted regressions and is conducted in 4 steps: (1) the propensity score of being treated are estimated (for instance with a probit model or a logit model), (2) common support is usually imposed to ensure comparable groups (treated and control). Common support implies that only individuals with a propensity score belonging to the intersection of treated and controls are included. For instance, if the propensity score ranges from 0.2 to 0.8 for treated and 0.1 to 0.7 for controls, only individuals with a propensity score of 0.2 to 0.7 are included, (3) if the focus of a study is to estimate the average treatment effect of the treated (e.g. the effect of a negative health shock on own retirement status), observations are given a sample weight equal to one for treated and:

$$\frac{\text{Propensity score}_i}{1 - \text{Propensity score}_i}$$

for each control i, and (4) the regression of interest is estimated (e.g. regressing retirement status on a negative health shock) using sample weights from the procedure described above (see Jones, 2007 or Nichols, 2008 for a further description).

Given the usage of PSM, no control variables are needed in the fourth step, if the estimation for propensity score is correctly specified. However, by including control variables in the fourth step, it is sufficient that either the estimation for propensity score or the regressions in the fourth step is correctly specified for the regression in the fourth step to be consistent. (e.g. Ho, Imai, King, & Stuart, 2007; Imbens & Wooldridge, 2009)

The propensity score for overnight hospitalization was estimated using a probit model[13]. Common support was imposed but no observations were lost as the propensity score ranges were identical for the treated and controls. Following Ho et al (2007), propensity scores was estimated using covariates,

[13] Estimates from a logit model yield similar results (not shown).

which later used for estimating the probability of retirement. However, some modifications were required to fulfill the PSM´s assumption. PSM is based on two assumptions: (1) balancing property, and (2) unconfoundedness. The balancing property implies that covariates should be balanced (similar) for treated and controls, given a certain propensity score. To test the balancing property, a test suggested by Dehjia & Wahba (1999) and Lee (2011) was implemented. Three modifications of the original specification were needed to ensure the balancing property: (1) overweight and obese were merged into one category, (2) retirement status of the spouse are not included, and (3) regarding marital status, only the variable for being married was included. The unconfoundedness assumption (there are no covariates beyond the included covariates) is not testable per se (Imbens & Wooldridge, 2009). Instead, this paper relies on theoretical justification for the choice of variables.

Even when using PSM, the coefficient for hospitalization may be biased if hospitalization affects the covariates used when estimating the propensity score (i.e. reversed causality). To minimize the risk of reversed causality, the covariates used when estimating the propensity score are measured before the potential hospitalization. This is further explained in the next paragraph.

Given the longitudinal dimension of the SHARE dataset, it is possible to estimate a difference-in-difference model (DiD). DiD relies on the parallel trend assumption, stating that the treated and controls should follow the same trends. Given that table 2 shows differences between the treated and controls, the parallel trend assumption may be violated and hence is DiD not used in the analysis[14].

[14] A DiD model was estimated for comparative reasons only (not shown). When only controlling for hospitalization, individuals experiencing hospitalization have 6 percent (statistically significant) higher likelihood of retiring, compared to those who do not experience hospitalization. The results turn statistically insignificant when control variables are included.

3.5 Results

3.5.1 Results-all countries

The estimations are done in two steps. First, only controls for hospitalization and an intercept. Second, control variables are added add other control variables. By doing the estimations in two steps, it is possible to get an indication of how sensitive the hospitalization variable is to adding control variables.

All estimations have sample weights originating from PSM. In table 3, the propensity score has been calculated for all countries jointly, whereas the propensity score is calculated separately for each country in table 4.

Table 3. Average effect of treatment on the probability of retiring. The dataset has been preprocessed with propensity score matching. Marginal effects are calculated from Probit estimations. Marginal effects are calculated from Probit estimations.

	(1)			(2)		
	M.E.		**Std. Err.**	***M.E.***		**Std. Err.**
Hospitalization	0.16	***	(0.02)	0.04	***	(0.01)
Married				0.01		(0.01)
Secondary school				0.01		(0.01)
Tertiary school				-0.03	***	(0.01)
HHsize3				-0.00		(0.01)
HHsize4+				0.00		(0.01)
Native				-0.00		(0.01)
Spouse retired				0.28	***	(0.00)
Age55-59				0.12	***	(0.02)
Age60-64				0.30	***	(0.02)
Age65-70				0.58	***	(0.02)
Male				0.04	***	(0.01)
Large town				-0.01	**	(0.01)
R squared	0.01			0.59		
Observations	8,538			8,538		

Note: Primary school and Age50-54 are baseline. Standard errors in parentheses. R squared originates from the probit estimation *** p<0.01, ** p<0.05 and * p<0.1

Turning to table 3, when only hospitalization is controlled for (column 1), those who have an overnight hospitalization have an 16 percent increase in the probability of retiring. More control variables are included in the second specification (column 2), where overnight hospitalization has a smaller, but still statistically significant positive effect on the probability of retiring (4 percent). Among the other control variables, having tertiary education has a statistically significant negative effect on the probability of retiring, whereas being overweight, obese or having a retired spouse has a statistically significant positive effect on the probability of retiring.

3.5.2 Results-separated by countries

Table 4 presents the results for each country separately. Propensity scores were been calculated for each country separately. As in Table 3, each country is analyzed twice, one output (a) for when only hospitalization is controlled for, and one output (b) where other control variables are controlled for. To make the results transparent, the output is limited to the hospitalization coefficient[15].

When only hospitalization is controlled for, column a, hospitalization has a statistically significant positive effect on the probability of retirement for all countries. The effect ranges in size between countries. For instance, hospitalization increases the probability of retirement by 25 percent in Sweden and 20 percent in France. The highest effect is observed in Greece (35 percent increase) and the result is statistically insignificant for Switzerland. Turning to column b where additional control variables are included, the overall pattern is that the effect decreases. For instance, for Spain, the coefficient for hospitalization decreases from 26 percentage points

[15] Full results are available upon request

Table 4 - Average effect of treatment of the treated on the probability of retiring. Each country has been preprocessed with propensity score matching separately. Marginal effects are calculated from Probit estimations.

	Austria		Germany		Sweden	
	(a)	(b)	(a)	(b)	(a)	(b)
Hospitalization	0.23 *** (0.06)	0.03 (0.03)	0.38 *** (0.03)	0.03 (0.02)	0.25 *** (0.05)	0.11 *** (0.03)
Additional controls*	No	Yes	No	Yes	No	Yes
R squared	0.01	0.67	0.02	0.76	0.00	0.61
n	622	622	833	833	1,118	1,118
	The Netherlands		**Spain**		**Italy**	
	(a)	(b)	(a)	(b)	(a)	(b)
Hospitalization	0.20 *** (0.06)	-0.00 (0.03)	0.26 *** (0.06)	0.10 ** (0.04)	0.12 ** (0.06)	0.02 (0.04)
Additional controls*	No	Yes	No	Yes	No	Yes
R squared	0.00	0.64	0.00	0.72	0.00	0.61
n	661	661	374	374	762	762
	France		**Denmark**		**Greece**	
	(a)	(b)	(a)	(b)	(a)	(b)
Hospitalization	0.20 *** (0.05)	0.02 (0.03)	0.10 (0.06)	-0.03 (0.04)	0.35 *** (0.06)	0.10 *** (0.04)
Additional controls*	No	Yes	No	Yes	No	Yes
R squared	0.00	0.66	0.00	0.65	0.00	0.60
n	971	971	624	624	946	946
	Switzerland		**Belgium**			
	(a)	(b)	(a)	(b)		
Hospitalization	0.13 (0.09)	0.01 (0.03)	0.14 *** (0.04)	0.01 (0.02)		
Additional controls*	No	Yes	No	Yes		
R squared	0.00	0.80	0.00	0.60		
n	246	246	1,263	1,263		

Note: *The full list of control variables is presented in table 2. Full results are found in Appendix A Standard errors in parentheses. R squared originates from the probit estimation. *** p<0.01, ** p<0.05 and * p<0.1or diseases Standard errors in parentheses. R squared originates from the probit estimation *** p<0.01, ** p<0.05 and * p<0.1parentheses. R squared originates from the probit estimation. *** p<0.01, ** p<0.05 and * p<0.1

to 10 percentage points when additional control variables are included. The coefficient for hospitalization is insignificant for 8 (Austria, Germany, The Netherlands,

3.5.3 Additional analysis

Several additional estimations were conducted. First, a regression (not shown) which includes, besides the control variables in column 2, table 3, interaction effects between education and hospitalization, as the theoretical framework stipulates that health and education may be interlinked. The interaction effects are not statistically significant (p-value>0.5) and the marginal effect of hospitalization on the probability of retiring is 4 percent (similar to column 2, table 3).

Second, the third wave (SHARELIFE) is included by adding early life conditions measured as different kinds of illnesses and diseases up until the individual is 15 years old. Although studies have shown that early life conditions are important for individual health (van den Berg, Doblhammer, & Christensen, 2009; Bengtsson & Mineau, 2009; Lindeboom, Portrait, & van den Berg, 2010; Tubeuf, Jusot, & Bricard, 2012), early life conditions are not included in the main analysis, as the sample size decreases when using three waves jointly. Due to the smaller sample size, only estimations on all the countries jointly are conducted. The results, contained in table 5, show that when only controlling for hospitalization, those who have experienced hospitalization increase their probability of retirement by 14 percent (column 1). When control variables are added, hospitalization increases the probability of retirement by 4 percent (column 2). It can also be noted that having allergies decreases the probability of retiring by 4 percent.

Table 5. Average treatment effect on the treated on the probability to retire. The dataset has been preprocessed with propensity score matching. Marginal effects are calculated from Probit estimations.

	(1)			(2)		
	M.E.		**Std. Err.**	***M.E.***		**Std. Err.**
Hospitalization	0.14	***	(0.02)	0.04	***	(0.01)
Married				0.01		(0.01)
Secondary school				0.00		(0.01)
Tertiary school				-0.03	***	(0.01)
HHsize3				0.00		(0.01)
HHsize4+				0.01		(0.01)
Native				-0.01		(0.01)
Spouse retired				0.28	***	(0.01)
Age55-59				0.13	***	(0.02)
Age60-64				0.31	***	(0.02)
Age65-70				0.59	***	(0.02)
Male				0.04	***	(0.01)
Large town				-0.02	**	(0.01)
Early life						
Polio				-0.04		(0.03)
Asthma				-0.02		(0.03)
Respiratory				0.03		(0.02)
Allergies				-0.04	**	(0.02)
Ear problems				-0.02		(0.02)
Headaches				-0.01		(0.02)
Appendicitis				0.01		(0.01)
R squared	0.01			0.60		
Observations	6,510			6,510		

Note: Primary school and Age50-54 are baseline. "Early life" refers to if the respondent had any of the listed illnesses or diseases Standard errors in parentheses. R squared originates from the probit estimation *** p<0.01, ** p<0.05 and * p<0.1

Third, those who are permanently sick or disabled (see table 1) are included in the category of retired, as individuals who experience hospitalization may end up in this group instead of the retirement group. Results (not shown) reveal that, when only controlling for hospitalization, those who have experienced hospitalization increase their probability of retirement by 17 percent. The corresponding figure, when adding control variables (similar to column 2, table 3), is 5 percent.

3.6 Discussion

The aim of this paper was to estimate how a negative health shock affects the retirement decision. When all the countries were analyzed together, the results suggest that hospitalization has a positive effect on the probability of retiring. This is in line with previous studies (e.g. McGeary, 2009; e.g. Renna & Thakur, 2010), which also find that worsened health has a positive effect on the probability of retiring.

The additional analyzes show that the results are robust; inclusion of early life conditions did not alter the size of the effect for hospitalization on the retirement decision, when other control variables were included (from 5 percent to 4 percent). Moreover, the result does not change (when control variables were included) when the definition of retired was extended to include those who were disabled or permanently sick.

Table 4 (separate regressions for each country) shows that, for all the countries, hospitalization has a statistically significant positive effect on the probability of retiring when no other control variables are included. When control variables are included, the positive effect of hospitalization on retirement decreases and turns statistically insignificant for some countries. This indicates that the effect of hospitalization on the retirement decision differs from country to country. As mentioned in section 2, the effect of hospitalization on the retirement decision may be influenced by access to rehabilitory care. Consequently, it is possible that some of the disparities between countries can be explained by disparities in access to rehabilitory care. It would be interesting to further analyze disparities in access to rehabilitory care and its effect on retirement decision. Unfortunately, such information is not available in the SHARE survey.

The results in this study are interesting from a policy perspective, as they indicate that the institutional structure is important for the effect of a negative health shock on the retirement status. On the one hand, policy makers might want to increase the labor force participation and create incentives to decrease the effect of a negative health shock on the retirement decision. On the other hand, policy makers might want to ease the burden for ill individuals and make it possible for them to retire after a negative health shock. Hence, depending on the policy makers' agenda, incentives can be aligned to achieve the desired results.

One limitation of this study is that the ease of which individuals are admitted to hospital might differ between countries. If some individuals are admitted to hospital for minor conditions in one country but not in other countries, the effect of hospitalization on the retirement decision will differ between countries. By only looking at overnight hospitalization, minor conditions may be avoided. To further analyze this issue, more detailed data on specific health conditions is needed.

