UNIVERSITA' DEGLI STUDI DI ROMA "TOR VERGATA"
FACOLTA' DI ECONOMIA

Dottorato di ricerca in Econometria ed Economia Empirica
XX ciclo

Essays in Applied Health Economics

Claudio Rossetti

Docente guida: prof. Franco Peracchi
Coordinatore: prof. Franco Peracchi

Abstract

Chapter 1 focuses on the issue of reporting bias in self-rated health. This chapter shows that gender and regional differences in self-rated health in Europe are only partly explained by differences in the prevalence of the various chronic conditions. However, a non-negligible part of these differences is due to other causes, which may include differences in reporting own health. The tool of "anchoring vignettes" is employed to understand whether and how women and men living in different regions differently report levels in a number of health components or domains. The analysis is based on Release 2 of the first (2004) wave of the Survey of Health, Ageing and Retirement in Europe (SHARE). This survey is ideal for the purpose because it contains information on subjective measures of health (such as self-rated health) and more objective measures (such as hospitalization and interviewer-measured grip strength), as well as detailed information on chronic health conditions. Release 2 of the data also includes the use of vignettes in self-administered questionnaires given to a randomly selected subsample of respondents. Vignettes are found to help identifying gender and regional differences in response scales. After correcting for these differences, both gender and regional variation in reported health is substantially reduced, although not entirely eliminated. The results suggest that differences in response styles should be taken into account when using self-assessment of health in socio-economic studies. Failing to do so may lead to misleading conclusions.

Focusing on a specific chronic condition, hypertension, Chapter 2 studies the relationship between medical compliance and health outcomes (hospitalization and mortality rates) using a large panel of patients residing in a local health authority in Italy. These data allow to follow individual patients through all their accesses to public health care services until they either die or leave the local health authority. The results show that health outcomes clearly improve when patients become more compliant to drug therapy. At the same time, it is possible to infer valuable information on the role that drug co-payment can have on compliance, and as a consequence on health outcomes, by exploiting the presence of two natural experiments during the period of analysis. The results show that drug co-payment has a strong effect on compliance, and that this effect is immediate.

Chapter 3 improves the analysis of the relationship between health and medical care provided in Chapter 2. In fact, looking at the raw correlation between medical care and health cannot be expected to give the right answer, because of simultaneity through the unobservable components of deterioration. In this chapter, it is used a dataset where very detailed information about medical drug use, hospitalization, and mortality, is collected over time for a sample of individuals suffering from hypertension, a chronic asymptomatic pathology affecting a large share of the adult population. All those variables are expected to be strongly dependent on each other. For analysing the amount of information embedded in such variables, a dynamic factor model is proposed, where medical treatments and mortality may all in principle be driven by latent individual stock of health. Dynamics is introduced by including the effects of lagged treatment on latent health. The model is estimated by Maximum Simulated Likelihood (MSL).

In line with findings provided so far in the literature, the results indicate that better health is associated to lower medical treatments. In addition, lagged medical drug use is found to have positive effects on current health. This is consistent with the fact that not taking the medication today may result in poorer health tomorrow. Nonetheless, taking more pills than needed cannot improve health. These findings have important policy implications. In fact, the results suggest that policies aimed at improving awareness of hypertensive diseases and the importance of the treatment of high blood pressure may help reduce cardiovascular risks, and consequent hospitalization and mortality. This is expected to have positive implications both for the large share of adult population suffering from hypertension and for the National Health Systems themselves.

Keywords: Self-rated health, health domains, anchoring vignettes, reporting bias, health policy reforms, co-payment, dynamic panel data models, factor models, simulated likelihood, latent variable models

JEL codes: C35, C81, I12, J14, D12, C33

Abstract

Il Capitolo 1 focalizza l'attenzione sui problemi di "reporting bias" legati all'indicatore di salute auto-riportato. Questo capitolo mostra che in Europa differenze di genere e differenze regionali possono solo parzialmente essere spiegate dalle differenze nella prevalenza delle varie condizioni croniche. Eppure, una parte non trascurabile di queste differenze è dovuta ad altre cause, che possono includere differenze nel modo in cui lo stato di salute viene riportato. Lo strumento delle "anchoring vignettes" è utilizzato per comprendere se e come le donne e gli uomini che vivono in diverse regioni d'Europa riportano differentemente il livello di salute relativo a vari "domini". L'analisi è basata sulla seconda Release della prima (2004) wave della Survey of Health, Ageing and Retirement in Europe (SHARE). Questa indagine è ideale per lo scopo in quanto contiene informazioni circa misure soggettive dello stato di salute e misure più oggettive (come ospedalizzazione e "grip strength"), come anche informazioni dettagliate circa condizioni croniche. La seconda Release dei dati contiene anche l'uso di "vignettes" in questionari assegnati ad un campione casuale di rispondenti. Le "vignettes" risultano essere utili per identificare differenze regionali e di genere nelle "response scales". Dopo aver corretto queste differenze, le variazioni regionali e di genere nelle stato di salute riportato risultano entrambe ridotte, seppure non del tutto eliminate. I risultati suggeriscono che le differenze nelle "response styles" devono essere prese in considerazione quando si utilizza lo salute auto-riportato in studi socio-economici. Non tenerne conto può condurre a risultati fuorvianti.

Focalizzando l'attenzione su una specifica condizione cronica, l'ipertensione, il Capitolo 2 studia la relazione tra compliance medica e outcome sanitari (ospedalizzazione e mortalità) utilizzando un panel di pazienti che risiedono in un'Autorità Sanitaria Locale italiana. Questi dati consentono di seguire i pazienti attraverso tutti i loro accessi ai servizi sanitari pubblici. I risultati mostrano che gli outcome sanitari migliorano decisamente quando i pazienti sono più "compliant" alla terapia. Inoltre, è possibile inferire importanti informazioni circa il ruolo che il co-payment ha sulla compliance, e di conseguenza sugli outcome sanitari, esplorando due esperimenti naturali verificatisi durante il periodo qui analizzato. I risultati mostrano che il co-payment ha forti effetti sulla compliance, e che questi effetti sono immediati.

Il Capitolo 3 estende l'analisi della relazione tra salute e trattamento sanitario fornita nel Capitolo 2. Infatti, considerando la semplice correlazione tra salute e trattamento sanitario non necessariamente fornisce la risposta adeguata, a causa della simultaneità nelle componenti inosservate del deterioramento della salute. In questo capitolo, si utilizza un dataset in cui informazioni molto dettagliate circa il consumo farmaceutico, l'ospedalizzazione e la mortalità sono collezionate nel tempo per un campione di individui affetti da ipertensione. L'ipertensione è una condizione cronica e asintomatica di cui soffre una larga parte della popolazione adulta. Tutte queste variabili sono fortemente dipendenti l'una dall'altra. Per analizzare l'informazione contenuta in tali variabili, viene proposto l'impiego di un modello a fattori dinamico, in cui il trattamento medico e la mortalità siano in principio tutti guidati dallo stato di salute latente. La dinamica viene introdotta nel modello includendo l'effetto del trattamento medico passato sullo stato di salute corrente. Il modello è stimato tramite Massima Verosimiglianza Simulata.

Coerentemente con i risultati presenti finora in letteratura, i risultati indicano che una migliore condizione di salute è associata con un minore trattamento medico. Inoltre, il consumo farmaceutico nel periodo precedente ha effetti positivi sullo stato di salute corrente. Questo è consistente con il fatto che non seguire la terapia medica oggi può risultare in una peggiore condizione di salute domani. Nonostante questo, assumere più pastiglie di quanto necessario non migliora ulteriormente la stato di salute. Questi risultati hanno importanti implicazioni in termini di policy. Infatti, i risultati suggeriscono che politiche mirate ad aumentare la consapevolezza delle malattie legate all'ipertensione e l'importanza della cura dell'alta pressione possono aiutare non poco a ridurre i rischi cardiovascolari, e la conseguente ospedalizzazione e mortalità. Ci si attende che questo abbia implicazioni positive sia per la larga parte di popolazione adulta affetta da ipertensione sia per gli stessi Servizi Sanitari Nazionali.

Contents

1 Chapter 1
Gender and regional differences in self-rated health in Europe

Abstract[1]: This paper shows that gender and regional differences in self-rated health in Europe are partly explained by differences in the prevalence of the various conditions. However, a non-negligible part of these differences is due to other causes, which may include differences in reporting own health. We employ the tool of "anchoring vignettes" to understand whether and how women and men living in different regions differently report levels in a number of health components or domains. We find that vignettes help identifying gender and regional differences in response scales. After correcting for these differences, both gender and regional variation in reported health is substantially reduced, although not entirely eliminated. Our results suggest that differences in response styles should be taken into account when using self-assessment of health in socio-economic studies. Failing to do so may lead to misleading conclusions.

Keywords: self-rated health, health domains, anchoring vignettes, reporting bias

JEL codes: C35, C81, I12, J14

[1] I wish to thank Anne Case and Arthur van Soest for helpful discussions. This chapter is based on the article "Gender and regional differences in self-rated health in Europe", written with Franco Peracchi.

1.1 Introduction

Self-rated health (SRH) tends to be worse for women than for men at all ages, although women are less likely to die and do not present higher hospitalization rates than men at ages when pregnancy-related hospitalization is no longer an issue. In Europe, not only gender differences, but also regional differences in SRH are observed. Both men and women living in Mediterranean countries tend to report worse health than those living in Continental and Scandinavian countries, but they are not more likely to be hospitalized or die.

This paradox could have different explanations, not necessarily mutually exclusive. One explanation is that gender and regional differences in SRH could be due to differences in the distribution of chronic conditions, for either biological or behavioural reasons. Suffering from conditions that are painful, but not life threatening, could lead to poorer SRH but need not imply higher hospitalization or mortality rates. Indeed, Case and Paxson (2005) show that the difference in SRH between women and men in the U.S. can almost entirely be explained by differences in the distribution of chronic conditions.

Another explanation is that there are gender and regional differences in the way people report their health status. This may depend on a different perception of health problems, or on a different mapping of true health status into SRH. In fact, since true health status and subjective thresholds may both vary across individuals, it is not possible, using answers to the subjective scale questions alone, to know how much of the individual rating on these scales reflects true objective differences among people and how much it reflects variation across people in their subjective thresholds. Several studies have focused their attention on heterogeneity of health reporting (see for example Sen 2002, Lindeboom and van Doorslaer 2004, Jürges 2008). Jürges (2007) shows that when differences in reporting styles are taken into account, cross-country variation in SRH in Europe are substantially reduced.

In this paper, we decompose gender and regional differences in morbidity into the contribution of differences in the distribution of chronic conditions and the contribution of the impact of such conditions. For this purpose, we compare men and women living in the same European region, as well as people of the same gender living in different regions, after controlling for differences in socio-demographic characteristics and other health measures, such as body mass and grip strength. The fact that differences in SRH between men and women living in different regions can partly be explained by differences in the distribution of chronic conditions does not exclude the possibility that these groups might use systematically different response scales. For this reason, we employ the tool of "anchoring vignettes" to correct self-assessment of health on six components or domains of health. The domains considered here are pain, mobility, sleeping problems, shortness of breath, concentration problems, and depression. Because reported general health can be regarded as a scalar summary that depends on the level in these different domains (Salomon et al. 2003), understanding whether and how men and women living in different regions differently report levels in these domains

may provide helpful insight into differences in SRH.

Anchoring vignettes have been developed as a new component of survey instruments that may be used to position self-reported responses on a common, interpersonally comparable scale. Respondents are first asked to evaluate their position on a scale in a given domain. They are then asked to evaluate the vignette on the same scale they used to rate their own position. Because the objective situation of the person described in the vignette is the same for all respondents, anchoring vignettes have the potential to identify individual variation in subjective thresholds. Vignette questions have been applied in works on international comparisons of health (Salomon, Tandon and Murray 2004, King and Wand 2007, D'Uva et al. 2008), political efficacy (King et al. 2004) and work disability (Kapteyn, Smith and van Soest 2007). In all these applications, subjective scales were used and significant differences were found across groups or countries in the subjective outcomes. Anchoring vignettes were employed to assess whether these groups also differed in their subjective thresholds. A validation study of the use of vignettes for correcting subjective response scales is provided by van Soest et al. (2007).

Our analysis is based on Release 2 of the first (2004) wave of the Survey of Health, Ageing and Retirement in Europe (SHARE). This survey is ideal for our purpose because it contains information on subjective measures of health (such as SRH) and more objective measures (such as hospitalization and interviewer-measured grip strength), as well as detailed information on chronic health conditions. Release 2 of the data also includes the use of vignettes in self-administered questionnaires given to a randomly selected subsample of respondents. For our purpose, the survey is better than other comparable surveys, such as the European Community Household Panel (ECHP), because the latter does not provide detailed information about chronic health conditions, contains little information on objective health measures, and does not include vignettes.

Our results indicate that the differences between men's and women's health are only partially explained by differences in the prevalence of the various conditions. A non-negligible part of the differences depends on unexplained factors, which may possibly include gender differences in reporting own health. Furthermore, most of the regional differences in the fraction reporting poor health is unexplained by the differences in health conditions and limitations, which again may possibly be due to differences in how people report their health. Socio-demographic characteristics turn out to be much less important than chronic conditions in explaining both gender and regional differences in SRH. We find that vignettes help identifying differences in how men and women living in different European regions report their health. We find that vignettes help identifying differences in how men and women living in different European regions report their health. Specifically, after correcting for response scales, both gender and regional variations in reported health are substantially reduced, although not eliminated. Our results suggest that differences in response styles should be taken into account when using self-assessment of health in socio-economic studies. Failing to do so may lead to misleading conclusions.

The remainder of this paper is organised as follows Section 1.2 describes the data used for this study. Section 1.2.3 provides preliminary evidence. Section 1.3 examines gender and regional differences in the relationship between chronic conditions and SRH. Section 1.4 examines gender and regional differences in self-assessment of health, using anchoring vignettes to correct for the possibility that different groups might use systematically different response scales. Finally, Section 1.5 offers some conclusions.

1.2 Data and descriptive statistics

1.2.1 Data

The data in this study are from Release 2 of the first (2004) wave of the Survey of Health, Ageing and Retirement in Europe (SHARE), a multidisciplinary and cross-national longitudinal survey on health, socio-economic status, and social and family networks. The target population of SHARE consists of individuals aged 50+ (born in 1954 or earlier), and their spouses/partners regardless of age, living in private households in Europe. Partners may be younger than 50, but must be living at the exact same address as the selected age-eligible respondent.

Eleven countries have contributed data to the 2004 SHARE baseline study. They are a representation of the various regions of Europe, ranging from Scandinavia (Denmark, Sweden) through Central Europe (Austria, France, Belgium, Germany, Netherlands, Switzerland) to the Mediterranean region (Greece, Italy, Spain). The survey has been administered by means of computer assisted personal interviews (CAPI) in the fall of 2004 to probability samples of individuals aged 50+ in the participating countries. For a detailed description, see Appendix A, and Börsch-Supan et al. (2005), and Börsch-Supan and Jürges (2005).

The survey collects information on health variables (SRH, physical functioning, cognitive functioning, health behavior, use of health care facilities, etc.), psychological variables (psychological health, life satisfaction, etc.), economic variables (current work activity, job characteristics, opportunities to work past retirement age, sources and composition of current income, wealth and consumption, housing, education), and social support variables (assistance within families, transfers of income and assets, social networks, volunteer activities, etc.).[2] The second release of SHARE 2004 also includes vignettes on health as self-administered questionnaires in Sweden, Belgium, Spain, France, Germany, Greece, Italy, and the Netherlands.

We restrict attention to men and women aged 50–90 for whom the vignette information is available. We remove all cases with missing data on any of the variables used. Note that, unlike the case of income or wealth, item nonresponse to health questions is negligible. Nonresponse to the SRH question is lower than 1% in all countries except France, where it is slightly higher than 2%. Even nonresponse to single vignette questions is lower than 1% in almost all countries. Nonetheless,

[2] The data may be downloaded by registered users from the SHARE website (http://www.share-project.org).

the fraction of respondents for whom answer to at least one vignette question is missing is a bit higher (on average 6%, ranging from about 2% in Greece to about 11% in Sweden).

Table 1 shows the composition of our final sample by country and gender. The subsample of respondents to which the vignettes questions were assigned represents about 18% of the full SHARE sample.

1.2.2 Variables

Our measure of morbidity is based on the European categorization of SRH into 5 categories:[3] 1="Very good", 2="Good", 3="Fair", 4="Bad", 5="Very bad". We use a dichotomization of SRH, namely a binary indicator equal to one if an individual reports herself to be in fair, bad or very bad health, and equal to zero otherwise. ¿From now on we refer to such binary indicator as "poor health".

SHARE also includes self-assessments and vignette questions on a set of health related concepts or domains, namely pain, mobility, sleeping problems, shortness of breath, concentration problems, depression, and work limitations. This set of health domains is sufficiently exhaustive to capture the common meaning of health. On the other hand, health domains provide a parsimonious description of health avoiding overlap and redundancy (Salomon et al. 2003). Respondents are asked to rate their own health problems in the six domains on an ordered qualitative scale. The five response categories are: (1) None, (2) Mild, (3) Moderate, (4) Severe, (5) Extreme. For parsimony, in the empirical work we merge the categories "Moderate", "Severe" and "Extreme" into a single one.[4] A detailed description of the self-assessment questions for all six domains is reported in Appendix B.

Two sets of covariates are used to model health outcomes. The first set includes indicators for diagnosed chronic conditions and illnesses, interviewer-measured grip strength, and a measure of relative body weight. The second set includes standard socio-demographic characteristics. The self-reported diagnosed conditions[5] considered are heart attack, high blood pressure, high blood cholesterol, stroke, diabetes, chronic lung disease, asthma, arthritis, osteoporosis, ulcer, Parkinson disease, cataracts, hip or femoral fracture, reproductive cancer, and other cancer. Illnesses which may be symptoms of diseases are pain in back, heart trouble, breathlessness, persistent cough, swollen legs, sleeping problems, falling down, fear of falling down, dizziness, stomach problems, incontinence, and other symptoms.

Grip strength is a core physical measure of health that potentially overcomes the measurement issues arising from subjectivity of SRH. Grip strength is also known to be a good predictor of

[3] We also carried out our analysis by using the US categorization of SRH. The results obtained by using the latter do not differ from the results obtained by using the EU categorization. For this reason we decided to report only the results from the EU categorization.

[4] Our main conclusions do not change if these categories are considered separately.

[5] These conditions are self-reports about medical diagnosis. In fact, the exact questions are "Has a doctor ever told you that you had" a certain condition.

future medical problems (Rantanen et al. 1999). It is measured here as the maximum of up to four measurements made by the interviewer, two on the left hand and two on the right hand. We use an indicator for the respondent's grip strength (normalized for height, weight and sex) being in the bottom quartile. We label such indicator as "low grip strength".

We include a measure of relative body weight to control for the effects of excessive body weight on physical health. Individuals are classified by relative weight based on their body mass index (BMI), computed from self-reported weight and height as weight (in kilograms) divided by the square of height (in meters). We use the evidence-based clinical guidelines for the classification of overweight and obesity in adults, published by the National Heart, Lung and Blood Institute of the National Institutes of Health (NIH) to classify the respondents into four weight classes: underweight (BMI < 18.5), normal weight ($18.5 \leq$ BMI < 25), overweight ($25 \leq$ BMI < 30), and obesity (BMI ≥ 30), (National Heart, Lung, and Blood Institute).

The set of socio-demographic characteristics includes a polynomial in age, the logarithm of per-capita household income, an indicatorsfor living with a spouse or a partner, and indicators for upper secondary and post-secondary completed education based on the international standard classification of education (ISCED). Household income, in Euros and before tax, is adjusted for purchasing power parity and is the sum of a number of income components that are asked separately in the questionnaire. For many observations, one or more of these components are missing. For observations with missing values, the SHARE data provide imputations largely based on the answers to the sequence of unfolding bracket questions asked to initial nonrespondents. We use the first of the five imputations available in SHARE. To adjust for household size, income is divided by the number of household members.

1.2.3 Descriptive statistics and preliminary evidence

Figure 1 shows the fraction reporting poor health by gender for each of the countries considered.[6] The fraction of women reporting poor health is always higher than the fraction of men, excepted in France and the Netherlands. The gender difference in SHARE is particularly high for Mediterranean countries (Greece, Italy, and Spain) and is much lower for non-Mediterranean countries (Belgium, France, Germany, Netherlands, and Sweden).

Figure 2 shows the fraction reporting poor health by region, gender and age. In Mediterranean countries, the fraction of women reporting poor health is higher than the fraction of men at almost all ages.

Figure 3 shows the histograms of self assessments for the health domains considered here by region and gender. For most health domains, women are more likely to report themselves to have

[6] We carried out a similar analysis using data from the ECHP. For both men and women, the fraction reporting poor health in the ECHP is systematically higher than in SHARE. Apart from this, the main conclusions about gender and regional differences in SRH are similar for the two survey. The results from the ECHP are available from the authors upon request.

moderate, severe or extreme health problems than men.

Figure 4 shows prevalence rates of some selected conditions by gender, age and region (non-Mediterranean countries in the top panel, Mediterranean countries in the bottom panel). Women are more likely to suffer from painful conditions such as arthritis, rheumatism, or osteoporosis than do men. On the other hand, men are more likely to suffer from life threatening conditions such as heart attack, or stroke.

Table 2 shows descriptive statistics of the variables used here by region and gender. In non-Mediterranean countries, about 36% of men and 39% of women report poor health. In Mediterranean countries, these percentages are about 35% and 54% respectively for men and women. For both men and women, the fraction of people with low hand grip strength is higher in non-Mediterranean countries. Average age varies little, from 63 to 65 years. The fraction with secondary and post-secondary completed education is always higher for men than for women, but people living in non-Mediterranean countries are on average more educated and have higher household income than people living in Mediterranean countries.

1.3 SRH and chronic conditions

In this section we analyze the relationship between the probability of reporting poor health on the one hand, and socio-economic characteristics and health problems and limitations on the other hand. To facilitate comparison with the results of Case and Paxson (2005) for the U.S., we largely follow their approach.

1.3.1 Model specification and estimation

We model the probability of reporting poor health ($H = 1$) as a linear function of a set of health problems and limitations C and a set of socio-economic characteristics W

$$\Pr\{H = 1|C, W\} = \alpha + \beta' C + \gamma' W. \tag{1}$$

The set of socio-economic characteristics includes age, age squared, the logarithm of per-capita household income, and indicators for educational attainments and for living with a spouse or a partner. The set of health problems and limitations depends on the specification of the model. In the first specification (Model 1), this set includes indicators for the presence of chronic conditions and symptoms, low grip strength and BMI. The second specification (Model 2) replaces the indicators for the presence of chronic conditions, low grip strength and BMI with a set of indicators for reported mild or moderate, and severe or extreme problems in the six health domains. The third specification (Model 3) contains all the regressors included in (Model 1) and (Model 2).

We estimate model (1) pooling data by country and gender, thus constraining coefficients to be the same for men and women living in different regions, and separately for four groups: non-

Mediterranean women (NW), non-Mediterranean men (NM), Mediterranean women (MW), and Mediterranean men (MM). In the second case, the two sets of covariates C and W include all the variables in the third specification (Model 3). The OLS estimates of β_{NW}, β_{NM}, β_{MW}, and β_{MM} provide information on regional and gender differences in how health problems and limitations map into health measures. Following Case and Paxson (2005), we use these estimates and the information about the prevalence of the various conditions and limitations to construct measures of "severity" and "prevalence" effects. Because the model is linear, we can decompose the differences in the probability of reporting poor health between any two groups, j and k ($j, k = NW, NM, MW, MM$), into a number of components. The first component is a "prevalence effect" (or endowments effect), capturing differences in the distributions of conditions and limitations. It is measured by the differences in prevalence rates weighted by a vector β_* of chronic condition's benchmark coefficients

$$\beta_*'(\bar{C}_j - \bar{C}_k).$$

The second component is a "severity effect" (or coefficients effect), due to differences in the impact of conditions and limitations

$$(\beta_j - \beta_*)'\bar{C}_j + (\beta_* - \beta_k)'\bar{C}_k'.$$

The other components are the endowment effects and the coefficient effects of the control variables in W, and a residual term which includes other regional differences (country dummies) and "unexplained" differences (the constant term). Alternative choices of benchmark coefficients are $\beta_* = \beta_j$, $\beta_* = \beta_k$, $\beta_* = (\beta_j + \beta_k)/2$, or β_* equal to the coefficients in the pooled sample of the two groups. To ensure comparison with Case and Paxson (2005), we set $\beta_* = (\beta_j + \beta_k)/2$.

1.3.2 Pooled data

Table 3 contains the estimated coefficients of the OLS regression for the probability of reporting poor health and our three different specifications using the pooled data.

In the first specification (Model 1), most of the indicators of chronic conditions and symptoms have a positive and statistically significant effect on the probability of reporting poor health. Low grip strength also has a positive and statistically significant coefficient, while the coefficients on the indicators for BMI turn out to be small and not statistically significant. There is a negative gradient in education, as the probability of poor health declines monotonically with educational attainments. The R^2 of this regression is about 30%.

In the second specification (Model 2), the use of indicators for reported problems in the six health domains achieves a similar fit as Model 1. Not surprisingly, pain and mobility problems have the highest impact on the probability of reporting poor health.

In the third specification (Model 3), most chronic conditions are still significant after controlling for problems in the health domains. Considering both sets of variables improves the R^2 from 30

to about 38%. This is interesting because it indicates that the six health domains are not just summaries of the information provided by the chronic conditions.

1.3.3 Gender differences

Table 4 shows gender differences in the impact of each condition on the probability of reporting poor health. Estimated OLS coefficients for the four groups are reported in Appendix D. In most cases, the differences in the coefficients between groups are not statistically significant. Further, the hypothesis that all the coefficients associated with conditions and limitations are the same for men and women cannot be rejected at conventional levels. This is consistent with the finding of Case and Paxson (2005) for the U.S. of no significant gender differences in how chronic conditions map into SRH.

Although we observe no gender differences in how conditions map into reported poor health, there are important gender differences in the prevalence of conditions. Table 5 shows excess prevalence of each condition and limitations in women relative to men. Women report significantly higher pain and have higher prevalence of painful conditions such as arthritis, rheumatism, osteoporosis, and other non-life-threatening problems such as sleeping problems and depression. Men, on the other hand, are significantly more likely to suffer from heart attack.

Table 6 shows the decomposition of gender differences in the probability of reporting poor health. The first column shows the decomposition of the differences between non-Mediterranean women and non-Mediterranean men. Women are only about 3% more likely to report poor health than men. The second column shows the decomposition of the differences between Mediterranean women and Mediterranean men. The former are about 19% more likely to report poor health than the latter. The difference between men's and women's health is partly explained by differences in the prevalence of the various conditions. Furthermore, estimated prevalence effects are much more important than severity effects. In particular, the latter explain only less than 3% of the differences. This is again consistent with the findings in Case and Paxson (2005). Nonetheless, a non negligible part of the differences is due to other causes, which may include gender differences in reporting own health.

1.3.4 Regional differences

The preliminary evidence in Section 1.2.3 showed that while SRH does not differ much by region for men, this is not true for women. In fact, women living in Mediterranean countries report themselves to be in poorer health than women living in non-Mediterranean countries, although the latter have lower life expectancy than the former. In this section we examine the relationship between regional differences in the probability of reporting poor health and regional differences in the prevalence of health conditions and limitations.

Table 7 shows regional differences in the impact of each condition on the probability of reporting poor health. In most cases, coefficients are not statistically different between groups and the hypothesis that the coefficients on conditions are the same for people living in non-Mediterranean and Mediterranean countries cannot be rejected at conventional levels. This suggests the absence of significant regional differences in how conditions and limitations map into reports of poor health.

On the other hand, Table 8 suggest that there are important regional differences in the prevalence of the various conditions. The table shows excess prevalence of each condition and limitations in women and men living in Mediterranean countries relative to women and men living in non-Mediterranean countries. Women living in Mediterranean countries have significantly higher rates of arthritis and osteoporosis than women living in non-Mediterranean countries. On the other hand, men living in non-Mediterranean countries are more likely to suffer of hearth attack or stroke than men living in Mediterranean countries.

Table 9 shows the decomposition of the regional differences in the probability of reporting poor health. Although very small for men, these differences are sizable for women. Mediterranean women are about 15% more likely to report poor health than non-Mediterranean women, but a large part of this regional difference remains unexplained. Consistently with the findings in Jürges (2007), this is possibly due to differences in how women living in different regions report their own health.

1.4 Anchoring vignettes

The results obtained thus far do not exclude the possibility that men and women living in different regions use systematically different response scales when reporting their health. In this section we employ the information contained in anchoring vignettes to check whether this is the case and to control for such differences.

Anchoring vignettes have been developed as a new component of survey instruments that may be used to position self-reported responses on a common, interpersonally comparable scale. Specifically, "an anchoring vignette is a description of a concrete level on a given health domain that respondents are asked to evaluate with the same questions and response scales applied to self-assessments on that domain. Vignettes fix the level of ability on a domain, so that variation in categorical responses is attributable to variation in response category cut-points " (Salomon et al. 2003). Because the same hypothetical situation is presented to each respondents, variability in vignette answers reveals lack of comparability. In practice, the self-assessment is usually asked first, followed by the vignettes randomly ordered. In SHARE, the names on each vignette are changed to match a respondent's gender and country.

1.4.1 A simple example

The following example illustrates how vignettes help identifying differences in response scale.

Suppose we want to characterize the amount of pain two groups of individuals have. Figure 5 presents the distribution of the density of the true but unobserved continuous level of pain for groups A and B. On average, people in group B have more pain than people in group A. However, people in the two groups use different response scales when asked whether or not they have pain on a three-point scale. The most common terminology for interpersonal incomparability is "differential item functioning" (DIF). The term originated in the educational testing literature, where a test question is said to have DIF if equally able individuals have unequal probabilities of answering the question correctly. In this example, pain is better tolerated by people in group A than by people in group B. The distribution of self-reports in the two groups suggests that people in A have more pain than those in B. This is in fact the opposite of the true distribution. Correcting for the differences in the response scales is essential to compare the actual level of pain in the two groups.

Vignettes can be used for this purpose. The hypothetical individual described in the vignettes is the same and its objective pain level is marked by the dashed line. This is evaluated as "Mild" by group A and as "None" by group B. Since the actual level of pain of the vignette person is the same, the difference in the evaluations by the two groups is likely to be due to DIF. Hence, vignette evaluations help identify differences in response scales. In fact, using the scales in one of the two groups as the benchmark, the distribution of evaluations in the other group can be adjusted by evaluating them on the benchmark scale. The corrected distribution of the evaluations can then be compared since they are now on the same scale.

1.4.2 Health on six domains and vignettes

Vignettes included in SHARE refer to the six health domains described in Section 1.2, namely pain, mobility, sleeping problems, shortness of breath, concentration problems, and depression, plus work limitations. We do not use the vignettes for work limitations because strictly speaking work limitations cannot be considered as a health domain. The reason why there are no vignettes for general health is that this is a multi-dimensional concept and therefore cannot be related to just one domain.

In this section, we use anchoring vignettes to correct for the lack of interpersonal comparability in reported health levels on each of the six domains. Although correction of reported health on the six domains does not offer a direct correction of self-rated general health, it may provide helpful insight into differences in how men and women living in different European regions report their own health.

For each of the six domains, three vignette questions were asked in a random order after the self-assessment question (Appendix C reports a detailed description of the vignettes questions).

For each vignette situation, respondents were asked to rate health problems of the hypothetical persons on the same five-point ordered scale ranging from "None" to "Extreme" used for the self-assessment question. As for self-assessments, we merge the categories "Moderate, "Severe" and "Extreme" into a single one. The health problems in the three hypothetical situations in each domain may be viewed as ordered from least to most severe.

Using anchoring vignettes to correct for self-assessment requires two key assumptions (King et al. 2004). The first ("response consistency") is the assumption that each individual uses the response categories for a particular survey question in the same way when providing self-assessment and when assessing each of the hypothetical situations in the vignettes. The second ("vignette equivalence") is the assumption that the level of the variable represented in each vignette is perceived by all respondents in the same way and on the same uni-dimensional scale, apart from random measurement error.