Acknowledgement

This paper used data from SHARE wave 4 release 1, as of November 30, 2012 or SHARE wave 1 and 2 release 2.5.0, as of May 24, 2011 or SHARELIFE release 1, as of November 24, 2010. The SHARE data collection has been primarily funded by the European Commission through the 5th Framework Program (project QLK6-CT-2001-00360 in the thematic program Quality of Life), through the 6th Framework Program (projects SHARE-I3, RII-CT-2006-062193, COMPARE, CIT5- CT-2005-028857, and SHARELIFE, CIT4-CT-2006-028812) and through the 7th Framework Program (SHARE-PREP, N° 211909, SHARE-LEAP, N° 227822 and SHARE M4, N° 261982). Additional funding from the U.S. National Institute on Aging (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, R21 AG025169, Y1-AG-4553-01, IAG BSR06-11 and OGHA 04-064) and the German Ministry of Education and Research as well as from various national sources is gratefully acknowledged (see www.share-project.org for a full list of funding institutions).

References

Alavinia, S. M. & Burdorf, A. (2008). Unemployment and retirement and ill-health: a cross-sectional analysis across European countries. *International Archives of Occupational and Environmental Health*, *82*(1), 39–45–45.

Anand, P., Mizala, A. & Repetto, A. (2008). *Using School Scholarships to Estimate the Effect of Private Education on the Academic Achievement of Low Income Students in Chile*. Education Working Paper Archive. Retrieved from http://ludwig.lub.lu.se/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED508954&site=eds-live&scope=site

Au, D. W. H., Crossley, T. F. & Schellhorn, M. (2005). The effect of health changes and long-term health on the work activity of older Canadians. *Health Economics*, *14*(10), 999–1018.

Bengtsson, T. & Mineau, G. P. (2009). Early-life effects on socio-economic performance and mortality in later life: A full life-course approach using contemporary and historical sources. *Social Science & Medicine*, *68*(9), 1561–1564.

Bingley, P. & Lanot, G. (2007). Public pension programmes and the retirement of married couples in Denmark. *Journal of Public Economics*, *91*(10), 1878–1901.

Bolin, K., Lindgren, A., Lindgren, B. & Lundborg, P. (2009). Utilisation of Physician Services in the 50+ Population: The Relative Importance of Individual versus Institutional Factors in 10 European Countries. *International Journal of Health Care Finance and Economics*, (1), 83.

Bound, J. (1991). Self-Reported Versus Objective Measures of Health in Retirement Models. *The Journal of Human Resources*, *26*(1), 106–138.

Bound, J., Stinebrickner, T. & Waidmann, T. (2010). Health, economic resources and the work decisions of older men. *Structural Models of Optimization Behavior in Labor, Aging, and Health*, *156*(1), 106–129.

Burkhauser, R. V. (1979). The Pension Acceptance Decision of Older Workers. *The Journal of Human Resources*, *14*(1), 63–75.

Börsch-Supan, A., Hank, K., Jürges, H. & Schröder, M. (2009). Introduction: empirical research on health, ageing and retirement in Europe. *Journal of European Social Policy*, *19*(4), 293 –300.

Christensen, B. J. & Kallestrup-Lamb, M. (2012). The Impact of Health Changes on Labor Supply: Evidence from Merged Data on Individual Objective Medical Diagnosis Codes and Early Retirement Behavior. *Health Economics*, *21*, 56–100.

Gupta, D. N. & Larsen, M. (1). (2007). Health shocks and retirement: The role of welfare state institutions. *European Journal of Ageing*, *4*(3), 183–190.

Gupta, D. N. & Larsen, M. (2010). The Impact of Health on Individual Retirement Plans: Self-Reported versus Diagnostic Measures. *Health Economics*, *19*(7), 792–813.

Dehejia, R. H. & Wahba, S. (1999). Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs. *Journal of the American Statistical Association*, *94*(448), 1053–1062.

Denaeghel, K., Mortelmans, D. & Borghgraef, A. (2011). Spousal influence on the retirement decisions of single-earner and dual-earner couples. *Advances in Life Course Research*, *16*(3), 112 – 123.

Disney, R., Emmerson, C. & Wakefield, M. (2006). Ill health and retirement in Britain: A panel data-based analysis. *Journal of Health Economics*, *25*(4), 621–649.

Dwyer, D. S. & Mitchell, O. S. (1999). Health problems as determinants of retirement: Are self-rated measures endogenous? *Journal of Health Economics*, *18*(2), 173–193.

Engelhardt, H. (2012). Late Careers in Europe: Effects of Individual and Institutional Factors. *European Sociological Review*, *28*(4), 550–563.

García Gómez, P. & López Nicolás, A. (2006). Health shocks, employment and income in the Spanish labour market. *Health Economics*, *15*(9), 997–1009.

García-Gómez, P. (2011). Institutions, health shocks and labour market outcomes across Europe. *Journal of Health Economics*, *30*(1), 200–213.

Gomez, G., Kippersluis, H. van, O'Donnell, O. & Doorslaer, E. (2013). Long Term and Spillover Effects of Health Shocks on Employment and Income. *Journal of Human Resources*.

Gortz, M. (2012). Early retirement in the day-care sector: the role of working conditions and health. *EUROPEAN JOURNAL OF AGEING*, *9*(3), 187–198.

Gruber, J. & Wise, D. (2000). Social security programs and retirement around the world. *Research in Labor Economics*, *18*.

Hamermesh, D. S. (1984). Life-Cycle Effects on Consumption and Retirement. *Journal of Labor Economics*, *2*(3), 353–370.

Ho, D. E., Imai, K., King, G. & Stuart, E. A. (2007). Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Political Analysis*, *15*(3), 199–236.

Imbens, G. W. & Wooldridge, J. M. (2009). Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature*, *47*(1), 5–86.

Jones, A. (2007). *Applied Econometrics for Health Economists* (Second.). Radcliffe Publishing. Retrieved from http://www.stata.com/bookstore/applied-econometrics-health-economists/index.html

Jürges, H. (2007). True health vs response styles: exploring cross-country differences in self-reported health. *Health Economics*, *16*(2), 163–178.

Kerkhofs, M., Lindeboom, M. & Theeuwes, J. (1999). Retirement, financial incentives and health. *Labour Economics*, *6*(2), 203–227.

Lee, W.-S. (2011). Propensity score matching and variations on the balancing test. *Empirical Economics*, 1–34.

Lindeboom, M., Portrait, F. & van den Berg, G. J. (2010). Long-run effects on longevity of a nutritional shock early in life: The Dutch Potato famine of 1846–1847. *Journal of Health Economics*, *29*(5), 617–629.

Mazzonna, F. & Peracchi, F. (2012). Ageing, cognitive abilities and retirement. *European Economic Review*, *56*(4), 691–710.

McClellan, M. B. (1998). Health Events, Health Insurance, and Labor Supply: Evidence from the Health and Retirement Survey. In D. A. Wise (Ed.), *Frontiers in the Economics of Aging* (pp. 301–350). University of Chicago Press. Retrieved from http://www.nber.org/chapters/c7304

McGarry, K. (2004). Health and Retirement: Do Changes in Health Affect Retirement Expectations? *The Journal of Human Resources*, *39*(3), 624–648.

McGeary, K. (2009). How do health shocks influence retirement decisions? *Review of Economics of the Household*, *7*(3), 307–321–321.

Nichols, A. (2008). Erratum and discussion of propensity-score reweighting. *The Stata Journal*, *8*(4), 532–539.

Renna, F. & Thakur, N. (2010). Direct and indirect effects of obesity on U.S. labor market outcomes of older working age adults. *Social Science & Medicine*, *71*(2), 405 – 413.

Rosenbaum, P. & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, *70*, 41–55.

Siegrist, J., Wahrendorf, M., von dem Knesebeck, O., Jürges, H. & Börsch-Supan, A. (2006). Quality of work, well-being, and intended early retirement of older employees—baseline results from the SHARE Study. *The European Journal of Public Health*, *17*(1), 62 –68.

Thoemmes, F. J. & Kim, E. S. (2011). A Systematic Review of Propensity Score Methods in the Social Sciences. *Multivariate Behavioral Research*, *46*(1), 90–118.

Tubeuf, S., Jusot, F. & Bricard, D. (2012). Mediating Role of Education and Lifestyles in the Relationship Between Early-Life Conditions And Health: Evidence from the 1958 British Cohort. *Health Economics*, *21*, 129–150.

Vaillant, N. G. & Wolff, F.-C. (2012). Retirement intentions of older migrant workers: Does health matter? *HAL : hal-00671527, version 1*.

Van den Berg, G. J., Doblhammer, G. & Christensen, K. (2009). Exogenous determinants of early-life conditions, and mortality later in life. *Social Science & Medicine*, *68*(9), 1591–1598.

Wang, H., Yip, W., Zhang, L. & Hsiao, W. C.. (2009). The impact of rural mutual health care on health status: Evaluation of a social experiment in rural China. *Health Economics*, *18*(SUPPL. 2), S65–S82.

Wim, G. (2000). Adaptation and scale of reference bias in self-assessments of quality of life. *Journal of Health Economics*, *19*(3), 403–420.

Appendix A. Average effect of treatment on the probability of retiring. The dataset has been preprocessed with propensity score matching. Marginal effects are calculated from Probit estimations. Marginal effects are calculated from Probit estimations.

	Austria						Germany					
	(1)			(2)			(3)			(4)		
	M.E.		**Std. Err.**	***M.E.***		**Std. Err.**	***M.E.***		**Std. Err.**	***M.E.***		**Std. Err.**
Hospitalization	0.23	***	(0.06)	0.03		(0.03)	0.38	***	(0.03)	0.03		(0.02)
Married				-0.02		(0.02)				0.01		(0.01)
Secondary school				-0.03		(0.03)				-0.05		(0.03)
Tertiary school				-0.07	**	(0.04)				-0.05	*	(0.03)
HHsize3				-0.03		(0.03)				-0.02		(0.02)
HHsize4+				0.03		(0.04)				-0.03		(0.02)
Native				0.05		(0.04)				-0.00		(0.01)
Spouse retired				0.23	***	(0.02)				0.14	***	(0.01)
Age55-59				0.33	***	(0.07)				0.05	*	(0.03)
Age60-64				0.52	***	(0.08)				0.19	***	(0.03)
Age65-70				0.69	***	(0.08)				0.58	***	(0.09)
Male				0.01		(0.03)				0.01		(0.01)
Large town				0.01		(0.02)				0.01		(0.01)
R squared	0.01			0.67			0.02			0.76		
Observations	622			622			833			833		

Note: Primary school and Age50-54 are baseline. Standard errors in parentheses. R squared originates from the probit estimation *** p<0.01, ** p<0.05 and * p<0.1

Appendix B. Average effect of treatment on the probability of retiring. The dataset has been preprocessed with propensity score matching. Marginal effects are calculated from Probit estimations. Marginal effects are calculated from Probit estimations.

	Sweden						The Netherlands					
	(1)			(2)			(3)			(4)		
	M.E.		**Std. Err.**	***M.E.***		**Std. Err.**	***M.E.***		**Std. Err.**	***M.E.***		**Std. Err.**
Hospitalization	0.25	***	(0.05)	0.11	***	(0.03)	0.20	***	(0.06)	-0.00		(0.03)
Married				-0.03		(0.02)				0.01		(0.03)
Secondary school				-0.00		(0.02)				-0.03		(0.03)
Tertiary school				-0.05	***	(0.02)				-0.01		(0.03)
HHsize3				0.02		(0.04)				-0.00		(0.04)
HHsize4+				-0.03		(0.04)				-0.08	**	(0.04)
Native				-0.06	**	(0.03)				0.00		(0.04)
Spouse retired				0.21	***	(0.02)				0.22	***	(0.02)
Age55-59				-0.02		(0.04)				0.01		(0.06)
Age60-64				0.17	***	(0.05)				0.15	**	(0.07)
Age65-70				0.59	***	(0.06)				0.50	***	(0.10)
Male				0.00		(0.02)				0.07	***	(0.02)
Large town				-0.00		(0.02)				0.02		(0.02)
R squared	0.00			0.61			0.00			0.64		
Observations	1,118			1,118			661			661		

Note: Primary school and Age50-54 are baseline. Standard errors in parentheses. R squared originates from the probit estimation *** p<0.01, ** p<0.05 and * p<0.1

Appendix C. Average effect of treatment on the probability of retiring. The dataset has been preprocessed with propensity score matching. Marginal effects are calculated from Probit estimations. Marginal effects are calculated from Probit estimations.

	Spain						Italy					
	(1)			(2)			(3)			(4)		
	M.E.		**Std. Err.**	***M.E.***		**Std. Err.**	***M.E.***		**Std. Err.**	***M.E.***		**Std. Err.**
Hospitalization	0.26	***	(0.08)	0.10	**	(0.04)	0.12	**	(0.06)	0.02		(0.04)
Married				0.00		(0.04)				0.01		(0.03)
Secondary school				-0.07	*	(0.04)				-0.03		(0.03)
Tertiary school				-0.05		(0.04)				-0.03		(0.04)
HHsize3				0.03		(0.03)				-0.00		(0.03)
HHsize4+				0.03		(0.03)				-0.05		(0.03)
Native				0.04		(0.05)				0.17	***	(0.06)
Spouse retired				0.23	***	(0.02)				0.27	***	(0.02)
Age55-59				0.17	***	(0.04)				0.39	***	(0.09)
Age60-64				0.27	***	(0.04)				0.59	***	(0.08)
Age65-70				0.71	***	(0.09)				0.71	***	(0.08)
Male				0.08	**	(0.03)				0.04		(0.03)
Large town				0.01		(0.03)				0.01		(0.03)
R squared	0.00			0.72			0.00			0.61		
Observations	374			374			762			762		

Note: Primary school and Age50-54 are baseline. Standard errors in parentheses. R squared originates from the probit estimation *** p<0.01, ** p<0.05 and * p<0.1

Appendix D. Average effect of treatment on the probability of retiring. The dataset has been preprocessed with propensity score matching. Marginal effects are calculated from Probit estimations. Marginal effects are calculated from Probit estimations.