1.4.3 The statistical model

King et al. (2004) propose a parametric approach for correcting interpersonal incomparability of self-assessed variables. Their approach is based on a parametric ordered probit model for the self-assessments where, under the assumption of response consistency, the individual specific thresholds depend on the same parameters as in the ordered probit model for the responses to the vignettes.

Consider one of the six health domains described above. The self-reports on that domain are assumed to be driven by an underlying latent index on a continuous scale

$$Y^* = \mu + U, \tag{2}$$

where μ is the actual level of health problems and U is a random measurement error with mean zero. Higher values of Y^* correspond to higher health problems. The actual level μ is modeled as a linear function of observed variables

$$\mu = \alpha_0 + \alpha_1 F + \alpha_2 M + \alpha_3 F \cdot M + \beta' C + \gamma' W \tag{3}$$

where $\alpha = (\alpha_0, \alpha_1, \alpha_2, \alpha_3)$, β and γ are vectors of unknown parameters, F is an indicators for being a female, M is an indicator for for living in Mediterranean countries, C is a set of indicators for the presence of chronic conditions, and W is a set of other controls. The conditional distribution of U given F, M, C and W is assumed to be $\mathcal{N}(0, \omega^2)$. What is observed is a categorical variable Y, which takes the value $l = 1, \ldots, L$ whenever $\tau_{l-1} < Y^* \leq \tau_l$, where the thresholds $\tau_0, \ldots, \tau_L$ are

given by

$$\begin{aligned}
&\tau_0 = -\infty \\
&\tau_1 = \phi_1 + \delta_1' X \\
&\tau_l = \tau_{l-1} + \exp(\phi_l + \delta_l' X), \quad l = 2, \ldots, L-1, \\
&\tau_L = \infty,
\end{aligned} \tag{4}$$

and the variables contained in X may include F, M, or some of the variables contained in C or W. Note that the nonlinearities in the threshold model (4) are introduced to ensure that thresholds of higher order are never smaller than thresholds of lower order. A test of homogeneity in response scales (no DIF) is a test of the hypothesis that $\delta_1 = \cdots = \delta_{L-1} = 0$.

Restrictions on the model parameters are needed for identifiability. First of all, the latent index must be assigned a location and a scale. Here, we fix the location by setting the constant of the first threshold ϕ_1 to be equal to 0. The scale is fixed by normalizing the variance of the measurement error ω^2 to 1. It is easy to verify that, if the variables in F, M, C and W are the same as those in X, and we use only the self-assessments, then the parameter vectors α, β and γ cannot be separately identified from the parameter vector $\delta = (\delta_1, \ldots, \delta_{L-1})$. In this case, identification only depends on the nonlinearities in the threshold model (4). Since identification by functional form is undesirable, strong identification requires using at least one vignette.

Responses to each of the $m = 1, 2, 3$ vignettes are also modeled using an ordered probit model

$$Z_m^* = \psi_m + V_m, \qquad m = 1, 2, 3,$$

where the ψ_m are the same for all respondents under the vignette equivalence assumption, and the V_m are assumed to be independently and identically distributed as $\mathcal{N}(0, \sigma^2)$ independently of F, M, C, W, X and U. The scale parameter σ^2 measures how well vignettes are understood. What is observed is a categorical variable Z_m, which takes the value $l = 1, \ldots, L$ whenever $\eta_{l-1} < Z_m^* \leq \eta_l$. Under the assumption of response consistency, the thresholds in the self-assessment and the vignette components of the model depend on the same parameters, which ensures identifiability of the entire parameter vector $\theta = (\alpha, \beta, \gamma, \phi, \delta, \psi, \sigma)$.

1.4.4 Model estimation

Although this model may be generalized by relaxing the normality assumption and by introducing time-invariant individual effects (Rossetti 2008), for simplicity here we confine ourselves to maximum likelihood (ML) estimation of a fully parametric version without individual effects. Given a random sample of n individuals (indexed by $i = 1, \ldots, n$), the sample likelihood for the self-

assessment component is

$$\mathcal{L}_s(\theta) = \prod_{i=1}^{n} w_i \prod_{l=1}^{L} [\Phi(\tau_{il} - \mu_i) - \Phi(\tau_{i,l-1} - \mu_i)]^{1\{Y_i = l\}},$$

where w_i is the survey weight for the ith individual, $1\{\cdot\}$ is the indicator function, and $\Phi(\cdot)$ is the standard normal distribution function. The sample likelihood for the vignette component is

$$\mathcal{L}_v(\theta) = \prod_{i=1}^{n} w_i \prod_{m=1}^{3} \prod_{l=1}^{L} \left[\Phi\left(\frac{\tau_{il} - \psi_m}{\sigma} \right) - \Phi\left(\frac{\tau_{i,l-1} - \psi_m}{\sigma} \right) \right]^{1\{Z_{im} = l\}}$$

Because the likelihood from the self-assessment and the vignette components share the parameter vectors ϕ and δ, they must be maximized jointly. Thus, a ML estimator of θ is obtained by maximizing the complete sample likelihood

$$\mathcal{L}(\theta) = \mathcal{L}_s(\theta) \cdot \mathcal{L}_v(\theta).$$

We estimate separate models for each of the six health domains.[7] The maximization routine is written in MATA, the matrix programming language of STATA, and is based on the Newton-Raphson algorithm, with numerical first and second derivatives.

We estimate the model after pooling data by gender and region, thus constraining the slope coefficients to be the same for men and women living in different regions. Nonetheless, the hypothesis that all the coefficients associated with conditions and limitations are the same for men and women living in different regions cannot be rejected at conventional levels.[8] The vector C_i includes indicators for chronic conditions, low grip strength and BMI. The vector W_i includes a set of socio-economic characteristics (age, age squared, the logarithm of household income, and indicators for educational attainments and for living with a spouse or a partner). The vector X_i in the threshold equation includes the same variables contained in C_i and W_i, the indicators for being a female F and for living in Mediterranean countries M, and their interaction.

Table 10 reports the estimated coefficients of both the ordered probit model with constant thresholds (the baseline) and the ordered probit model with individual specific thresholds for the three domains which have the highest impact on SRH, namely pain, mobility and concentration.[9] For parsimony, only the coefficients of the gender and regional dummies and their interaction are reported. Complete parameter estimates of the vignettes model are reported in Appendix E. A likelihood ratio (LR) test rejects the hypothesis of no DIF (constant thresholds) in favor of the

[7] We also estimated a model with common thresholds for all six domains, but such model is rejected against the model with different response scales for each of the six domains.

[8] Given the small sample size, a model with full heterogeneity in the parameters (both in the latent index and in the thresholds) could not be estimated due to problems of convergence.

[9] Results for the other health domains are available from the authors upon request.

model with individual specific thresholds for all health domains. The results of the ordered probit model with constant thresholds appear in the first numerical column of Table 10. After controlling for chronic conditions, grip strength, BMI and socio-economic characteristics, people living in Mediterranean countries report significantly lower health problems in each domain (the coefficient of the indicator for living in Mediterranean countries is negative and significant for each domain). Female respondents report significantly higher pain.

The second numerical column of Table 10 presents the parameter estimates using the vignettes to correct for differences in thresholds among respondents. First of all, the estimates of the actual values of the three vignettes for each domain turn out to be ordered in exactly the way we expected (from least to most health problems in each domain). This also provides some evidence that each concept being measured is likely to be unidimensional. For most health domains the estimated coefficient of the dummy for living in Mediterranean countries substantially reduce in magnitude compared to the model with constant thresholds. Furthermore, for pain such dummy is no longer significant. For concentration problems, the interaction between the dummy for female and the dummy for living in Mediterranean countries is no longer significant. Finally, for pain the female dummy reduces in magnitude compared to the model with constant thresholds. The explanation for these differences in the estimated coefficients between the model with individual specific thresholds and the model with constant thresholds is given by the estimates of threshold parameters. In fact, significant shifts in the thresholds are observed both by gender and region for all considered domains. This indicates that there are both gender and regional differences in response scales.

1.4.5 Decomposition of gender and regional differences

Analogously to the decomposition exercise computed for SRH in the first part of this paper, we now decompose gender and regional differences in the level of health problems in each domain. Because the latent model (3) is linear in such level, we can decompose the differences between any two groups, j and k, into a number of components. The first component is a "prevalence effect", capturing differences in the distributions of conditions

$$\beta'(\bar{C}_j - \bar{C}_k).$$

The "severity effect" is zero under the assumption that coefficients are the same for men and women living in different regions, The second component is the "endowment effect" of the socio-economic characteristics W

$$\gamma'(\bar{W}_j - \bar{W}_k).$$

The last component is a residual term which includes differences in the health measure that cannot be explained neither by differences in the distributions of conditions, nor by differences in the distributions of the socio-economic characteristics. Specifically, the unexplained difference between

non-Mediterranean women and non-Mediterranean men is α_1. The unexplained difference between women and men living in Mediterranean countries is $\alpha_1 + \alpha_3$. The unexplained difference between Mediterranean women and non-Mediterranean women is $\alpha_2+\alpha_3$. Finally, the unexplained difference between Mediterranean men and non-Mediterranean men is α_2.

Table 11 shows the decomposition of gender differences in the level of health problems for selected domains. The decomposition is reported for both the ordered probit model with constant thresholds and the ordered probit model with individual specific thresholds. The top panel of Table 11 shows the decomposition of the differences between non-Mediterranean women and non-Mediterranean men. Non-Mediterranean women have a higher level of health problems than non-Mediterranean men. Unexplained differences in pain are reduced from 70% to about 61% when correcting for differences in response scales. Unexplained differences in mobility and concentration problems are instead increased when correcting for differences in response scales. The bottom panel of Table 11 shows the decomposition of the differences between Mediterranean women and Mediterranean men. Mediterranean women have much higher level of health problems than Mediterranean men. Unexplained differences in pain, mobility and concentration are all reduced when correcting for differences in response scales.

Table 12 shows the decomposition of the regional differences in the level of health problems for selected domains. All regional differences are substantially reduced when correcting for differences in response scales.

1.5 Conclusions

In this paper we looked at gender and regional differences in SHR using data from Release 2 of the first (2004) wave of SHARE. Our results indicate that the difference between men's and women's health is partly explained by differences in the prevalence of the various conditions. However, a non negligible part of the difference is due to "other causes", which may possibly include gender differences in reporting own health. Furthermore, most of the regional differences in the fraction reporting poor health is unexplained by differences in health conditions and limitations or by socio-demographics characteristics. Again, this is possibly due to differences in how people report their health.

We employ the tool of "anchoring vignettes" for correcting response scales in the self-assessment of health on six domains: pain, mobility, sleeping problems, shortness of breath, concentration problems, and depression. Understanding whether and how women and men living in different regions differently report levels in these domains can give us helpful insight into differences in SRH. We find that vignettes help identifying both gender and regional differences in how respondents report their health. In particular, the fraction of gender differences in the level of health which cannot explained by chronic conditions nor by socio-economic characteristics is substantially decreased after correcting for response scales. Furthermore, after correcting for differences in response scales,

regional differences in the level of health are substantially reduced, although not entirely eliminated. Our results suggest that differences in response styles should be taken into account when using self-assessment of health in socio-economic studies. Failing to do so may lead to misleading conclusions.

References

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

Börsch-Supan A., Jürges H. (2005). *The Survey of Health, Aging, and Retirement in Europe Methodology*, Mannheim: Mannheim Research Institute for the Economics of Aging.

Case A., Paxson C. (2005). "Sex differences in morbidity and mortality", *Demography*, 42: 189–214.

D'Uva T.B., Van Doorslaer E., Lindeboom M., O'Donnell O. (2008). "Does reporting heterogeneity bias the measurement of health disparities?", *Health Economics*, 17: 351–375.

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

Jürges H. (2008). "Self-assessed health, reference levels, and mortality", *Applied Economics*, 40: 569–582.

Kapteyn,A., Smith J., van Soest A. (2007). "Vignettes and self-reports of work disability in the United States and the Netherlands", *American Economic Review*, 97: 461–473.

Kerkhofs M., Lindeboom M. (1995). "Subjective health measures and state department reporting errors", *Health Economics*, 4: 221–235.

King G., Murray C.J.L., Salomon J.A., Tandon A. (2004). "Enhancing the validity and cross-cultural comparability of measurement in survey research", *American Political Science Review*, 98: 567-583.

King G., Wand J. (2007). "Comparing incomparable survey responses: evaluating and selecting anchoring vignettes", *Political Analysis*, 15: 46–66.

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

Rantanen T., Guralnik J.M., Foley D., Masaki K., Leveille S.G., Curb J.D., et al. (1999), "Midlife hand grip strength as a predictor of old age disability", *Journal of the American Medical Association*, 21: 558–560.

Rossetti C. (2008). "Non-parametric estimation of ordered probit model with anchoring vignettes", mimeo.

Salomon J.A., Mathers C.D., Chatterji S., Sadana R., Üstün T.B., Murray C.J.L. (2003). "Quantifying individual levels of health: definitions, concepts and measurement issues", in Murray C.J.L. and Evans D.B. (eds.), *Health Systems Performance Assessment: Debates, Methods and Empiricism*, Geneva, World Health Organization.

Salomon J.A., Tandon A., Murray C.J.L. (2004). "Comparability of self rated health: cross sectional multi-country survey using anchoring vignettes", *British Medical Journal*, 328: 258–260.

Sen A. (2002). "Health: perception versus observation", *British Medical Journal*, 324: 860–861.

van Soest A., Delaney L., Harmon C., Kapteyn A., Smith J.P. (2007). "Validating the use of vignettes for subjective threshold scales", mimeo.

Table 1: Final sample by country and gender.

Country	Men			Women			Total		
	Full	Vign.	%	Full	Vign.	%	Full	Vign.	%
Germany	1,263	175	13.9	1,391	220	15.8	2,654	395	14.9
Sweden	1,313	144	11.0	1,410	154	10.9	2,723	298	10.9
Netherlands	1,256	223	17.8	1,365	221	16.2	2,621	444	16.9
Spain	870	179	20.6	1,113	212	19.0	1,983	391	19.7
Italy	1,002	153	15.3	1,183	191	16.1	2,185	344	15.7
France	1,159	318	27.4	1,412	390	27.6	2,571	708	27.5
Greece	1,102	294	26.7	1,181	275	23.3	2,283	569	24.9
Belgium	1,621	209	12.9	1,758	249	14.2	3,379	458	13.6
Total	9,586	1,695	17.7	10,813	1,912	17.7	20,399	3,607	17.7

Table 2: Descriptive statistics.

	Non-Mediterranean				Mediterranean			
	Men		Women		Men		Women	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Poor health	0.360	0.480	0.387	0.487	0.352	0.478	0.541	0.499
Heart attack	0.155	0.362	0.087	0.283	0.120	0.325	0.095	0.294
High blood pressure	0.305	0.461	0.320	0.467	0.284	0.451	0.402	0.491
High blood cholesterol	0.221	0.415	0.181	0.386	0.245	0.430	0.210	0.407
Stroke	0.045	0.207	0.020	0.141	0.012	0.107	0.026	0.160
Diabetes	0.080	0.272	0.100	0.301	0.114	0.319	0.103	0.304
Chronic lung disease	0.051	0.220	0.047	0.212	0.057	0.232	0.084	0.278
Asthma	0.043	0.204	0.030	0.171	0.046	0.211	0.036	0.187
Arthritis	0.139	0.346	0.211	0.408	0.176	0.381	0.379	0.485
Osteoporosis	0.015	0.122	0.090	0.286	0.010	0.099	0.150	0.357
Ulcer	0.073	0.261	0.037	0.189	0.081	0.274	0.042	0.200
Parkinson disease	0.007	0.085	0.002	0.044	0.002	0.039	0.002	0.041
Cataracts	0.051	0.220	0.060	0.238	0.063	0.243	0.100	0.300
Hip or femoral fracture	0.019	0.135	0.013	0.112	0.003	0.050	0.015	0.123
Reproductive cancer	0.017	0.131	0.047	0.212	0.005	0.067	0.022	0.146
Other cancer	0.036	0.185	0.027	0.162	0.051	0.220	0.024	0.153
Pain in back	0.500	0.500	0.562	0.496	0.426	0.495	0.622	0.485
Heart trouble	0.097	0.297	0.063	0.243	0.047	0.212	0.079	0.270
Breathlessness	0.110	0.313	0.086	0.280	0.081	0.273	0.106	0.308
Persistent cough	0.044	0.206	0.046	0.210	0.037	0.188	0.067	0.251
Swollen legs	0.040	0.197	0.157	0.364	0.076	0.266	0.228	0.420
Sleeping problems	0.143	0.350	0.224	0.417	0.087	0.283	0.258	0.438
Falling down	0.017	0.129	0.031	0.174	0.026	0.160	0.071	0.256
Fear of falling down	0.032	0.177	0.104	0.306	0.034	0.183	0.129	0.335
Dizziness	0.047	0.212	0.069	0.254	0.070	0.256	0.136	0.343
Stomach problems	0.124	0.330	0.151	0.359	0.110	0.314	0.203	0.403
Incontinence	0.021	0.143	0.038	0.191	0.020	0.141	0.088	0.283
Other symptoms	0.045	0.208	0.031	0.175	0.042	0.200	0.056	0.229
Low grip strength	0.274	0.446	0.161	0.368	0.349	0.477	0.330	0.471
Underweight	0.000	0.020	0.011	0.105	0.004	0.064	0.013	0.114
Overweight	0.513	0.500	0.374	0.484	0.510	0.500	0.408	0.492
Obese	0.155	0.362	0.173	0.379	0.193	0.395	0.196	0.397
Pain: mild	0.347	0.476	0.381	0.486	0.366	0.482	0.353	0.478
Pain: mod/sev/extr	0.295	0.456	0.391	0.488	0.226	0.419	0.400	0.490
Sleeping problems: mild	0.240	0.427	0.309	0.462	0.278	0.448	0.275	0.447
Sleeping problems: mod/sev/extr	0.267	0.442	0.341	0.474	0.172	0.378	0.369	0.483
Mobility problems: mild	0.235	0.424	0.268	0.443	0.177	0.382	0.201	0.401
Mobility problems: mod/sev/extr	0.218	0.413	0.233	0.423	0.124	0.330	0.252	0.435
Concentration problems: mild	0.359	0.480	0.369	0.483	0.297	0.457	0.308	0.462
Concentration problems: mod/sev/extr	0.202	0.401	0.231	0.422	0.149	0.356	0.310	0.463
Shortness of breath: mild	0.199	0.399	0.233	0.423	0.144	0.351	0.147	0.354
Shortness of breath: mod/sev/extr	0.155	0.362	0.142	0.349	0.078	0.268	0.118	0.323
Depression: mild	0.252	0.435	0.316	0.465	0.240	0.427	0.290	0.454
Depression: mod/sev/extr	0.175	0.380	0.218	0.413	0.119	0.324	0.323	0.468
Age	63.6	9.4	64.7	10.0	63.7	8.8	64.9	10.2
Living with spouse or partner	0.796	0.403	0.599	0.490	0.782	0.413	0.540	0.499
Secondary education	0.471	0.499	0.399	0.490	0.217	0.413	0.130	0.337
Post-secondary education	0.250	0.433	0.183	0.387	0.091	0.288	0.060	0.237
Log HH income	9.73	0.99	9.69	1.02	9.10	0.99	8.84	1.24
Observations	1,069		1,234		626		678	

Notes: weighted results

Table 3: Estimated coefficients of the OLS regression for poor health. (* significant at 5%; ** significant at 1%).

	Model 1	Model 2	Model 3
Heart attack	0.178 **	.	0.158 **
High blood pressure	0.070 **	.	0.064 **
High blood cholesterol	0.023	.	0.024
Stroke	0.150 **	.	0.108 **
Diabetes	0.200 **	.	0.146 **
Chronic lung disease	0.018	.	0.006
Asthma	0.028	.	0.006
Arthritis	0.157 **	.	0.105 **
Osteoporosis	0.177 **	.	0.143 **
Ulcer	0.073 *	.	0.061 *
Parkinson disease	0.379 **	.	0.287 *
Cataracts	-0.061 *	.	-0.039
Hip or femoral fracture	-0.155 *	.	-0.117 *
Reproductive cancer	0.115 **	.	0.069
Other cancer	0.303 **	.	0.258 **
Pain in back	0.074 **	.	0.004
Heart trouble	0.081 **	.	0.060 *
Breathlessness	0.219 **	.	0.178 **
Persistent cough	0.081 *	.	0.077 *
Swollen legs	0.011	.	-0.018
Sleeping problems	0.068 **	.	0.032
Falling down	-0.029	.	-0.036
Fear of falling down	0.104 **	.	0.057 *
Dizziness	0.056 *	.	0.020
Stomach problems	0.047 *	.	0.023
Incontinence	-0.065	.	-0.087 *
Other symptoms	0.085 *	.	0.050
Low grip strength	0.086 **	.	0.046 **
Underweight	-0.087	.	-0.116
Overweight	-0.011	.	-0.020
Obese	-0.017	.	-0.044 *
Pain: Mild	.	0.121 **	0.106 **
Pain: Mod/sev/extr	.	0.300 **	0.228 **
Sleeping problems: Mild	.	0.008	0.003
Sleeping problems: Mod/sev/extr	.	0.041 *	0.024
Mobility problems: Mild	.	0.121 **	0.097 **
Mobility problems: Mod/sev/extr	.	0.208 **	0.154 **
Concentration problems: Mild	.	-0.042 *	-0.043 **
Concentration problems: Mod/sev/extr	.	0.051 *	0.037
Shortness of breath: Mild	.	0.049 **	0.028
Shortness of breath: Mod/sev/extr	.	0.104 **	0.006
Depression: Mild	.	0.007	0.007
Depression: Mod/sev/extr	.	0.054 *	0.044 *
Age - 55	0.005 **	0.011 **	0.006 **
(Age - 55) squared /100	-0.005	-0.016 *	-0.014 *
Secondary education	-0.053 **	-0.039 *	-0.045 **
Post-secondary education	-0.097 **	-0.072 **	-0.077 **
Living with spouse or partner	-0.030	0.008	-0.010
Log HH income - log(12763)	-0.034 **	-0.035 **	-0.035 **
Constant	0.230 **	0.095 **	0.112 **
Observations	3,607	3,607	3,607
R^2	0.318	0.310	0.388

Notes: weighted results, country dummies omitted

Table 4: Gender differences in the impact of conditions on the probability of reporting poor health. Model 3 (* significant at 2%).

	Non-Medit.	Medit.
Heart attack[a)]	-0.037	-0.052
High blood pressure	-0.046	0.138 *
High blood cholesterol	0.115 *	0.030
Stroke	-0.039	-0.184
Diabetes	0.025	-0.092
Chronic lung disease	0.102	-0.190
Asthma	-0.036	-0.068
Arthritis	0.046	0.054
Osteoporosis	0.059	0.043
Ulcer	0.197 *	-0.206
Parkinson disease	-0.106	0.277
Cataracts	-0.185 *	0.061
Hip or femoral fracture	-0.112	0.269
Reproductive cancer	-0.128	-0.102
Other cancer	0.236 *	-0.032
Pain in back	-0.064	-0.093
Heart trouble	0.005	-0.257
Breathlessness	0.018	0.114
Persistent cough	0.082	-0.169
Swollen legs	-0.109	0.071
Sleeping problems	-0.004	-0.073
Falling down	-0.157	0.016
Fear of falling down	-0.013	-0.023
Dizziness	0.104	-0.020
Stomach problems	-0.054	0.168
Incontinence	0.216	-0.057
Other symptoms	0.093	-0.118
Low grip strength	-0.104	-0.042
Underweight	-0.258	-0.221
Overweight	0.074	-0.018
Obese	0.192 *	-0.101
Pain: Mild	0.037	0.025
Pain: Mod/sev/extr	-0.019	0.081
Sleeping problems: Mild	-0.017	0.015
Sleeping problems: Mod/sev/extr	0.024	-0.049
Mobility problems: Mild	0.005	-0.040
Mobility problems: Mod/sev/extr	0.060	-0.184
Concentration problems: Mild	-0.034	0.042
Concentration problems: Mod/sev/extr	-0.069	0.265 *
Shortness of breath: Mild	0.006	-0.052
Shortness of breath: Mod/sev/extr	-0.109	0.042
Depression: Mild	0.002	0.077
Depression: Mod/sev/extr	-0.004	-0.016
All conditions[b)]	0.907	0.972

Notes:
a) significance from t-tests of the hypothesis that the coefficients of each condition on poor health are identical for men and women
b) F-tests of the hypothesis that all the coefficients of chronic conditions on poor health are identical for men and women

Table 5: Excess prevalence of conditions in women (* significant at 2%).

	Non-Medit.[a)]	Medit.[b)]
Heart attack	-0.068 *	-0.036
High blood pressure	-0.003	0.116 *
High blood cholesterol	-0.039	-0.038
Stroke	-0.029 *	0.018
Diabetes	0.004	-0.019
Chronic lung disease	-0.008	0.016
Asthma	-0.013	-0.008
Arthritis	0.072 *	0.189 *
Osteoporosis	0.078 *	0.138 *
Ulcer	-0.039 *	-0.032 *
Parkinson disease	-0.005	0.001
Cataracts	0.004	0.018
Hip or femoral fracture	-0.007	0.014 *
Reproductive cancer	0.032 *	0.019 *
Other cancer	-0.006	-0.029 *
Pain in back	0.055 *	0.163 *
Heart trouble	-0.045 *	0.016
Breathlessness	-0.029	0.025
Persistent cough	-0.002	0.024
Swollen legs	0.112 *	0.137 *
Sleeping problems	0.091 *	0.179 *
Falling down	0.011	0.043 *
Fear of falling down	0.060 *	0.086 *
Dizziness	0.020	0.055 *
Stomach problems	0.028	0.077 *
Incontinence	0.014	0.057 *
Other symptoms	-0.012	0.005
Low grip strength	-0.132 *	-0.043
Underweight	0.012 *	0.008
Overweight	-0.135 *	-0.077 *
Obese	-0.001	-0.008
Pain: Mild	0.044	-0.013
Pain: Mod/sev/extr	0.074 *	0.151 *
Sleeping problems: Mild	0.068 *	-0.003
Sleeping problems: Mod/sev/extr	0.085 *	0.202 *
Mobility problems: Mild	0.026	0.034
Mobility problems: Mod/sev/extr	-0.012	0.097 *
Concentration problems: Mild	0.021	0.017
Concentration problems: Mod/sev/extr	-0.006	0.126 *
Shortness of breath: Mild	0.041	0.011
Shortness of breath: Mod/sev/extr	-0.025	0.029
Depression: Mild	0.067 *	0.046
Depression: Mod/sev/extr	0.024	0.182 *

Notes: Excess prevalence coefficients are the coefficients on an indicator that the respondent is female
a) in the sample of non-Mediterr. countries,
b) in the sample of Mediterr. countries,
in OLS regression for each condition, which also includes a set of control variables W

Table 6: Gender differences. Decomposition of the probability of poor health.

	Non-Mediterranean	Mediterranean
Men	0.360	0.352
Women	0.387	0.541
Difference (women - men)	0.027	0.189
Decomposition of the difference (%)		
Prevalence effect	131.6	62.0
Severity effect	30.0	2.6
Socio-dem. char.: Endowments effect	29.4	10.6
Socio-dem. char.: Coefficients effect	30.2	-41.8
Residual difference	-121.2	66.6

Table 7: Regional differences in the impact of conditions on the probability of reporting poor health. Model 3 (* significant at 2%).

	Women	Men
Heart attack[a)]	-0.074	-0.059
High blood pressure	0.080	-0.104
High blood cholesterol	0.002	0.088
Stroke	-0.000	0.145
Diabetes	-0.108	0.009
Chronic lung disease	-0.149	0.143
Asthma	0.074	0.106
Arthritis	-0.007	-0.015
Osteoporosis	-0.006	0.010
Ulcer	-0.189	0.214 *
Parkinson disease	-0.208	-0.591
Cataracts	0.148	-0.098
Hip or femoral fracture	0.380 *	-0.000
Reproductive cancer	0.212	0.186
Other cancer	0.018	0.286 *
Pain in back	-0.037	-0.008
Heart trouble	-0.098	0.165
Breathlessness	-0.072	-0.168
Persistent cough	-0.107	0.145
Swollen legs	0.187 *	0.008
Sleeping problems	0.066	0.135
Falling down	-0.106	-0.279
Fear of falling down	0.094	0.104
Dizziness	-0.086	0.038
Stomach problems	0.014	-0.208 *
Incontinence	0.002	0.275
Other symptoms	-0.187	0.024
Low grip strength	0.066	0.005
Underweight	-0.251	-0.289
Overweight	0.013	0.105
Obese	-0.096	0.197 *
Pain: Mild	0.060	0.071
Pain: Mod/sev/extr	-0.008	-0.107
Sleeping problems: Mild	-0.038	-0.070
Sleeping problems: Mod/sev/extr	-0.121	-0.049
Mobility problems: Mild	-0.017	0.028
Mobility problems: Mod/sev/extr	-0.089	0.155
Concentration problems: Mild	0.089	0.013
Concentration problems: Mod/sev/extr	0.052	-0.281 *
Shortness of breath: Mild	-0.067	-0.009
Shortness of breath: Mod/sev/extr	0.111	-0.040
Depression: Mild	0.111	0.037
Depression: Mod/sev/extr	-0.004	0.007
All conditions[b)]	1.034	1.392

Notes:
a) significance from t-tests of the hypothesis that the coefficients of each condition on poor health are identical for people living in Medit. and non-Medit. countries
b) F-tests of the hypothesis that all the coefficients of chronic conditions on poor health are identical for people living in Medit. and non-Medit. countries

Table 8: Excess prevalence of conditions in Mediterranean countries (* significant at 2%).

	Women[a)]	Men[b)]
Heart attack	-0.006	-0.006
High blood pressure	0.015	-0.042
High blood cholesterol	0.034	0.021
Stroke	0.013	-0.054 *
Diabetes	-0.043	0.007
Chronic lung disease	0.063 *	-0.006
Asthma	-0.001	0.008
Arthritis	0.260 *	0.129 *
Osteoporosis	0.115 *	-0.008
Ulcer	0.017	0.014
Parkinson disease	0.000	-0.008
Cataracts	0.026	0.004
Hip or femoral fracture	-0.015	-0.019
Reproductive cancer	-0.029	-0.011
Other cancer	0.002	0.014
Pain in back	0.052	-0.115 *
Heart trouble	0.005	-0.059 *
Breathlessness	0.016	-0.015
Persistent cough	0.022	-0.025
Swollen legs	0.071 *	0.056 *
Sleeping problems	0.044	-0.048
Falling down	0.059 *	0.002
Fear of falling down	0.003	-0.014
Dizziness	0.077 *	0.016
Stomach problems	0.095 *	0.003
Incontinence	0.078 *	-0.011
Other symptoms	-0.002	-0.027
Low grip strength	0.185 *	0.030
Underweight	0.022 *	0.009 *
Overweight	-0.070	-0.016
Obese	-0.066 *	-0.038
Pain: Mild	0.100 *	0.121 *
Pain: Mod/sev/extr	-0.066	-0.099 *
Sleeping problems: Mild	-0.008	0.047
Sleeping problems: Mod/sev/extr	0.101 *	-0.040
Mobility problems: Mild	-0.115 *	-0.115 *
Mobility problems: Mod/sev/extr	-0.047	-0.209 *
Concentration problems: Mild	0.049	0.001
Concentration problems: Mod/sev/extr	0.005	-0.117 *
Shortness of breath: Mild	-0.075 *	-0.048
Shortness of breath: Mod/sev/extr	-0.037	-0.151 *
Depression: Mild	0.078 *	-0.018
Depression: Mod/sev/extr	0.053	-0.114 *

Notes: Excess prevalence coefficients are the coefficients
on an indicator that the respondent lives in Medit. countries
a) in the sample of women,
b) in the sample of men,
in OLS regression for each condition, which also includes
a set of control variables W

Table 9: Regional differences. Decomposition of the probability of poor health.

	Women	Men
non-Med.	0.387	0.360
Medit.	0.541	0.352
Difference (Medit. - non-Med.)	0.154	-0.008
Decomposition of the difference (%)		
Prevalence effect	29.3	762.5
Severity effect	24.9	-832.2
Socio-dem. char.: Endowments effect	31.0	-666.5
Socio-dem. char.: Coefficients effect	-13.0	-666.4
Residual difference	27.8	1502.6

Table 10: Ordered probit model with constant thresholds, and ordered probit model with individual specific thresholds for selected health domains (* significant at 5%; ** significant at 1%).

		Pain		Mobility problems		Concentration problems	
Equation	Variable	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds
SA	medit	-0.154 *	-0.098	-0.449 **	-0.293 **	-0.294 **	-0.144 *
	female	0.277 **	0.215 **	0.016	0.069	0.001	0.005
	medit*female	-0.096	-0.172	0.142	0.042	0.272 **	0.109
	Constant	-0.309 **	-0.462 **	-0.891 **	-0.964 **	-0.103	-0.139
Thres. 1	medit	.	-0.036	.	0.153 **	.	0.179 **
	female	.	-0.151 **	.	0.027	.	-0.017
	medit*female	.	0.068	.	-0.067	.	-0.127 *
	Constant	0.000	0.000	0.000	0.000	0.000	0.000
Thres. 2	medit	.	0.158 **	.	-0.015	.	-0.100 *
	female	.	0.145 **	.	0.080	.	0.052
	medit*female	.	-0.230 **	.	-0.108	.	-0.046
	Constant	1.203 **	-0.070	0.827 **	-0.410 **	1.031 **	-0.073
Vign. 1	Constant	.	0.477 **	.	0.762 **	.	0.503 **
Vign. 2	Constant	.	1.612 **	.	1.470 **	.	1.246 **
Vign. 3	Constant	.	2.120 **	.	1.657 **	.	1.884 **
ln σ	Constant	.	-0.146 **	.	-0.192 **	.	-0.151 **
Obs.			3,607		3,607		3,607
LR test			439.7		413.0		394.4
p(LR test)			0.0		0.0		0.0

Notes: weighted results, only the coefficients of the gender and regional dummies are reported for parsimony. LR test is a Likelihood ratio test of the hypothesis of no DIF (constant thresholds)

Table 11: Gender differences. Decomposition of health level in selected domains.

Non-Mediterranean countries						
	Pain		Mobility problems		Concentration problems	
	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds
Men	0.512	0.212	-0.109	-0.486	0.156	-0.003
Women	0.907	0.562	0.007	-0.265	0.266	0.109
Difference (women - men)	0.395	0.350	0.117	0.221	0.110	0.112
Decomposition of the difference (%)						
Prevalence effect	24.5	22.1	45.4	24.5	26.5	3.5
Socio-dem. char.	5.5	16.3	40.5	44.1	73.0	91.7
Residual difference	70.0	61.5	14.1	31.4	0.5	4.8
Mediterranean countries						
	Pain		Mobility problems		Concentration problems	
	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds
Men	0.324	0.150	-0.573	-0.704	-0.123	-0.084
Women	0.928	0.569	-0.074	-0.271	0.414	0.236
Difference (women - men)	0.605	0.420	0.499	0.433	0.537	0.320
Decomposition of the difference (%)						
Prevalence effect	66.1	75.3	56.7	51.1	32.2	26.3
Socio-dem. char.	4.0	14.3	11.7	23.1	17.1	37.7
Residual difference	29.9	10.4	31.7	25.7	50.7	36.0

Table 12: Regional differences. Decomposition of health level in selected domains.

Women						
	Pain		Mobility problems		Concentration problems	
	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds
non-Med.	0.907	0.562	0.007	-0.265	0.266	0.109
Medit.	0.928	0.569	-0.074	-0.271	0.414	0.236
Difference (Medit. - non-Med.)	0.021	0.008	-0.081	-0.005	0.148	0.126
Decomposition of the difference (%)						
Prevalence effect	1307.6	2755.4	-280.0	-3335.5	58.7	41.8
Socio-dem. char.	-14.2	843.4	0.3	-1305.0	56.0	85.3
Residual difference	-1193.4	-3498.7	379.7	4740.5	-14.7	-27.0
Men						
	Pain		Mobility problems		Concentration problems	
	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds
non-Med.	0.512	0.212	-0.109	-0.486	0.156	-0.003
Medit.	0.324	0.150	-0.573	-0.704	-0.123	-0.084
Difference (Medit. - non-Med.)	-0.188	-0.062	-0.463	-0.217	-0.280	-0.081
Decomposition of the difference (%)						
Prevalence effect	15.2	42.6	0.7	-4.3	20.4	33.7
Socio-dem. char.	2.8	-99.3	2.4	-30.5	-25.3	-110.8
Residual difference	82.0	156.7	96.8	134.7	105.0	177.1

Figure 1: Fraction reporting poor health by country and gender.

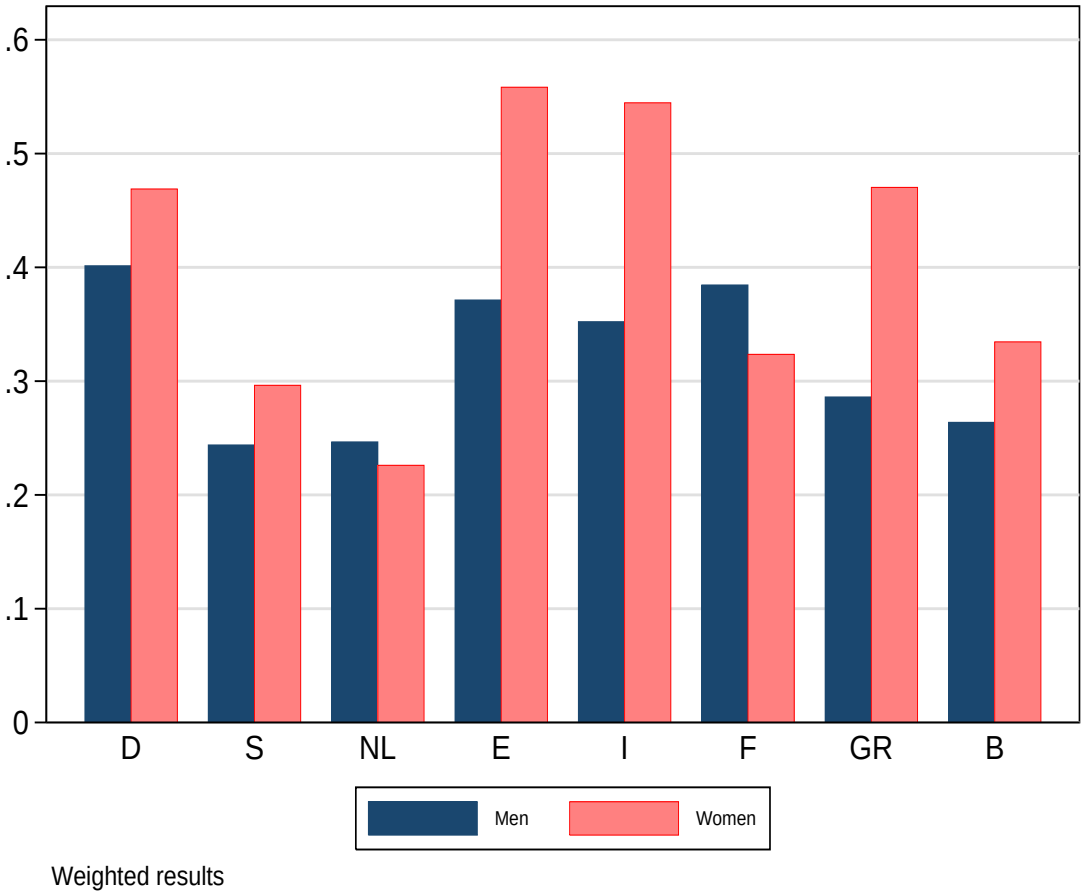
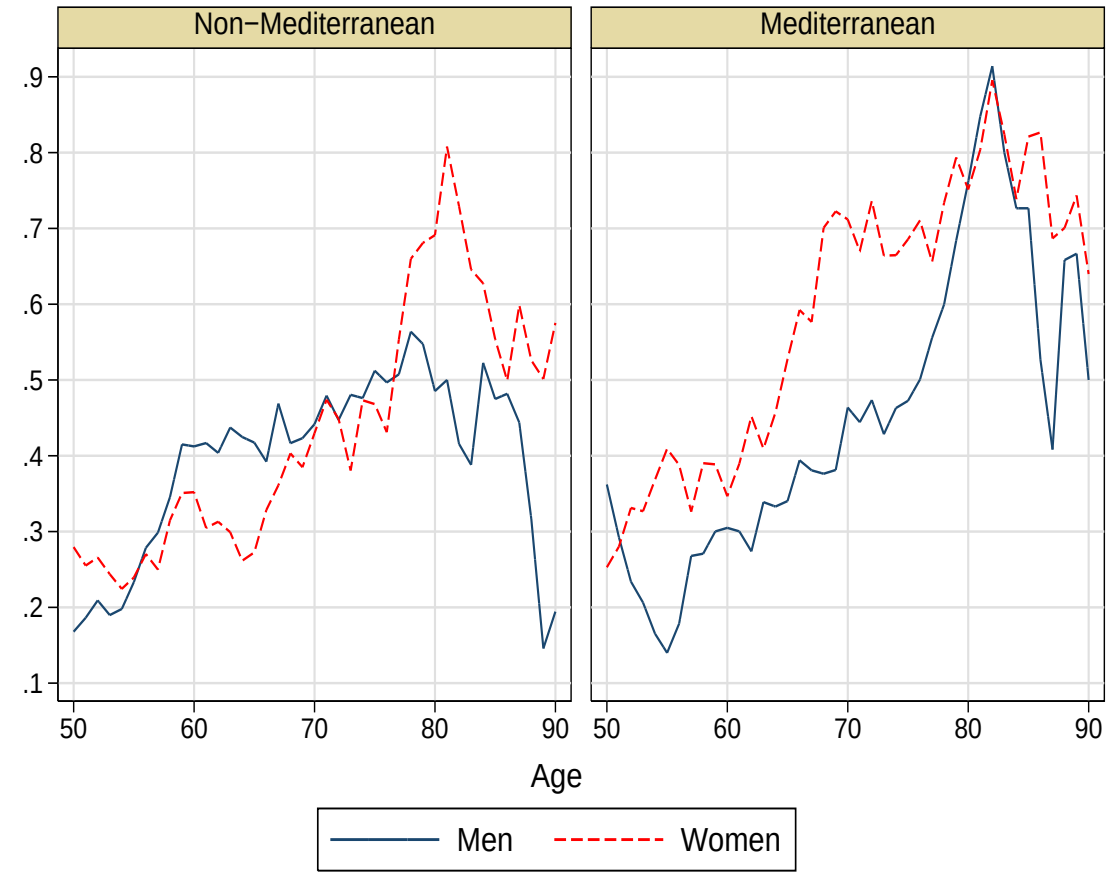
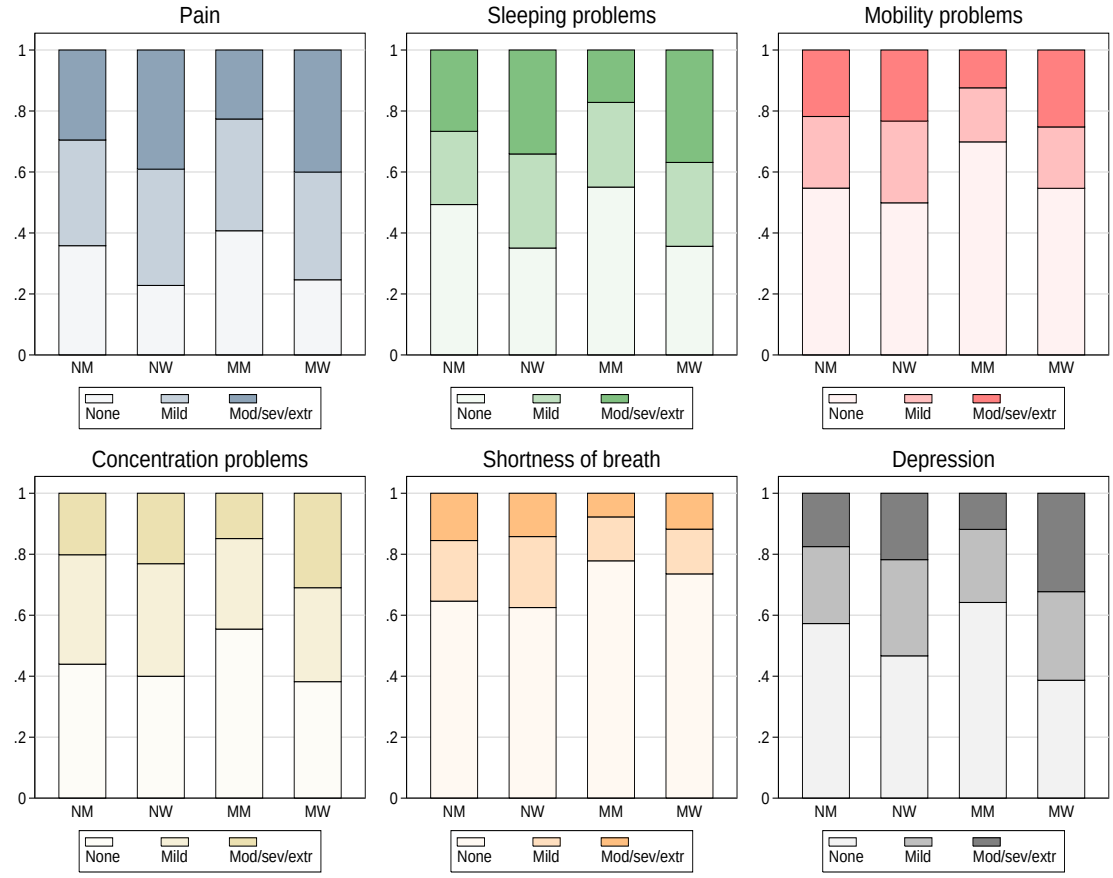


Figure 2: Fraction reporting poor health by age, region and gender.



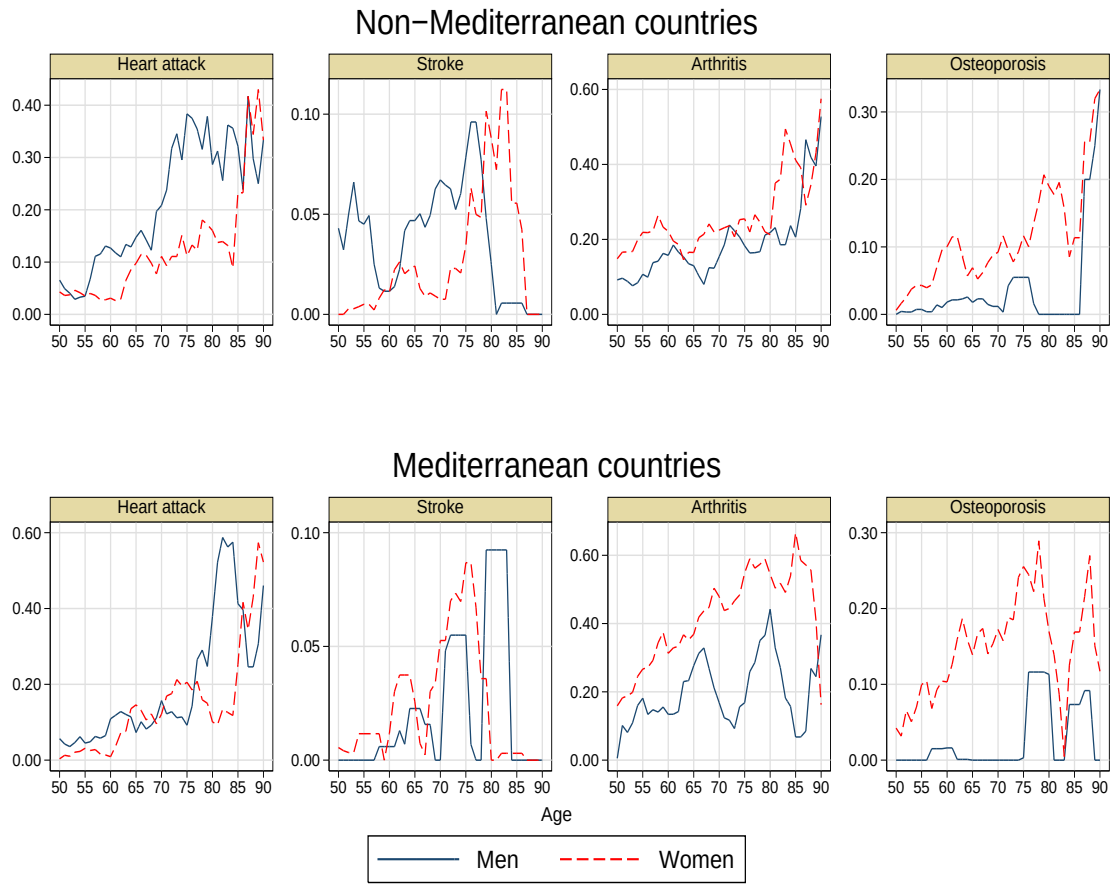
Weighted results

Figure 3: Histograms of self assessments for health domains by region and gender.



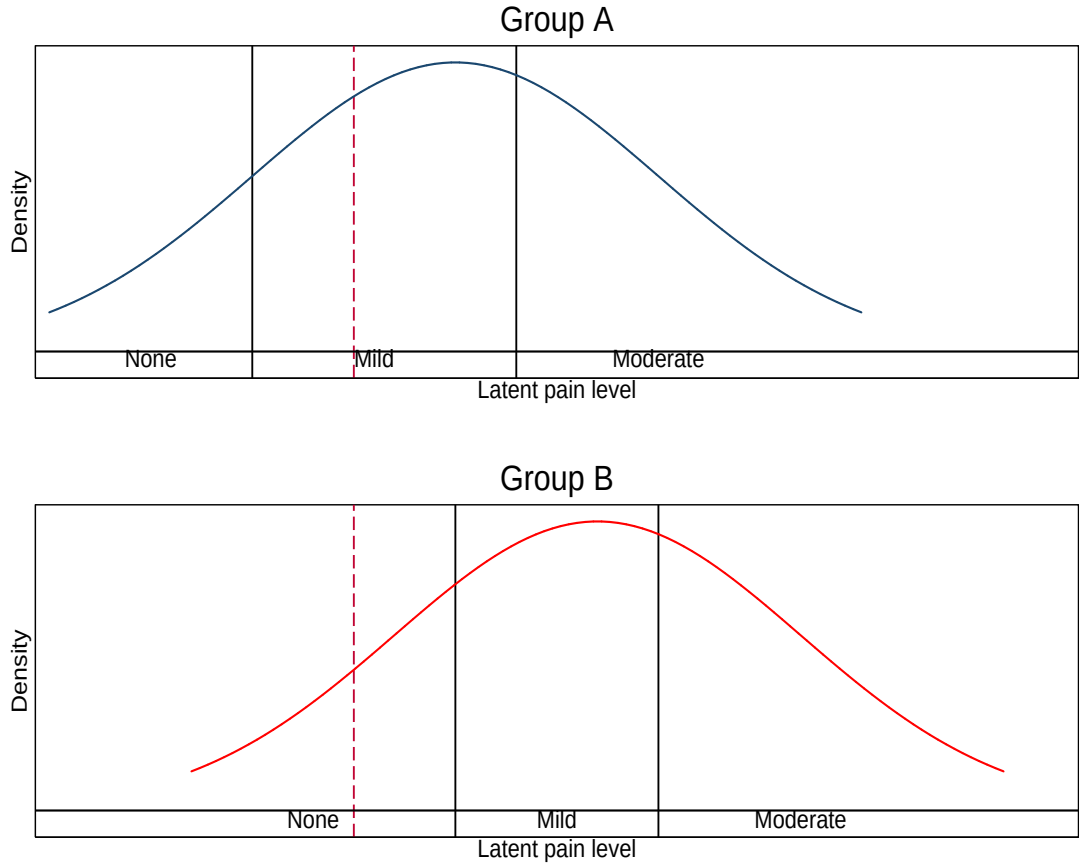
Weighted results

Figure 4: Prevalence rates of some selected conditions by gender and region.



Weighted results

Figure 5: Example of comparison of self-assessed pain in two groups in case of differences in response scales.



2 Chapter 2
Medical drug compliance, co-payment and health outcomes

Abstract[10]: This paper studies the relationship between medical compliance and health outcomes – hospitalization and mortality rates – using a large panel of patients residing in a local health authority in Italy. These data allow us to follow individual patients through all their accesses to public health care services until they either die or leave the local health authority. We adopt a disease specific approach, concentrating on hypertensive patients treated with ACE-inhibitors. Our results show that medical compliance has a clear effect on both hospitalization and mortality rates: health outcomes clearly improve when patients become more compliant to drug therapy. At the same time, we are able to infer valuable information on the role that drug co-payment can have on compliance, and as a consequence on health outcomes, by exploiting the presence of two natural experiments during the period of analysis. Our results show that drug co-payment has a strong effect on compliance, and that this effect is immediate.

Keywords: Health outcomes, compliance, health policy reforms, prescription charges, co-payment, natural experiments, panel data.

JEL classification codes: C35, C81, D12, I12.

[10] Financial support from Pfizer is gratefully acknowledged. I also thank Alessandro Chinellato for many valuable comments and ideas regarding the clinical aspects of the research involved, and Alan Garber, David Meltzer and Arthur van Soest for their comments on previous drafts of this paper, and participants to the NBER Summer Institute 2005. This chapter is based on the article "Drug compliance, co-payment and health outcomes: Evidence from a panel of Italian patients", written with Vincenzo Atella, Franco Peracchi, and Domenico Depalo.

2.1 Introduction

In developed countries, the increase in the cost of health care services has produced a vast concern among policy makers, who have enforced restrictive measures to contain those trends. This phenomenon has been particularly relevant for drug costs, who have recorded higher increases (in both volumes and prices) compared to other major components of healthcare spending (Jacobzone, 2000). Health economists have extensively studied the effects of such policies on drug expenditure, and a large literature on this subject is available.[11]

Much less is known, however, about the effect of cost containment measures on drug compliance and, as a consequence, on health outcomes. Not complying with medication, possibly because of affordability issues, can have serious consequences for health. For example, Dracup and Meleis (1982) report evidence that 80% compliance to a medication regimen for hypertension lowers blood pressure to normal, whereas 50% compliance is ineffective. Shaw *et al.* (1995) report evidence that poor adherence to drug therapy decreases the effectiveness of antihypertensive treatment, whereas IMS Health (2000) shows that even interrupting hypertensive treatment by just seven days can increase the risk of stroke.

When a co-payment is established, patients must contribute towards the cost of their medication and health care use. Several empirical studies have found that the demand for prescription drugs is reduced by a direct contribution from the patient, although the overall impact appears to be quite limited, with estimated price elasticities ranging from -.1 to -.6. Unfortunately, as pointed out by Freemantle and Bloor (1996), drug reimbursement may reduce the use of both essential and non-essential drugs. Although the reduction in the use of non-essential drugs has been shown to be greater (McManus *et al.*,1996), the concern remains that essential medication may be affected.

Following this line of research, Atella *et al.* (2005) investigate the role that increasing out-of-the-pocket expenditure can have on consumers' attitudes to adopt strategies to contain the cost of medication. Using micro-data from two surveys, conducted in Italy and the UK respectively, they find a tendency for both British and Italian patients suffering from hypertension and dyspepsia to use cost reducing strategies which are strongly influenced by income and drug affordability problems. Reduction in compliance (defined as strategies that either induce patients to not obtain their medication at all, or to select fewer prescribed drugs or lower their dosage) is one of the main strategies used. Piette, Heisler and Wagner (2004) find similar evidence in the USA, suggesting that cost remains a significant barrier to health care for many adults, especially among the uninsured and the low-income elderly population.

Further evidence has been provided by Case *et al.* (2004), who explore directly the relationship between income level and medical compliance for hypertensive patients through an *ad hoc* survey carried out in an urban township of South Africa. They find that the fraction of hypertensive

[11] Main studies on the topic include Leibowitz *et al.* (1985), Soumerai *et al.* (1987), O'Brien (1989), Harris *et al.* (1990), Ryan and Birch (1991), Hughes and McGuire (1995) and Atella (1999), Atella (2003).

patients who report to be low compliant is about 47% at the top income quintile, but it jumps to 75% at the bottom the income quintile.

Due to the cross-sectional nature of their data, these studies are unable to study the link between compliance and health outcomes. Our goal is to fill this gap by using a unique longitudinal data set collected for an Italian province and covering the period from 1997 to 2002. Following Atella *et al.* (2005) and Case *et al.* (2004), our analysis is disease specific. We are able to obtain some evidence on the relationship between co-payment, compliance and health outcomes by exploiting the changes introduced by the Italian government, in January 2001, September 2001 and March 2002, to the rules regarding co-payment on drugs provided by the National Health Service (NHS) through GP prescriptions. These changes may be regarded as "natural experiments" that allow us to use a difference-in-difference approach to detect statistically significant differences in the behavior of "high compliant" versus "low compliant" patients "before" and "after" the changes.

The remainder of this paper is organized as follows. Section 2.2 describes the data. Section 2.3 describes our drug-specific approach. Section 2.4 discusses our indicator of compliance. Section 2.5 looks at the relationship between compliance and health outcomes. Section 2.6 investigates if and how health policy changes affect compliance. Finally, Section 2.7 offers some conclusions.

2.2 The data

Our data comes from three administrative registries maintained by the Pharmaceutical Service Department of ULSS 9, the public health agency covering the southern part of the Italian province of Treviso. The first registry is the drug prescription database, which contains records of patient prescriptions, including date of dispensing, amount and Anatomical Terapeutical and Clinical Classification (ATC) code of substance dispensed, unit price and number of packages dispensed. It also includes gender and date of birth of the patient receiving the medications, a unique anonymized patient identifier, a unique anonymized identifier of the practitioner who prescribed the medication, and gender, date of birth and typology–whether general practitioner (GP) or specialist (SP)–of the practitioner. The second is the hospitalization registry, which contains records of each single hospitalization, including date of entry and dismissal, primary Diagnosis Related Groups (DRG), and cost of hospitalization. Unfortunately, this registry contains no information on drugs dispensed during a hospitalization period. Through the anonymized personal identifiers, we were able to link patient prescription and hospitalization information to the third registry, the death and transfer registry. The resulting dataset allows us to follow individual patients through all their accesses to public health care services until they either die or leave the local health authority. Data are available from 1993 for drug prescriptions and from 1997 for hospitalizations. For a detailed description of the administrative data used here, see Appendix F.

2.3 Hypertensive patients

Patients may behave differently in terms of compliance depending on the pathology they suffer from or the treatment they receive. For example, a chronic "asymptomatic" pathology (such as hypertension) leads to patterns of compliance that are different from those involved in acute "painful" pathologies (such as headache). Focusing on specific pathologies or specific drug treatments offers the advantage of exploring consumer decision-making in relation to specific clinical conditions and, subsequently, allowing us to derive more precise conclusions concerning the determinants of compliance and the effects of compliance on health outcomes.

In this paper we focus on patients treated with active ingredients in the ATC class C09AA, corresponding to Angiotensin Converting Enzyme inhibitors (ACE-inhibitors).[12] These active ingredients are the most important for the Italian NHS in terms of expenditure, accounting for about 9% of total public drug expenditure in 2003, and are mainly employed in the treatment of hypertension. According to evidence gathered by Health Search,[13] in 2003 about 80% of the prescriptions of ACE-inhibitors were issued for treating hypertension[14] and the remaining 20% for other uses.

Hypertension is a chronic asymptomatic pathology affecting a large share of the Italian population. According to ISTAT (2001), about 20% of the Italian adult population suffers of hypertension and its prevalence increases with age (37% at age 55–64, 50% at age 65–7 4 years, and 67% at age 75+). Because hypertension is an asymptomatic condition, patients do not generally feel ill because of high blood pressure. In this case, compliance with anti-hypertensives is often problematic (McInnes, 1999). Hypertension treatment is generally long-term, and this may have important economic implications as patients receive regular prescriptions, thus incurring regular costs. The large prevalence also affects the public budget. Finally, hypertension is an interesting condition to study from the viewpoint of health outcomes. In fact, left untreated, it can have serious consequences in terms of hospitalization and mortality.

[12] ACE-inhibitors block conversion of Angiotensin I into Angiotensin II. Angiotensin II is a very powerful chemical which causes the muscles surrounding blood vessels to contract and thereby narrows the blood vessels. The narrowing of the vessels increases the pressure within the vessels and can cause high blood pressure (hypertension). Angiotensin II is formed from Angiotensin I by the "angiotensin converting enzyme" (ACE). ACE-inhibitors are medications that slow (inhibit) the activity of such enzyme, which then reduces the production of Angiotensin II. As a result, blood vessels can dilate and blood pressure is reduced. Lower blood pressure makes it easier for the heart to pump blood, thus reducing the probability of heart failure. In addition, the progression of kidney disease due to high blood pressure or diabetes is slowed.

[13] Health Search is a network of Italian GPs that records information on drug prescriptions and related pathology. In 2003, the network had 320 member GPs covering 465,200 patients (for a total of 3,826,000 prescriptions).

[14] This is in contrast with the experience from other countries. According to OSMED (2005), "contrary to what has emerged in the most recent studies of hypertension, ..., the prescription of amlodipin, doxazosin, ACE inhibitors and angiotensin II inhibitors continues to increase [in Italy]. The prescriptive behavior of Italian clinics seems to be guided mostly by the European guidelines regarding the therapy for arterial hypertension, as opposed to the American behavior whose priority is to obtain a reduction in the pressure values rather than recommend a specific pharmacological choice".

2.4 Drug compliance

Drug compliance may be defined as "the extent to which the patient's actual history of drug administration corresponds to the prescribed regimen" (Vermeire *et al.* 2001, p. 332). Thus, at the individual level, an ideal index of drug compliance would be

$$c_{ij}^* = \frac{C_{ij}}{P_{ij}},$$

where C_{ij} is the amount of substance (active ingredient) j consumed by patient i in a given period, P_{ij} is the amount of substance j prescribed for the same period to patient i by her physician given her health characteristics, and C_{ij} and P_{ij} are measured in the same suitably defined unit.

Although such an indicator is in principle straightforward, the question of how to measure the numerator and the denominator has vexed many researchers. Drug consumption is typically hard to measure and most datasets only contain information on drug purchased or dispensed, whereas the actual amount of drug prescribed to a particular patient is typically unavailable. Generally available are only guidelines that specify the amount of active ingredients recommended for the typical or average treatment of a specific pathology. For example, guidelines for the treatment of hypertension have been published by the WHO (1999). Of course, "average dosage" or "international standards" represent an imperfect measure in the construction of an indicator of compliance, as physicians may decide to prescribe different dosages for specific patients under specific conditions.

As an alternative definition, Vermeire *et al.* (2001) propose that "compliance is the extent to which a person's behavior in terms of taking medications ... coincides with medical and health advice". Operationally, medical advice could be approximated by international guidelines or national standards. Therefore, instead of c_{ij}^*, one may work with

$$c_{ij} = \frac{D_{ij}}{\bar{P}_j},$$

where D_{ij} is the amount of substance j purchased by patient i in a given time period and $\bar{P}_j$ is the average amount of substance j that should be prescribed to a patient for the same period according to international guidelines or national standards. The relationship between the measured index c_{ij} and the ideal index c_{ij}^* is

$$c_{ij} = c_{ij}^* \frac{D_{ij}}{C_{ij}} \frac{P_{ij}}{\bar{P}_j}.$$

It is plausible to assume that $D_{ij} \geq C_{ij}$, so $c_{ij} \geq c_{ij}^*$ whenever $P_{ij} \geq \bar{P}_j$. The term $P_{ij}/\bar{P}_j$ represents an important source of unobserved heterogeneity that is unlikely to be independent of a patient's observable characteristics.

With regard to the choice of measurement unit, the WHO adopts the Defined Daily Dose (DDD), which represents the average maintenance dose per day for a substance used in its main indication

on adults. A DDD is not a recommended dose and may not represent a real dose. However, being a measurement unit, DDDs can be added and compared across different products.[15] As a consequence, compliance across groups of drugs may be compared between patients, practices, health authorities, and regions. This allows us to derive compliance indicators for different active ingredients that are themselves comparable and additive. We can therefore measure the compliance of a single patient without having to distinguish between active ingredients used. For the same reason, we can account for multi-therapies.