	France						Denmark				
	(1)			(2)			(3)		(4)		
	M.E.		**Std. Err.**	***M.E.***		**Std. Err.**	***M.E.***	**Std. Err.**	***M.E.***		**Std. Err.**
Hospitalization	0.20	***	(0.05)	0.02		(0.03)	0.10	(0.06)	-0.03		(0.04)
Married				0.02		(0.02)			-0.01		(0.02)
Secondary school				0.01		(0.02)			0.04		(0.04)
Tertiary school				-0.04		(0.03)			-0.04		(0.04)
HHsize3				-0.03		(0.04)			0.00		(0.00)
HHsize4+				-0.03		(0.03)			0.15	**	(0.06)
Native				-0.02		(0.03)			0.11	**	(0.05)
Spouse retired				0.25	***	(0.02)			0.20	***	(0.02)
Age55-59				0.18	***	(0.04)			-0.03		(0.05)
Age60-64				0.58	***	(0.06)			0.37	***	(0.07)
Age65-70				0.74	***	(0.07)			0.70	***	(0.07)
Male				0.15	***	(0.03)			0.01		(0.02)
Large town				-0.05	**	(0.02)			0.02		(0.02)
R squared	0.00			0.66			0.00		0.65		
Observations	971			971			624		624		

Note: Primary school and Age50-54 are baseline. Standard errors in parentheses. R squared originates from the probit estimation *** p<0.01, ** p<0.05 and * p<0.1

Appendix E. Average effect of treatment on the probability of retiring. The dataset has been preprocessed with propensity score matching. Marginal effects are calculated from Probit estimations. Marginal effects are calculated from Probit estimations.

	Greece						Switzerland					
	(1)			(2)			(3)			(4)		
	M.E.		**Std. Err.**	***M.E.***		**Std. Err.**	***M.E.***		**Std. Err.**	***M.E.***		**Std. Err.**
Hospitalization	0.35	***	(0.06)	0.10	***	(0.04)	0.35	***	(0.06)	0.10	***	(0.04)
Married				0.03		(0.02)				0.03		(0.02)
Secondary school				0.03	*	(0.02)				0.03	*	(0.02)
Tertiary school				-0.01		(0.02)				-0.01		(0.02)
HHsize3				-0.02		(0.03)				-0.02		(0.03)
HHsize4+				0.03		(0.03)				0.03		(0.03)
Native				0.07		(0.07)				0.07		(0.07)
Spouse retired				0.29	***	(0.01)				0.29	***	(0.01)
Age55-59				0.05	*	(0.03)				0.05	*	(0.03)
Age60-64				0.14	***	(0.03)				0.14	***	(0.03)
Age65-70				0.34	***	(0.05)				0.34	***	(0.05)
Male				-0.05	***	(0.02)				-0.05	***	(0.02)
Large town				-0.02		(0.02)				-0.02		(0.02)
R squared	0.00			0.60			0.00			0.60		
Observations	946			946			246			246		

Note: Primary school and Age50-54 are baseline. Standard errors in parentheses. R squared originates from the probit estimation *** p<0.01, ** p<0.05 and * p<0.1

Appendix F. Average effect of treatment on the probability of retiring. The dataset has been preprocessed with propensity score matching. Marginal effects are calculated from Probit estimations. Marginal effects are calculated from Probit estimations.

	Belgium					
	(1)			(2)		
	M.E.		**Std. Err.**	***M.E.***		**Std. Err.**
Hospitalization	0.14	***	(0.04)	0.01		(0.02)
Married				0.05	***	(0.02)
Secondary school				-0.01		(0.02)
Tertiary school				-0.02		(0.02)
HHsize3				-0.08	***	(0.03)
HHsize4+				-0.10	***	(0.03)
Native				0.04		(0.03)
Spouse retired				0.23	***	(0.01)
Age55-59				0.17	***	(0.04)
Age60-64				0.41	***	(0.05)
Age65-70				0.69	***	(0.06)
Male				0.04	**	(0.02)
Large town				-0.02		(0.02)
R squared	0.14			0.60		
Observations	1,263			1,263		

Note: Primary school and Age50-54 are baseline. Standard errors in parentheses. R squared originates from the probit estimation *** p<0.01, ** p<0.05 and * p<0.1

4 Pharmaceutical-based health investment differences between immigrants and natives in Sweden

Abstract

This paper (co-authored with Sofie Gustafsson) analyzed the pharmaceutical-based health investments differences between immigrants and natives in Sweden. This paper used an interview survey combined with a registered database containing all prescribed pharmaceuticals. Immigrants were divided into the groups according to their region of origin: Nordic, Western and non-Western.

The results show that there are differences in the utilization of prescribed pharmaceuticals between immigrants and natives. For instance, when looking at all pharmaceuticals analyzed together, males with a non-Western origin are less likely to access prescribed pharmaceuticals, compared with native males. Turning to specific pharmaceutical groups, the results show that immigrants, compared with natives, are less likely to have access to first-line treatments for high blood pressure, heart failure and kidney diseases.

4.1 Introduction

Worldwide, differences in healthcare utilization (Rue et al. 2008; Wamala et al. 2007) and health (Denktaş et al. 2010; Morgan et al. 2011) has been observed between immigrants and natives. To reduce financial barriers to healthcare, the Swedish health insurance covers all (legal) residents. For prescribed pharmaceuticals, the county council fully subsidizes individual payments exceeding SEK 2,200 (~USD 330) on a 12-month rolling basis. Nevertheless, differences in the utilization of prescribed pharmaceuticals (Nordin, Dackehag, and Gerdtham 2013; Sundquist 1993), as well as health care utilization in general (Westin et al. 2004; Wamala et al. 2007), are observed between natives and immigrants in Sweden (an immigrant is defined as someone who resides in Sweden but was born elsewhere). These results can imply that the goal of Swedish health policy - that all (legal) residents should have equal access to medical care according to need - is unachieved.

Pharmaceutical treatment is the dominating medical intervention available for several health conditions, and is often immensely important for the course of the disease. For instance, pharmaceuticals substantially reduce cardiovascular related morbidity and mortality (WHO 2012). Thus, disparities between populations regarding pharmaceutical utilization may have significant public health consequences.

Despite the fact that immigrants constituted 15% of the Swedish population in 2011 (Statistics Sweden 2011), and that immigrants' health has considerable consequences for general public health and healthcare expenditures, little has been written to explore differences in pharmaceutical utilization between immigrants and natives in Sweden.

A previous study using Swedish data (Sundquist 1993) and broad overall measures of pharmaceutical utilization, but not controlling for health or socioeconomic differences between populations made it less suited as a basis for designing public health policies, as disparities in utilization partly reflect differences in need. By combining data from the Swedish Survey of Living Conditions (ULF) and the Swedish Prescribed Drug Register (SPDR), our paper analyzes differences in prescribed pharmaceutical utilization between immigrants and natives in Sweden. In other words, we analyzed pharmaceutical-based health investments between immigrants and natives in Sweden. The detailed individual-level data enables us not only to explore overall differences in pharmaceutical-based health investments between natives and immigrants, but also to disentangle differences related to health and socioeconomic status from other factors related to immigration.

This paper proceeds as follows. Section 4.2 provides a brief background of immigrants in Swedish. Section 4.3 introduces the demand-for-health framework. The dataset and the pharmaceutical classification system are presented in Section 4.4. Section 4.5 describes the empirical specification. The results are presented in Section 4.6 and Section 4.7 concludes the paper.

4.2 Heterogeneity among immigrants in Sweden: a brief background

The main reason for migration to Sweden has changed over time. Between the Second World War and up to the early 1970s, Sweden experienced a labor shortage. As immigration was predominantly labor-related, immigrants were largely employed in the Swedish labor market. Since the 1970s, the Swedish labor shortage has decreased and the composition of immigrants has changed into mainly comprising refugees and family-related immigrants. In contrast to labor-related immigrants, refuges more often face harsh labor-market opportunities in Sweden. (Ekberg 2011)

"The healthy immigrant effect" (Marmot et al. 1984) – where immigrants, due to self-selection, are on average healthier than natives, has been proposed as an explanation for observed disparities in healthcare utilization between immigrants and natives. However, this has been refuted by findings that immigrants are disadvantaged in several health-related aspects. For instance, compared with natives, immigrants have (1) lower self-rated health (Lindström, Sundquist, and Östergren 2001), (2) higher risk of cardiovascular diseases (Gadd et al. 2005), (3) higher overall mortality (Sundquist and Johansson 1997), (4) higher rate of suicide (Johansson et al. 1997; Ferrada-Noli 1997), and (5) higher prevalence of psychiatric illness (Bayard-Burfield, Sundquist, and Johansson 2001).

4.3 Theoretical framework

We will specify and interpret the empirical analysis within the demand-for-health framework (Grossman 1972), which is the dominating economic theory of individual health-related behavior. It is formulated as a neo-classical model in the same tradition as Becker´s human-capital theory (1964). Its novelty lies in the notion that health-capital differs from other human-capital components, as it affects the individual´s time-budget constraint (the investment aspect of health), whereas human-capital affects the individual´s productivity. Therefore, good health is desirable not only because it is enjoyable, but also because it affects the time available for market earnings. Consequently, the individual´s monetary pay-off regarding own health investments largely depends on his or her success in the labor market and, hence, the individual´s demand for goods and services to be used for health investments depends on expected labor-market opportunities.

The demand-for-health framework assumes that individuals simultaneously invest and produce health by engaging in healthy activities by combining own time and market goods. Thus, the demand for market goods, such as pharmaceuticals, is derived from the underlying demand for health, which in turn depends on labor-market opportunities. In addition, health is assumed to depreciate with time at an increasing rate. Therefore, the demand for health investments is likely to increase with age.

The demand-for-health framework also assumes that educationally advantaged individuals are more efficient producers of their own health. More specifically, more educated individuals may both be better informed about pharmaceutical treatment options and more capable at maneuvering in the healthcare system, making a specific pharmaceutical-based health investment less costly than for less educated individuals. Education should be interpreted broadly as incorporating not only formal education, but also country-specific knowledge such as the ability to speak the native language. Priebe et al. (2011) report that immigrants' linguistic difficulty and unfamiliarity with the healthcare system are two barriers to healthcare. To exemplify, patients´ linguistic difficulties may obstruct physicians' diagnostic efforts, leading to additional consultations or fewer prescriptions. Thus, pharmaceutical-based health investments may be more costly for immigrants than for natives. Moreover, Wamala et al. (2007) show that some immigrants refrain from seeking healthcare due to perceived ethnic discrimination, which adds to the cost of pharmaceutical-based health investments. In addition, Carlsson and Rooth (2007) showed that immigrants are discriminated against in the Swedish labor market, which reduces the monetary pay-off from the labor market and, hence, reduces the monetary pay off from pharmaceutical-based health investments.

4.4 Data

We used the HILDA (Health and Individuals Longitudinal Data and Analysis) dataset, which combines the Swedish Survey of Living Conditions (ULF) and the Swedish Prescribed Drug Register (SPDR). The ULF survey asks about, for instance, socioeconomic and labor-market status and the respondents' answers were complemented with national registry data on such things as taxes and monetary transfers. As the ULF survey years 2004 and 2005 had a special health focus, the survey also included many health-related questions. Compared with population wide register data only, the ULF data includes a wide array of health-related information. As health is strongly correlated to pharmaceutical utilization, health must be adjusted for when estimating disparities in the pharmaceutical-based health investments of population groups, otherwise differences in investments will only partly reflect differences in health between population groups. The inclusion of health variables makes the utilizations of survey data, complemented with register data, superior to using population wide register data, which normally does not include measures of health. ULF comprises respondents aged 16 or older and is representative for the Swedish population. The 2004 and 2005 ULF-survey had a 75% response rate, which resulted in 10,179 respondents (Statistics Sweden 2012). Because we are interested in individuals with completed education, we restrict our sample to respondents aged 25 and older, resulting in a working sample of 8,488 respondents (48% males and 11% immigrants).

The SPDR registers *all* dispensed prescribed pharmaceuticals in Sweden, and includes patient identities since July 2005. It also contains detailed information on the patient (e.g. age, sex and personal identification number), the prescriber (e.g. profession and practice) and the dispensed pharmaceutical (e.g. substance, volume, costs, reimbursement, prescribed dosage regime, date of prescribing and dispensing). For a more detailed description of the SPDR register, see Wettermark et al. (2007). At the time of data collection, SPRD was available for July 2005 through November 2007.

Pharmaceutical classification system

As we are interested in the volume of pharmaceutical-based health investments, we utilize the standardized pharmaceutical measurement unit, defined daily doses (DDD). The WHO Collaborating Centre for Drug Statistics Methodology (WHOCC) assigns pharmaceutical substances a DDD which by definition is the assumed daily maintenance dose when used for its main therapeutic use by adults. Compared with other measurement units like the number of unique pharmaceuticals, the DDD is superior for our purpose as it measures the volume of active pharmaceutical substances. To illustrate, the volume of attended gym classes may be a better measurement of health investments than the number of memberships in health clubs.

In addition to the volume of pharmaceutical-based health investments, we are interested in the type of pharmaceutical-based health investments. To distinguish between pharmaceutical classes, we utilize the Anatomical, Therapeutic and Chemical (ATC) classification system. The WHOCC hierarchically assigns pharmaceutical substances a five-level ATC-code for its main therapeutic use. The first and second levels are the directed anatomical system and the therapeutic subgroup. These levels are broad and comprise pharmaceuticals with diverse indications. The third level – the pharmacological subgroup – is narrower and comprises pharmaceuticals with similar therapeutic use. To illustrate, the often interchangeably used painkillers, acetylsalicylsyra (Aspirin®) and paracetamol (Alvedon®), have similar therapeutic use and belong to the same third ATC-level. Therefore, we use the third ATC-level to distinguish between different types of pharmaceutical-based health investments. Due to sample size restrictions, we refrain from distinguishing between pharmaceutical classes in the fourth and fifth and ATC-levels.