Prescription practices in individual countries may differ significantly from international standards because of both the existence of different indications for the same drug[16] and different prescribing habits of GPs compared to international standards. As an example, Table 13 shows, for each active ingredients in the class of ACE-inhibitors, the differences between the DDDs provided by the WHO and the average daily dosages according to the Italian drug prescription practice (for short, ADD). The main differences are for Enalapril, Lisinopril and Ramipril, for which the Italian ADDs are twice the WHO DDDs. Notice that these three substances represent more than half of total dispensing of ACE-inhibitors in Italy.

Taking the Italian ADDs as the measurement unit and the year as the time unit, our index of drug compliance for unit i is

$$\bar{c}_{ij} = \frac{\sum_{t=1}^{T_i} D_{ijt}}{ADD_j \times T_i} = \frac{\bar{D}_{ij}}{ADD_j},$$

where $\sum_{t=1}^{T_i} D_{ijt}$ is total amount of doses of substance j dispensed to patient i over the T_i days for which she is observed during a year, and $\overline{D}_{ij} = T_i^{-1} \sum_{t=1}^{T_i} D_{ijt}$ is the average daily dosage of substance j dispensed to patient i during a year. Thus, our index of annual compliance is simply the ratio between the average daily purchase and the Italian ADD.

Since a therapy may include more than one active ingredient, adding over all J active ingredients in the ACE-inhibitor class gives our measure of annual compliance for the ith patient

$$\bar{c}_i = \sum_{j=1}^{J} I_{ij}\, \bar{c}_{ij} = \frac{\sum_{j=1}^{J} I_{ij}\, \bar{D}_{ij}}{\sum_{j=1}^{J} I_{ij}\, ADD_j},$$

where I_{ij} is equal to 1 if substance j is included in patient i's therapy and is equal to zero otherwise.

Problems arise when patients undergo therapy only for certain periods, based on physician advice. Consider for example the case of a patient with recorded prescriptions only for the first half of the year. Should this patient be considered "fully" or "half" compliant? Similarly, when the therapy is interrupted for a long period, we may wonder whether this reflects non-compliance

[15] We can add up DDDs of different active ingredients prescribed and dispensed to the same individual because our analysis is based only on plain active ingredients, thus excluding drugs with combinations of active ingredients, such as drugs composed by "diuretics" and "ACE-inhibitors".

[16] For example, the DDD for quinine is based on the dose used for malaria prophylaxis (1200mg) whereas in England its main indication is the treatment of leg cramps (300mg).

by patients or perfect adherence to medical advice of stopping the therapy. Unfortunately, our data records patient information only if they interact with the system. We therefore decided to drop from our sample all those patients who present missing values for one year or more over the observation period.

A further problem is the fact that, when patients are hospitalized, drugs are dispensed directly by the hospital pharmacy and are not recorded in the pharmaceutical registry. This would lead to underestimate compliance. We therefore correct the doses purchased by hospitalized patients by assuming that, when hospitalized, they are treated according to the Italian ADD. We then add this amount to the doses purchased through pharmacies. The importance of this correction is larger for older patients, as hospitalization rates tend to increase with age.

2.5 Drug compliance and health outcomes

This section looks at the relationship between drug compliance and health outcomes. After describing our sample selection criteria, we analyze the variability of compliance across socio-demographic groups. We then consider how compliance and other socio-demographic characteristics help predict health outcomes such as hospitalization and mortality rates.

2.5.1 Sample selection and descriptive statistics

We start with all patients, born between 1910 and 1960, who were prescribed at least one drug in the ACE-inhibitor class at any time during the period 1993–2002. We restrict attention to these cohorts because they comprise the bulk of the population suffering of hypertension. Reliable data on hospitalization is only available from 1997, and so we focus on the 6-year period from 1997 to 2002. This results in an unbalanced panel of 40,168 patients and 159,959 observations. We drop patients with compliance greater than 2 (505 patients and 1,827 observations dropped) because they might be outliers or cases of co-morbidity, and patients who were hospitalized for renal diseases but not for cardiovascular diseases (1,270 patients and 4,943 observations dropped). Although we may miss hypertensive patients who are not treated with ACE inhibitors, we are confident that we avoid selecting non-hypertensive patients. Our final sample consists of an unbalanced panel of 38,393 patients and 153,189 observations, with an average of 4 annual observations per patient.

Figure 6 shows hospitalization and mortality rates by age and gender in our sample. Patients treated with ACE-inhibitors present higher hospitalization rates than those treated with other cardiovascular drugs. In either case, hospitalization rates are higher for men than for women at almost all ages. Mortality rates are close to zero until about age 55 for men and about age 60 for women. It is only after age 55 that men experience significantly higher mortality rates than women. After age 65, patients treated with ACE inhibitors tend to have higher mortality rates than those treated with other cardiovascular drugs.

Table 14 reports summary statistics of the variables in our sample by gender. We split the sample by gender and consumption pattern, with patients classified as "regular" if their drug purchases are strictly positive in each year since they entered the sample, and as "occasional" if their purchases are zero for at least one year after they entered the sample. This distinction is important given the "chronic" nature of hypertension, as patients affected by this disease are supposed to be under continuous treatment.

Variables `y1997-y2002` are dummy variables for the years from 1997 to 2002. The average age of male patients is about 66 years, while the average age of female patients is about 70 years. This reflects the higher life expectancy of women. The variable `large pack size` is a dummy variable equal to one for a large pack size (28-pill package) and equal to zero for a normal pack size (14-pill package). According to our data, about 60% of regular patients purchase large packages. The share of occasional patients who purchase large packages is slightly lower (about 55%). Notice that, other things being equal, purchasing a large package means halfing the time spent meeting the practitioner to get a prescription and visiting a pharmacy to cash the prescription. Variable `No Enalapril` is a dummy variable equal to one if a patient is treated with only one ACE-inhibitor different from Enalapril and equal to zero otherwise. Average age of prescribing physicians is about 48 years, and over 80% of them are males. Patients whose prescription were written directly by a specialist, rather than a GP, are less than 1%. This does not mean that specialists have a marginal role in Italy, but rather that it is uncommon for a specialist to write out a prescription herself. For this reason we decided not to use this variable in our empirical analysis. Both hospitalization and mortality rates are higher for men than for women. Regular patients have higher hospitalization and mortality rates than occasional patients. This is consistent with the fact that regular patients have higher probability of being affected by cardiovascular diseases. Average compliance is slightly higher for men than for women. For regular patients average compliance is three times higher than for occasional patients. Occasional patients have positive purchases for only half of the years they are in the sample and, when they purchase, they tend to purchase less than regular patients.

Figure 7 shows the histogram of our measure of annual compliance, $\bar{c}_i$. The histogram peaks at values equal to .25, .50, .75, and 1. Notice that the number of patients with compliance values above 1.5 is only 1 percent.

As a summary of our data, the first two columns of Table 15 report the fraction of regular patients by gender and age group, while the other four columns report the sample mean and sample standard deviation of compliance of regular patients by gender, age group and pack size (small or large). Average compliance differs little by gender. For both men and women, the relationship between age and average compliance has an inverse U-shape, reaching a maximum in the 60–69 age range. On the other hand, buying a large (28-pill) package increases a patient's compliance considerably.

2.5.2 Modeling the probability of hospitalization and mortality

We now present the results of fitting simple parametric models for the probability of hospitalization and mortality in year $t+1$ as functions of annual compliance in year t, controlling for demographic and other characteristics. To reduce the amount of unobserved heterogeneity in the data, we further select the sample by dropping patients with annual compliance below .1, as they may be affected by mild hypertension that could be treated simply by a healthy diet and by reducing stress factors. This further selection produces an unbalanced panel of 18,626 patients and 65,956 observations, with an average of 3.5 annual observations per patient.

The basic model for the probability of future hospitalization (Model 1) is a logit model whose covariates include cubic polynomials in age and annual compliance in the current year, indicators for using more than one ACE-inhibitor and for patients who have also been prescribed other cardiovascular drugs (multi-therapy), and a set of time dummies. The dummy for multi-therapy is included to control for confounding effects for which we do not have adequate information.[17]

In addition to the basic model, we consider two specifications that include a richer set of covariates. The first (Model 2) controls for the type of ACE-inhibitor by including a dummy for consuming an ACE-inhibitor different from Enalapril. The second (Model 3) also controls for the gender and the age of the GP.[18]

Table 16 presents the coefficients of the estimated models, separately for men and women. The bottom part of the table presents likelihood ratio test statistics for the joint significance of certain covariates or subsets of covariates (age, compliance, GP characteristics and calendar time).[19] The intercept of the basic model corresponds to the log-odds for the baseline case, namely a person aged 55, observed in 1998, with annual compliance equal to 1, under mono-therapy and taking only one kind of ACE-inhibitor. For the baseline case, women have a slightly lower probability to be hospitalized than men. For both men and women, the effect of compliance is highly statistically significant and implies a U-shaped effect of compliance on hospitalization rates. The coefficients on the indicator for the use of more than one active ingredient and on the indicator of multi-therapy are both positive and highly significant. Notice that these dummy variables may be proxies for a patient's poor health. The negative and significant coefficients on the dummy for the year 2002 will be discussed more thoroughly in Section 2.6. Hospitalization rates tend to be significantly lower for patients treated with ACE-inhibitors other than Enalapril (Model 2). Age and gender of the practitioner, instead, are only weak predictors of health outcomes (Model 3). In general, Model 2 fits the data better than Model 1, whereas Model 3 represents only a marginal improvement over Model 2.

[17] We also fitted the models separately for patients under mono- and multi-therapy obtaining very similar results.

[18] Random effects versions of all three models shows little evidence of unobserved heterogeneity.

[19] Statistical significance is based on asymptotic standard errors that are robust to heteroskedasticity and clustering arising from the panel structure of the data.

The top panels of Figure 8 compare the hospitalization rates actually observed by gender and compliance level with the fitted probabilities from Model 2. The shaded areas are (asymptotic) 2-standard error bands for the fitted probabilities. We do not show fitted probabilities for levels of compliance greater than 1.55 because the fraction of patients with annual compliance above this value is only 1 percent. The graphs indicate a U-shaped relationship between hospitalization and compliance, with a minimum around the value of 1 (the "optimal value" of our index of compliace). For both men and women, the probability of future hospitalization for cardiovascular problems falls as current compliance moves toward the value of 1. In particular, for male patients, the probability of future hospitalization falls from about 12% when compliance is near .15 to 8% when compliance is close to 1. For female patients the reduction is less pronounced, but the lowest hospitalization rate again corresponds to compliance near 1.

We fitted similar logit models for the probability of future mortality, including an additional dummy variable for hospitalization for cardiac illness in the current year in order to account for the different health of the patients. As before, all models were fitted separately by gender.

Table 17 presents the estimated coefficients. The intercept of the basic model corresponds to the log-odds for the baseline case, namely a person aged 55, observed in 1998, with current compliance equal to 1, taking only one kind of ACE-inhibitor, under mono-therapy, and not hospitalized in the current year. The goodness of fit is always higher than for the hospitalization model, mainly because of the higher predictive power of age in explaining mortality. For the baseline case, the probability of death is twice as high for men than for women. As for the hospitalization model, the estimated effect of annual compliance is U-shaped. The coefficient on hospitalization in the current year is large and positive, and is slightly larger for women than for men.

The bottom panels of Figure 8 compare the mortality rates actually observed by gender and compliance level with the fitted probabilities from Model 2. Focusing on male patients, we estimate that increasing current compliance from .15 to 1 would reduce future mortality by half.

2.6 Health policy changes and compliance

The results in Section 2.5.2 indicate that compliance helps predict future health outcomes. We now investigate the link between health policy changes and compliance. If the relationship between compliance and health outcomes may be interpreted as causal, then the existence of such link may have important implications for public policy, because it implies that health policy changes may affect health outcomes by changing the level of compliance.

Alan *et al.* (2002), (2005) and Poirier *et al.* (1998) analyze the effect of public prescription drug programs on out-of-pocket household drug expenditure in Canada. For Italy, Atella and Rosati (2002) find that the drug policy reforms during the 1990s and in 2001–although effective in controlling public expenditure–caused undesired redistributive effects, by penalizing mostly the frailest groups in the population. All these studies only evaluate the impact of policy changes

on out-of-pocket expenditure, and do not assess their effects on drug compliance and therefore on health outcomes. We try to fill this gap by exploiting the fact that our data span three major policy changes that may be regarded as "natural experiments", whose effects on medical compliance and health outcomes can be evaluated using a difference-in-difference (DID) specification.

2.6.1 The policy changes

The three policy changes are: (i) the abolition of the co-payment on drug prescriptions, on January 1, 2001, (ii) the reduction from 6 to 3 of the maximum number of packages for each prescription, on September 30, 2001, and (iii) the reintroduction of the co-payment, on March 1, 2002. Until January 2001, patients were subject to a flat charge of about 1.5 Euros on each prescription received by their physician. This prescription charge, known to Italians as the "ticket", applied equally to all packages, irrespective of pack size, dosage or pharmaceutical form. After its abolition in January 2001, the ticket was reintroduced in March 2002 as a flat charge of 1 Euro per prescription.

The co-payment was expected to reduce both public expenditure (financial concern) and unnecessary consumption (clinical concern). Patients were exempt from the ticket either because of low income or disability, or because they suffered from specific chronic or rare pathologies diagnosed by specialists. About 93% of the adult population in our sample pays the ticket, the percentage being lower for older people. This percentage falls to zero in 2001, when the ticket was abolished.

Being a fixed amount, the ticket has an intrinsic regressive structure affecting mostly low income patients suffering from chronic conditions. From an empirical point of view, many studies confirm the role of co-payment in reducing the level of drug consumption of low income patients (see among others Freemantle and Bloor (1996), Lundberg *et al.* (1998), and Atella, Rosati and Rossi (2005)). In this section, we present a simple framework that allows us to interpret such evidence from an economic point of view. For expository reasons, assume that patients choose between drugs (all products in the ACE-inhibitor class) and all other goods. As ACE-inhibitors are provided by the Italian NHS, their price is equal to the ticket. Therefore, a change in the ticket leads to a change in drug consumption, and therefore drug compliance, for each given level of consumption of the other goods.

Figure 9 shows the budget lines and the indifference curves over drugs and other goods for two types of patients – poor (BC_1) and rich (BC_2). For simplicity, we assume that drug consumption cannot exceed the recommended level corresponding to full compliance ($Y = 1$). For positive drug prices, poor patients would reach full compliance at the point A_3, where consumption of all other goods is equal to F_0. This point, however, need not be chosen by poor patients. In fact, if the asymptomatic nature of hypertension leads them to underestimate the long-run utility of consuming an adequate level of hypertensive drugs, then the slope of the indifference curve at A_3 may be greater than the slope of the budget line. In turn, this would lead poor patients to trade off drugs for higher quantities of other goods, moving down the budget line until reaching points

like A_2 or even A_1. Atella, Rosati and Rossi (2005) and Huttin *et al.* (2003) present empirical evidence supporting such behavior by poor patients. On the contrary, rich patients are more likely to choose the recommended level of compliance (point E). For rich patients, the trade-off between drugs and other goods is less relevant, as their income allows purchasing the desired level of other goods without sacrificing drug consumption.

The first policy change lowers the price of ACE-inhibitors to zero, resulting in the new budget lines BC_1' for the poor and BC_2' for the rich.[20] This enables poor patients to move to a higher level of compliance (from $Y = .4$ to $Y = 1.0$ in the figure), while rich patients do not change their compliance as they are already full compliants.

2.6.2 The DID specification

The model in Section 2.6.1 suggests that, after the first policy change (abolition of the ticket), we should observe a higher increase in compliance for patients who were "low compliant" at the beginning of the period relative to those who were "high compliant". After the second policy change (decrease in the maximum number of packages), we should observe a higher decrease in compliance for patients who were "high compliant" at the beginning of the period relative to those who were "low compliant". Similarly, after the third policy change (reintroduction of the ticket), we should observe a higher decrease in compliance for patients who were "low compliant" at the beginning of the period relative to those who were "high compliant".

To verify this, we divide time into four periods (period 0 corresponding to the period before the first policy change, period 1 to the period between the first and the second policy change, period 2 to the period between the second and the third policy change, and period 3 to the period after the third policy change) and consider the following model for the level of compliance Y_{it} of individual i in period t

$$Y_{it} = \alpha_0 + \sum_{j=1}^{3} \alpha_j D_{jt} + \beta_0 C_i + \sum_{j=1}^{3} \beta_j D_{jt}\, C_i + U_{it}, \qquad t = 0, 1, 2, 3,$$

where D_{jt} is a time dummy equal to 1 for period j and to 0 otherwise, C_i is equal to 1 for patients with initial compliance above a certain threshold, and U_{it} is a regression error with the usual properties. According to this model, average compliance of high compliants is equal to $\mu_0^H = \alpha_0 + \beta_0$ in period 0 and to $\mu_j^H = \alpha_0 + \alpha_j + \beta_0 + \beta_j$ in period $j = 1, 2, 3$, whereas average compliance of the low compliants is equal to $\mu_0^L = \alpha_0$ in period 0 and to $\mu_j^L = \alpha_0 + \alpha_j$ in period $j = 1, 2, 3$. Thus, the average change in compliance after the first policy change is equal to $\Delta\mu_{01}^H = \mu_1^H - \mu_0^H = \alpha_1 + \beta_1$ for the high compliants and to $\Delta\mu_{01}^L = \mu_1^H - \mu_0^H = \alpha_1$ for the low compliants, the average change

[20] If we take into account the time costs to obtain the drug prescription from the physician and then go to the pharmacy to get the drugs dispensed, the budget constraint need not be vertical even when the ticket is zero. Accounting for these costs does not change the qualitative conclusions of our analysis.

in compliance after the second policy change is equal to $\Delta\mu_{12}^H = \mu_2^H - \mu_1^H = \alpha_2 - \alpha_1 + \beta_2 - \beta_1$ for the high compliants and to $\Delta\mu_{12}^L = \mu_2^L - \mu_1^L = \alpha_2 - \alpha_1$ for the low compliants, whereas the average change in compliance after the third policy change is equal to $\Delta\mu_{23}^H = \mu_3^H - \mu_2^H = \alpha_3 - \alpha_2 + \beta_3 - \beta_2$ for the high compliants and to $\Delta\mu_{23}^L = \mu_3^L - \mu_2^L = \alpha_3 - \alpha_2$ for the low compliants. Finally, $\beta_1 = \Delta\mu_{01}^H - \Delta\mu_{01}^L$ is the DID parameter for the first policy change (the difference in the average change in compliance between high compliants and low compliants from period 0 to period 1), $\beta_2 - \beta_1 = \Delta\mu_{12}^H - \Delta\mu_{12}^L$ is the DID parameter for the second policy change (the difference in the average change in compliance between high compliants and low compliants from period 1 to period 2), whereas $\beta_3 - \beta_2 = \Delta\mu_{23}^H - \Delta\mu_{23}^L$ is the DID parameter for the third policy change (the difference in the average change in compliance between high compliants and low compliants from period 2 to period 3). The model was estimated by OLS, after dropping patients who entered the panel after January 2001 or left the panel before March 2002.

To check the robustness of our results, we considered four modification of the basic model. The first considers two different subsamples, respectively covering the periods 2000–2002 and 1997–2002. The second adds a vector of demographic variables. The third adds a set of individual specific effects and uses the fixed-effect (within-group) estimator instead of OLS. The fourth uses different thresholds to classify patients as high compliant.

After fitting the various models separately by gender, we found no significant difference in the estimated coefficients. Thus, we simply report the results for the specification that only includes a gender dummy. Table 18 shows, for the two sub-samples (2000–2002 and 1997–2002), the OLS estimates of the model without demographic variables when patients are classified as high compliant if their initial indicator of compliance is greater or equal to .55. All parameters are statistically significant. In particular, the DID parameters are highly statistically significant. The negative estimate of β_1 means that, after the first policy change, low compliants increased their compliance more than high compliants. On the other hand, the negative estimate of $\beta_2 - \beta_1$ means that, after the second policy change, high compliants decreased average compliance more than low compliants. This finding reflects the fact that, in September 2001, the maximum number of packages allowed in a single prescription was lowered from 6 to 3, thereby increasing transaction costs. This affected mainly high compliants, whose mean number of packages per prescription was higher than 3 before September 2001 (it was lower than 3 for low compliants). The positive estimate of $\beta_3 - \beta_2$ means that, after the third policy change, low compliants decreased average compliance more than high compliants.

These results are robust to alternative specifications and estimation procedures. In particular, the magnitude and statistical significance of the DID parameters are unaffected by adding a vector of demographic variables or a set of individual specific effects to the basic model.

As a further robustness check, we re-estimated the model using different thresholds to classify patients as high compliants. Figure 10 presents the estimated DID parameter for the first policy

change under different values of the threshold. The DID estimates are fairly stable at around -25% for thresholds ranging from .5 to .75. At about .80, we observe a noticeable increase of the estimates (in absolute terms). From .90 to 1.15, the negative slope becomes even steeper. Our finding that the DID parameter is higher (in absolute terms) the higher the threshold is a simple consequence of the fact that patients who initially are high compliants have little room to further increase their level of compliance after the abolition of the ticket.

Overall, our results provide strong support for the argument that changes in compliance associated with changes in prescription charges tend to be greater for low compliant patients than for high compliant patients. Further, the fact that high compliants react to policy changes affecting the number of packages that can be prescribed, whereas low compliants mostly react to policy changes affecting co-payment, may be taken as an indicator that low compliants are low income patients whereas high compliants are high income patients.

2.6.3 Speed of adjustment to policy changes

How responsive are changes in compliance to changes in the co-payment structure? To answer this question, we re-parameterize the model in Section 2.6.2 to capture changes in compliance over time and estimate the resulting model at quarterly rather than annual frequency.[21] Interacting all coefficients with quarter dummies, we are able to estimate average quarterly compliance for both high compliants and low compliants. Figure 11 reports these estimates.

It is clearly seen all three policy changes had an effect on compliance. Further, this effect was almost immediate. In particular, the abolition of the ticket in January 2001 increased the average compliance of "low compliant" patients, the new equilibrium being reached within one quarter. The reintroduction of the ticket fifteen months later lowered the equilibrium level, again within one quarter. However, the average level of compliance was higher than before January 2001. Because the new co-payment was below its level prior to January 2001, "low compliant" patients faced a less stringent constraint than before January 2001. As predicted by our simple theoretical model, this allowed them to maintain higher levels of compliance than before.

2.6.4 Policy changes and health outcomes

Our empirical analysis shows clear evidence of causality running from co-payment to compliance and from compliance to health outcomes. In order to provide a quantitative measure of the effect on health outcomes of changes in co-payment, we adopt a two-step procedure.

In the first step, we use Model 2, estimated in Section 2.6.2, to predict compliance for both high and low compliants under four policies–Policy 0 (pre 2001), Policy 1 (abolition of the ticket) Policy 2 (decrease in the number of maximum packages per prescription), and Policy 3 (reintroduction of the

[21] The monthly frequency was not used to avoid problems of infrequency of purchase.

ticket)–all else being held constant. In the second step, we used the predicted values of compliance to feed the hospitalization and mortality models. More precisely, for each of the four policies, predicted compliance for the ith patient refers to year $t = 2000$ and is computed as $\mu(X_{it}) = E[\bar{c}_{it} \mid X_{it}]$, where $\bar{c}_{it}$ is annual compliance at time t, X_{it} is a vector of patient's characteristics and the expected value is estimated using the model in Section 2.6.2. Predicted hospitalization and mortality rates for the ith patient are then computed, respectively, as $E[H_{i,t+1} \mid \bar{c}_{it} = \mu(X_{it}), X_{it}]$ and $E[M_{i,t+1} \mid \bar{c}_{it} = \mu(X_{it}), X_{it}, H_{it}]$, where $H_{i,t+1}$ is the binary indicator of hospitalization at time $t+1$, $M_{i,t+1}$ is the binary indicator of mortality at time $t+1$, and the expected values are computed using the estimates from Model 2 in Section 2.5.2.

The average values of predicted outcomes (compliance, hospitalization and mortality), averaged over all units in our sample, are reported in Table 19 for each of the four policies. Notice that, for low compliants, Policy 1 implies a drop of .8 percentage points (from 7.9% to 7.0%) in the hospitalization rate and a drop of .2 percentage points (from 3.4% to 3.2%) in the mortality rate, relative to Policy 0. On the other hand, for high compliants, the differences between the two policies are not statistically significant. Overall, the predicted differences in health outcomes between Policy 3 and Policy 0 are negative and statistically significant for low compliants but not for high compliants. These effects are more clearly shown in Figure 12.

2.7 Conclusions

Our results show that compliance to anti-hypertensive drug treatment matters for health outcomes. For male patients, we estimate that the probability of future hospitalization for cardiovascular problems falls from about 12% when current compliance is near .15 to 8% when current compliance is close to its "optimal" value of 1. For female patients the reduction is less pronounced, but the lowest hospitalization rate is still observed when current compliance is near 1. Similar conclusions hold for mortality. Focusing on male patients, we estimate that increasing current compliance from .15 to 1 reduces future mortality rate by half. These results are robust to different econometric specifications and to sample selection.

Changes in the co-payment structure appear to have a strong effect on the average compliance of previously low compliant patients, while leaving almost unchanged the average compliance of previously high compliant patients. Further, the fact that high compliants react to policy changes affecting the number of packages that can be prescribed, whereas low compliants mostly react to policy changes affecting co-payment, may be taken as an indicator that low compliants are low income patients and high compliants are high income patients.

The speed of adjustment appears to be extremely rapid. This is consistent with the view that policy makers should operate through changes in co-payments whenever they want to achieve rapid effects on demand. Finally, the average level of compliance of previously low compliant patients after the reintroduction of a reduced co-payment is higher than their initial level.

Our results have two important policy implications. First, although the Italian NHS spends a large amount of money on drugs, the low level of compliance observed for a substantial fraction of patients may generate negligible returns in terms of improved health outcomes. Second, as long as co-payment affects drug consumption, it will consequently affect drug compliance and therefore health outcomes. Our calculations show that, all else being constant, abolishing the prescription charge affects health outcomes for low compliants in our sample by reducing the hospitalization rate by .8 percentage point (from 7.9% to 7.0%) and the mortality rate by .2 percentage points (from 3.4% to 3.2%). This implies that the expected additional drug expenditure could be at least partly offset by the cost reduction associated with lower hospitalization and mortality rates.

References

Alan S, Crossley TF, Grootendorst P, Veall MR. The effects of drug subsidies on out-of-pocket prescription drug expenditures by seniors: Regional evidence from Canada. *J Health Econ* 2002; **21**: 805–826.

Alan, S, Crossley TF, Grootendorst P, Veall MR. Out-of-pocket prescription drug expenditures and public prescription drug programs. *Can J Econ* 2005; **38**: 128–148.

Atella V. Drug cost containment policies in Italy: are they really effective in the long-run? The case of minimum reference price. *Health Policy* 1999; **50**: 197–218.

Atella V. Physicians' prescription behavior and cost to the patients in Italy: first evidence based on focus group techniques and prescription data. In *Patient Charges and Decision-Making Behaviors of Consumers and Physicians*, Huttin C (eds). Amsterdam: IOS Press, 2003.

Atella V, Schafheutle EI, Noyce PR, Hassell K. Affordability of medicines and patients' cost reduction behaviors: empirical evidence based on SUR estimates from Italy and the United Kingdom. *Appl Health Econ Health Policy* 2005; **4**: 23–35.

Atella V, Rosati FC. Spesa, equit e politiche sanitarie in Italia negli anni '90: quali insegnamenti per il futuro. *Politiche Sanitarie* 2002; **3**: 3–17.

Atella V, Rosati FC, Rossi MC. Precautionary saving and health risk. Evidence from the Italian households using a time series of cross sections. *Rivista di Politica Economica* 2005, forthcoming.

Case A, Le Roux I, Menendez A. Medical compliance and income-health gradients. *American Economic Association – Papers and Proceedings* 2004; **94**: 331–335.

Dracup KA, Meleis AI. Compliance: an interactionist approach. *Nurs Res* 1982; **31**: 31–36.

Freemantle N, Bloor K. Lessons from international experience in controlling pharmaceutical expenditure. I: influencing patients. *Br Med J* 1996; **312**: 1469–1471.

Harris BL, Stergachis A, Ried LD. The effect of drug co-payments on utilization and cost of pharmaceuticals in a health maintenance organization. *Med Care* 1990; **28**: 907–917.

Hughes D, McGuire A. Patient charges and the utilization of NHS prescription medicines: some estimates using a co-integration procedure. *Health Economics* 1995; **4**: 213–220.

Huttin C, Atella V, Brenner G, Hedvall MB, Mechtler R, Noyce P, Schafheutle E. Patient survey: International hypertension results. In *Patient Charges and Decision-Making Behaviors of Consumers and Physicians*, Huttin C (eds). Amsterdam: IOS Press, 2003.

IMS Health. Antihypertensive compliance and stroke risk. *The Pharmaceutical Journal* 2000; **265**: 642.

ISTAT. Indagine multiscopo sulle famiglie. Rome, Condizioni di salute 2001.

Jacobzone S. Pharmaceutical policies in OECD countries: Reconciling social and industrial goals. OECD Labour Market and Social Policy - Occasional Paper n. 40. 2000.

Leibowitz A, Manning W, Newhouse J. The demand for prescription drugs as a function of cost sharing. *Social Science and Medicine* 1985; **21**: 1063–1069.

Lundberg L, Johannesson M, Isacson DGL, Borgquist L. Effects of user charges on the use of prescription medicines in different socio-economic groups. *Health Policy* 1998; **44**: 123–134.

McInnes GT. Integrated approaches to management of hypertension: promoting treatment acceptance. *Am Heart J* 1999; **138**: S252–S255.

McManus P, Donnelly N, Henry D, Hall W, Primrose J, Lindner J. Prescription drug utilization following patient co-payment changes in Australia. *Pharmacoepidemiol Drug Saf* 1996; **5**: 385-392.

O'Brien B. The effects of patient charges on the utilization of prescription medicines. *J Health Econ* 1989; **8**: 109–131.

OSMED. L'Uso dei Farmaci in Italia. Ministero della Salute. Roma, 2003.

OSMED. L'Uso dei Farmaci in Italia. Ministero della Salute. Roma, 2005.

Piette J, Heisler M, Wagner TH. Cost-related medication underuse: Do patients with chronic illnesses tell their doctors?. *Arch Intern Med* 2004; **164**: 1749–1755.

Poirier S, LeLorier J, Pag V, Lacour A. The effect of a $2 co-payment on prescription refill rates of Quebec elderly and its relationship to socio-economic status. *Can Pharm J* 1998; **131**: 30–34.

Ryan M, Birch S. Charging for health care: Evidence on the utilization of NHS prescribed medicines. *Social Science and Medicine* 1991; **33**: 681–687.

Shaw E, Anderson JG, Maloney M, Jay SJ, Fagan D. Factor associated with noncompliance of patients taking antihypertensive medications. *Hosp Pharm* 1995; **30**: 201–207.

Soumerai SB, Avorn J, Ross-Degnan D, Gortmaker S. Payment restrictions for drugs under Medicaid: Effects on therapy, cost and equity. *N Engl J Med* 1987; **317**: 550-556.

Vermeire E, Hearnshaw H, van Royen P, Denekens J. Patients adherence to treatment: Three decades of research. A comprehensive review. *Journal of Clinical Pharmacy and Terapeutics* 2001; **26**: 331–342.

WHO. Guideline Sub-Committee. Geneva, 1999.

Table 13: WHO DDDs and Italian ADDs by substance.

ATC code	Active ingredient	1995 WHO DDDs	Italian ADDs	Ratio DDD/ADD
C09AA01	Captopril	50	50	1
C09AA02	Enalapril	10	20	0.5
C09AA03	Lisinopril	10	20	0.5
C09AA04	Perindopril	4	4	1
C09AA05	Ramipril	2.5	5	0.5
C09AA06	Quinapril	15	15	1
C09AA07	Benazepril	7.5	10	0.75
C09AA08	Cilazapril	2.5	5	0.5
C09AA09	Fosinopril	15	15	1
C09AA10	Trandolapril	2	2	1
C09AA11	Spirapril	6	6	1
C09AA12	Delapril	30	30	1
C09AA13	Moexipril	15	15	1
C09AA15	Zofenopril	30	30	1

* Calculations based on WHO (1999) and OSMED (2000).

Table 14: Descriptive statistics. Final sample: Patients born 1910–1960 filling ACE-inhibitor prescriptions (38,393 patients and 153,189 observations).

	Occasional				Regular			
	Men		Women		Men		Women	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Year 1997	.102	.302	.101	.301	.106	.307	.104	.305
Year 1998	.139	.346	.138	.344	.124	.329	.124	.330
Year 1999	.170	.376	.168	.374	.147	.354	.145	.352
Year 2000	.191	.393	.190	.392	.169	.375	.168	.374
Year 2001	.203	.402	.205	.404	.202	.401	.201	.401
Year 2002	.195	.396	.199	.399	.252	.434	.257	.437
Age	65.7	11.8	69.3	12.0	66.3	11.2	69.9	11.4
Year of birth	1934	12	1931	12	1934	11	1930	11
Large pack size	.579	.494	.555	.497	.613	.487	.591	.492
Female GP	.165	.371	.186	.389	.155	.362	.180	.385
Year of birth of GP	1951	7	1951	7	1951	7	1952	7
Age of GP	47.9	7.5	47.6	7.4	48.5	7.0	48.2	6.8
Specialist	.004	.060	.003	.054	.006	.076	.005	.068
No Enalapril	.539	.498	.544	.498	.508	.500	.520	.500
More than 1 ACE-inhibitor	.039	.194	.036	.187	.042	.200	.033	.180
More than 1 card. drug	.689	.463	.697	.460	.634	.482	.598	.490
Hospital. rate for cardiov. DRG	.099	.299	.072	.259	.125	.330	.083	.277
Mortality rate	.033	.179	.023	.148	.042	.201	.028	.165
Years in the sample	5.0	1.3	5.0	1.2	4.6	1.7	4.7	1.7
Years with nonzero purchases	2.3	1.4	2.3	1.4	4.6	1.7	4.7	1.7
Compliance	.218	.334	.203	.322	.643	.354	.619	.342
Compliance when purchasing	.468	.350	.451	.345	.643	.354	.619	.342
Observations	34,586		44,614		35,828		38,161	
Patients	7,688		9,700		10,273		10,732	

Table 15: Descriptive statistics. Average annual compliance of regular patients by age group, gender and pack size (small and large) for regular patients. For each cell we report, in order, the sample mean and standard deviation of compliance, and the cell sample size.

Age group	Share of regular patients		Average compliance of regular patients			
	Men	Women	Men		Women	
			Small	Large	Small	Large
40–49	.505	.473	.564	.626	.516	.570
	(0.500)	(0.499)	(0.348)	(0.341)	(0.335)	(0.340)
	2,093	1,612	1,166	1,635	937	983
50–59	.573	.503	.591	.662	.580	.634
	(0.495)	(0.500)	(0.342)	(0.347)	(0.330)	(0.341)
	4,013	3,561	3,041	4,459	2,618	3,083
60–9	.568	.530	.604	.720	.598	.673
	(0.495)	(0.499)	(0.354)	(0.355)	(0.344)	(0.341)
	5,168	5,131	3,983	6,587	3,961	5,520
70–9	.597	.535	.555	.716	.573	.683
	(0.491)	(0.499)	(0.340)	(0.348)	(0.342)	(0.334)
	4,718	6,331	3,970	6,574	4,880	7,916
80+	.592	.546	.435	.671	.471	.653
	(0.492)	(0.498)	(0.304)	(0.345)	(0.320)	(0.330)
	1,969	3,797	1,706	2,707	3,202	5,061

Table 16: Estimated coefficients of the logit model for hospitalization (* significant at 10%; ** significant at 5%; *** significant at 1%).

	Model 1		Model 2		Model 3	
	Men	Women	Men	Women	Men	Women
Age	-.515**	-.090	-.506**	-.082	-.490**	-.133
Age^2/100	.893***	.191	.879***	.178	.853***	.254
Age^3/10000	-.466***	-.091	-.461***	-.085	-.447***	-.122
Compliance	-1.480**	-.529	-1.369**	-.450	-1.505**	-.662
Compliance^2	.871	-.363	.760	-.455	.956	-.163
Compliance^3	-.022	.466	.018	.511	-.055	.399
More than 1 ACE-inhibitor	.723***	.593***	.579***	.468***	.572***	.482***
Multi-therapy	1.079***	1.148***	1.066***	1.134***	1.080***	1.129***
No Enalapril			-.313***	-.272***	-.331***	-.275***
Female GP					.040	.031
Age of GP					.008**	.003
Year 1999	-.191**	-.047	-.195***	-.049	-.211***	-.082
Year 2000	-.136*	-.055	-.142*	-.055	-.139*	-.069
Year 2001	-.230***	-.097	-.235***	-.095	-.254***	-.116
Year 2002	-.336***	-.353***	-.331***	-.344***	-.345***	-.373***
Constant	-3.712***	-4.281***	-3.520***	-4.103***	-3.509***	-4.098***
No. obs.	22,891	24,439	22,891	24,439	22,484	23,856
Pseudo R^2	.0701	.0691	.0729	.0711	.0748	.0725
Log-lik.	-6550.8	-5387.9	-6530.9	-5376.2	-6373.0	-5225.8
Joint significance tests						
Age	200.72***	164.14***	179.88***	145.77***	178.50***	147.80***
Compliance	37.56***	27.01***	35.56***	27.71***	34.05***	25.19***
GP characteristics	.	.	.	.	4.20	.53
Year	22.84***	23.44***	22.08***	22.12***	23.87***	23.67***

Table 17: Estimated coefficients of the logit model for mortality (* significant at 10%; ** significant at 5%; *** significant at 1%).

	Model 1		Model 2		Model 3	
	Men	Women	Men	Women	Men	Women
Age	-.215	-.128	-.190	-.118	-.159	-.109
Age^2/100	.366	.200	.329	.183	.285	.168
Age^3/10000	-.142	-.038	-.125	-.030	-.105	-.022
Compliance	-.480	-.287	-.279	-.147	-.305	.007
Compliance^2	-.039	-.058	-.263	-.225	-.224	-.345
Compliance^3	.161	.135	.246	.211	.239	.213
More than 1 ACE-inhibitor	.457***	.232	.309**	.113	.275**	.137
Multi-therapy	.606***	.537***	.593***	.521***	.615***	.572***
No Enalapril			-.387***	-.314***	-.391***	-.293***
Hospitalized at $t-1$	.684***	.758***	.647***	.731***	.633***	.716***
Female GP					.079	.011
Age of GP					.007	.005
Year 1999	-.331***	-.120	-.337***	-.122	-.350***	-.100
Year 2000	-.393***	-.373***	-.401***	-.373***	-.425***	-.395***
Year 2001	-.479***	-.587***	-.486***	-.585***	-.495***	-.579***
Year 2002	-.573***	-.802***	-.570***	-.796***	-.580***	-.818***
Constant	-4.767***	-5.658***	-4.533***	-5.452***	-4.536***	-5.469***
No. obs.	22,891	24,439	22,891	24,439	22,484	23,856
Pseudo R^2	.125	.135	.128	.138	.129	.138
Log-lik.	-3514.3	-2545.7	-3500.5	-2539.4	-3427.5	-2463.8
Joint significance tests						
Age	565.14***	450.19***	513.13***	423.25***	500.76***	406.80***
Compliance	8.03**	2.03	7.01*	1.83	6.79*	1.21
GP characteristics	.	.	.	.	2.37	.70
Year	32.30***	47.95***	32.43***	47.05***	32.55***	48.24***

Table 18: OLS estimates of the DID model without demographic variables. Comparison across subsamples (* significant at 10%; ** significant at 5%; *** significant at 1%).

	2000–2002	1997–2002
$\alpha_0 = \mu_0^L$	.356 ***	.391 ***
$\alpha_1 = \Delta\mu_{01}^L$	.214 ***	.166 ***
$\alpha_2 = \Delta\mu_{02}^L$	.175 ***	.129 ***
$\alpha_3 = \Delta\mu_{03}^L$	.123 ***	.072 ***
$\beta_0 = \mu_0^H - \mu_0^L$	.567 ***	.498 ***
$\beta_1 = \Delta\mu_{01}^H - \Delta\mu_{01}^L$	-.237 ***	-.196 ***
$\beta_2 = \Delta\mu_{02}^H - \Delta\mu_{02}^L$	-.282 ***	-.238 ***
$\beta_3 = \Delta\mu_{03}^H - \Delta\mu_{03}^L$	-.259 ***	-.210 ***
$\beta_2 - \beta_1 = \Delta\mu_{12}^H - \Delta\mu_{12}^L$	-.045 ***	-.042 ***
$\beta_3 - \beta_2 = \Delta\mu_{23}^H - \Delta\mu_{23}^L$	.023 ***	.028 ***

Table 19: Predicted compliance, hospitalization and mortality under alternative policies.

	Compliance	Hospitalization	Mortality
	Low compliants		
Policy 0	.356	.079	.034
Policy 1	.570	.070	.032
Policy 2	.532	.072	.032
Policy 3	.481	.073	.033
Policy 1 - Policy 0	.215	-.008	-.002
Policy 2 - Policy 1	-.038	.001	.000
Policy 3 - Policy 2	-.051	.002	.000
Policy 3 - Policy 0	.125	-.005	-.001
	High compliants		
Policy 0	.923	.069	.027
Policy 1	.901	.068	.027
Policy 2	.817	.069	.027
Policy 3	.789	.069	.027
Policy 1 - Policy 0	-.022	-.000	.000
Policy 2 - Policy 1	-.084	.000	.000
Policy 3 - Policy 2	-.028	.000	.000
Policy 3 - Policy 0	-.134	.001	.000

Figure 6: Hospitalization and mortality rates by age and gender.

Figure 7: Histogram of annual compliance.

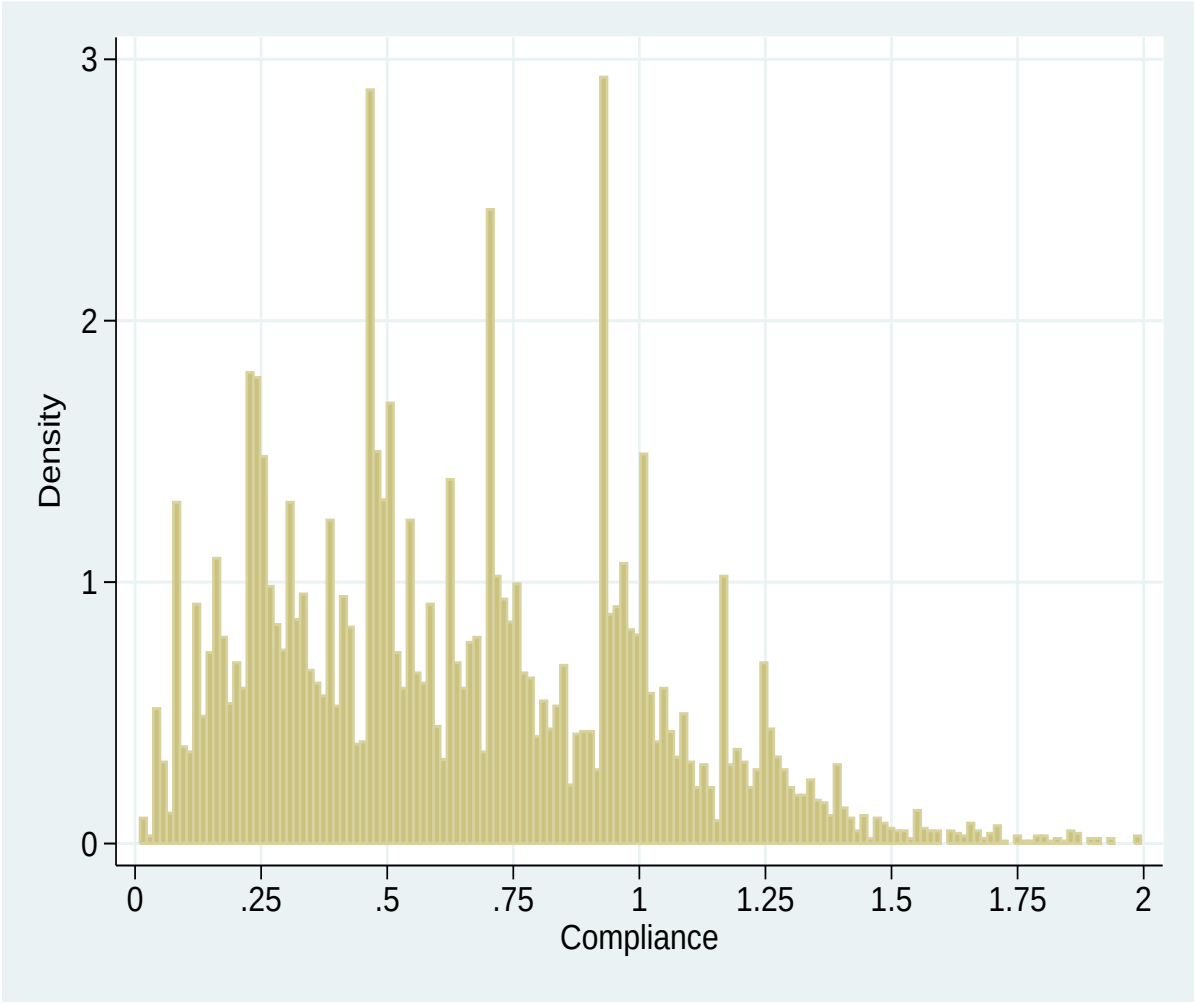


Figure 8: Observed and fitted hospitalization and mortality rates by gender and compliance level. Fitted hospitalization and mortality rates are based on Model 2.

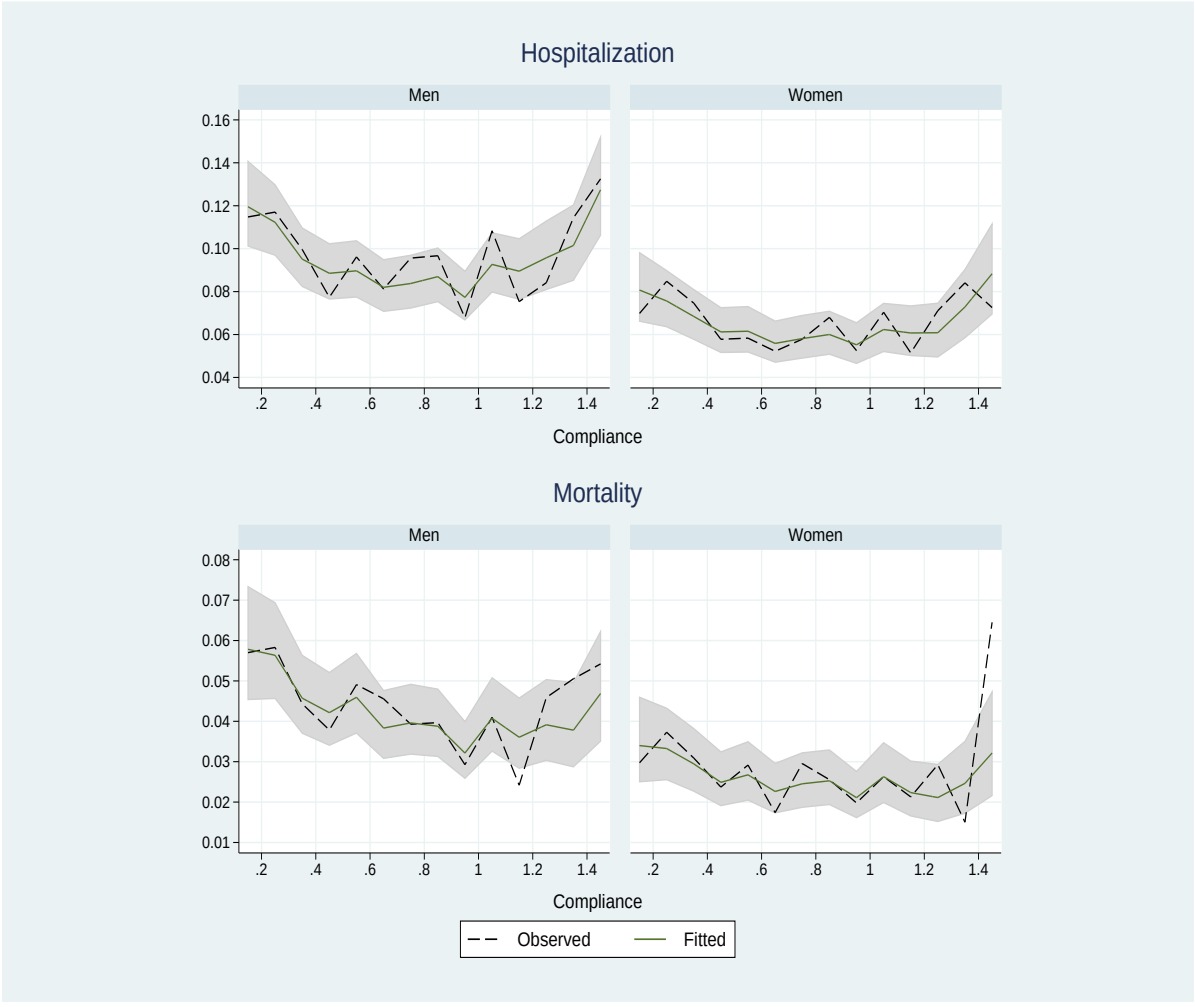


Figure 9: The effect of co-payment abolition on utility maximization for “high compliant” and “low compliant” patients.

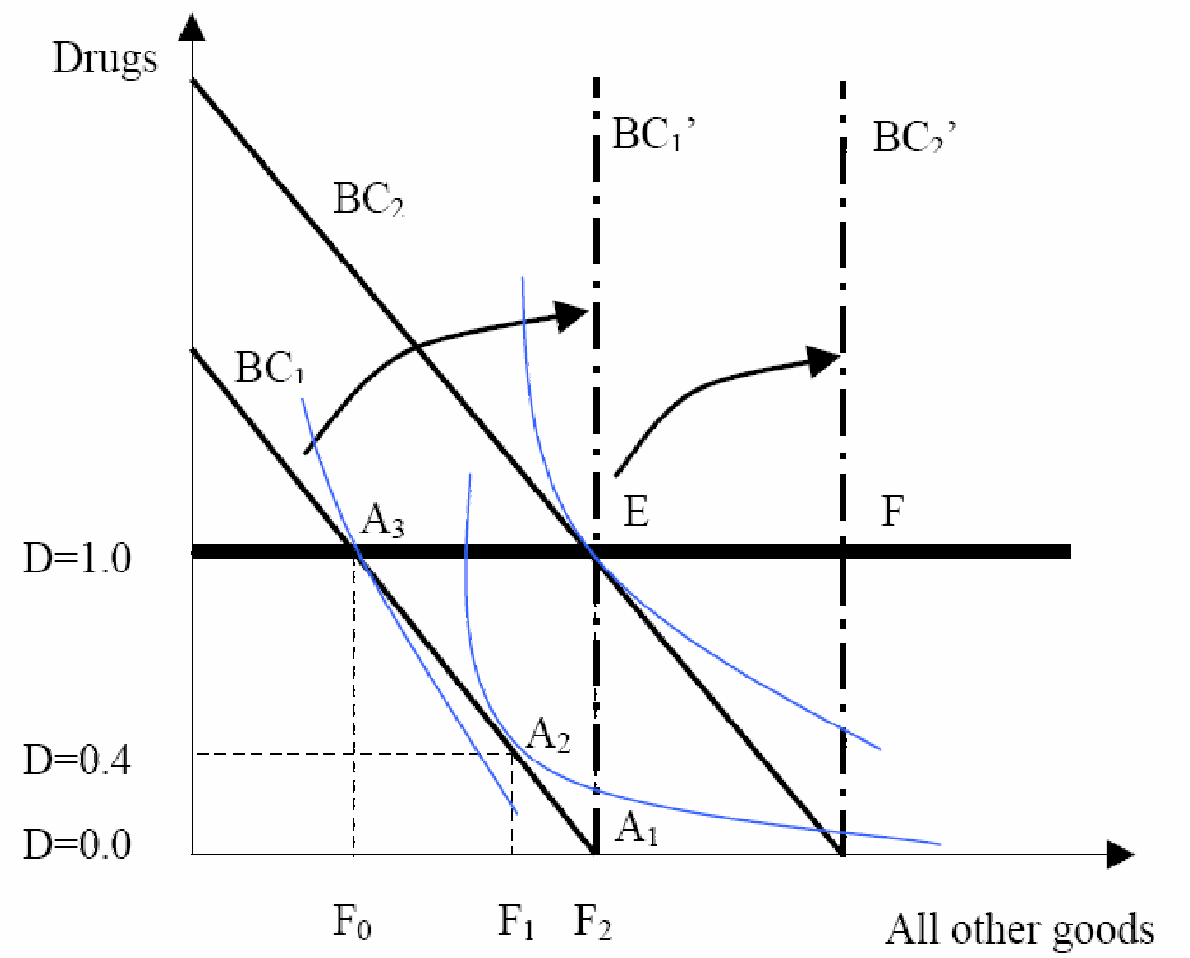


Figure 10: Estimated DID coefficient for the first policy change under different values for the threshold used to distinguish between "low" and "high" compliance.

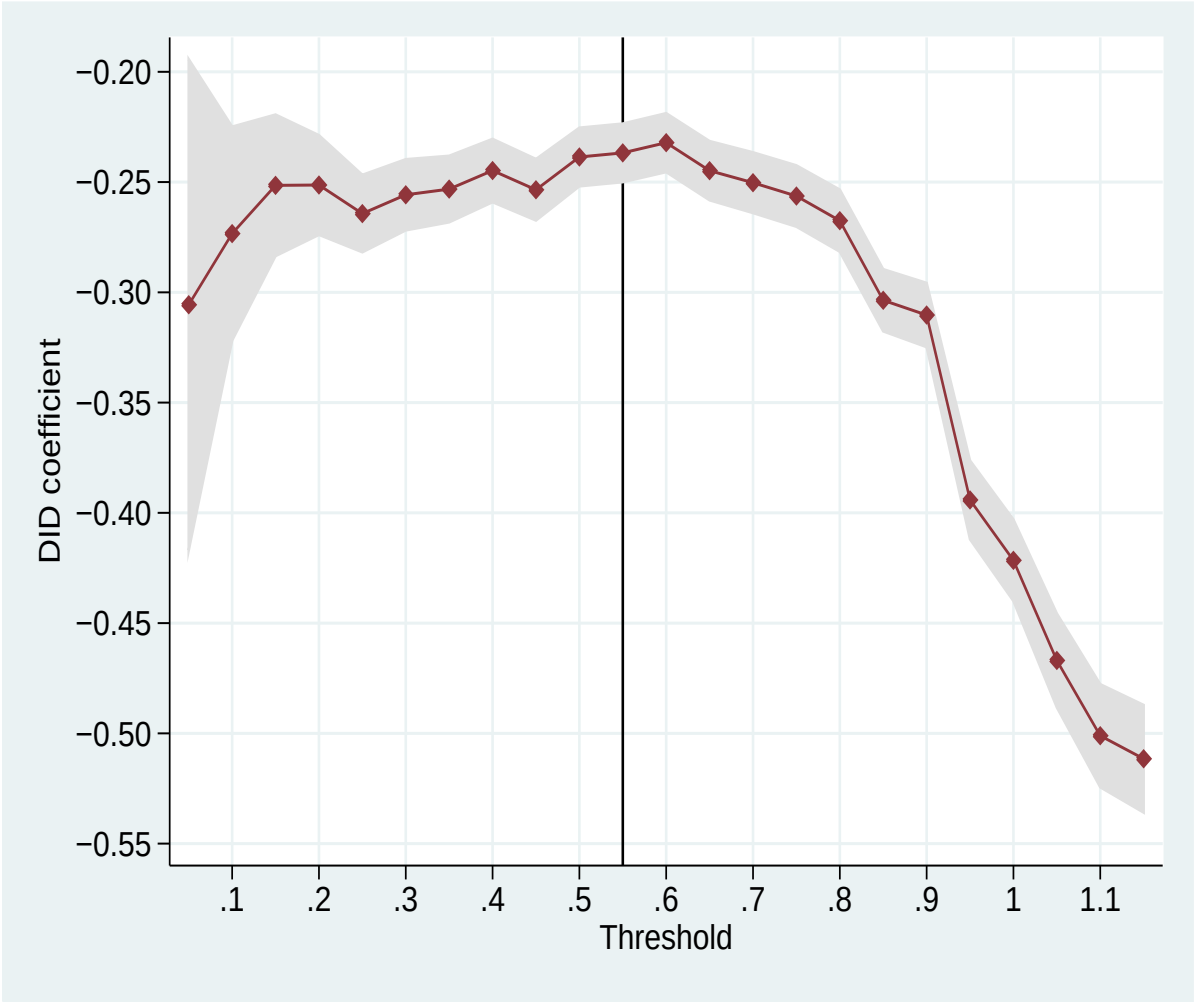
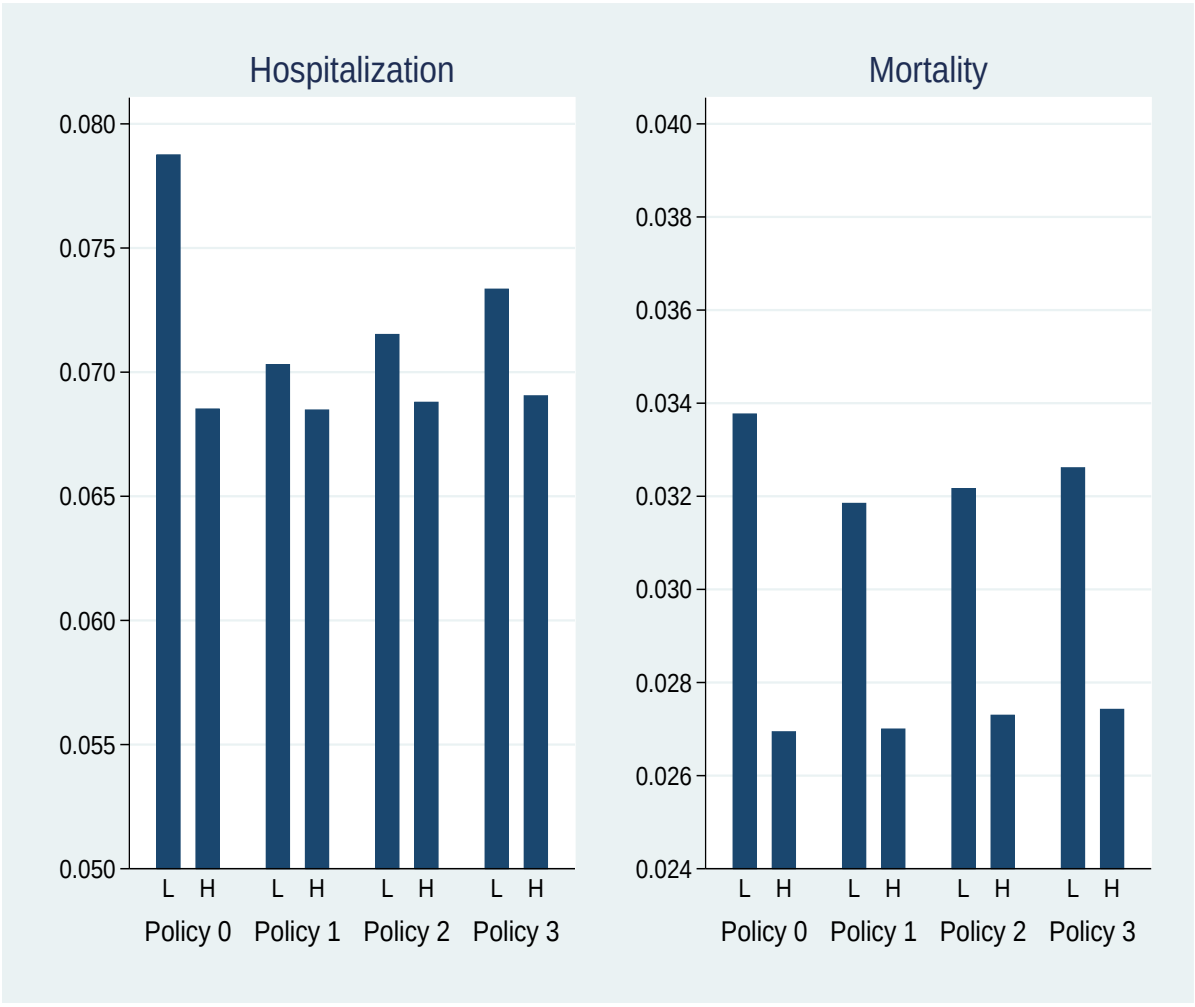


Figure 11: Average quarterly compliance for high compliants and low compliants.

Figure 12: Predicted hospitalization and mortality rates based on compliance in $t-1$ for high (H) and low (L) compliants.



3 Chapter 3
Health and medical care. A dynamic factor approach using a panel of Italian patients

Abstract[22]: Analysing the relationship between health and medical care turns out to be a critical issue. In fact, looking at the raw correlation between medical care and health cannot be expected to give the right answer, because of simultaneity through the unobservable components of deterioration. In this paper, we use a dataset where very detailed information about medical drug use, hospitalization, and mortality, is collected over time for a sample of individuals suffering from hypertension, a chronic asymptomatic pathology affecting a large share of the adult population. All those variables are expected to be strongly dependent on each other. For analysing the amount of information embedded in such variables, we employ a dynamic factor model where medical treatments and mortality may all in principle be driven by latent individual stock of health. We introduce dynamics by including the effects of lagged treatment on latent health. We estimate our model by Maximum Simulated Likelihood (MSL). In line with findings provided so far in the literature, our results indicate that better health is associated to lower medical treatments. In addition, we find that lagged medical drug use has positive effects on current health. This is consistent with the fact that not taking the medication today may result in poorer health tomorrow. Nonetheless, taking more pills than needed cannot improve health.

Keywords: Dynamic panel data models, factor models, simulated likelihood, latent variable models

JEL codes: C33, C35, C81, D12, I12, J14

[22] I wish to thank Arthur van Soest for helpful discussions. This chapter is based on the article "Health and medical care. A dynamic factor approach using a panel of Italian patients", written with Vincenzo Atella and Franco Peracchi.

3.1 Inroduction

Analysing the relationship between health and medical care turns out to be a critical issue. On the one hand, medical care is supposed to have positive effects on health (Grossman model, Grossman (1972)). On the other hand, in Grossman's original work, and in several papers since, authors have found a negative correlation between the stock of health and medical care. Maybe this is not a surprise, "given that people tend to seek medical care when they are sick, not when they are well" (Case and Deaton, 2004). Nonetheless, as Grossman (2000) himself has argued, looking at the raw correlation between medical care and health cannot be expected to give the right answer, because of simultaneity through the unobservable components of deterioration. Hence, these findings could be possibly attributed to problems with the empirical implementation of the model.

In this paper, we focus on a sample of patients who were prescribed medical drugs employed against hypertension in the local health authority covering the southern part of the Italian province of Treviso during the period 1997–2002. Due to the peculiarity of the pathology considered here, we believe that focusing on such patients can give us helpful insights into the relation between health and medical care. In fact, hypertension is a chronic asymptomatic pathology affecting a large share of the adult population. Because hypertension is a chronic pathology, patients should take their medication regularly. Nonetheless, since it is asymptomatic they generally do not feel ill because of high blood pressure, and hence following the medical prescription is more likely not to be done. Furthermore, if left untreated, hypertension can have serious consequences in terms of mortality. Finally, focusing on specific pathologies or specific drug treatments offers the advantage of exploring consumer decision-making in relation to specific clinical conditions and allows us to derive more precise conclusions concerning the relation between medical care utilization and health.

Specifically, we use here a dataset where very detailed information about medical drug use, hospitalization, and mortality, is collected over time for a selected sample of individuals. These variables are expected to be strongly dependent on each other. VAR models seem to be not suitable for analysing the amount of information embedded in such variables because of the large number of parameters to estimate. A dynamic factor model is much better suited because it is both flexible and parsimonious. In a factor model, each observed response variable is represented as the sum of two components. The first is a common component, a term depending on a small number of unobserved common factors. The second is an idiosyncratic component, which is orthogonal both to the common factors and to the idiosyncratic components of all the other variables.