Starting with the individual's first dispensed pharmaceutical; each reimbursement period lasts 12 months. On reaching the cap for patient pharmaceutical co-payment, the individual is exempted from further payments within the reimbursement period. As stockpiling pharmaceuticals is more common among individuals exempted from payments than among individuals with co-payments (Krigsman et al. 2007), dispensed pharmaceuticals may deviate from the utilized amount, especially over short periods. Given that our dataset enables us to measure dispensed pharmaceuticals from July 2005 through November 2007, we utilize the whole period when analyzing pharmaceutical-based health investments.

Dependent variables

We measured pharmaceutical-based health investments by the access-no access dichotomy and, if access, the accessed number of DDDs.

Independent variables

We defined immigrant as a Swedish resident born abroad to non-Swedish-born parents. We used two immigration measures: (1) *immigrant,* which is a dummy variable taking the value 1 if immigrant and 0 otherwise, and (2) immigrant region of origin i.e. *Nordic origin*, *Western origin* and *non-Western origin,* which for immigrants is a set of mutually exclusive dummy variables; *Nordic origin* takes the value 1 if born in Denmark, Finland, Iceland or Norway and 0 otherwise; *Western origin* takes the value 1 if born in Belgium, France, Greece, Ireland, Italy, Great Britain and Northern Ireland, Germany, Austria, USA, Portugal, The Netherlands, Switzerland or Spain and 0 otherwise; *non-Western origin* takes the value 1 if immigrant region of origin is neither *Nordic* nor *Western* and 0 otherwise.

On a conceptual level, individual characteristics like age, education and health-state influence the "price" and "benefit" for health and accordingly the demand for prescribed pharmaceuticals. Thus, we include controls for age, education and health-state (e.g. self-assessed health, life style factors and the occurrence of specific medical conditions). Because financial and time budget constraints also influence such demand, we control for disposable household income, full-time work and children in the household. Moreover, we control for additional factors (married or cohabitant, residency in larger cities and 2nd generation immigrant) that may be associated with health-related behavior and attitude to pharmaceutical utilization. The control variables are described more in depth in Appendix A.

Descriptive statistics

Table 1 reports the descriptive statistics. Our working sample includes 8,488 individuals (48% males and 11% immigrants. For immigrants, the most common region of origin is non-Western (54%), followed by Nordic (33%) and Western (12%), whereas the most common countries or birth are Finland (22%), former Yugoslavia (12%), Iraq (6%) and Iran (5%).

In terms of pharmaceutical access and average accessed number of DDDs, the respondents in our sample are similar to the population in general[16]. Comparing population groups within our sample, the descriptive statistics indicate some differences between immigrants and natives. For instance, unconditional on access, female immigrants accessed on average 1,650 DDDs whereas native females accessed 1,896 DDDs. Contrastingly, immigrants in general reported lower self-assessed health than natives.

[16] In 2005, the Swedish national mean was 1,542 DDDs per 1000 individuals and per day (The Swedish National Board of Health and Welfare 2008), which corresponds to 1,434 DDDs per individual and per 31 months. As older people generally utilize more pharmaceuticals than younger people, the figure is presumably larger for individuals aged 25 or older.

Table 1. Descriptive statistics for males and females

		Females N=4410				Males N=4078			
		Natives N=3907		Immigrants N=503		Natives N=3639		Immigrants N=439	
Variables	Description	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Dependent variables (June 2005-November 2007)									
DDD	*Daily Defined Doses*	1896.3	3343.9	1650.0	3150.8	1414.4	2866.9	1455.4	2736.458
Independent variables ULF survey 2004-2005									
Age	*Age in years*	52.7	17.3	50.4	14.9	51.3	16.2	49.2	14.5
lndispinc	*log of disposable income*	7.4	0.5	7.3	0.8	7.4	0.5	7.1	0.9
		Females percent				Males percent			
	Self-assessed health								
SAH 1	Very good health	34.0%		29.2%		39.4%		27.3%	
SAH 2	Good health	36.3%		34.5%		37.4%		40.8%	
SAH 3	Fair or worse health	29.5%		35.9%		23.1%		31.9%	
Underweight	BMI<18.5	5.2%		5.8%		1.2%		1.4%	
Normal weight	18.5≤BMI≤25	56.5%		50.6%		42.8%		39.2%	
Overweight	BMI>25	38.2%		43.7%		56.0%		59.5%	
	Frequency of exercise								
Exercise 1	Never exercises	9.6%		15.1%		10.7%		15.9%	
Exercise 2	Less than once a week	32.4%		33.5%		31.4%		34.9%	
Exercise 3	Once a week	13.0%		12.3%		14.8%		13.4%	
Exercise 4	More than once a week	44.9%		39.1%		43.1%		35.8%	
Smoker	*Currently smokes*	17.5%		21.0%		13.0%		29.2%	
Mobility impairment	Mobility impairment	23.2%		25.6%		13.8%		17.3%	
Self_care	Self-care ability	93.6%		92.9%		97.8%		97.5%	
	Medical diagnose								
Phsych_illness	Psychical illness	5.5%		5.4%		3.2%		6.4%	
Neurology	Neurologic disease	7.8%		7.1%		7.4%		5.0%	
Cardiovasc	Cardiovascular disease	18.5%		17.9%		18.5%		15.7%	
Respiratory	Problem with respiratory system	8.1%		7.5%		7.1%		5.2%	
Skeleton	Problem with your skeleton	24.5%		27.0%		15.8%		21.0%	
Pain	Pain	59.3%		64.6%		50.8%		58.5%	
Other_disease_1	One other diagnose	21.1%		21.8%		17.0%		14.6%	
Other disease_1+	At least two other diagnoses	8.4%		6.0%		7.9%		9.3%	
2nd gen immigrant	Born in Sweden to foreign born parents	1.9%		0%		1.9%		0%	
	Educational level								
Primary school	Up to primary school	21.4%		21.6%		20.4%		19.8%	
Secondary	Completed secondary school	42.9%		39.5%		45.8%		42.8%	
Higher education ≤2	Up to two years of higher education	16.0%		14.5%		15.0%		13.9%	
Higher education >2	At least two years of higher education	19.7%		22.4%		18.8%		23.0%	
Married_cohab	Married or cohabiting	67.5%		65.9%		73.1%		77.7%	
Child_1	Have one child	12.0%		21.0%		11.6%		15.3%	
Child_2+	Have two or more children	22.3%		22.8%		20.5%		25.7%	
Large_city	Live in any of the three biggest cities	32.2%		51.0%		32.0%		50.3%	
Work_full	Work full time	39.4%		39.7%		62.2%		53.8%	

Note: Conditional on access, the accessed amount, (in DDDs) was on average for native females: 2,146.4, immigrant females: 1,929.4, native males: 1,861.9 and immigrant males: 1,836.0.

4.5 Empirical specification

The distribution of the dependent variable has two key characteristics: (1) the presence of non-users (18% had no pharmaceutical access), and (2) a long right-tail (10% of those who accessed most, accessed 25% of the DDDs). As the presence of non-users causes the standard Ordinary Least Squares (OLS) to yield inconsistent results (Cameron and Trivedi 2005), we turn to alternative models.

There are three obvious empirical candidate models to address the presence of non-users: the Tobit model, the Two-part model and the Heckman model. The Tobit model is a standard econometric model addressing dependent variables taking a mixture of zero and positive values. The Tobit model assumes that the same mechanism generates non-zeros and the size of the positive values for non-zeros. In our case, this would imply that the mechanism for pharmaceutical access would equal the mechanism for the accessed amount. This assumption may be violated as people are more or less reluctant to seek healthcare. Thus, when seeking care, people who are more reluctant to seek healthcare may be sicker and, accordingly, they may be prescribed more pharmaceuticals than their less reluctant counterparts. Differently from the Tobit model, the more general Two-part model assumes that two different mechanisms generate the non-zeros and the size of the positive values for non-zeros. A restriction of the Two-part model is if the residuals from the two parts, after control for independent variables, are correlated, in which case selection on unobservables causes selection bias. As the residuals from our analysis are correlated, such selection bias prevents us from using the Two-part model. The Heckman two-step model handles such selection by using an exclusion-criterion (Heckman 1979). For our analysis, an exclusion-criterion would be a variable that influences pharmaceutical access but not the accessed amount. As our dataset lacks such a variable, we are unable to meet the Heckman model exclusion-criterion.

Given that the three described empirical models for addressing the presence of non-users rely on assumptions we may violate, we estimate pharmaceutical utilization in two separate steps: (1) the probability of pharmaceutical access using a probit model , and (2) conditional on access, the accessed amount of DDDs. One way to handle the long right-tailed distribution of DDD in the second step is to estimate log DDD with OLS, which yields log-scaled coefficients. To simplify inference drawing, the log-scaled coefficients are often retransformed to un-scaled coefficients with, for instance, the Duan´s smearing factor (Duan 1983). Given that the retransformation is biased if the residuals are heteroscedastic to the explanatory variables (Ai and Norton 2000), we test for such heteroscedasticity. As the White-test (White 1980) shows that our residuals are heteroscedastic, we refrain from using log OLS and turn to the Generalized Linear Model (GLM). As GLM, via a link function, yields un-scaled coefficients, retransformation is unnecessary. In our case, the long right-tailed distribution of DDD calls for a log-link function. Before using GLM, we used the Park test (see Manning and Mullahy 2001 for a description) which identifies the distribution of our dependent variable, DDD, as a member of the Gaussian family. Accordingly, we use GLM, with Gaussian distribution family and log-link function, to estimate the second step i.e., conditional on access, the accessed amount of DDDs. For comparability reasons only, we also estimate the unconditional OLS model (including non-users), the Tobit model and the Two-part model with the joint effect. STATA version 12.2 was used for all analyses.

4.5 Results

The main results are presented in tables 2 to 6. Tables 2 and 3 analyze overall pharmaceutical-based health investments and table 4 analyzes specific pharmaceutical-based health investments with the 20 most commonly dispensed pharmaceutical subgroups individually. Extensions of the models for overall pharmaceutical-based health investments are presented in tables 5 to 8.

4.5.1 Main results

Overall pharmaceutical-based health investments

Table 2 (for males) and table 3 (for females) examine overall pharmaceutical-based health investments in two parts: the likelihood of accessing pharmaceuticals with probit models (columns 1 to 4) and, conditional on access, the accessed number of DDDs with GLM (columns 5 to 8). The columns in table 2 and table 3, stepwise include control variables. Columns 1 and 5 control for immigrant; columns 2 and 6 control for immigrant region of origin i.e. Nordic, Western or non-Western; columns 3 and 7 add controls for health and columns 4 and 8 add controls for socioeconomic status.

Table 2 shows the results for males. Controlling for immigrant region of origin only, column 2 shows that non-Western male immigrants are 5% more likely to access pharmaceuticals than native males. This difference persists after controlling for health (column 3) and socioeconomic status (column 4). Column 6 shows that, conditional on access, Western male immigrants accessed fewer DDDs than native males.

Table 3 shows the results for females. Controlling for immigrant region of origin only, column 6 shows that, conditional on access, non-Western female immigrants accessed fewer DDDs than native females. This difference persists after controlling for health (column 7) and socioeconomic status (column 8).

Table 2. Marginal effects (M.E.)of the probability of accessing pharmaceuticals and conditional on access, accessed number of defined daily doses (DDD). Results for males.

	Probit on access											
	(1)			(2)			(3)			(4)		
	M.E.		**Std. Err.**	***M.E.***		**Std. Err.**	***M.E.***		**Std. Err.**	***M.E.***		**Std. Err.**
Native Swedish	*reference*			*reference*			*reference*			*reference*		
All immigrants	0.03		0.02									
Nordic origin				0.01		0.04	-0.01		0.04	-0.00		0.04
Western origin				-0.01		0.06	0.02		0.05	0.02		0.05
non-Western origin				0.05	**	0.03	0.06	**	0.02	0.07	***	0.02
Controls for health	*No*			*No*			*Yes*			*Yes*		
Controls for human capital	*No*			*No*			*No*			*Yes*		
No. of observations	4,078			4,078			4,078			4,078		

	GLM on DDD|access											
	(5)			(6)			(7)			(8)		
	M.E.		**Std. Err.**	***M.E.***		**Std. Err.**	***M.E.***		**Std. Err.**	***M.E.***		**Std. Err.**
Native Swedish	*reference*			*reference*			*reference*					
All immigrants	-87.71		175.67									
Nordic origin				520.31		324.56	82.32		263.3	190.20		279.92
Western origin				-915.82	***	237.81	-196.31		391.79	-191.29		411.57
non-Western origin				-171.26		222.28	188.82		290.23	115.78		254.1
Controls for health	*No*			*No*			*Yes*			*Yes*		
Controls for human capital	*No*			*No*			*No*			*Yes*		
No. of observations	3,114			3,114			3,114			3,114		

Note: Controls for health include: self-assessed health, BMI, frequency of exercise, smoking status, mobility impairment, self-care ability, medical diagnoses and age. Controls for human capital include: 2^{nd} generation immigrant, educational level, whether married or cohabitant, children in household, residency in a large city and work full time. Appendix C presents full results for the models in columns 4 and 8. Full results for the remaining models are available from the authors upon request. ***p-value<0.01, **p-value<0.05 and *p-value<0.1.

Table 3. Marginal effects (M.E.) of the probability of accessing pharmaceuticals and conditional on access, accessed number of defined daily doses (DDD). Results for females.