Thus, we employ a dynamic factor model where medical drug use, hospitalization and mortality may all in principle be driven by latent (possibly multidimensional) individual stock of health. We introduce dynamics by including the effects of lagged treatment on latent stock of health. Our dynamic model aims of providing a better tool for investigating the effects of medical treatment against health deterioration. We estimate the model by Maximum Simulated Likelihood (MSL). The method of Maximum Simulated Likelihood for dynamic factor analysis employed here may

be potentially extended to include a general framework including factor model for mixed outcome variables and mixed latent variables. Hence, the estimation methods used here evidently accommodate a great variety of data. Thus, they do not just apply to special cases, but enable the study of dynamic models with unobserved common factors in a wider framework.

In line with findings provided so far in the literature, our results indicate that better health is associated to lower medical treatments. In fact, estimated factor loadings have all negative sign. In addition, from our dynamic model we find a reversed U-shaped relation between lagged medical drug use and current health. This is consistent with the fact that not taking the medication today may result in poorer health tomorrow. Stated differently, our results indicate that taking the medication may help reduce health deterioration and avoid its negative consequences. Nonetheless, taking more pills than needed cannot improve health.

The remainder of this paper is organized as follows. Section 3.2 describes our disease-specific approach. Section 3.3 describes the data (samples and variables) used for this study. Section 3.4 provides descriptive statistics and some preliminary evidence. Section 3.5 introduces a simple static factor model for unobserved health and reports estimation results. Section 3.6 introduces our dynamic factor model and the estimation methods, and reports estimation results. Finally, Section 3.7 offers some conclusions.

3.2 A disease-specific approach

In this paper we focus on patients treated with active ingredients in five "therapeutic main groups" of the Anatomical Terapeutical and Clinical (ATC) Classification System. These therapeutic groups are Antihypertensives (AH), Diuretics (D), Beta blocking agents (BBA), Calcium channel blockers (CCB), and Agents acting on the renin-angiotensin system (ARA).[23] The substances belonging to these groups are mainly employed against hypertension. Detailed information about the chemical substances (or active ingredients) used in this paper are reported in Appendix F.

Hypertension is a chronic asymptomatic pathology affecting a large share of the Italian population and tends to have long-term health implications. Over the last 30 years an increase of the use of antihypertensive drugs has been observed in the developed countries (see for example Gross, Wise and Knapp, 1989). According to ISTAT (2001), about 20% of the Italian adult population suffers of hypertension and its prevalence increases with age (37% at age 55–64, 50% at age 65–74 years, and 67% at age 75+). Because hypertension is an asymptomatic condition, patients do not generally feel ill because of high blood pressure. Hypertension is an interesting condition to study from the viewpoint of health. In fact, left untreated, it can have serious consequences in terms of mortality. Finally, focusing on specific pathologies or specific drug treatments also offers the advantage of exploring consumer decision-making in relation to specific clinical conditions.

[23] Patients treated with Enalapril 5mg are not included because Enalapril 5mg is not used against hypertension.

3.3 The data

Our data comes from three administrative registries maintained by the Pharmaceutical Service Department of ULSS 9, the local health authority covering the southern part of the Italian province of Treviso. The first registry is the drug prescription database, which contains records of patient prescriptions, including the date the prescription is made, the ATC code of the substance prescribed, the number of packages prescribed, the unit price of the package. The unit price of the package allows us to recover information about the number of pills contained in each package. The registry also includes gender and date of birth of the patient receiving the medications, a unique anonymized patient identifier, a unique anonymized identifier of the practitioner who prescribed the medication, and gender and date of birth of the practitioner.

The second is the hospitalization registry, which contains records of each single hospitalization episode, including date of entry and dismissal, primary Diagnosis Related Groups (DRG), and cost of hospitalization. Through the anonymized personal identifiers, we were able to link patient prescription and hospitalization information to the last registry, the death and transfer registry. The resulting dataset allows us to follow individual patients through all their accesses to public health care services until they either die or leave the local health authority. Complete data are available from 1993 to 2002 for drug prescriptions, and from 1997 to 2002 for hospitalizations For this reason, we focus on the period going from 1997 to 2002. For a detailed description of the administrative data used here, see Appendix F.

Relative to survey data, these administrative data have both advantages and disadvantages. An important advantage is that they do not present problems which are typical of survey data, namely unit and item non-response, measurement errors and bias effects due to interaction with interviewers. Another advantage is that they contain extremely rich information on health care services received by patients. The main disadvantage is that they contain little information on patients' socio-economic characteristics. In particular, no information on income and education is available.

3.3.1 Variables

Medical drug use is measured as the logarithm of average daily purchase.[24] Average daily purchase of medical drug is computed as the total number of pills purchased during the period divided by the number of days

$$C_{it} = \frac{1}{T_{it}} \sum_{j=1}^{J_{it}} N_{ij} \cdot P_{ij},$$

[24] Note that patients may not necessarily consume all the quantity of medical drug they purchase. Furthermore, in principle, they could happen to buy medical drug in a pharmacy out of territory covered by the local health authority considered here. Although we believe that these events turn out to be uncommon, this is a source of measurement error, which is taken into account in our model.

where J_{it} is the number of prescriptions filled to patient i during the period t, P_{ij} and N_{ij} are respectively the number of pills per pack, and the number of packs contained in the j^{th} prescription, and T_{it} is the number of days the patient is observed during period t. Obviously, patients cannot buy medical drugs when they are hospitalized or when they happen to be dead. Hence, the number of days T_{it} the patient is observed during period t is given by the total number of days the patient is alive minus the number of days the patient is hospitalized in such period.

We distinguish by hospitalization for cardiovascular DRGs, and hospitalization for any other DRG. Hospitalization for cardiovascular DRGs is a dummy variable for an individual being hospitalized with a cardiovascular DRG in period t. Analogously, hospitalization for other DRGs is a dummy variable for an individual being hospitalized with any DRG but cardiovascular DRGs in period t. Finally, we observe whether an individual is alive in period t. Individual health status cannot be measured directly. Hence, we employ a factor model where medical drug use, hospitalization and mortality may all in principle be driven by latent health.

The set of control variables used here includes age, gender, dummies for type of substance used ("therapeutic main group" based on ATC codes), first order interactions between drugs belonging to different therapeutic main groups, and the characteristics (age and gender) of the practitioner who prescribed the drugs.[25] Furthermore, we also use information about medical treatment prior to the sample period. Such piece of information is available only for medical drug use for the period going from 1993 to 1996. Specifically, we include the logarithm of average medical drug use for the period 1993-1996, and the number of years a patient has been under treatment (i.e. she has been purchasing medical drugs) from 1993 to 1996.

3.3.2 Sample selection

We start with all patients born 1920-1940 (aged 60-80 in year 2000),[26] who have at least one prescription filled every year,[27] in 1997, and with at least one prescription between 1993 and 1996.[28] This corresponds to an initial sample of 21,208 patients. Since some of the substances considered here are not only employed against hypertension, but they are also employed against kidney failure, we drop patients who were hospitalized for renal diseases but never for cardiovascular diseases (577 patients dropped). We then drop patients with missing data on any of the variables used (729 patients dropped). Finally, we drop patients with average drug use greater than 5 pills per day

[25] Because during the same time period more than one physician may possibly prescribe drugs to a patient, we follow the common practice of including the characteristics of the physician who made most of the prescriptions.

[26] We restrict attention to these cohorts because they comprise the bulk of the population suffering of hypertension.

[27] Patients who do not have prescriptions for more than 365 days may be possibly affected by mild hypertension that could be treated simply by a healthy diet and by reducing stress factors.

[28] We consider patients with at least one prescription before the sample period (between 1993 and 1996) in order to ensure a minimum period of treatment before the period of interest. This period allows the physician to define the treatment "suitable" for the patient. This is an important issue for the choice of the therapy against hypertension, which is still a very empirical matter.

(612 patients dropped).[29] The final sample consists of 19,290 different patients. Figure 13 shows the distribution of patients included in our final sample by gender and year of birth.

3.4 Descriptive statistics and preliminary evidence

Because we expect a possible different relation between health and medical care by birth cohort and gender, we carry out our analysis distinctively for four groups of patients. The groups are given by birth cohort-gender combination. We consider two birth cohorts, namely 1920-1930, and 1931-1940. Table 20 shows descriptive statistics. Perhaps not surprisingly, medical drug use is higher for older patients, and it is higher for men than for women for any birth cohort. Female physicians are slightly more likely to prescribe to female patients. Agents acting on the renin-angiotensin system are the most commonly used substances, and they are more likely to be prescribed to younger cohorts. The most common first-order interaction between drugs belonging to different ATC classes is between Calcium channel blockers and Agents acting on the renin-angiotensin system. The number of years a patient has been under treatment during the period 1993-1996 (prior to the sample period) is on average 2.

Figure 14 shows the histogram of medical drug use. The black vertical lines correspond to a half, one and two pills. The histogram peaks at about 1 pill per day and half a pill per day. Figure 15 shows average drug use, hospitalization, and mortality by age and gender. Medical drug use increases with age and is higher for men at any age. Men are more likely to be hospitalized (both with cardiovascular DRGs and other DRGs). The gender difference in cardiovascular related hospitalization is higher than the difference in hospitalization due to any other cause. Finally, male patients are more likely to die than female patients at any age, and the gender difference in mortality rates increases with age.

3.5 A static factor model for unobserved health

In this section, we start with a simple static factor model with no unobserved heterogeneity for unobserved health. In this model, mortality, hospitalization, and medical drug use are all assumed to be driven by latent (possibly multidimensional) individual stock of health. In the next section we extend to the case of a dynamic factor model with unobserved heterogeneity.

3.5.1 The statistical model

Consider J latent response or outcome variables $Y_{it}^* = (Y_{it1}^*, \ldots, Y_{itJ}^*)'$. In our case $J = 4$. Specifically, Y_{it1}^* is the logarithm medical drug use, Y_{it2}^* is the propensity to be hospitalized with any but cardiovascular DRGs, and Y_{it3}^* is the propensity to be hospitalized with cardiovascular DRGs, Y_{it4}^* is some measure of "frailty". We can think of all these variables as being driven by possibly

[29] They might be cases of co-morbidity or outliers.

multidimensional latent or unobserved health η_{it}. Hence, $\eta_{it} = (\eta_{it1}, \ldots, \eta_{itq})'$ is a vector of q common factors. Specifically, we assume a linear relation between the latent response variables Y_{it}^* and unobserved health η_{it} and a set of observed regressors X_{it}

$$Y_{it}^* = AX_{it} + \Gamma\eta_{it} + \xi_{it}, \tag{5}$$

where A is a matrix of regression coefficients, Γ is a matrix of factor loadings, and ξ_{it} are residuals or measurement errors. The term AX_{it} contains the intercepts. In addition, AX_{it} also contains the effect of the physician's characteristics and of the type of substance on medical drug use. Specifically, in the medical drug equation it also includes age and gender of physician, the dummies for the ATC class of the medical drug, and the interactions for using drugs belonging to two different ATC classes.

The relation between η_{it} and X_{it} is the following

$$\eta_{it} = BX_{it} + v_{it}, \tag{6}$$

where B is a matrix of regression coefficients, v_{it} is a vector of random errors or disturbances, with $v_{it} \perp v_{is} \forall s \neq t$. The term BX_{it} contains the intercepts and age.

The variables Y_{it}^* are not necessarily observable. What we observe is the vector Y_{it}, where

$$Y_{itj} = \begin{cases} Y_{itj}^* & j = 1 \\ 1\{Y_{itj}^* > \tau_j\} & j = 2, 3, 4 \end{cases}$$

where $1(\cdot)$ is the indicator function.

Define $\Xi \equiv Cov(\xi_{it})$, and $\Psi \equiv Cov(v_{it})$. As usual in factor analysis, we assume that the response variables are conditionally independent given the common factors. Hence, Ξ is a diagonal matrix. The matrix Ψ is unrestricted. v_{it} can be conveniently expressed as a linear combination of uncorrelated random variables ε_{it}. So, $v_{it} = C\varepsilon_{it}$, where C is a lower triangular matrix. C is the cholesky root of Ψ, so that $\Psi = CC'$. Finally, we assume that the residuals and measurement errors are normally distributed.

Models with latent variables usually require many restrictions to obtain identification. Our model is no exception. For medical drug use no normalizations are necessary, because of the identity $Y_{it1} = Y_{it1}^*$. For the binary dependent variables, we fix the scales by normalizing the variances of the residuals to 1, and fix the locations by normalizing the thresholds τ_j to 0. Also each element of η must be assigned a location and a scale. The most common normalization is to assign a reference variable from the list of response variables to each latent variable. In this paper, we restrict to only one health dimension. So $q = 1$. Because mortality turns out to be the most objective measure of health among the ones we observe, we use mortality as the reference variable. Then, the constant

in the equation of mortality is normalized to 0, and the factor loading is normalized to -1 (because we obviously expect a negative relation between the probability of dying and latent health).

We estimate the factor model described so far by maximum likelihood. Let $h(Y_{it}|X_{it}, \varepsilon_{it})$ be the conditional distribution of Y_{it} given X_{it} and ε_{it}. Given the fact that ε_{it} is not observed, ε_{it} is integrated out, and inference is based on

$$h(Y_{it}|X_{it}) = \int_{R_\varepsilon} h(Y_{it}|X_{it}, \varepsilon_{it}) g(\varepsilon_{it}) d\varepsilon_{it}, \tag{7}$$

where the density $g(\varepsilon_{it})$ is the product of the prior distributions of the q elements of ε_{it}, and R_ε is the range space of ε (omitted from now on). In our case, $q = 1$ and $g(\cdot) = \phi(\cdot)$, where $\phi(\cdot)$ is the standard normal density function. Under the assumption that the correlation between the observed response variables is only due to the common factors, the conditional distribution of Y_{it} may be written as

$$h(Y_{it}|X_{it}, \varepsilon_{it}) = \prod_{j=1}^{J} h_j(Y_{itj}|X_{it}, \varepsilon_{it}), \tag{8}$$

where $h_j(Y_{itj}|X_{it}, \varepsilon_{it})$ is the conditional distribution of the j-th response variable. The previous assumption is also referred as the "conditional independence" assumption, because it implies that after conditioning on the latent variables, the observed response variables are independent. Under the "conditional independence" assumption the distribution (7) may be written as

$$h(Y_{it}|X_{it}) = \int \prod_{j=1}^{J} h_j(Y_{itj}|X_{it}, \varepsilon_{it}) g(\varepsilon_{it}) d\varepsilon_{it}. \tag{9}$$

Let $\mu_{itj} = (A + \Gamma B)X_{it} + \Gamma C \varepsilon_{it}$. Then, for medical drug use

$$h_j(Y_{itj}|X_{it}, \varepsilon_{it}) = \frac{1}{\sigma} \phi\Big(\frac{Y_{itj} - \mu_{itj}}{\sigma}\Big),$$

where $\phi(\cdot)$ is the standard normal density function. For the binary indicators for hospitalization and mortality

$$h_j(Y_{itj}|X_{it}, \varepsilon_{it}) = \big[\Phi(\mu_{itj})\big]^{Y_{jit}} \big[1 - \Phi(\mu_{itj})\big]^{1 - Y_{itj}},$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function.

The loglikelihood of the sample is then

$$l = \sum_{i=1}^{n} \sum_{t=1}^{T_i} \log \left(\int \Big\{ \prod_{j=1}^{J} h_j(Y_{itj}|X_{it}, \varepsilon_{it}) \Big\} g(\varepsilon_{it}) d\varepsilon_{it} \right). \tag{10}$$

The integrals in (10) do not have a closed form solution. To solve this problem, we use max-

imum simulated likelihood (MSL). The idea of simulated likelihood is the following. The log-likelihood (10) can be approximated by $\tilde{l}$ obtained by drawing R random variables ε_{it}^r from the distribution $g(\cdot)$

$$\tilde{l} = \sum_{i=1}^{n} \sum_{t=1}^{T_i} \log \left(\frac{1}{R} \sum_{r=1}^{R} \left\{ \prod_{j=1}^{J} h_j(Y_{itj}|X_{it}, \varepsilon_{it}^r) \right\} \right). \tag{11}$$

The integrals of (10) are approximated through summations over $r = 1, \ldots, R$ draws from the distribution of ε instead of being solved numerically. From the strong law of large numbers $\tilde{l} \to l$ as $R \to \infty$. Therefore the simulated likelihood (11) is a consistent simulator of the likelihood (10). As long as the approximation is close enough, MSL provides estimators which are consistent and asymptotically equivalent to the MLE (Gouriéroux and Montfort, 1997).

Instead of using pseudo-random draws to obtain ε^r we follow Train (2003) and base the simulation on Halton sequences. Halton sequences generate quasi random draws that provide a more systematic coverage of the domain of integration than independent random draws. Quasi-random sequences usually go along with a lower integration error and faster convergence rates and therefore require clearly less number of draws compared to pseudo-random sequences. Furthermore, since Halton sequences are deterministic, following Wang and Hickernell (2000) we introduce randomness to the Halton sequences by using a random start procedure. Specifically, we draw an integer randomly between 0 and some large K and label the draw N_0. Then, we create a Halton sequence starting at integer N_0 in step 1 above. Note that the (simulated) likelihood of factor models like the one we propose here may suffer from local optima. For this reason we used multiple random starting values and checked that we ended up with the same parameter estimates. The maximization routine is written in MATA, the matrix programming language of STATA, and is based on the Newton-Raphson algorithm, with numerical first and second derivatives.

3.5.2 Results

Table 21 shows SML estimates of the static factor model by birth cohort and gender. In the bottom part of Table 21 we report a likelihood-ratio (LR) test of our factor model against a model with no common factor η (and thus Γ, B and C equal to zero). From now on we refer to the latter model as to the "independence model", because it implies that all response variables are independent of each other. The LR test strongly rejects the hypothesis of independence in favour of our factor model.

The factor loadings of all medical treatments are statistically significant and, as expected, have all negative sign. This is consistent with the fact that a higher level of – unobserved – health involves lower medical drug use and lower probability of being hospitalized (both with cardiovascular DRGs and with other DRGs). There are a few differences between the cohorts and between the sexes. Younger cohorts have slightly higher factor loadings than older cohorts in absolute values. Men

have lower factor loadings than women. This shows a slightly higher impact of health on medical drug use and on hospitalization among younger males.

In the factor equation, age is reported to have a negative effect on latent health. This is consistent with a depreciation of health as individuals grow older. Finally, the constant terms in the factor equation reveal that, on average, women have higher levels of – unobserved – health than men. This gender difference is particularly high for the older cohorts. Estimation results from the static factor model actually confirm the presence of a negative correlation between current levels of health and current medical treatment. This is consistent with the fact that better health is associated with less medical care. Nonetheless, this static model does not tell us much about the effects of medical treatment on health.

3.6 A dynamic factor model for unobserved health

In this section we extend the simple static model presented in the previous section to the case of a dynamic factor model for unobserved health. Our dynamic model aims of providing a better tool for investigating the effects of medical treatment against health deterioration.

3.6.1 The statistical model

We introduce dynamics in our model by including the effects of lagged medical treatment on latent stock of health. In addition, we allow for a direct effect of lagged values of each response variable on current values of the same variable. These direct effects are meant to capture "pure" state dependence in medical treatment. In fact, because we are considering individuals suffering from a chronic pathology, state dependence is expected to be particularly important in our case, and not taking it into account could result in misleading conclusions. We also include individual specific effects in the latent health equation, which are meant to capture time-invariant unobserved heterogeneity in the individual stock of health. The model for the latent response variables Y_{it}^* in (5) becomes

$$Y_{it}^* = DY_{i,t-1} + AX_{it} + \Gamma\eta_{it} + \xi_{it}, \tag{12}$$

where A and D are matrices of regression coefficients, Γ is a matrix of factor loadings, and ξ_{it} are residuals or measurement errors. The term AX_{it} is the same as in the static case. The term $DY_{i,t-1}$ includes the direct effects of lagged medical drug use on current medical drug use, and the direct effects of lagged hospitalization on current hospitalization. Hence, D is a diagonal matrix with the element d_{44} equal to 0.

The model for latent health η_{it} in (6) becomes

$$\eta_{it} = EY_{i,t-1} + BX_{it} + a_i + v_{it}, \tag{13}$$

where E and B are matrices of regression coefficients, a_i is a time-invariant individual component, and v_{it} is a vector of random errors or disturbances. The term BX_{it} is the same as in the static case. The term $EY_{i,t-1}$ includes lagged treatment, i.e. lagged medical drug use and lagged hospitalization.

Like in the static case, it is convenient to rewrite v_{it} as a linear combination of uncorrelated random variables ε_{it}, $v_{it} = C\varepsilon_{it}$. The conditional distribution of $Y_{i1}, \ldots, Y_{iT_i}$ given $Y_{i,t-1}$, X_{it}, ε_{it} and a_i is then

$$\prod_{t=1}^{T_i} h(Y_{it}|Y_{i,t-1}, X_{it}, \varepsilon_{it}, a_i).$$

As usual in dynamic nonlinear models with unobserved individual effects, we have two alternatives for dealing with the fact that a_i is not observed. One possibility is to treat the n unobserved effects a_i as parameters to be estimated. This approach avoids specifying the distribution of a_i. Nonetheless, with fixed time periods, it suffers from an incidental parameters problem.

Alternatively, we can treat a_i as random effects and integrate them out of the distribution. The need to integrate a_i out raises the issue of how to deal with initial treatment, i.e. initial observations Y_{i0}. This issue is usually referred as the initial conditions problem. We deal with the initial conditions problem by modelling the distribution of the unobserved effect conditional on the initial treatment and any exogenous explanatory variables (Wooldridge, 2005). Furthermore, to increase explanatory power, we also include information about medical treatment prior to the sample period. Such piece of information is available only for medical drug use for the period going from 1993 to 1996. Note that specifying a model for the conditional distribution of the unobserved effect is an approximation and is a potential source of specification error.[30] Nonetheless, such an approximation allows avoiding the strong assumption of no correlation between a_i and Y_{i0}.

Here we specify the conditional distribution of the unobserved effect as $N(FY_{i0} + GX_i, \sigma_u)$. The term FY_{i0} contains medical treatment (both medical drug use and hospitalization) in the first year of the sample period. FY_{i0} also includes the logarithm of average medical drug use and the number of years a patient has been under treatment (i.e. she has been purchasing medical drugs) prior to the sample period (from 1993 to 1996). In principle all leads and lags of the exogenous variables could be included in the term GX_i. In practice, it may be found that a limited summary as the average over the sample period works well. For this reason, here the term GX_i contains the average over the sample period of the observations on the exogenous variables.

This corresponds to writing the individual effects as

$$a_i = FY_{i0} + GX_i + u_i, \tag{14}$$

where u_i is assumed to be $N(0, \sigma_u^2)$ distributed and independent of ε_{it}, ξ_{it}, and the X variables.

[30] An alternative approach to the initial conditions problem would be to specify a model for the distribution of the initial treatment conditional on any exogenous explanatory variables and the unobserved effect (Heckman, 1981). This approach can also be viewed as an approximation, and hence is subject to specification error too.

Substituting (14) into (13) gives a model for η_{it} that has a random effects structure, with the regressors at time t augmented to include initial treatment, medical drug use prior to the sample period and a summary of the exogenous variables over the sample period.

Thus, by substituting (14) and (13) into (12) the model for the latent response variables Y_{it}^* can be written as

$$Y_{it}^* = (D + \Gamma E)Y_{i,t-1} + (A + \Gamma B)X_{it} + \Gamma F Y_{i0} + \Gamma G X_i + \Gamma u_i + \Gamma C \varepsilon_{it} + \xi_{it}. \tag{15}$$

Thus, in our final model both direct and indirect (through the common factor) effects of lagged values of each response variable on current values of the same variable are allowed. The roles of direct and indirect effects are the following. First, direct effects are meant to capture "pure" state dependence in medical treatment. As discussed before, because we are considering individuals suffering from a chronic pathology, state dependence is expected to be particularly important in our case. Secondly, medical treatment in period $t-1$ may possibly affect medical treatment in period t through its effects on latent health.

Coherently with our factor model, only indirect (through the common factor) effects of lagged values of each response variable on current values of all the other response variables are allowed. This is consistent with the "conditional independence" assumption. Hence, medical drug use at time $t-1$ can only affect hospitalization and mortality at time t through its effects on unobserved health at time t. Analogously, hospitalization at time $t-1$ can only affect medical drug use and mortality at time t through unobserved health at time t.

Separate identification of direct and indirect effects is provided by the restriction of no direct effects of lagged variables on current values of all the other variables. Specifically, indirect effects of lagged medical drug use are identified from the equations of all the other response variables. Then, direct effects are identified from the equation of medical drug use. The same is true for direct and indirect effects of lagged hospitalization.

Under the "conditional independence" assumption, the distribution of $Y_{i1}, \ldots, Y_{iT_i}$ is

$$\prod_{t=1}^{T_i} \prod_{j=1}^{J} h_j(Y_{itj} | Y_{i,t-1}, X_{it}, \varepsilon_{it}, u_i).$$

Because ε_{it} and u_i are not observed, they are both integrated out of the distribution. Our final model can then be viewed as a "multilevel" model. In particular, this model is a three-level model, where for each individual i we observe J response variables over T_i time periods. In such a model, the contribution to the likelihood from a three-level unit is found by integrating the product of contributions from the level-2 units inside the level-3 unit over the level-3 random

effects distribution. The loglikelihood of our final model is then

$$l = \sum_{i=1}^{n} \log \left(\int \left\{ \prod_{t=1}^{T_i} \left\{ \int \left\{ \prod_{j=1}^{J} h_j(Y_{itj}|Y_{i,t-1}, X_{it}, u_i, \varepsilon_{it}) \right\} g_\varepsilon(\varepsilon_{it}) d\varepsilon_{it} \right\} \right\} g_u(u_i) du_i \right). \tag{16}$$

Like in the static case, the log-likelihood (16) can be approximated by $\tilde{l}$ obtained by drawing R random variables ε_{it}^r and S random variables u_i^s from the distributions $g_\varepsilon(\cdot)$ and $g_u(\cdot)$ respectively

$$\tilde{l} = \sum_{i=1}^{n} \log \left(\frac{1}{S} \sum_{s=1}^{S} \left\{ \prod_{t=1}^{T_i} \frac{1}{R} \sum_{r=1}^{R} \left\{ \prod_{j=1}^{J} h_j(Y_{itj}|Y_{i,t-1}, X_{it}, u_i^s, \varepsilon_{it}^r) \right\} \right\} \right). \tag{17}$$

Given our distributional assumptions $g_\varepsilon(\cdot) = g_u(\cdot) = \phi(\cdot)$, where $\phi(\cdot)$ is the standard normal density function.[31] Again, instead of using pseudo-random draws to obtain ε^r and u^s we base the simulation on Halton sequences. Different bases are used to generate the sequences for ε and u. Specifically, we use base 3 for ε and base 5 for u. Note that estimation of this model is computationally demanding. In fact, evaluation of the marginal likelihood requires summing $S \times R$ terms.

3.6.2 Results

Table 22 and Table 23 show SML estimates of the dynamic factor model with unobserved heterogeneity by birth cohort and gender. Like in the static case, a LR test strongly rejects the "independence model" in favour of our factor model. Table 22 reports medical drug use, hospitalization and mortality equations. As expected, there is strong state dependence in medical drug use and in hospitalization.

Table 23 reports the common factor equation. Like in the static case, estimated factor loadings of all medical treatments are statistically significant and, as expected, have all negative sign. A depreciation of latent health as age increases is reported. Estimated second order polynomials of $Y_{1,t-1}$ on current health implies a reversed U-shaped effects of lagged medical drug use on current health. Being hospitalized in period $t-1$ implies a negative shock in current health. In fact, the coefficients of $Y_{2,t-1}$ and $Y_{3,t-1}$ are negative, although not significant for younger women. Coefficients of the initial period observations have the expected sign. In fact, medical drug use observed in the first sample period is reported to be negatively related to latent health. Analogously, being hospitalized in the first period is associated to lower health. Although not always significant, medical drug use and the number of years under therapy during the years 1993-1996 (prior to the sample period) are both associated with lower health too. A LR test strongly rejects the hypothesis of no unobserved heterogeneity in latent health ($\sigma_u = 0$). For men and women

[31] In this paper we assumed that residuals, measurement errors, and unobserved individual effects are all normally distributed. We leave the issue of investigating the sensitivity of our results to different distributional assumptions to future research.

born 1920-1930 respectively 37% and 33% of the error variance in latent health is attributable to unobserved heterogeneity, as measured by the intra-class correlation coefficient $\rho = \sigma_u^2/(c_{11}^2 + \sigma_u^2)$. For younger cohorts, the error variance in latent health due to unobserved heterogeneity is about 17% and 67% respectively.

Figure 16 shows estimated polynomials of lagged medical drug use in the factor equation from the dynamic model, with pointwise 95% confidence intervals. These polynomials implies a reversed U-shaped effects of lagged medical drug use on current health. In fact, these polynomials are increasing from low levels of drug use until 0 (1 pill per day), then they become flat and start to decrease slowly. This means that taking less than a half pill or less than one pill per day today may result in poorer health tomorrow. This is consistent with the fact that not taking the correct medication can indeed have negative consequences on health, but taking more pills than needed cannot improve health. Thus, in addition to the negative correlation between current levels of health and current medical treatment found also in our static model, estimation results from the dynamic factor model tell us more about the effects of medical treatment on health.

3.7 Conclusions

In this paper, we look at the relationship between health and medical care using a dataset where very detailed information about medical drug use, hospitalization, and mortality, is collected over time for a selected sample of individuals. In particular, we focus on patients who were prescribed medical drugs employed against hypertension, a chronic asymptomatic pathology affecting a large share of the adult population. Due to the peculiarity of the pathology considered here, focusing on such patients offers the possibility of having helpful insights into the relation between health and medical care.

Previous analyses of the relationship between health and medical care have used simple empirical models that look at the raw correlation between medical care and health. This is not expected to give the right answer, because of simultaneity through the unobservable components of deterioration. In fact, authors have found a negative correlation between the stock of health and medical care, when medical care is supposed to have positive effects on health (Grossman, 1972). We proposed a dynamic factor model where medical drug use, hospitalization and mortality may all in principle be driven by latent individual stock of health. Our dynamic model provides a tool for investigating the effects of medical treatment against health deterioration. Our model also allows for persistence in the medical treatment due to state dependence and unobserved individual heterogeneity in the latent stock of health. We estimated the model by Maximum Simulated Likelihood.

In line with previous findings, our results indicate that better health is associated to lower medical treatments. In fact, estimated factor loadings have all negative sign. In addition, from our dynamic model we found a reversed U-shaped relation between lagged medical drug use and

current health. This is consistent with the fact that not taking the medication today may result in poorer health tomorrow. Stated differently, our results indicate that taking the medication may help reduce health deterioration and avoid its negative consequences. Nonetheless, taking more pills than needed cannot improve health.

Our findings have important policy implications. In fact, our results suggest that policies aimed at improving awareness of hypertensive diseases and the importance of the treatment of high blood pressure may help reduce cardiovascular risks, and consequent hospitalization and mortality. This is expected to have positive implications both for the large share of adult population suffering from hypertension and for the National Health Systems themselves.

References

Bartholomew DJ, Knott M. (1999). *Latent Variable Models and Factor Analysis*. London: Arnold.

Case A, Deaton A. (2004). Broken down by work and sex: how our health declines. Center for Health and Wellbeing Working Paper n. 25. **CHW WP#25**.

Chamberlain G. (1984). Panel data, in *Handbook of Econometrics* , volume 1, chapter 22, pp 1247-1318, edited by Griliches Z. and Intriligator M. D., North Holland: Amsterdam.

Gouriéroux C, Monfort A. (1997). *Simulation based econometric methods*. New York, NY: Oxford University Press.

Gross TP, Wise RP, Knapp DE (1989). Antihypertensive drug use. Trends in the United States from 1973 to 1985. *Hypertension*; **13(5suppl)**: I113-I118.

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

Grossman M. (2000). The human capital model, in *Handbook of Health Economics*, edited by Culyer A. J. and Newhouse J. P., Amsterdam: Elsevier.

Heckman JJ. (1981). The incidental parameters problem and the problem of initial conditions in estimating a discrete timediscrete data stochastic process. In CF Manski, D McFadden (eds.) *Structural Analysis of Discrete Data with Econometric Applications*, pp. 179-195, MIT Press: Cambridge, MA.

Heckman JJ, Singer B. (1984). A method of minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica*, **52**: 271–320.

Sandor Z, Train K. (2004). Quasi-Random Simulation of Discrete Choice Models. *Transportation Research*, **Part B, Vol. 38**: 313–327.

Snijders TAB, Bosker RJ. (1999). *Multilevel Analysis*. London: Sage.

Train K. (2003). *Discrete Choice Models using Simulation*. Cambridge University Press, Cambridge: UK. Pre-print version available from http://emlab.berkeley.edu/users/train/distant.html.

Wagstaff A., (1993). The demand for health: an empirical reformulation of the Grossman model. *Health Economics*, **2(2)**: 189–98.

Wang X, Hickernell F (2000). Randomized halton sequences. *Mathematical and Computer Modeling*, **32**: 887–899.

Wooldridge, J.M. (2000). A framework for estimating dynamic, unobserved effects panel data models with possible feedback to future explanatory variables. *Economics Letters*, **68**: 245–250.

Woolridge, J. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics*, **20**: 39-54.

Table 20: Descriptive statistics.

	Men				Women			
	1920-1930		1931-1940		1920-1930		1931-1940	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Log. of drug use	0.101	0.664	0.071	0.633	0.025	0.661	-0.019	0.635
Frac. hospit. (other DRGs)	0.207	0.405	0.130	0.336	0.162	0.368	0.105	0.307
Frac. hospit. (card. DRGs)	0.099	0.299	0.060	0.238	0.059	0.235	0.037	0.189
Mortality rate	0.041	0.199	0.015	0.122	0.017	0.129	0.006	0.075
Age	74.0	3.5	64.0	3.3	74.5	3.5	64.3	3.3
Age of physician	48.2	7.5	48.3	7.4	48.1	7.2	48.0	7.2
Female phys.	0.163	0.369	0.151	0.358	0.178	0.382	0.164	0.370
AH	0.079	0.269	0.084	0.278	0.083	0.276	0.080	0.271
D	0.330	0.470	0.209	0.406	0.343	0.475	0.258	0.437
BBA	0.219	0.414	0.300	0.458	0.246	0.431	0.328	0.469
CCB	0.507	0.500	0.461	0.498	0.417	0.493	0.338	0.473
ARA	0.583	0.493	0.614	0.487	0.591	0.492	0.601	0.490
AH & D	0.031	0.173	0.021	0.145	0.031	0.174	0.025	0.155
AH & BBA	0.016	0.124	0.022	0.147	0.019	0.135	0.022	0.146
AH & CCB	0.038	0.192	0.037	0.188	0.032	0.175	0.025	0.155
AH & ARA	0.049	0.215	0.055	0.229	0.051	0.220	0.050	0.218
D & BBA	0.064	0.245	0.058	0.234	0.071	0.257	0.069	0.253
D & CCB	0.143	0.350	0.092	0.289	0.130	0.336	0.082	0.275
D & ARA	0.165	0.371	0.113	0.316	0.158	0.365	0.115	0.318
BBA & CCB	0.093	0.291	0.108	0.311	0.081	0.273	0.082	0.274
BBA & ARA	0.107	0.309	0.137	0.344	0.113	0.316	0.136	0.343
CCB & ARA	0.235	0.424	0.223	0.416	0.199	0.399	0.161	0.368
Log. of drug use (1993-1996)	-0.252	0.741	-0.215	0.705	-0.325	0.733	-0.329	0.725
Years (1993-1996)	2.311	0.935	2.205	0.956	2.331	0.934	2.250	0.964
Observations	23,771		20,456		38,950		25,771	
Patients	4,509		3,581		6,829		4,371	

Table 21: Estimated coefficients of the static factor model by birth cohort and gender.

	Men		Women	
	1920-1930	1931-1940	1920-1930	1931-1940
Equation 1. Log. of drug use				
Age of phys./10	-0.009	0.002	-0.012 **	0.013 **
Female phys.	-0.008	0.011	-0.012	-0.009
AH	0.628 **	0.625 **	0.622 **	0.609 **
D	0.374 **	0.419 **	0.162 **	0.133 **
BBA	0.154 **	0.040 **	0.113 **	0.049 **
CCB	0.149 **	0.123 **	0.174 **	0.186 **
AH & D	-0.145 **	-0.234 **	-0.057 *	-0.089 **
AH & BBA	-0.157 **	-0.121 **	-0.128 **	-0.100 **
AH & CCB	-0.104 **	-0.116 **	-0.201 **	-0.164 **
AH & ARA	0.056 *	0.071 **	0.084 **	0.077 **
D & BBA	-0.025	-0.013	0.106 **	0.105 **
D & CCB	-0.014	-0.085 **	0.128 **	0.102 **
D & ARA	0.157 **	0.107 **	0.238 **	0.251 **
BBA & CCB	0.054 **	0.142 **	0.044 **	0.105 **
BBA & ARA	0.283 **	0.391 **	0.325 **	0.390 **
CCB & ARA	0.449 **	0.444 **	0.376 **	0.357 **
Constant	-0.149 **	-0.049	-0.159 **	0.292 **
ln σ	-0.622 **	-0.692 **	-0.617 **	-0.665 **
Equation 2. Hospitalization (other DRGs)				
Constant	1.251 **	1.248 **	1.166 **	0.914 **
Equation 3. Hospitalization (card. DRGs)				
Constant	-0.254 **	0.034	-0.396 **	-0.236
Equation 4. Mortality				
Constant	0.000	0.000	0.000	0.000
Factor Equation				
$(Age/10)$	-0.320 **	-0.344 **	-0.451 **	-0.382 **
Constant	2.109 **	2.415 **	2.751 **	2.711 **
ln c_{11}	-0.304 **	-0.636 **	-0.025	-0.920 **
Factor loadings				
γ_1	-0.089 **	-0.103 **	-0.072 **	-0.227 **
γ_2	-1.064 **	-1.034 **	-0.866 **	-0.816 **
γ_3	-0.521 **	-0.687 **	-0.464 **	-0.585 **
γ_4	-1.000	-1.000	-1.000	-1.000
Obs.	23,771	20,456	38,950	25,771
Log Like.	-42,498	-28,960	-52,247	-33,364
LR test	993	384	850	302
p(LR test)	0	0	0	0

* Significant at 5%; ** Significant at 1%.

Table 22: Estimated coefficients of the dynamic factor model by birth cohort and gender. Medical drug use, hospitalization and mortality equations.

	Men		Women	
	1920-1930	1931-1940	1920-1930	1931-1940
Equation 1. Log. of drug use				
$Y_{1,t-1}$	0.534 **	0.532 **	0.552 **	0.512 **
Age of phys./10	-0.006	-0.006	-0.009 **	0.005
Female phys.	-0.004	-0.008	-0.009	-0.006
AH	0.328 **	0.284 **	0.297 **	0.274 **
D	0.180 **	0.228 **	0.087 **	0.072 **
BBA	0.091 **	0.019	0.070 **	0.021 *
CCB	0.062 **	0.053 **	0.086 **	0.081 **
AH & D	-0.125 **	-0.121 **	-0.016	-0.088 **
AH & BBA	-0.094 **	-0.055 *	-0.070 **	-0.068 **
AH & CCB	-0.058 **	-0.034	-0.072 **	-0.068 **
AH & ARA	0.044	0.063 **	0.059 **	0.084 **
D & BBA	0.009	-0.023	0.046 **	0.041 **
D & CCB	0.032 *	-0.028	0.078 **	0.056 **
D & ARA	0.115 **	0.068 **	0.136 **	0.136 **
BBA & CCB	0.026	0.074 **	0.021	0.068 **
BBA & ARA	0.146 **	0.210 **	0.176 **	0.207 **
CCB & ARA	0.260 **	0.240 **	0.211 **	0.198 **
Constant	0.027	0.108 **	0.006	1.149 **
ln σ	-0.904 **	-0.978 **	-0.898 **	-0.972 **
Equation 2. Hospitalization (other DRGs)				
$Y_{2,t-1}$	0.362 **	0.465 **	0.336 **	0.693 **
Constant	0.178 **	0.074	0.568 **	-0.070
Equation 3. Hospitalization (card. DRGs)				
$Y_{3,t-1}$	0.700 **	0.748 **	0.693 **	0.869 **
Constant	-0.556 **	-0.578 **	-0.402 **	0.407
Equation 4. Mortality				
Constant	0.000	0.000	0.000	0.000

* Significant at 5%; ** Significant at 1%.

Table 23: Estimated coefficients of the dynamic factor model by birth cohort and gender. Health equation.

	Men		Women	
	1920-1930	1931-1940	1920-1930	1931-1940
Factor Equation				
$(Age/10)$	-2.229 **	-1.275 **	-1.194 **	-0.248 **
$(Y_{1,t-1})$	0.098 **	0.079	-0.021	0.045
$(Y_{1,t-1})^2$	-0.189 **	-0.151 **	-0.143 **	-0.122 **
$Y_{2,t-1}$	-0.242 **	-0.387 **	-0.255 **	-0.004
$Y_{3,t-1}$	-0.323 **	-0.217 **	-0.245 **	-0.053
Average age	2.114 **	1.127 **	1.011 **	0.185 **
Y_{10}	-0.276 **	-0.274 **	-0.078 **	-0.172 **
Y_{20}	-0.471 **	-0.325 **	-0.482 **	-0.091 **
Y_{30}	-0.333 **	-0.415 **	-0.477 **	-0.074 *
Y_1 (1993-1996)	0.018	-0.063	-0.060 **	-0.080 **
Years (1993-1996)	-0.048 *	-0.068 **	-0.046 **	-0.003
Constant	2.932 **	3.151 **	3.049 **	2.622 **
ln c_{11}	-0.232 **	-0.193	-0.437 **	-2.205 **
ln σ_u	-0.497 **	-0.997 **	-0.776 **	-1.843 **
Factor loadings				
γ_1	-0.066 **	-0.076 **	-0.057 **	-0.489 **
γ_2	-0.450 **	-0.464 **	-0.620 **	-0.494 **
γ_3	-0.363 **	-0.398 **	-0.487 **	-0.886 **
γ_4	-1.000	-1.000	-1.000	-1.000
Obs.	19,262	14,734	32,121	21,400
Log Like.	-28,202	-16,267	-39,706	-21,070
LR test	1,497	518	1,552	806
p(LR test)	0	0	0	0
LR test $\sigma_u = 0$	145	18	138	67
p(LR test)	0	0	0	0
Intra-class correlation				
$\rho = \sigma_u^2/(c_{11}^2 + \sigma_u^2)$	0.370 **	0.167 **	0.337 **	0.674 **
Joint significance tests				
$Y_{1,t-1}$	68.336 **	22.193 **	59.949 **	21.774 **

* Significant at 5%; ** Significant at 1%.

Figure 13: Distribution of patients by gender and year of birth.

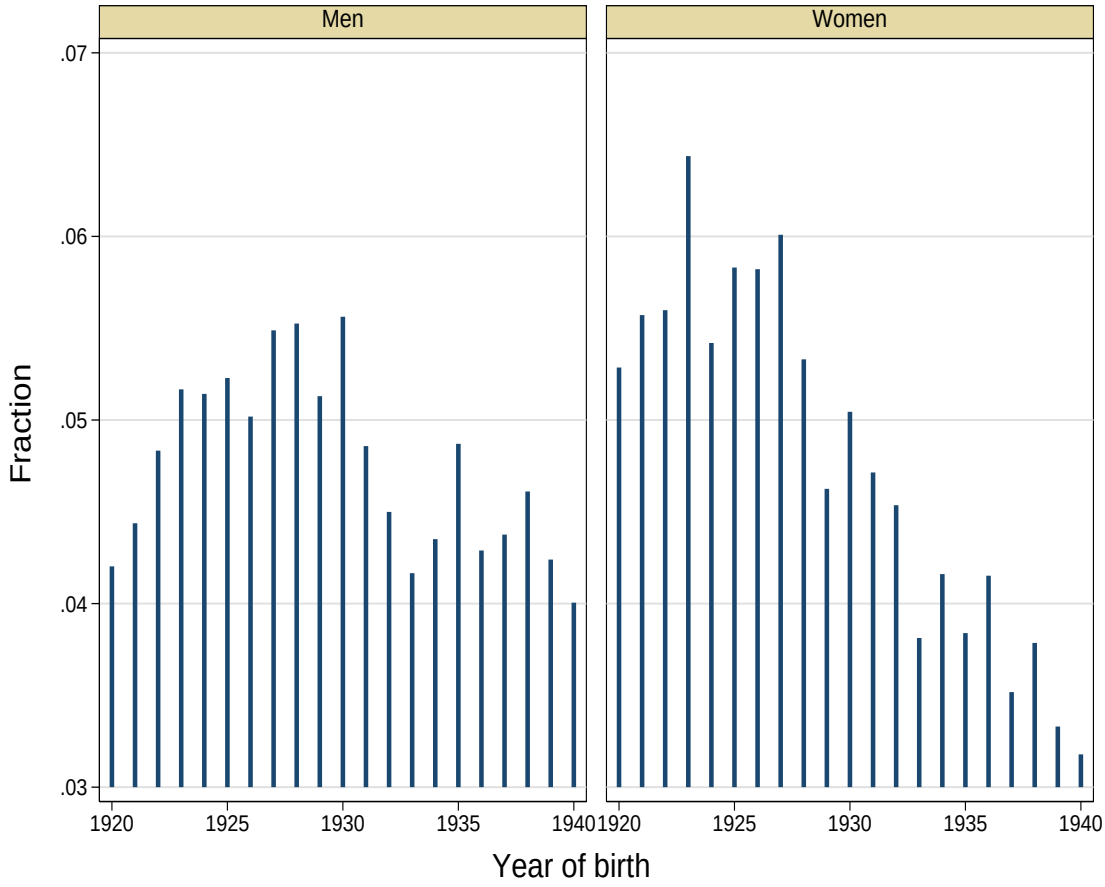


Figure 14: Histogram of medical drug use.

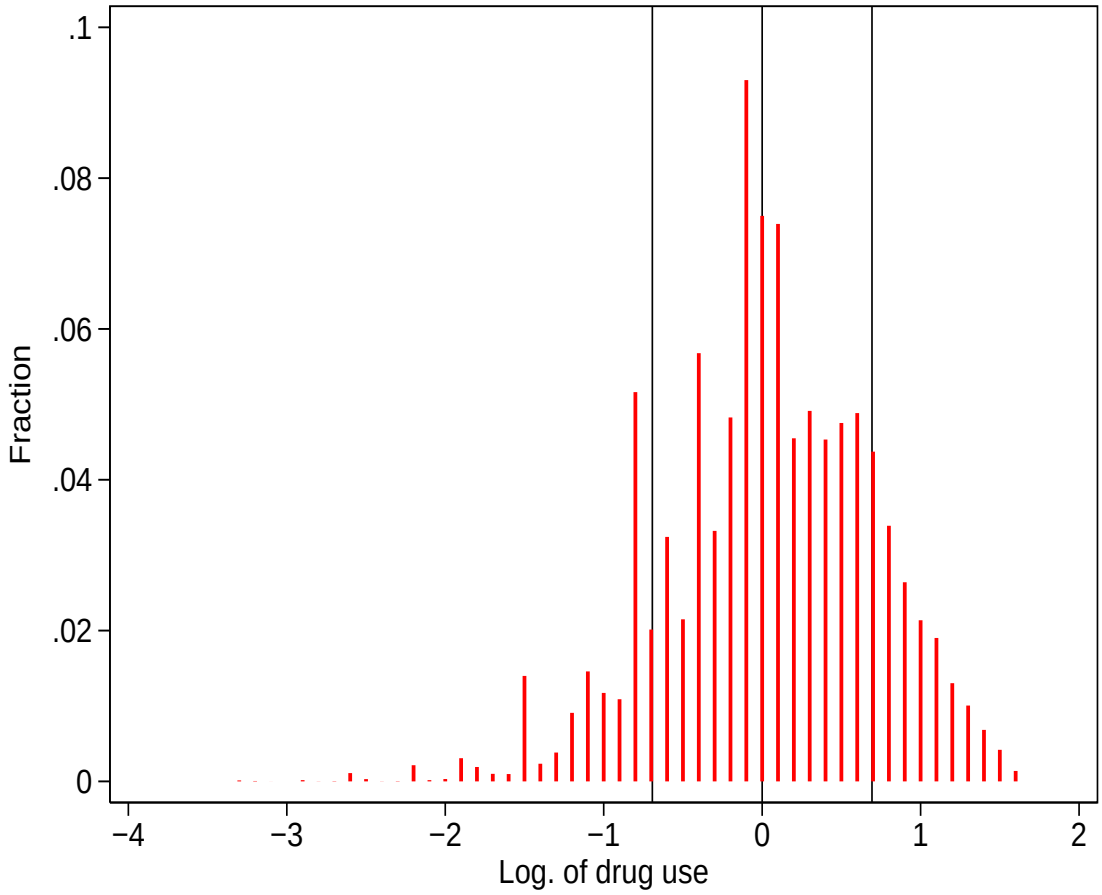
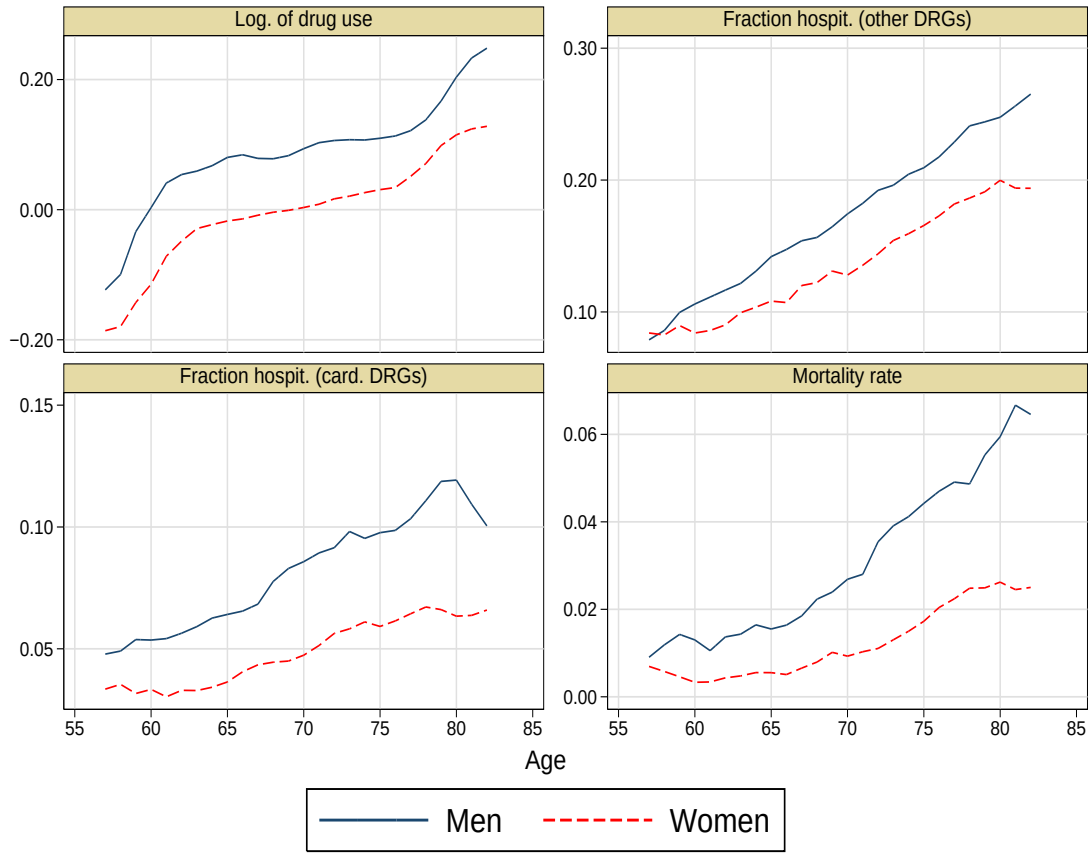
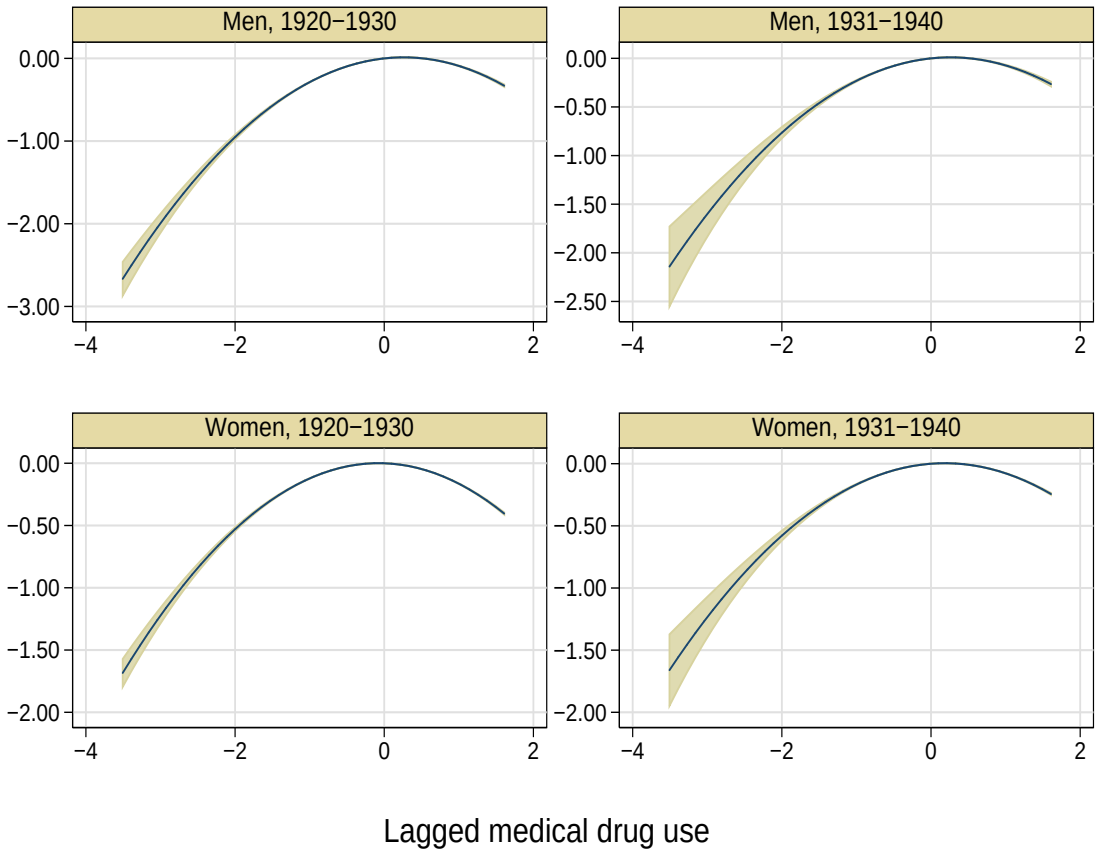


Figure 15: Average drug use, hospitalization, and mortality by age and gender.



MA(3) smoothing

Figure 16: Estimated polynomials of lagged medical drug use in the Factor equation, with pointwise 95% confidence intervals. Dynamic model.



4 APPENDICES

A The Survey of Health, Ageing and Retirement in Europe

The Survey of Health, Ageing and Retirement in Europe (SHARE) is a multidisciplinary and cross-national longitudinal survey on health, socio-economic status, and social and family networks.[32] Eleven countries have contributed data to the 2004 SHARE baseline study. They are a representation of the various regions of Europe, ranging from Scandinavia (Denmark, Sweden) through Central Europe (Austria, France, Belgium, Germany, Netherlands, Switzerland) to the Mediterranean region (Greece, Italy, Spain). The data used in this study are from Release 2 of the first (2004) wave. Table A1 gives a breakdown of all wave 1 2004 samples (release 2) by country, gender and age groups. Complete documentation and methodological details about SHARE are contained in the SHARE methodology volume (see Börsch-Supan and Jürges 2005).

The target population of SHARE consists of all people born in 1954 or earlier, plus their (possibly younger) spouses/partners. Partners may be younger than 50, but must be living at the exact same address as the selected age-eligible respondent. The target population is further restricted by excluding people who currently do not reside at the sampled address, or died before the starting of the field period, or are unable to speak the specific language of the national questionnaire, or are physically or mentally unable to participate to the survey. These people are referred to as nonsample persons. Households where all members are nonsample persons are instead referred to as nonsample households. Notice that, people living in institutions for elderly are included in the target population, while people living in prisons or similar institutions are not.

The institutional conditions and regulations with respect to sampling of the countries involved in the SHARE project were so different that using a common sampling frame and sampling design for all countries was infeasible. In most countries, suitable sampling frames for the target population investigated by SHARE either do not exist, or could not be used. Thus, national sampling frames were selected depending on what was already available in each country.

The national sampling frames used in the first wave of SHARE can be distinguished between two types of sampling frames. In one group of countries (Denmark, Germany, Italy, Netherlands, Spain, and Sweden), the sampling frame is a population register of households or individuals. In the other group of countries, the sampling frame is either a telephone register (Austria, Belgium, Greece, and Switzerland) or a register of dwellings (France). A detailed description of the frame adopted in each SHARE country can be found in Klevmarken *et al.* (2005).

For most of the SHARE countries the sample consists of two parts, the main sample and the vignette sample. The main sample is the original part of the sample. The vignette sample is an additional part of the sample that was added in eight countries (Belgium, France, Germany, Greece,

[32] The data may be downloaded by registered users from the SHARE website (http://www.share-project.org).

Italy, Netherlands, Spain and Sweden). Main and vignette samples differ because for the vignette sample the latest a part of the self-completion questionnaire is replaced by a section with anchoring vignette questions. These questions were designed to improve cross-country comparability of the data. Vignettes have been used as a tool to overcome measurement errors in self-reported measures. In fact, when asking self-reported health questions, people belonging to different socio-economic groups may assign different answers to the same situation. The idea of vignettes is to ask people not only about their own situation, but to ask them also to evaluate that of some hypothetical person, which is the same for every respondent. These additional measures can be used to adjust answers for their systematic differences in response scales. Finally, a supplementary sample was added in Sweden in order to increase the low number of achieved interviews.

The interview mode adopted by SHARE is Computer Assisted Personal Interview (CAPI), supplemented by show-cards and a self-administered paper and pencil questionnaire (the drop-off questionnaire). The CAPI interview, which is known to be one of the most effective interview modes, represents the largest part of the interview. The drop-off questionnaire is instead used to ask more sensitive questions, like questions on social and psychological well-being, health-care, religiosity and political affiliation, and to expand topics not fully accommodated in the CAPI interview. Respondents of the main sample received the same version of the questionnaire, while respondents of the vignette sample received one of two different versions of the vignette questionnaire which were randomized by interviewer. The CAPI interview is organized in twenty modules which cover a wide range of topics including demographics, family composition, physical and cognitive functioning, mental health, health care services, self-reported and objective health measures, well-being, labor force participation, income, consumption, assets, financial transfers, social relations and expectations.

SHARE provides three types of weights that can be used to compensate for unequal selection probabilities and unit nonresponse errors. They are the sampling design weights, the calibrated household weights, and the calibrated individual weights. Sampling design weights are equal to the inverse of the selection probability of each sampled unit. Calibrated household and individual weights are instead adjustments of the sampling design weights that help reduce the bias generated by unit nonresponse. For countries including a vignette sample, each type of weight has been computed both for the main sample and the vignette sample separately, and the two samples combined. As discussed in Klevmarken *et al.* (2005), the selection probability of each sample unit has been computed separately by country according to the specific sampling design. Given that the probability of selecting a household is the same for all eligible household members, the sampling design weight of each eligible household member is the same as the sampling design weight of the corresponding household.

References

Börsch-Supan A., H. Jürges, O. Lipps (2003), SHARE: Building a Panel Survey on Health, Ageing and Retirement in Europe, *MEA Discussion Paper Series, 32-03.*

Börsch-Supan A., and H. Jürges (2005), The Survey of Health, Ageing and Retirement in Europe - Methodology, MEA, Mannheim.

Klevmarken A., Hesselius P., and Swensson B., (2005), The SHARE Sampling Procedures and Calibrated Designs Weights, in *The Survey of Health, Ageing and Retirement in Europe - Methodology, edited by Börsch-Supan A., and H. Jürges.*

Table A1: Breakdown of all wave 1 2004 samples (release 2) by country, gender and age groups.

Country	Total	Male	Female	>50	50 - 64	65-74	<75	Household Response Rate	Individual Response Rate
Austria	1,893	782	1,111	44	949	544	356	55,6%	87,5%
Belgium	3,827	1,739	2,088	178	1,991	986	672	39,2%	90,5%
Denmark	1,707	771	936	92	916	369	330	63,2%	93,0%
France	3,193	1,386	1,807	155	1,648	759	631	81,0%	93,3%
Germany	3,008	1,380	1,628	65	1,569	886	486	63,4%	86,2%
Greece	2,898	1,244	1,654	229	1,458	712	499	63,1%	91,8%
Italy	2,559	1,132	1,427	51	1,342	785	381	54,5%	79,7%
Netherlands	2,979	1,368	1,611	102	1,693	713	459	61,6%	87,8%
Spain	2,396	994	1,402	42	1,079	701	573	53,0%	73,7%
Sweden	3,053	1,414	1,639	56	1,589	816	592	46,9%	84,6%
Switzerland	1,004	462	542	42	505	251	204	38,8%	86,9%
Total	31,115	13,811	17,304	1,198	16,155	8,212	5,530	61,6%	85,3%

B Description of self-assessments on health domains

Self-assessment questions for each health domain are the following.

Pain

"Overall in the last 30 days, how much of bodily aches or pains did you have?".

Sleeping problems

"In the last 30 days, how much difficulty did you have with sleeping such as falling asleep, waking up frequently during the night or waking up too early in the morning?".

Mobility

"Overall in the last 30 days, how much of a problem did you have with moving around?".

Concentration problems

"Overall in the last 30 days how much difficulty did you have with concentrating or remembering things?".

Shortness of breath

"In the last 30 days, how much of a problem did you have because of shortness of breath?".

Depression

"Overall in the last 30 days, how much of a problem did you have with feeling sad, low, or depressed?".

C Description of vignette hypothetical situations

The three vignette hypothetical situations for each health domain are the following.