	Probit on access											
	(1)			(2)			(3)			(4)		
	M.E.		**Std. Err.**	***M.E.***		**Std. Err.**	***M.E.***		**Std. Err.**	***M.E.***		**Std. Err.**
Native Swedish	*reference*			*reference*			*reference*			*reference*		
All immigrants	-0.02		0.01									
Nordic origin				-0.03		0.03	-0.04		0.03	-0.04		0.03
Western origin				0.02		0.04	0.02		0.04	0.02		0.04
non-Western origin				-0.03		0.02	-0.03		0.02	-0.03		0.02
Controls for health	*No*			*No*			*Yes*			*Yes*		
Controls for human capital	*No*			*No*			*No*			*Yes*		
No. of observations	*4,410*			*4,410*			*4,410*			*4,410*		

	GLM on DDD|access											
	(5)			(6)			(7)			(8)		
	M.E.		**Std. Err.**	***M.E.***		**Std. Err.**	***M.E.***		**Std. Err.**	***M.E.***		**Std. Err.**
Native Swedish	*reference*			*reference*			*reference*			*reference*		
All immigrants	-248.28		188.75									
Nordic origin				-49.13		296.22	-83.41		354.29	-137.57		373.37
Western origin				291.98		413.05	-252.67		316.44	-224.58		328.2
non-Western origin				-482.24	**	220.89	-629.55	*	355.49	-797.07	**	323.09
Controls for health	*No*			*No*			*Yes*			*Yes*		
Controls for human capital	*No*			*No*			*No*			*Yes*		
No. of observations	3,877			3,877			3,877			3,877		

Note: Controls for health include: self-assessed health, BMI, frequency of exercise, smoking status, mobility impairment, self-care ability, medical diagnoses and age. Controls for human capital include: 2nd generation immigrant, educational level, whether married or cohabitant, children in household, residency in a large city and work full time. Appendix C presents full results for the models in columns 4 and 8. Full results for the other models are available from the authors upon request. ***p-value<0.01, **p-value<0.05 and *p-value<0.1.

Appendix A shows the full specification for the models analyzing overall pharmaceutical-based health investments (columns 4 and 8 in tables 2 and 3, respectively) and that: (1) low health and high age increase the likelihood of accessing pharmaceuticals and, conditional on access, accessed number of DDDs for males and females, (2) high disposable household income increases the likelihood of accessing pharmaceuticals for males (column 1) and, conditional on access, accessed number of DDDs for females (column 2), (3) full-time work and children in the household decrease the likelihood of accessing pharmaceuticals and, conditional on access, accessed number of DDDs for males and females, and (4) respondents with some, but no more than two years of higher education, are more likely to access pharmaceuticals than respondents with education up to primary school.

4.5.2 Pharmaceutical-based health investments with the 20 most dispensed pharmaceutical subgroups

Using a probit model with controls for immigrant region of origin, health and socioeconomic status, we estimate the likelihood of accessing the 20 most dispensed pharmaceutical subgroups (in the third ATC-level) separately for males and females. Table 4 shows the pharmaceutical subgroups where immigrants and natives have statistically different (on 10% level) probabilities of access (see Appendix B for full results).

For males, immigrants and natives have significantly different probabilities of accessing 6 of the 20 analyzed pharmaceutical subgroups: compared with native males, (1) male Nordic immigrants are less likely to access pharmaceuticals with ATC-codes A02B (for e.g. peptic ulcer) and D07A (corticosteroids) but more likely to access C07A (beta blockers for e.g. cardiovascular diseases) and A10B (anti-diabetics, excluding insulin), (2) male Western immigrants are less likely to access pharmaceuticals with ATC-codes C09A (ACEs for cardiovascular diseases), C07A (beta blockers for e.g. cardiovascular diseases) and C01D (vasodilators for cardiac diseases) and (3) male non-Western immigrants are more likely to access pharmaceuticals with ATC-codes A02B (for e.g. peptic ulcer), A10B (anti-diabetics, excluding insulin) and C01D (vasodilators for cardiac diseases).

Table 4 - Marginal effects on the probability to access the 20 most dispensed pharmaceuticals, compared with natives, for men and females respectively. Results are only shown when at least one of the regions of origin immigrant coefficient is statistically significant.

Males - Probit on access									
ATC		**M.E.**		**Std. Err.**	**ATC**		**M.E.**		**Std. Err.**
	Nordic origin	-0.01		-0.03		Nordic origin	0.04	*	-0.02
C09A	Western origin	-0.07	***	-0.03	*A10B*	Western origin	0.05		-0.04
	non-Western origin	0		-0.02		non-Western origin	0.03	**	-0.02
	Nordic origin	0.09	***	-0.03		Nordic origin	-0.06	**	-0.03
C07A	Western origin	-0.09	**	-0.04	*D07A*	Western origin	0.02		-0.05
	non-Western origin	-0.02		-0.02		non-Western origin	-0.02		-0.02
A02B	Nordic origin	-0.05	**	-0.03		Nordic origin	0.03		-0.02
	Western origin	-0.06		-0.04	*C01D*	Western origin	-0.04	*	-0.02
	non-Western origin	0.05	*	-0.03		non-Western origin	0.05	**	-0.02
Females - Probit on access									
ATC		**M.E.**		**Std. Err.**	**ATC**		**M.E.**		**Std. Err.**
	Nordic origin	-0.05	***	-0.02		Nordic origin	-0.01		-0.03
D02A	Western origin	0.01		-0.04	*N02B*	Western origin	0.01		-0.06
	non-Western origin	0.03		-0.02		non-Western origin	0.05	*	-0.03
	Nordic origin	-0.01		-0.02		Nordic origin	-0.01		-0.02
C09A	Western origin	-0.07	***	-0.02	*A06A*	Western origin	0.05		-0.05
	non-Western origin	-0.02		-0.02		non-Western origin	0.07	**	-0.03
	Nordic origin	-0.03		-0.03		Nordic origin	-0.02		-0.02
A02B	Western origin	0.07		-0.06	*R03A*	Western origin	-0.04		-0.02
	non-Western origin	0.07	**	-0.03		non-Western origin	-0.04	**	-0.02
	Nordic origin	-0.01		-0.03		Nordic origin	-0.02	*	-0.01
G03A	Western origin	0.05		-0.05	*C03A*	Western origin	-0.01		-0.03
	non-Western origin	-0.06	***	-0.02		non-Western origin	0		-0.02
	Nordic origin	0		-0.04					
M01A	Western origin	0		-0.07					
	non-Western origin	0.1	***	-0.03					

Note: Health variables includes: self-assessed health, BMI, frequency of exercise, smoking status, mobility impairment, self-care ability, medical diagnose, age and age squared. Human capital variables includes: 2nd generation immigrant, educational level, whether married or cohabitant, children in household, residency in a large city and full time work. Full results are available upon request. *** p<0.01, ** p<0.05, * p<0.1.

For females, immigrants and natives have significantly different probabilities of accessing 9 of the 20 analyzed pharmaceutical subgroups: compared with native females (1) female Nordic immigrants are less likely to access pharmaceuticals with ATC-codes D02A (for e.g., dry skin) and C03A (thiazide diuretics e.g., cardiovascular diseases), (2) female Western immigrants are less likely to access pharmaceuticals with the ATC-code C09A (ACEs for cardiovascular diseases), and (3) female non-Western immigrants are less likely to access pharmaceuticals with ATC-codes G03A (oral hormonal contraceptives) and R03A (adrenergic inhalants for obstructive airways) but more likely to access A02B (for e.g., peptic ulcer), M01A (for e.g. rheumatism), N02B (e.g. analgesics) and A06A (laxatives for constipation).

4.5.3 Extensions

Analyzing overall pharmaceutical-based health investments with models controlling for immigration status, health, and socioeconomic status, tables 5 to 8 present estimates from four different extensions. Table 5 controls for years of immigrants' Swedish residency among and table 6 omits oral hormonal contraceptives for women. Table 7 omits respondents with up to primary schooling and table 8 estimates the models with three alternative econometric approaches – the unconditional OLS (including non-users), Tobit and Two-part models.

Years of Swedish residency

Country specific-knowledge, such as the ability to speak the native language, affects the individual´s productivity. While residing in Sweden, immigrants acquire these country-specific skills, which influence the relative costs and benefits of pharmaceutical-based health investments. On a conceptual level, such skills make immigrants' costs and pay-offs from pharmaceutical-based health-investments converge towards natives´ costs and pay-offs, which in turn reduces the disparity in demand for pharmaceutical-based health investments between immigrants and natives. To empirically explore if years of Swedish residency influences the pharmaceutical-based health investments of immigrants, we group immigrants according to years since immigration to Sweden: 0 to 5, 6 to 10, 11 to 20 and more than 20 years, and estimate the models. Table 5 presents the resulting estimates from models controlling for years of Swedish residency, health and socioeconomic status. For males, column 1 shows that immigrants with 11 to 20 years of residency are 8% more likely to access pharmaceuticals than natives, and column 2 shows that immigrants with up to 5 years of residency, conditional on access, access 1,252 fewer DDDs than natives. Correspondingly for females, column 3 shows that immigrants with up to 5 years of residency are 19% less likely to access pharmaceuticals than natives, and column 3 shows that immigrants with 6 to 10 years of residency, conditional on access, access 1,115 fewer DDDs than natives, whereas immigrants with more than 20 years of residency, conditional on access, access 513 fewer DDDs than natives. The discrepancy in pharmaceutical-based health investments is larger between natives and immigrants with up to 10 years of Swedish residency, than for immigrants with more than 10 years of Swedish residency.

Excluding oral hormonal contraceptives for women

As the main therapeutic use for contraceptives is to prevent pregnancy and not to improve health per se, contraceptives differ from other pharmaceuticals. Moreover, differences in cultural norms may influence women's demand for contraceptives. Thus, we omit oral hormonal contraceptives and re-estimate the models to analyze if the discrepancy in pharmaceutical-based health investments between female immigrants and natives is due to oral hormonal contraceptives. Table 6 presents the resulting estimates from models controlling for immigration status, health and socioeconomic status. Column 2 shows that non-Western immigrants, conditional on access, access 788 fewer DDDs than natives. As the estimates from the models omitting contraceptives are qualitatively and quantitatively similar to the estimates from the models with contraceptives (table 3, columns 4 and 8), the discrepancy in pharmaceutical-based health investments between female immigrants and natives is not driven by oral hormonal contraceptives.

Table 5. Marginal effects (M.E.) of the probability of accessing pharmaceuticals and conditional on access, accessed number of defined daily doses (DDD). Results for males and females.

	Males						**Females**					
	(1)			(2)			(3)			(4)		
	Probit on access			**GLM on DDD|access**			**Probit on access**			**GLM on DDD|access**		
	M.E.		**Std. Err.**	**M.E.**		**Std. Err.**	**M.E.**		**Std. Err.**	**M.E.**		**Std. Err.**
Native Swedish	*reference*			*reference*			*reference*			*reference*		
In Sweden -5	0.03		0.05	-1,251.62	***	245.6	-0.19	***	0.06	2,415.43		1,838.60
In Sweden 6-10	0.03		0.06	218.75		456.32	-0.08		0.05	-1,114.96	***	385.87
In Sweden 11-20	0.08	***	0.03	-289.09		289.69	-0.01		0.03	-436.80		460.51
In Sweden 21-	0.01		0.03	208.62		209.99	-0.00		0.02	-512.64	**	213.72
Control for health	*Yes*			*Yes*			*Yes*			*Yes*		
Control for human capital	*Yes*			*Yes*			*Yes*			*Yes*		
No. of observations	4,078			3,114			4,410			3,877		

Note: Controls for health include: self-assessed health, BMI, frequency of exercise, smoking status, mobility impairment, self-care ability, medical diagnoses and age. Controls for human capital include: 2nd generation immigrant, educational level, whether married or cohabitant, children in household, residency in a large city and work full time. See appendix D for full results. ***p-value<0.01, **p-value<0.05 and *p-value<0.1.

Table 6. Marginal effects ((M.E.) of the probability of accessing pharmaceuticals and conditional on access, accessed number of defined daily doses (DDD) after exclusion of hormonal contraceptive (ATC: G03A). Results for females.

	(1)			(2)		
	Probit on access			**GLM on DDD|access**		
	M.E.		**Std. Err.**	***M.E.***		**Std. Err.**
Native Swedish	*reference*			*reference*		
Nordic origin	-0.04		0.03	-135.49		375.03
Western origin	0.04		0.04	-242.59		326.03
non-Western origin	-0.00		0.02	-787.73	**	329.11
Controls for health	*Yes*			*Yes*		
Controls for human capital	*Yes*			*Yes*		
No. of observations	4,410			3,774		

Note: Controls for health include: self-assessed health, BMI, frequency of exercise, smoking status, mobility impairment, self-care ability, medical diagnoses and age. Controls for human capital include: 2nd generation immigrant, educational level, whether married or cohabitant, children in household, residency in a large city and work full time. See appendix F for full results. ***p-value<0.01, **p-value<0.05 and *p-value<0.1.

Excluding respondents with no more than primary school

In 1936, 7 years of education became mandatory for children in Sweden. Thus, most natives have at least 7 years of education. Contrastingly, some immigrants may have substantially less education. To analyze if the observed patterns of pharmaceutical-based health investments of immigrants and natives is driven by respondents belonging to the lowest educational group, we omit respondents with up to primary school and re-estimate the models. Table 7 presents the resulting estimated marginal effects for models controlling for immigration status, health and socioeconomic status. For males, column 1 shows that non-Western immigrants are 7% more likely to access pharmaceuticals than natives (column 1). For comparison, no such effect is found in the model including primary school (table 2, column 4).