Pain

1. *"Paul/Karen has a headache once a month that is relieved after taking a pill. During the headache he/she can carry on with his/her day-to-day affairs."*
2. *"Henri/Maria has pain that radiates down his/her right arm and wrist during his/her day at work. This is slightly relieved in the evenings when he/she is no longer working on his/her computer."*
3. *"Charles/Alice has pain in his/her knees, elbows, wrists and fingers, and the pain is present almost all the time. Although medication helps, he/she feels uncomfortable when moving around, holding and lifting things."*

Sleeping problems

1. *"Charles/Alice falls asleep easily at night, but two nights a week he/she wakes up in the middle of the night and cannot go back to sleep for the rest of the night."*
2. *"Paul/Karen wakes up almost once every hour during the night. When he/she wakes up in the night, it takes around 15 minutes for his/her to go back to sleep. In the morning he/she does not feel well-rested."*
3. *"Henri/Maria takes about two hours every night to fall asleep. He/she wakes up once or twice a night feeling panicked and takes more than one hour to fall asleep again."*

Mobility

1. *"Rob/Eve is able to walk distances of up to 200 metres without any problems but feels tired after walking one kilometre or climbing more than one flight of stairs. He/she has no problems with day-to-day activities, such as carrying food from the market."*
2. *"Kevin/Lisa does not exercise. He/she cannot climb stairs or do other physical activities because he/she is obese. He/she is able to carry the groceries and do some light household work."*
3. *"Tom/Sue has a lot of swelling in his/her legs due to his/her health condition. He/she has to make an effort to walk around his/her home as his/her legs feel heavy."*

Concentration problems

1. *"Kevin/Lisa can concentrate while watching TV, reading a magazine or playing a game of cards or chess. Once a week he/she forgets where his/her keys or glasses are, but finds them within five minutes."*
2. *"Tom/Sue is keen to learn new recipes but finds that he/she often makes mistakes and has to reread several times before he/she is able to do them properly."*

3. *"Rob/Eve cannot concentrate for more than 15 minutes and has difficulty paying attention to what is being said to his/her. Whenever he/she starts a task, he/she never manages to finish it and often forgets what he/she was doing. He/she is able to learn the names of people he/she meets."*

Shortness of breath

1. *"Mark/Karen has no problems with walking slowly. He/she gets out of breath easily when climbing uphill for 20 meters or a flight of stairs."*
2. *"Paul/Karen suffers from respiratory infections about once every year. He/she is short of breath 3 or 4 times a week and had to be admitted in hospital twice in the past month with a bad cough that required treatment with antibiotics."*
3. *"Henri/Maria has been a heavy smoker for 30 years and wakes up with a cough every morning. He/she gets short of breath even while resting and does not leave the house anymore. He/she often needs to be put on oxygen."*

Depression

1. *"Paul/Karen enjoys his/her work and social activities and is generally satisfied with his/her life. He/she gets depressed every 3 weeks for a day or two and loses interest in what he/she usually enjoys but is able to carry on with his/her day-to-day activities."*
2. *"Henri/Maria feels nervous and anxious. He/she worries and thinks negatively about the future, but feels better in the company of people or when doing something that really interests his/her. When he/she is alone he/she tends to feel useless and empty."*
3. *"Mark/Anna feels depressed most of the time. He/she weeps frequently and feels hopeless about the future. He/she feels that he/she has become a burden on others and that he/she would be better dead."*

D Parameter estimates of the OLS regression for poor health by region and gender

This appendix presents parameter estimates of the OLS regression for poor health separately for non-Mediterranean women, non-Mediterranean men, Mediterranean women, and Mediterranean men (* significant at 5%; ** significant at 1%).

	Mediterranean countries		Non-Mediterranean countries	
	Men	Women	Men	Women
Heart attack	0.197 **	0.160 **	0.139 *	0.086
High blood pressure	0.086 **	0.040	-0.018	0.120 **
High blood cholesterol	-0.056	0.059 *	0.032	0.062
Stroke	0.093	0.054	0.238	0.054
Diabetes	0.147 **	0.172 **	0.156 **	0.064
Chronic lung disease	-0.052	0.050	0.091	-0.100
Asthma	0.010	-0.027	0.116	0.048
Arthritis	0.068	0.113 **	0.053	0.107 **
Osteoporosis	0.086	0.145 **	0.096	0.139 **
Ulcer	-0.043	0.154 *	0.170 **	-0.036
Parkinson disease	0.392 **	0.286	-0.198	0.078
Cataracts	0.066	-0.118 *	-0.032	0.029
Hip or femoral fracture	-0.116	-0.228 *	-0.116	0.153
Reproductive cancer	0.150	0.022	0.336	0.234 *
Other cancer	0.088	0.324 **	0.374 **	0.342 **
Pain in back	0.054 *	-0.011	0.045	-0.048
Heart trouble	0.065	0.070	0.230 *	-0.027
Breathlessness	0.217 **	0.235 **	0.049	0.163 *
Persistent cough	0.013	0.094	0.157	-0.012
Swollen legs	0.017	-0.092 **	0.024	0.095 *
Sleeping problems	0.012	0.008	0.148 *	0.074
Falling down	0.140	-0.017	-0.139	-0.123
Fear of falling down	0.018	0.005	0.122	0.099
Dizziness	-0.030	0.075	0.009	-0.011
Stomach problems	0.085 *	0.031	-0.123 *	0.045
Incontinence	-0.248 **	-0.032	0.027	-0.030
Other symptoms	0.014	0.107	0.038	-0.080
Low grip strength	0.089 **	-0.014	0.094 *	0.052
Underweight	0.235	-0.023	-0.053	-0.274
Overweight	-0.080 **	-0.006	0.026	0.007
Obese	-0.161 **	0.031	0.036	-0.064
Pain: Mild	0.063 *	0.099 **	0.134 **	0.159 **
Pain: Mod/sev/extr	0.246 **	0.228 **	0.139 *	0.220 **
Sleeping problems: Mild	0.022	0.005	-0.047	-0.032
Sleeping problems: Mod/sev/extr	0.030	0.054	-0.019	-0.068
Mobility problems: Mild	0.099 **	0.104 **	0.127 *	0.087
Mobility problems: Mod/sev/extr	0.131 **	0.191 **	0.286 **	0.102
Concentration problems: Mild	-0.040	-0.074 **	-0.027	0.015
Concentration problems: Mod/sev/extr	0.121 **	0.052	-0.161 **	0.104 *
Shortness of breath: Mild	0.048	0.054	0.039	-0.013
Shortness of breath: Mod/sev/extr	0.035	-0.074	-0.005	0.037
Depression: Mild	-0.018	-0.016	0.018	0.095 *
Depression: Mod/sev/extr	0.057	0.053	0.064	0.048
Age - 55	0.011 **	0.001	0.007	0.009 *
(Age - 55) squared /100	-0.057 **	0.014	-0.008	-0.025
Secondary education	-0.045	-0.047	-0.066	-0.066
Post-secondary education	-0.051	-0.107 **	-0.050	0.021
Living with spouse or partner	0.022	0.001	0.042	-0.065
Log HH income - log(12763)	-0.043 **	-0.025 *	-0.039 *	-0.033 *
Constant	0.110 *	0.127 **	-0.009	0.119 *
Observations	1,069	1,234	626	678
R^2	0.426	0.445	0.373	0.439

Notes: weighted results, country dummies omitted

E Parameter estimates of an ordered probit model with individual specific thresholds for each health domain

This appendix presents all parameter estimates of the ordered probit model with constant thresholds, and ordered probit model with individual specific thresholds for selected health domains, including parameters in Table 10 (* significant at 5%; ** significant at 1%).

		Pain		Mobility problems		Concentration problems	
Equation	Variable	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds
SA	Heart attack	0.251 **	0.217 *	0.117	0.009	0.068	-0.026
	High blood pressure	0.069	0.060	0.023	0.005	0.060	0.044
	High blood cholesterol	0.033	-0.045	-0.076	-0.098	0.122 *	0.096
	Stroke	0.349 **	0.285	0.354 **	-0.020	0.384 **	0.098
	Diabetes	0.366 **	0.516 **	0.510 **	0.664 **	0.276 **	0.343 **
	Chronic lung disease	0.125	0.226	0.045	0.130	0.035	0.137
	Asthma	0.351 **	0.331 *	0.068	0.128	0.013	0.123
	Arthritis	0.543 **	0.468 **	0.131 *	0.091	0.066	0.004
	Osteoporosis	0.287 **	0.151	0.215 *	0.161	0.064	-0.034
	Ulcer	0.124	0.172	0.084	0.005	0.168	0.144
	Parkinson disease	0.869 *	0.448	0.992 *	0.442	0.458	0.334
	Cataracts	-0.218 *	-0.235 *	-0.033	0.057	-0.233 **	-0.083
	Hip or femoral fracture	0.002	-0.085	-0.132	-0.261	-0.271	-0.491 *
	Reproductive cancer	0.257	0.271	0.266 *	0.265	0.149	0.185
	Other cancer	0.441 **	0.415 **	0.354 **	0.161	-0.172	-0.192
	Pain in back	0.625 **	0.670 **	0.497 **	0.556 **	0.248 **	0.286 **
	Heart trouble	0.198 *	0.173	0.295 **	0.409 **	0.078	0.065
	Breathlessness	0.168 *	0.159	0.259 **	0.270 **	0.398 **	0.352 **
	Persistent cough	0.179	0.058	0.027	-0.358 **	0.090	-0.213
	Swollen legs	0.127	0.125	0.295 **	0.289 **	0.107	0.120
	Sleeping problems	0.181 **	0.216 **	0.106	0.013	0.209 **	0.140 *
	Falling down	0.067	0.038	0.020	0.248	-0.018	0.025
	Fear of falling down	0.317 **	0.014	0.479 **	0.162	0.116	-0.084
	Dizziness	0.213 *	0.162	0.314 **	0.190 *	0.175 *	0.050
	Stomach problems	0.257 **	0.143 *	0.061	0.088	0.076	0.097
	Incontinence	0.210	0.011	0.076	-0.151	0.146	-0.110
	Other symptoms	0.169	-0.017	0.327 **	-0.009	0.530 **	0.297 **
	Low grip strength	0.224 **	0.170 **	0.409 **	0.376 **	-0.023	0.035
	Underweight	0.492 *	0.270	-0.222	-0.225	-0.165	-0.339
	Overweight	0.083	0.070	0.148 **	0.080	0.025	0.082
	Obese	0.270 **	0.295 **	0.345 **	0.194 **	-0.199 **	-0.107
	Age - 55	-0.005	0.002	0.006	0.009	0.011 *	0.004
	(Age - 55) squared /100	0.044 *	0.038	0.047 *	0.046	0.038	0.060 *
	Secondary education	0.006	-0.065	0.062	-0.059	-0.121 *	-0.098
	Post-secondary education	-0.081	-0.283 **	-0.054	-0.383 **	-0.203 **	-0.317 **
	Living with spouse or partner	-0.046	-0.096	-0.135 **	-0.219 **	-0.161 **	-0.248 **
	Log HH income - log(12763)	0.020	-0.003	0.005	0.016	-0.011	-0.025
	medit	-0.154 *	-0.098	-0.449 **	-0.293 **	-0.294 **	-0.144 *
	female	0.277 **	0.215 **	0.016	0.069	0.001	0.005
	medit*female	-0.096	-0.172	0.142	0.042	0.272 **	0.109
	Constant	-0.309 **	-0.462 **	-0.891 **	-0.964 **	-0.103	-0.139

Thres. 1	Heart attack	.	-0.022	.	-0.128 *	.	-0.101
	High blood pressure	.	-0.053	.	-0.036	.	-0.005
	High blood cholesterol	.	-0.071	.	-0.017	.	-0.023
	Stroke	.	0.071	.	-0.446 **	.	-0.325 **
	Diabetes	.	0.313 **	.	0.192 **	.	0.038
	Chronic lung disease	.	0.118	.	0.074	.	0.163 *
	Asthma	.	-0.155	.	0.046	.	0.055
	Arthritis	.	-0.194 **	.	-0.052	.	-0.109 **
	Osteoporosis	.	-0.166 *	.	-0.014	.	-0.082
	Ulcer	.	0.111	.	-0.185 *	.	-0.049
	Parkinson disease	.	-0.725	.	-0.686	.	-0.284
	Cataracts	.	0.012	.	0.122	.	0.131 *
	Hip or femoral fracture	.	-0.360	.	-0.224	.	-0.407 **
	Reproductive cancer	.	-0.128	.	-0.067	.	0.061
	Other cancer	.	-0.158	.	-0.316 **	.	-0.026
	Pain in back	.	0.004	.	0.058	.	0.020
	Heart trouble	.	-0.085	.	0.088	.	0.014
	Breathlessness	.	0.068	.	0.032	.	0.008
	Persistent cough	.	-0.339 **	.	-0.428 **	.	-0.419 **
	Swollen legs	.	0.056	.	-0.058	.	0.000
	Sleeping problems	.	0.085	.	-0.079	.	-0.039
	Falling down	.	-0.076	.	0.297 **	.	0.020
	Fear of falling down	.	-0.420 **	.	-0.419 **	.	-0.237 **
	Dizziness	.	-0.209 **	.	-0.061	.	-0.130 *
	Stomach problems	.	-0.117 *	.	0.053	.	0.004
	Incontinence	.	-0.487 **	.	-0.158	.	-0.168 *
	Other symptoms	.	-0.183 *	.	-0.471 **	.	-0.228 **
	Low grip strength	.	0.017	.	-0.013	.	0.105 **
	Underweight	.	-0.170	.	0.026	.	-0.358
	Overweight	.	-0.034	.	-0.080 *	.	0.057
	Obese	.	0.108 *	.	-0.170 **	.	0.114 **
	Age - 55	.	0.013 **	.	0.003	.	-0.012 **
	(Age - 55) squared /100	.	-0.045 *	.	0.003	.	0.044 **
	Secondary education	.	-0.126 **	.	-0.151 **	.	0.017
	Post-secondary education	.	-0.318 **	.	-0.408 **	.	-0.146 **
	Living with spouse or partner	.	-0.093 *	.	-0.113 **	.	-0.125 **
	Log HH income - log(12763)	.	-0.033 *	.	0.009	.	-0.015
	medit	.	-0.036	.	0.153 **	.	0.179 **
	female	.	-0.151 **	.	0.027	.	-0.017
	medit*female	.	0.068	.	-0.067	.	-0.127 *
	Constant	0.000	0.000	0.000	0.000	0.000	0.000
Thres. 2	Heart attack	.	-0.022	.	0.074	.	-0.001
	High blood pressure	.	0.062 *	.	0.069	.	-0.010
	High blood cholesterol	.	-0.005	.	-0.015	.	-0.014
	Stroke	.	-0.150	.	0.194	.	0.079
	Diabetes	.	-0.205 **	.	-0.078	.	0.073
	Chronic lung disease	.	-0.046	.	0.025	.	-0.108
	Asthma	.	0.122	.	0.017	.	0.069
	Arthritis	.	0.131 **	.	0.026	.	0.096 **
	Osteoporosis	.	0.018	.	-0.103	.	-0.029
	Ulcer	.	-0.059	.	0.265 **	.	0.072
	Parkinson disease	.	0.272	.	0.127	.	0.158
	Cataracts	.	-0.006	.	-0.091	.	0.041
	Hip or femoral fracture	.	0.290 **	.	0.188	.	0.381 **
	Reproductive cancer	.	0.147 *	.	0.132	.	-0.051
	Other cancer	.	0.111	.	0.279 **	.	0.024
	Pain in back	.	0.076 **	.	-0.004	.	0.059 *
	Heart trouble	.	0.078	.	0.003	.	-0.041
	Breathlessness	.	-0.076	.	-0.067	.	-0.080
	Persistent cough	.	0.211 **	.	0.125	.	0.190 **
	Swollen legs	.	-0.058	.	0.134 *	.	0.028
	Sleeping problems	.	-0.078 *	.	-0.062	.	-0.075 *
	Falling down	.	0.066	.	-0.153	.	0.017
	Fear of falling down	.	0.115 *	.	0.218 **	.	0.089
	Dizziness	.	0.161 **	.	-0.140	.	0.002
	Stomach problems	.	-0.001	.	-0.054	.	0.020
	Incontinence	.	0.253 **	.	-0.107	.	-0.112
	Other symptoms	.	-0.021	.	0.367 **	.	0.012
	Low grip strength	.	-0.106 **	.	-0.055	.	-0.113 **
	Underweight	.	-0.051	.	-0.178	.	0.325 *
	Overweight	.	0.022	.	0.019	.	0.003
	Obese	.	-0.132 **	.	0.036	.	-0.058
	Age - 55	.	-0.009 **	.	-0.002	.	0.010 **
	(Age - 55) squared /100	.	0.052 **	.	0.002	.	-0.044 **
	Secondary education	.	0.086 **	.	0.090 *	.	0.012
	Post-secondary education	.	0.199 **	.	0.269 **	.	0.092 *
	Living with spouse or partner	.	0.064 *	.	0.108 **	.	0.098 **
	Log HH income - log(12763)	.	0.019	.	-0.000	.	0.004
	medit	.	0.158 **	.	-0.015	.	-0.100 *
	female	.	0.145 **	.	0.080	.	0.052
	medit*female	.	-0.230 **	.	-0.108	.	-0.046
	Constant	1.203 **	-0.070	0.827 **	-0.410 **	1.031 **	-0.073
Vign. 1	Constant	.	0.477 **	.	0.762 **	.	0.503 **
Vign. 2	Constant	.	1.612 **	.	1.470 **	.	1.246 **
Vign. 3	Constant	.	2.120 **	.	1.657 **	.	1.884 **
ln σ	Constant	.	-0.146 **	.	-0.192 **	.	-0.151 **
Obs.			3,607		3,607		3,607
LR test			439.7		413.0		394.4
p(LR test)			0.0		0.0		0.0

Notes: weighted results

F Administrative data from the Pharmaceutical Service Department of Treviso

Our data comes from three administrative registries maintained by the Pharmaceutical Service Department of ULSS 9, the local health authority (LHA) covering the southern part of the Italian province of Treviso. Figure F1 shows the distribution of Italian population and population of the Italian province of Treviso in 2000. No major demographic differences between Italy and the province of Treviso are observed. The first maintained registry is the drug prescription database, which contains records of patient prescriptions. The second is the hospitalization registry, which contains records of each single hospitalization episode. Through the anonymized personal identifiers, we were able to link patient prescription and hospitalization information to the last registry, the death and transfer registry. The resulting dataset allows us to follow individual patients through all their accesses to public health care services until they either die or leave the local health authority. Complete data are available from 1993 to 2002 for drug prescriptions, and from 1997 to 2002 for hospitalizations

Relative to survey data, these administrative data have both advantages and disadvantages. An important advantage is that they do not present problems which are typical of survey data, namely unit and item non-response, measurement errors and bias effects due to interaction with interviewers. Another advantage is that they contain extremely rich information on health care services received by patients. The main disadvantage is that they contain little information on patients' socio-economic characteristics. The whole database includes 46,140 male and 61,428 female patients with at least one prescription filled at any time between 1993 and 2002. Figure F2 shows the distribution of patients included in the database by gender and year of birth.

The prescription database

The drug prescription database contains records of patient prescriptions, including the date the prescription is made, the Anatomical Terapeutical and Clinical Classification (ATC) code of the substance prescribed, the number of packages prescribed, the unit price of the package. The unit price of the package allows us to recover information about the number of pills contained in each package. The registry also includes gender and date of birth of the patient receiving the medications, a unique anonymized patient identifier, a unique anonymized identifier of the practitioner who prescribed the medication, and gender and date of birth of the practitioner.

In the ATC system drugs are classified in groups at five different levels. The drugs are divided into fourteen main groups (1st level), with one pharmacological/therapeutic subgroup (2nd level). The 3rd and 4th levels are chemical/pharmacological/therapeutic subgroups and the 5th level is the chemical substance. The 2nd, 3rd and 4th levels are often used to identify pharmacological subgroups when that is considered more appropriate than therapeutic or chemical subgroups. To

illustrates the structure of the code, Table F1 shows the complete classification of Amlodipine, which is given the ATC code C08CA01.

Our drug prescription database contains prescriptions of active ingredients in the "therapeutic main groups" of the ATC Classification System mainly employed against hypertension, namely Antihypertensives (AH), Diuretics (D), Beta blocking agents (BBA), Calcium channel blockers (CCB), Agents acting on the renin-angiotensin system (ARA). All the substances used here are mainly employed against hypertension. Hypertension is a chronic asymptomatic pathology affecting a large share of the adult population and tends to have long-term health implications.

Table F2 reports the complete list of the chemical substances (or active ingredients) contained in the database. The table contains indication of both codes and description of therapeutic main groups and chemical substances. Table F3 shows the total number of prescriptions by therapeutic main group and gender. Agents acting on the renin-angiotensin system are by far the most commonly prescribed drugs, followed by Calcium channel blockers and Diuretics.

The hospitalization database

The hospitalization registry contains records of each single hospitalization, including date of entry and dismissal, primary Diagnosis Related Groups (DRG), and cost of hospitalization. Diagnosis-related group (DRG) is a system to classify hospital cases into one of approximately 500 groups, also referred to as DRGs, expected to have similar hospital resource use, developed for Medicare[33] as part of the prospective payment system. DRGs are assigned by a program based on the International Statistical Classification of Diseases and Related Health Problems (ICD) diagnoses, procedures, age, sex, and the presence of complications or comorbidities. DRGs have been used since 1983 to determine how much Medicare pays the hospital, since patients within each category are similar clinically and are expected to use the same level of hospital resources. DRGs may be further grouped into 25 mutually exclusive diagnosis areas, called Major Diagnostic Categories (MDCs). In our analysis we distinguish between hospitalization with cardiovascular DRGs and hospitalization with any other DRG. Table F4 reports the complete list of cardiovascular DRGs, including codes, description and type.

The death and transfer registry

Through the anonymized personal identifiers, we were able to link patient prescription and hospitalization information to the death and transfer registry. This registry contains records about whether individuals either die or leave the LHA. The date of those events is also recorded. During the period 1993-2002 the fraction of individuals who left the LHA is negligible. Specifically, less than 0.21% of the patients included in our database left the LHA. Most of them are aged between

[33] Medicare is a social insurance program administered by the United States government, providing health insurance coverage to people who are either age 65 and over, or who meet other special criteria.

30 and 40. The transfer rate of individuals relevant for our analysis (i.e. older than 40) is almost zero.

Table F1: The complete classification of Amlodipine (ATC code C08CA01).

	(1st level, anatomical main group)
C	Cardiovascular system
	(2nd level, therapeutic subgroup)
C08	Calcium channel blockers
	(3rd level, pharmacological subgroup)
C08C	Selective calcium channel blockers with mainly vascular effects
	(4th level, chemical subgroup)
C08CA	Dihydropyridine derivatives
	(5th level, chemical substance)
C08CA01	Amlodipine

Table F2: List of active ingredients.

Therapeutic main group		Chemical substance	
C02 ANTIHYPERTENSIVES			
C02A	ANTIADRENERGIC AGENTS, CENTRALLY ACTING	C02AB01	Methyldopa (levorotatory)
		C02AC01	Clonidine
C02C	ANTIADRENERGIC AGENTS, PERIPHERALLY ACTING	C02CA04	Doxazosin
		C02CA06	Urapidil
C03 DIURETICS			
C03A	LOW-CEILING DIURETICS, THIAZIDES	C03AA03	Hydrochlorothiazide
C03B	LOW-CEILING DIURETICS, EXCL. THIAZIDES	C03BA04	Chlortalidone
		C03BA08	Metolazone
		C03BA10	Xipamide
		C03BA11	Indapamide
		C03BA13	Fenquizone
C03C	HIGH-CEILING DIURETICS	C03CA01	Furosemide
		C03CA02	Bumetanide
		C03CA03	Piretanide
		C03CA04	Torasemide
		C03CC01	Etacrynic acid
		C03CX01	Etozolin
C03D	POTASSIUM-SPARING AGENTS	C03DA01	Spironolactone
		C03DA02	Potassium canrenoate
		C03DA03	Canrenone
C03E	DIURETICS AND POTASSIUM-SPARING AGENTS IN COMBINATION	C03EA01	Hydrochlorothiazide
		C03EA14	Butizide
		C03EB01	Furosemide
C07 BETA BLOCKING AGENTS			
C07A	BETA BLOCKING AGENTS	C07AA02	Oxprenolol
		C07AA03	Pindolol
		C07AA05	Propranolol
		C07AA06	Timolol
		C07AA07	Sotalol
		C07AA12	Nadolol
		C07AA14	Mepindolol
		C07AB02	Metoprolol
		C07AB03	Atenolol
		C07AB04	Acebutolol
		C07AB05	Betaxolol
		C07AB07	Bisoprolol
		C07AB08	Celiprolol
		C07AB12	Nebivolol
		C07AG01	Labetalol
		C07AG02	Carvedilol
C07B	BETA BLOCKING AGENTS AND THIAZIDES	C07BB02	Metoprolol and thiazides
C07C	BETA BLOCKING AGENTS AND OTHER DIURETICS	C07CA02	Oxprenolol and other diuretics
		C07CB02	Metoprolol and other diuretics
		C07CB03	Atenolol and other diuretics

Therapeutic main group		Chemical substance	
C08 CALCIUM CHANNEL BLOCKERS			
C08C	SELECTIVE CALCIUM CHANNEL BLOCKERS WITH MAINLY VASCULAR EFFECTS	C08CA01	Amlodipine
		C08CA02	Felodipine
		C08CA03	Isradipine
		C08CA04	Nicardipine
		C08CA05	Nifedipine
		C08CA06	Nimodipine
		C08CA07	Nisoldipine
		C08CA08	Nitrendipine
		C08CA09	Lacidipine
		C08CA11	Manidipine
		C08CA13	Lercanidipine
C08D	SELECTIVE CALCIUM CHANNEL BLOCKERS WITH DIRECT CARDIAC EFFECTS	C08DA01	Verapamil
		C08DA02	Gallopamil
		C08DB01	Diltiazem
C09 AGENTS ACTING ON THE RENIN-ANGIOTENSIN SYSTEM			
C09A	ACE INHIBITORS, PLAIN	C09AA01	Captopril
		C09AA02	Enalapril
		C09AA03	Lisinopril
		C09AA04	Perindopril
		C09AA05	Ramipril
		C09AA06	Quinapril
		C09AA07	Benazepril
		C09AA08	Cilazapril
		C09AA09	Fosinopril
		C09AA10	Trandolapril
		C09AA11	Spirapril
		C09AA12	Delapril
		C09AA13	Moexipril
		C09AA15	Zofenopril
C09B	ACE INHIBITORS, COMBINATIONS	C09BA01	Captopril and diuretics
		C09BA02	Enalapril and diuretics
		C09BA03	Lisinopril and diuretics
		C09BA04	Perindopril and diuretics
		C09BA05	Ramipril and diuretics
		C09BA06	Quinapril and diuretics
		C09BA07	Benazepril and diuretics
		C09BA08	Cilazapril and diuretics
		C09BA09	Fosinopril and diuretics
		C09BA12	Delapril and diuretics
		C09BA13	Moexipril and diuretics
C09C	ANGIOTENSIN II ANTAGONISTS, PLAIN	C09CA01	Losartan
		C09CA02	Eprosartan
		C09CA03	Valsartan
		C09CA04	Irbesartan
		C09CA06	Candesartan
		C09CA07	Telmisartan
C09D	ANGIOTENSIN II ANTAGONISTS, COMBINATIONS	C09DA01	Losartan and diuretics
		C09DA03	Valsartan and diuretics
		C09DA04	Irbesartan and diuretics
		C09DA06	Candesartan and diuretics

Table F3: Total number of prescriptions by therapeutic main group and gender.

Therapeutic main group	Men	Women	Total
Antihypertensives	59,279	70,522	129,801
Diuretics	248,686	382,289	630,975
Beta blocking agents	161,339	219,502	380,841
Calcium channel blockers	370,316	404,816	775,132
Agents acting on the renin-angiotensin system	445,954	563,019	1,008,973
Total	1,285,574	1,640,148	2,925,722

Table F4: List of cardiovascular DRGs.

DRG	Description	DRG type
103	Heart Transplant	Surgical
104	Cardiac Valve Proc w Cardiac Cath	Surgical
105	Cardiac Valve Proc w/o Cardiac Cath	Surgical
106	Coronary Bypass w Cardiac Cath	Surgical
107	Coronary Bypass w/o Cardiac Cath	Surgical
108	Other Cardiothoracic Procedures	Surgical
110	Major Cardiovascular Procedures w CC	Surgical
111	Major Cardiovascular Procedures w/o CC	Surgical
112	Percutaneous Cardiovascular Procedures	Surgical
113	Amput for Circ Disor Exc Uppr Limb & Toe	Surgical
114	Uppr Limb & Toe Amput for Circ Disor	Surgical
115	Perm Pacemkr Impl w Ami, Heart Fail, Shck	Surgical
116	Perm Pacemkr Impl w/o Ami, Heart Fail, Shck	Surgical
117	Card Pacemkr Revision Exc Device Replace	Surgical
118	Cardiac Pacemaker Device Replacement	Surgical
119	Vein Ligation & Stripping	Surgical
120	Other Circulatory System O.R. Procedures	Surgical
121	Circ Disor w Ami & Cv Comp Disch Alive	Medical
122	Circ Disor w Ami w/o Cv Comp Disch Alive	Medical
123	Circ Disor w Ami, Expired	Medical
124	Circ Dis Ex Ami w Card Cath & Complx Dx	Medical
125	Circ Dis Ex Ami w Card Cath Wo Complx Dx	Medical
126	Acute & Subacute Endocarditis	Medical
127	Heart Failure & Shock	Medical
128	Deep Vein Thrombophlebitis	Medical
129	Cardiac Arrest, Unexplained	Medical
130	Peripheral Vascular Disorders w CC	Medical
131	Peripheral Vascular Disorders w/o CC	Medical
132	Atherosclerosis w CC	Medical
133	Atherosclerosis w/o CC	Medical
134	Hypertension	Medical
135	Card Congen & Valv Disor Age ¿17 w CC	Medical
136	Card Congen & Valv Disor Age ¿17 w/o CC	Medical
137	Card Congen & Valv Disor Age 0-17	Medical
138	Card Arrhythmia & Conductn Disor w CC	Medical
139	Card Arrhythmia & Conductn Disor w/o CC	Medical
140	Angina Pectoris	Medical
141	Syncope & Collapse w CC	Medical
142	Syncope & Collapse w/o CC	Medical
143	Chest Pain	Medical
144	Oth Circulatory System Diagnoses w CC	Medical
145	Oth Circulatory System Diagnoses w/o CC	Medical
478	Other Vascular Procedures w CC	Surgical
479	Other Vascular Procedures w/o CC	Surgical

Figure F1: Distribution of Italian population and population of the Italian province of Treviso in 2000.

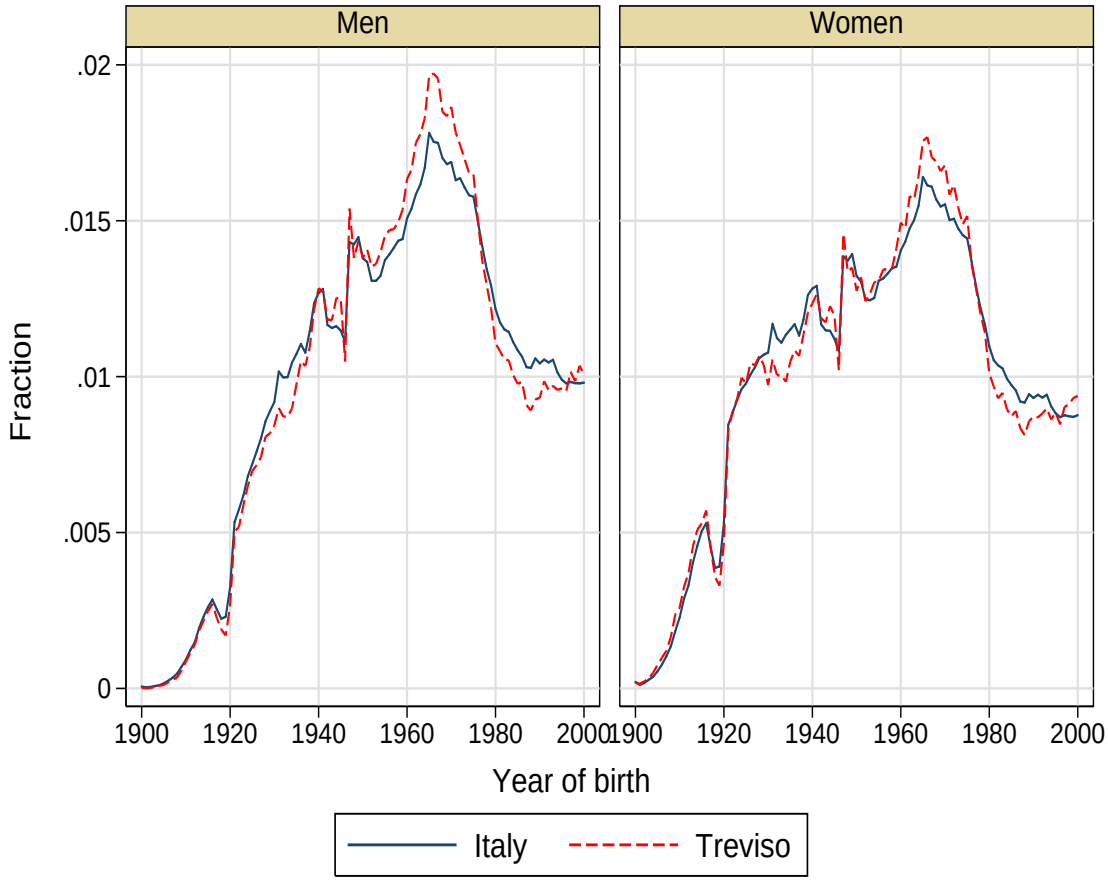


Figure F2: Distribution of patients by gender and year of birth.