Table 7. Marginal effects (M.E.) of the probability of accessing pharmaceuticals and conditional on access, accessed number of defined daily doses (DDD). Results for males and females when individuals with up to primary education are excluded.

	Males					Females			
	(1)			(2)		(3)		(4)	
	Probit on access			**GLM on DDD|access**		**Probit on access**		**GLM on DDD|access**	
	M.E.		**Std. Err.**	***M.E.***	**Std. Err.**	***M.E.***	**Std. Err.**	***M.E.***	**Std. Err.**
Native Swedish	*reference*			*reference*		*reference*		*reference*	
Nordic origin	-0.03		0.05	545.47	347.69	-0.03	0.03	453.58	552.47
Western origin	0.02		0.05	-20.41	409.44	0.00	0.04	69.21	431.69
non-Western origin	0.07	***	0.03	204.10	242.42	-0.03	0.03	-441.70	445.33
Controls for health	*Yes*			*Yes*		*Yes*		*Yes*	
Controls for human capital	*Yes*			*Yes*		*Yes*		*Yes*	
No. of observations	3,245			2,442		3,457		3,053	

Note: Controls for health include: self-assessed health, BMI, frequency of exercise, smoking status, mobility impairment, self-care ability, medical diagnoses and age. Controls for human capital include: 2nd generation immigrant, educational level, whether married or cohabitant, children in household, residency in a large city and work full time. See appendix G for full results.***p-value<0.01, **p-value<0.05 and *p-value<0.1.

For females, although not statistically significant, column 4 shows that non-Western immigrants, conditional on access, access fewer DDDs than natives. For comparison, the equivalent marginal effect of the model including primary school is quantitatively larger and statistically significant (table 3, column 8). Therefore, the differences in pharmaceutical-based health investments between female immigrants and natives may partly arise from immigrant females belonging to the lowest educational group.

Alternative econometric approaches

For comparability (see discussion in the empirical specification), we estimate overall pharmaceutical-based health investments with three additional econometric models: (1) unconditional OLS (including non-users), (2) Tobit, and (3) the joint two-part. The results are shown in table 8. The estimates for males (columns 1 to 3) reveal no difference in pharmaceutical-based health investments between natives and immigrants with Nordic, Western or non-Western origin. In contrast, the estimates for females (columns 4 to 6) show – independent of the econometric model – that non-Western female immigrants have lower pharmaceutical utilization than native females.

Table 8. The combined marginal effect (M.E.) of access and conditional on access, accessed number of defined daily doses (DDD) estimated with OLS (columns 1 and 4), Tobit (columns 2 and 5) and the Two-part model (columns 3 and 6). Results for males and females.

	Males								
	(1) OLS DDD			(2) Tobit DDD			(3) Two-part model DDD		
	M.E.		**Std. Err.**	**M.E.**		**Std. Err.**	**M.E.**		**Std. Err.**
Native Swedish	*reference*			*reference*			*reference*		
Nordic origin	268.24		234.75	261.66		286.63	147.77		46383.65
Western origin	-66.19		154.16	60.13		230.71	-134.61		539.71
non-Western origin	-57.29		143.88	119.64		171.43	184.06		2652.26
Controls for health	*Yes*			*Yes*			*Yes*		
Controls for human capital	*Yes*			*Yes*			*Yes*		
No. of observations	4,078			4,078			3,114		

	Females								
	(4) OLS DDD			(5) Tobit DDD			(6) Two-part model DDD		
	M.E.		**Std. Err.**	**M.E.**		**Std. Err.**	**M.E.**		**Std. Err.**
Native Swedish	*reference*			*reference*			*reference*		
Nordic origin	-306.96		230.17	-390.54		255.41	-251.35		219.50
Western origin	7.54		341.37	-70.48		387.98	-139.23		363.10
non-Western origin	-479.35	***	173.99	-627.25	***	200.32	-531.58	***	175.23
Controls for health	*Yes*			*Yes*			*Yes*		
Controls for human capital	*Yes*			*Yes*			*Yes*		
No. of observations	4,410			4,410			3,877		

4.6 Discussion

Compared with previous studies, our results regarding the effect of age and education on health-related investments are in line with findings in previous studies (Merlo et al. 2003; Bolin, Lindgren, and Rössner 2006; Nordin, Dackehag, and Gerdtham 2013). Our results show that immigrants and natives differ in (1) access to prescribed pharmaceuticals and (2) accessed number of DDDs, conditional on access. For females, the disparity between natives and non-Western immigrants is substantial, conditional on access – female non-Western immigrants access about 800 fewer DDDs. For comparison, conditional on access, the average female in our sample accesses 1,866 DDDs.

Pharmaceutical access differs between immigrants and natives across several of the 20 most dispensed pharmaceutical subgroups (on the 3rd ATC-level). The directions of the disparity in pharmaceutical access between natives and immigrants from different regions of origin are mixed – natives have higher access to some pharmaceuticals whereas immigrants have higher access to others. Focusing on first-line pharmaceuticals listed by evidence based treatment guidelines (e.g. Janus 2006; Läkemedelsrådet 2006), a uniform pattern emerges – immigrants are less likely than natives to access the first-line pharmaceuticals: thiazide-diuretics, ACE inhibitors and adrenergic inhalators. As these first-line pharmaceuticals make up the medical foundation in the treatment of several cardiovascular-related diseases (high blood pressure, heart failure and kidney diseases) and asthmatic diseases, immigrants may be less likely than natives to have adequate pharmaceutical treatment. As cardiovascular-related morbidity and mortality are leading public health concerns in Sweden, and pharmaceuticals substantially reduce such related morbidity and mortality (WHO 2012), disparities in access to these pharmaceuticals between immigrants and natives may have significant public health consequences. Furthermore, female non-Western immigrants are less likely to access oral hormonal contraceptives than native women. This finding may be related to that immigrants from non-Western countries are overrepresented among women with induced abortion in Sweden (Helström et al. 2003)

The empirical results regarding pharmaceutical-based health investments are consistent with the predictions obtained within the demand-for-health framework. In brief, low health and high age are both positively related to pharmaceutical-based health investments, whereas monetary and time constraints (measured as full-time work and children in the household) are negatively related. During residency in a host country, immigrants may acquire country-specific knowledge, such as the ability to master the native language. Such knowledge can be regarded as educational capital, which makes pharmaceutical-based health investments less costly. In line with theory, the years of Swedish residency of immigrants are positively related to pharmaceutical-based health investments. Compared with immigrants with fewer years of Swedish residency, immigrants with more years of residency are more similar to natives with regard to pharmaceutical-based health investments.

Limitations

The nature of our data creates two caveats that should be mentioned. First, cultural differences, when perceived self-assessed health is used as a proxy for true health, may yield biased estimates. To reduce such bias, we complement self-assessed health with more objective health measures (e.g. mobility impairment and the occurrence of cardiovascular disease). Second, sample selection may arise due to systematic differences in characteristics between respondents and non-respondents. Our sample comprises 11% immigrants, whereas the corresponding figure was 14% in 2004 (Statistics Sweden 2013). A study of non-respondents in the ULF survey for the year 2000 (Statistics Sweden 2003) shows that the majority of non-respondent immigrants are outside the labor force. Given that immigrants are selected into the labor force on, for instance, the basis of country-specific skills such as mastering the native language, immigrants inside and outside the labor force may differ in terms of such skills. As educational capital, including country-specific skills, is positively related to pharmaceutical-based health investments, we may underestimate the discrepancy in pharmaceutical-based health investments between immigrants and natives.

Policy implications and future studies

Our results indicate that the Swedish health policy goal – that all (legal) residents should have equal access to medical care according to need – is unachieved. When creating policies for tackling inequalities in the utilization of prescribed pharmaceuticals, policymakers can either address general socioeconomic inequalities or specific vulnerable groups (e.g. female immigrants with low educational level).

Differences in pharmaceutical utilization between population groups can reinforce present and future health inequalities. To ascertain the consequences of differences in pharmaceutical utilization between immigrants and natives for public health, more research is needed. More specifically, given that physicians are gatekeepers to prescribed pharmaceuticals; future research should address physicians' prescribing patterns regarding immigrants and natives. Ultimately, the individual alone decides his or her degree of adherence to the treatment. Therefore, future research should also address immigrants' and natives' adherence to prescribed pharmaceutical regimens.

References

Ai, Chunrong, and Edward C. Norton. 2000. “Standard Errors for the Retransformation Problem with Heteroscedasticity.” *Journal of Health Economics* 19 (5) (September): 697–718.

Bayard-Burfield, L., J. Sundquist, and S.-E. Johansson. 2001. “Ethnicity, Self Reported Psychiatric Illness, and Intake of Psychotropic Drugs in Five Ethnic Groups in Sweden.” *Journal of Epidemiology and Community Health* 55 (9) (September 1): 657–664.

Becker, Gary S. 1964. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. Chicago, University of Chicago Press.

Bolin, K, B Lindgren, and S Rössner. 2006. “The Significance of Overweight and Obesity for Individual Health Behaviour: An Economic Analysis Based on the Swedish Surveys of Living Conditions 1980-81, 1988-89, and 1996-97.” *Scandinavian Journal of Public Health* 34 (4) (August): 422–431. c8h.

Cameron, Adrian Colin, and Pravin K. Trivedi. 2005. *Microeconometrics : Methods and Applications*. New York, N.Y. : Cambridge University Press, 2005. http://library.lu.se/cgi-bin/ipchk/http://search.ebscohost.com/login.aspx?direct=true&db=cat01310a&AN=lovisa.001730410&site=eds-live.

Carlsson, Magnus, and Dan-Olof Rooth. 2007. “Evidence of Ethnic Discrimination in the Swedish Labor Market Using Experimental Data.” *Labour Economics* 14: 716–729.

Denktaş, Semiha, Gerrit Koopmans, Erwin Birnie, Marleen Foets, and Gouke Bonsel. 2010. “Underutilization of Prescribed Drugs Use Among First Generation Elderly Immigrants in the Netherlands.” *BMC Health Services Research* 10 (1) (June 22): 176.

Duan, Naihua. 1983. “Smearing Estimate: A Nonparametric Retransformation Method.” *Journal of the American Statistical Association* 78 (383) (September 1): 605–610.

Ekberg, J. 2011. “Will Future Immigration to Sweden Make It Easier to Finance the Welfare System?” *European Journal of Population* 27 (1): 103–124.

Feng, Jin, Bei Qin, and Xiabo Yu. 2010. “Wealth, Education and Demand for Medical Care: Evidence from Rural China.” In *Investing In Human Capital For Economic Development In China*, edited by Gordon G. Liu, Shufang Zhang, and Zongyi Zhang. World Scientific.

Ferrada-Noli, M. 1997. “A Cross-cultural Breakdown of Swedish Suicide.” *Acta Psychiatrica Scandinavica* 96 (2) (August): 108–116.

Gadd, Malin, Jan Sundquist, Sven-Erik Johansson, and Per Wändell. 2005. "Do Immigrants Have an Increased Prevalence of Unhealthy Behaviours and Risk Factors for Coronary Heart Disease?" *European Journal of Cardiovascular Prevention and Rehabilitation* 12 (6) (December 1): 535–541.

Grossman, Michael. 1972. "On the Concept of Health Capital and the Demand for Health." *Journal of Political Economy* 80 (2) (March 1): 223–255.

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47 (1) (January): 153–161. bth.

Helström, Lotti, Viveca Odlind, Catharina Zätterström, Monica Johansson, Fredrik Granath, Nestor Correia, and Anders Ekbom. 2003. 'Abortion Rate and Contraceptive Practices in Immigrant and Native Women in Sweden'. *Scandinavian Journal of Public Health* 31 (6) (January 12): 405–410.

Janus. 2006. "Läkemedelsnytt [In Swedish]". Karolinska University Hospital. http://www.janusinfo.se/Global/Tidskrifter/L%C3%A4kemedelsnytt/lmnytt1_06.pdf.

Johansson, Leena Maria, Jan Sundquist, Sven-Erik Johansson, B.O. Bergman, Jan Qvist, and Lil Träskman-Bendz. 1997. "Suicide Among Foreign-born Minorities and Native Swedes: An Epidemiological Follow-up Study of a Defined Population." *Social Science & Medicine* 44 (2) (January): 181–187.

Krigsman, Kristin, Arne Melander, Anders Carlsten, Anders Ekedahl, and J.LarsG. Nilsson. 2007. "Refill Non-adherence to Repeat Prescriptions Leads to Treatment Gaps or to High Extra Costs." *Pharmacy World & Science* 29 (1) (February 1): 19–24.

Lindström, M, J Sundquist, and P-O Östergren. 2001. "Ethnic Differences in Self-Reported Health in Malmö in Southern Sweden." *Journal of Epidemiology and Community Health* 55 (2) (February 1): 97–103.

Läkemedelsrådet. 2006. "Skånelistan 2006 [in Swedish]". Skåne Regional Council. http://www.skane.se/sv/Webbplatser/Lakemedelsradet/Rekommendationer_riktlinjer/Skanelistan/.

Manning, Willard G, and John Mullahy. 2001. "Estimating Log Models: To Transform or Not to Transform?" *Journal of Health Economics* 20: 461–494. edselp.

Marmot, M. G., Abraham Manie Adelstein, L. Bulusu, and Great Britain Office of Population Censuses and Surveys. 1984. *Immigrant Mortality in England and Wales, 1970-78: Causes of Death by Country of Birth*. H.M.S.O.

Merlo, Juan, John W. Lynch, Min Yang, Martin Lindström, Per Olof Östergren, Niels Kristian Rasmusen, and Lennart Råstam. 2003. "Effect of Neighborhood Social Participation on Individual Use of Hormone Replacement Therapy and Antihypertensive Medication: A Multilevel Analysis." *American Journal of Epidemiology* 157 (9) (May 1): 774–783.

Morgan, Steven G., Gillian Hanley, Colleen Cunningham, and Hude Quan. 2011. "Ethnic Differences in the Use of Prescription Drugs: a Cross-sectional Analysis of Linked Survey and Administrative Data." *Open Medicine* 5 (2) (May 16): 87–93.

Nordin, Martin, Margareta Dackehag, and Ulf-G. Gerdtham. 2013. "Socioeconomic Inequalities in Drug Utilization for Sweden: Evidence from Linked Survey and Register Data." *Social Science & Medicine*. edselp. http://library.lu.se/cgi-bin/ipchk/http://search.ebscohost.com/login.aspx?direct=true&db=edselp&AN=S0277953612007629&site=eds-live.

Priebe, Stefan, Sima Sandhu, Sónia Dias, Andrea Gaddini, Tim Greacen, Elisabeth Ioannidis, Ulrike Kluge, et al. 2011. "Good Practice in Health Care for Migrants: Views and Experiences of Care Professionals in 16 European Countries." *BMC Public Health* 11 (1) (March 25): 187.

Rue, Montserrat, Maria-Catalina Serna, Jorge Soler-Gonzalez, Anna Bosch, Maria-Cristina Ruiz-Magaz, and Leonardo Galvan. 2008. "Differences in Pharmaceutical Consumption and Expenses Between Immigrant and Spanish-born Populations in Lleida, (Spain): a 6-months Prospective Observational Study." *BMC Health Services Research* 8: 35. doi:10.1186/1472-6963-8-35.

Statistics Sweden. 2003. "2003:8 Studie Av Bortfallet i 2000-års Undersökning Av Levnadsförhållanden (ULF) [in Swedish]." http://www.scb.se/statistik/BE/OV9999/2002I03/BE96ST0308.pdf.

———. 2011. "Den utrikes födda befolkningen ökar [in Swedish]." http://www.scb.se/Statistik/BE/BE0101/2011A01C/Den%20utrikes%20f%C3%B6dda%20befolkningen%20%C3%B6kar.pdf

———. 2012. "Beskrivning Av Statistiken - Statistiska Centralbyrån [in Swedish]." http://www.scb.se/Pages/ProductDocumentations____19231.aspx.

———. 2013. "Befolkningsstatistik [in Swedish]." http://www.scb.se/Pages/SSD/SSD_TablePresentation____340486.aspx?rxid=6b565665-f0f6-4050-abe7-fbcd5f84d8a2.

Sundquist, Jan. 1993. "Ethnicity as a Risk Factor for Consultations in Primary Health Care and Out-patient Care." *Scandinavian Journal of Primary Health Care* 11 (3) (January 1): 169–173. doi:10.3109/02813439308994825.

Sundquist, Jan, and Sven-Erik Johansson. 1997. "Long-term Illness Among Indigenous and Foreign-born People in Sweden." *Social Science & Medicine* 44 (2) (January): 189–198.

The Swedish National Board of Health and Welfare. 2008. *Läkemedel - Statistik För År 2007*. http://www.socialstyrelsen.se/publikationer2008/2008-46-1.

TLV. 2012. "TLV". Text. http://www.tlv.se/in-english-old/medicines-new/the-swedish-high-cost-threshold/how-it-works/.

Wamala, S. (1, J. (3) Merlo, G. (4) Boström, and C. (4) Hogstedt. 2007. "Perceived Discrimination, Socioeconomic Disadvantage and Refraining from Seeking Medical Treatment in Sweden." *Journal of Epidemiology and Community Health* 61 (5): 409–415. edselc.

Westin, Marcus, Annika Åhs, Kristina Bränd Persson, and Ragnar Westerling. 2004. "A Large Proportion of Swedish Citizens Refrain from Seeking Medical Care—lack of Confidence in the Medical Services a Plausible Explanation?" *Health Policy* 68 (3) (June): 333–344. doi:10.1016/j.healthpol.2003.10.008.

Wettermark, Björn, Niklas Hammar, C. MichaelFored, Andrejs Leimanis, Petra Otterblad Olausson, Ulf Bergman, Ingemar Persson, Anders Sundström, Barbro Westerholm, and Måns Rosén. 2007. "The New Swedish Prescribed Drug Register—Opportunities for Pharmacoepidemiological Research and Experience from the First Six Months." *Pharmacoepidemiology and Drug Safety* 16 (7): 726–735. doi:10.1002/pds.1294.

White, Halbert. 1980. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica* 48 (4) (May): 817–838. bth.

WHO. 2012. "WHO | The Top 10 Causes of Death." *WHO*. http://www.who.int/mediacentre/factsheets/fs310/en/index.html.

Appendix A – Variables

The variables in the analyses are specified as follows:

SAH 1, SAH 2, and SAH 3 are mutually exclusive dummy variables that take the value 1 if the respondents' self-assessed health is (1) very good, (2) good, and (3) low, respectively. In all cases, the variables are 0 otherwise.

Underweight, Normal weighing and *overweight* are mutually exclusive dummy variables that takes the value 1 if the respondent is (1) underweight (BMI<18.5), (2) normal weighing (18.5≤BMI<25), and (3) overweight (BMI≥25), respectively. In all cases, the variables are 0 otherwise.

Exercise 1, Exercise 2, Exercise 3, and *Exercise 4* are mutually exclusive dummy variables and takes the value 1 if the respondent reports(1) never exercise, (2) exercise, but less than once a week, (3) exercise once a week, and (4) exercise more than once a week, respectively. In all cases, the variables are 0 otherwise.

Smoker is a dummy variable which takes the value 1 if the respondent reports that he or she smokes daily and 0 otherwise.

Mobile impairment is a dummy variable which takes the value 1 if the respondent reports having a mobility impairment and 0 otherwise.

Pain is a dummy variable which takes the value 1 if the respondent reportsexperienced pain (e.g. back pain), and 0 otherwise.

Low self-care ability is a dummy variable which takes the value 1 if the respondent reports low self-care ability and 0 otherwise.

In addition, the respondents were asked questions about specific diseases. The ULF survey categorizes the occurrence of a disease according to the WHO's ICD-9 classification.

Mental is a dummy variable which takes the value 1 if the respondent reports mental illness corresponding to at least one diagnosis in the interval 290.0-316.9 in WHO´s ICD-9, and 0 otherwise.

Neurology is a dummy variable which takes the value 1 if the respondent reports neurological disease corresponding to at least one diagnosis in the interval 320.0-389.9 in WHO´s ICD-9, and 0 otherwise.

Cardiovasc is a dummy variable which takes the value 1 if the respondent reports cardiovascular disease corresponding to at least one diagnosis in the intervals 390.0-405.9 or 410.0-429.9 in WHO´s ICD-9, and 0 otherwise.

Respiratory is a dummy variable which takes the value 1 if the respondent reports respiratory disease corresponding to at least one diagnosis in the interval 460.0-519.9 according to WHO´s ICD-9, and 0 otherwise.

Skeleton is a dummy variable which takes the value 1 if the respondent reports skeletal disease corresponding to at least one diagnosis in the interval 710.0-739.9 in WHO´s ICD-9, and 0 otherwise.

Other_disease_1 is a dummy variable which takes the value 1 if the respondent reports diseases corresponding to one of the following categories: (1) diabetes, infections, tumors, (2) diseases of the eye, ear, skin, joints, (3) diseases in the blood, digestive, congenital, endocrine and urogenital systems, or (4) morbidity from external causes, and 0 otherwise.

Other_disease_2+ is a dummy variable which takes the value 1 if the respondent reports diseases corresponding to two or more of the following categories: (1) diabetes, infections, tumors, (2) diseases of the eye, ear, skin, joints, (3) diseases in the blood, digestive, congenital, endocrine and urogenital systems, or (4) morbidity from external causes, and 0 otherwise.

Age is the respondent's age in year 2006 and *Age2* is *Age* squared. *Age2* controls for potential nonlinear relationships between age and the dependent variables.

2^{nd} *gen immigrant.* The dummy variable takes the value 1 if the respondent was born in Sweden, but neither of the respondent´s parents was born in Sweden, and 0 otherwise.

Primary school, Secondary school, higher education≤2 and *higher education>2* are mutually exclusive dummy variables that take the value 1 if the respondent's education level is (1) up to primary school, (2) up to secondary school, (3) some, but less than two years, higher education, and (4) at least two years of higher education, respectively. In all cases, the variables are 0 otherwise.

Married_cohab is a dummy variable which takes the value 1 if the respondent is married or cohabiting, and 0 otherwise.

ln(dispinc) is the respondent´s logged disposable income after taxes and social transfers. For married or cohabiting respondents, *ln(dispinc)* is the mean of the household's disposable income after taxes and social transfers. Due to data limitations, we do not include hourly wage in the analysis. The non-response rate for hourly wage is roughly 40% and a closer analysis of responders and non-responders shows systematical differences of hourly wage between the responders and non-responders.

Child_1 and Child_2+ are mutually exclusive dummy variables. Child_1 takes the value 1 if the respondent has one child in the household, and Child_2+ takes the value 1 if the respondent has two or more children in the household. In all cases, the variables are 0 otherwise.

Large_city is a dummy variable which takes the value 1 if the respondent lives in Stockholm, Gothenburg or Malmo, and 0 otherwise.

Work_full_time is a dummy variable which takes the value 1 if the respondent reports working 40 hours a week or more, and 0 otherwise.

Appendix B. The 20 most dispensed pharmaceuticals (third level ATC) in Sweden 2006 for males and females respectively.

Males		Females	
ATC	pharmaceutical subgroups	ATC	pharmaceutical subgroups
B03B	Vitamin B12 and folic acid	B03B	Vitamin B12 and folic acid
C10A	Lipid modifying agents	D02A	Emollients and protectives
D02A	Emollients and protectives	G03A	Hormonal contraceptives for systemic use
B01A	Antithrombotic agents	N06A	Antidepressants
C09A	ACE inhibitors	C10A	Lipid modifying agents
C07A	Beta blocking agents	B01A	Antithrombotic agents
N06A	Antidepressants	N05C	Hypnotics and sedatives
C08C	Selective calcium channel blockers with mainly vascular effects	C07A	Beta blocking agents
C03C	High-ceiling diuretics	C03C	High-ceiling diuretics
N05C	Hypnotics and sedatives	C09A	ACE inhibitors, plain
A02B	Agents for peptic ulcer and gastro-oesophageal reflux	M01A	Non-steroid anti-inflammatory and antirheumatic agents
C09C	Angiotensin II antagonists	A02B	Agents for peptic ulcer and gastro-oesophageal reflux
A10A	Insulins and analogues	C08C	Selective calcium channel blockers with mainly vascular effects
M01A	Non-steroid anti-inflammatory and antirheumatic agents	H03A	Thyroid preparations
A10B	Blood glucose lowering agents, excluding insulin	N02B	Analgesics and antipyretics
D07A	Corticosteroids	A06A	Laxatives
R03A	Adrenergic inhalants	R03A	Adrenergic inhalants
C01D	Vasodilators used in cardiac diseases	C09C	Angiotensin II antagonists
A06A	Laxatives	R06A	Antihistamines for systemic use
G04C	Drugs used in benign prostatic hypertrophy	C03A	Low-ceiling diuretics, thiazides

Source: National Board of Health and Welfare, 2007

Appendix C. Marginal effects (M.E.) of the probability of accessing pharmaceuticals and conditional on access, accessed number of defined daily doses (DDD). Results for males and females.

	Males						Females					
	(1)			(2)			(3)			(4)		
	Probit on access			GLM on DDD|access			Probit on access			GLM on DDD|access		
	M.E.		**Std. Err.**	**M.E.**		**Std. Err.**	**M.E.**		**Std. Err.**	**M.E.**		**Std. Err.**
Native Swedish		*reference*			*reference*			*reference*			*reference*	
Nordic origin	-0.00		0.04	190.20		279.92	-0.04		0.03	-137.57		373.37
Western origin	0.02		0.05	-191.29		411.57	0.02		0.04	-224.58		328.2
non-Western origin	0.07	***	0.02	115.78		254.1	-0.03		0.02	-797.07	**	323.09
SAH 1		*reference*			*reference*			*reference*			*reference*	
SAH 2	0.06	***	0.02	351.34	**	139.29	0.03	***	0.01	184.89		146.48
SAH 3	0.08	***	0.02	633.14	***	192.57	0.07	***	0.01	792.83	***	194.52
Underweight	-0.05		0.06	1,083.82		751.83	-0.04	*	0.02	35.72		303.69
Normal weight		*reference*			*reference*			*reference*			*reference*	
Overweight	0.03	**	0.01	318.43	**	129.8	0.01		0.01	68.90		180.88
Exercise 1		*reference*			*reference*			*reference*			*reference*	
Exercise 2	0.01		0.02	29.42		177.49	0.02		0.02	-383.68		360.51
Exercise 3	0.04		0.03	-214.15		209.98	0.01		0.02	-189.04		438.44
Exercise 4	0.05	*	0.02	-22.44		186.55	0.02		0.02	-403.56		381.1
Smoker	0.01		0.02	-62.31		174.51	-0.00		0.01	15.03		359.29
Mobility impairment	-0.01		0.03	459.96	***	152.55	-0.02		0.01	710.40	***	178.26
Self_care	0.19	***	0.05	-334.26		323.17	0.09	***	0.02	-661.52	**	289.8
Phsych_illness	0.06		0.04	820.68	***	304.05	0.04	*	0.02	92.53		334.05
Neurology	0.07	*	0.03	-392.14	*	209.84	0.01		0.02	53.76		285.02
Cardiovasc	0.21	***	0.02	1,269.72	***	156.96	0.10	***	0.01	1,099.98	***	173.6
Respiratory	0.08	***	0.03	483.94	**	232.81	0.06	**	0.02	627.90	**	277.78
Skeleton	0.06	**	0.02	-151.74		139.33	0.02		0.01	159.96		169.56
Pain	0.02		0.01	-56.81		142.36	0.02	**	0.01	226.31		161.39
Other_disease_1	0.10	***	0.02	340.72	**	159.06	0.06	***	0.01	184.45		159.2
Other disease_2+	0.12	***	0.03	1,035.47	***	233.67	0.06	***	0.02	698.06	**	338.64
ageULF	0.00	***	0	25.70	***	7	-0.00	***	0	5.20		9.88
2nd gen immigrant	0.12	**	0.05	-1,259.78	**	511.22	-0.03		0.03	214.27		1,130.07
Primary education		*reference*			*reference*			*reference*			*reference*	

Secondary education	0.02		0.02	155.17		149.03	0.02	*	0.01	33.48		172.95
Higher education ≤2	0.05	**	0.02	257.44		248.53	0.02		0.02	454.20		671.77
Higher education >2	0.03		0.02	448.72		282.35	-0.01		0.02	406.65		349.37
lndispinc	0.04	***	0.01	221.71		177.22	0.02	***	0.01	-212.45	*	114.46
Married_cohab	0.05	***	0.02	-146.59		134.24	0.02	**	0.01	-125.19		186.97
Child_1	0.01		0.02	-57.94		338.35	0.01		0.01	-318.95		305.25
Child_2+	-0.05	***	0.02	-174.45		375.03	-0.05	***	0.02	-554.72	**	252.27
Large_city	0.00		0.01	-81.16		146.18	0.04	***	0.01	161.40		180.93
Work_full_time	-0.02		0.02	-463.84	***	154.81	-0.01		0.01	-272.53		216.78
No. of observations	4,078			3,114			4,410			3877		

Note: Conditional on access, the average accessed number of DDDs was 2,122 DDDs for females and 1,859 DDDs for males. ***p-value<0.01, **p-value<0.05 and *p-value<0.1.

Appendix D. Marginal effects (M.E.) of the probability of accessing pharmaceuticals and conditional on access, accessed number of defined daily doses (DDD). The models control for years of Swedish residency among immigrants. Results for males and females.

	Males						Females					
	(1)			(2)			(3)			(4)		
	Probit on access			**GLM on DDD|access**			**Probit on access**			**GLM on DDD|access**		
	M.E.		**Std. Err.**	**M.E.**		**Std. Err.**	**M.E.**		**Std. Err.**	**M.E.**		**Std. Err.**
Native Swedish	*reference*			*reference*			*reference*			*reference*		
In Sweden -5	0.03		0.05	-1,251.62	***	245.6	-0.19	***	0.06	2,415.43		1,838.60
In Sweden 6-10	0.03		0.06	218.75		456.32	-0.08		0.05	-1,114.96	***	385.87
In Sweden 11-20	0.08	***	0.03	-289.09		289.69	-0.01		0.03	-436.80		460.51
In Sweden 21-	0.01		0.03	208.62		209.99	-0.00		0.02	-512.64	**	213.72
SAH 1	*reference*			*reference*			*reference*			*reference*		
SAH 2	0.06	***	0.02	346.74	**	140.12	0.03	***	0.01	150.45		145.06
SAH 3	0.08	***	0.02	625.46	***	194.17	0.07	***	0.01	769.36	***	197.08
Underweight	-0.05		0.06	1,036.17		751.24	-0.04	*	0.02	-41.27		317.29
Normal weight	*reference*			*reference*			*reference*			*reference*		
Overweight	0.03	**	0.01	319.76	**	130.18	0.00		0.01	105.04		186.84
Exercise 1	*reference*			*reference*			*reference*			*reference*		
Exercise 2	0.01		0.02	39.29		176.13	0.02		0.02	-360.74		374.46
Exercise 3	0.04		0.03	-216.21		209.5	0.01		0.02	-110.51		444.77
Exercise 4	0.04	*	0.02	-19.76		186.13	0.02		0.02	-372.87		389.01
Smoker	0.01		0.02	-50.06		176.83	-0.01		0.01	47.78		359.38
Mobility impairment	-0.01		0.03	464.77	***	152.76	-0.02		0.01	703.49	***	172.67
Self_care	0.19	***	0.05	-320.86		330.11	0.09	***	0.02	-619.17	**	277.66
Phsych_illness	0.06		0.04	885.96	***	315.76	0.05	*	0.02	78.69		318.87
Neurology	0.07	*	0.03	-401.67	*	209.7	0.01		0.02	50.11		290.66

Cardiovasc	0.21	***	0.02	1,250.40	***	160.36	0.10	***	0.01	1,089.94	***	172.96
Respiratory	0.07	***	0.03	489.13	**	236.85	0.05	**	0.02	637.24	**	273.72
Skeleton	0.06	**	0.02	-158.80		139.89	0.01		0.01	158.34		179.15
Pain	0.02		0.01	-47.93		143.27	0.03	**	0.01	202.11		161.31
Other_disease_1	0.10	***	0.02	347.47	**	158.58	0.06	***	0.01	211.62		162.26
Other disease_2+	0.12	***	0.03	1,058.69	***	238.56	0.06	***	0.02	763.66	**	349.14
ageULF	0.00	***	0	24.67	***	7.02	-0.00	***	0	9.22		10.38
2nd gen immigrant	0.12	**	0.05	-1,261.20	**	498.8	-0.03		0.03	208.50		1,145.57
Primary education		*reference*			*reference*			*reference*			*reference*	
Secondary education	0.02		0.02	149.92		150.87	0.02		0.01	33.90		170.84
Higher education ≤2	0.05	**	0.02	258.67		247.15	0.02		0.02	437.96		678.86
Higher education >2	0.03		0.02	442.61		280.89	-0.01		0.02	303.17		343.75
lndispinc	0.04	***	0.01	204.46		175.38	0.02	**	0.01	29.51		121.01
Married_cohab	0.05	***	0.02	-138.89		132.73	0.02	**	0.01	-108.62		198.96
Child_1	0.01		0.02	-72.27		345.05	0.01		0.01	-395.40		311.61
Child_2+	-0.05	***	0.02	-154.34		389.46	-0.05	***	0.02	-693.96	**	293.25
Large_city	0.01		0.01	-79.56		147.1	0.04	***	0.01	67.49		194.24
Work_full_time	-0.02		0.02	-472.42	***	155.45	-0.01		0.01	-266.69		205.29
No. of observations	4,078			3,114			4,410			3,877		

Note: Years of Swedish residency are dummy variables. For instance, "In Sweden -5" takes on the value 1 for immigrants with up to 5 years of Swedish residency and 0 otherwise. ***p-value<0.01, **p-value<0.05 and *p-value<0.1.

Appendix F. Marginal effects (M.E) of the probability to access pharmaceuticals and conditional on access, accessed number of defined daily doses (DDD) after exclusion of hormonal contraceptive (ATC: G03A). Results for females.

	(1)			(2)		
	Probit on access			**GLM on DDD|access**		
	M.E.		**Std. Err.**	***M.E.***		**Std. Err.**
Native Swedish		*reference*			*reference*	
Nordic origin	-0.04		0.03	-135.49		375.03
Western origin	0.04		0.04	-242.59		326.03
non-Western origin	-0.00		0.02	-787.73	**	329.11
SAH 1		*reference*			*reference*	
SAH 3	0.09	***	0.02	800.18	***	198.11
Underweight	-0.05	*	0.03	31.76		305.56
Normal weight		reference			reference	
Overweight	0.00		0.01	62.94		185.42
Exercise 1		*reference*			*reference*	
Exercise 2	0.05	**	0.02	-381.22		366.76
Exercise 3	0.04		0.02	-188.58		443.91
Exercise 4	0.04	**	0.02	-400.99		387.3
Smoker	0.01		0.01	28.53		362
Mobility impairment	-0.02		0.02	707.67	***	180.64
Self_care	0.10	***	0.03	-670.80	**	290.48
Phsych_illness	0.05	*	0.03	82.23		340.03
Neurology	0.01		0.02	51.93		289.91
Cardiovasc	0.11	***	0.02	1,096.55	***	174.45
Respiratory	0.08	***	0.02	614.63	**	282.6
Skeleton	0.02		0.02	168.90		172
Pain	0.02	*	0.01	225.60		165.77
Other_disease_1	0.08	***	0.01	175.62		159.89
Other disease_2+	0.07	***	0.02	688.10	**	343.04
ageULF	0.00		0	7.37		10.79
2nd gen immigrant	-0.02		0.04	188.22		1,194.22
Primary education		*reference*			*reference*	
Secondary education	0.02		0.01	31.90		175.83
Higher education ≤2	0.03		0.02	467.31		694.8
Higher education >2	-0.01		0.02	411.84		353.52
lndispinc	0.02	**	0.01	-207.14	*	116.31
Married_cohab	0.03	***	0.01	-127.78		190.37
Child_1	0.00		0.01	-273.13		333.75
Child_2+	-0.03	**	0.02	-551.21	**	274.86
Large_city	0.05	***	0.01	159.56		185.73
Work_full_time	-0.01		0.01	-245.76		228.77
No. of observations	4,410			3,774		

Note: ***p-value<0.01, **p-value<0.05 and *p-value<0.1.

Appendix G. Marginal effects of the probability of accessing pharmaceuticals and conditional on access, accessed number of defined daily doses (DDD). Results for males and females when individuals with up to primary education are excluded.

	Males						**Females**					
	(1)			(2)			(3)			(4)		
	Probit on access			**GLM on DDD|access**			**Probit on access**			**GLM on DDD|access**		
	M.E.		**Std. Err.**	***M.E.***		**Std. Err.**	***M.E.***		**Std. Err.**	***M.E.***		**Std. Err.**
Native Swedish		*reference*			*reference*			*reference*			*reference*	
Nordic origin	-0.03		0.05	545.47		347.69	-0.03		0.03	453.58		552.47
Western origin	0.02		0.05	-20.41		409.44	0.00		0.04	69.21		431.69
non-Western origin	0.07	***	0.03	204.10		242.42	-0.03		0.03	-441.70		445.33
SAH 1		*reference*			*reference*			*reference*			*reference*	
SAH 2	0.06	***	0.02	462.98	***	136.52	0.03	**	0.01	9.93		169.09
SAH 3	0.10	***	0.02	659.18	***	169.5	0.08	***	0.02	540.55	**	216.37
Underweight	-0.02		0.09	1,594.16	**	700.48	-0.03		0.03	44.36		377.06
Normal weight		*reference*			*reference*			*reference*			*reference*	
Overweight	0.04	**	0.01	-27.27		167.35	0.00		0.01	344.71	**	170.84
Exercise 1		*reference*			*reference*			*reference*			*reference*	
Exercise 2	-0.00		0.03	-45.57		261.19	-0.02		0.02	-635.01	*	328.76
Exercise 3	0.02		0.03	-178.81		281.4	-0.03		0.02	-41.86		517.64
Exercise 4	0.03		0.03	3.80		239.87	-0.01		0.02	-668.42	**	328.67
Smoker	0.01		0.02	269.64		236.4	-0.01		0.01	233.98		315.47
Mobility impairment	0.01		0.04	412.86	**	186.78	-0.02		0.02	667.94	***	214.86
Self_care	0.31	***	0.08	-677.12	**	297.12	0.11	***	0.03	-492.29		364.35
Phsych_illness	0.09	*	0.05	591.10		416.53	0.05		0.03	-23.76		437.12
Neurology	0.08	*	0.04	-36.02		283.58	0.01		0.02	-358.74		282.78
Cardiovasc	0.21	***	0.03	1,353.66	***	152.01	0.14	***	0.02	1,185.48	***	163.35
Respiratory	0.08	***	0.03	269.05		204.56	0.08	***	0.03	533.63	*	313.46
Skeleton	0.05	*	0.03	-173.35		158.27	0.01		0.02	221.58		190.39
Pain	0.01		0.02	31.44		189.63	0.02	*	0.01	0.16		157.14
Other_disease_1	0.11	***	0.02	180.21		156.8	0.06	***	0.01	396.18	**	183.02
Other disease_2+	0.17	***	0.03	1,042.20	***	305.48	0.06	***	0.02	1,294.94	***	336.26
ageULF	0.00	***	0	23.52	***	6.41	-0.00	***	0	1.95		7.37
2^{nd} gen immigrant	0.10	*	0.05	-545.43		369.6	-0.03		0.03	-1,053.81		806.69
Secondary education		*reference*			*reference*			*reference*			*reference*	
Higher education ≤2	0.03	*	0.02	95.63		193.92	0.01		0.01	503.30		371.65
Higher education >2	0.02		0.02	170.80		153.62	-0.02	*	0.01	286.63		306.99
lndispinc	0.03	***	0.01	292.35	*	150.96	0.03	***	0.01	-10.41		126.81

Married_cohab	0.04	**	0.02	-71.18		172.73	0.01		0.01	-109.74		173.99
Child_1	0.02		0.02	-29.95		428.45	0.01		0.01	-324.91		314.66
Child_2+	-0.05	**	0.02	-244.80		365.17	-0.04	***	0.02	-666.24	***	224.93
Large_city	0.00		0.02	-36.68		189.26	0.02	**	0.01	20.08		168.98
Work_full_time	0.00		0.02	-400.03	**	167.93	-0.00		0.01	-263.77		210.62
No. of observations	3,245			2,442			3,457			3,053		

Note: ***p-value<0.01, **p-value<0.05 and *p-value<0.1.