HEDG

HEALTH, ECONOMETRICS AND DATA GROUP

WP 19/13

Explaining the causal effect of adherence to medication on cholesterol through the marginal patient

Domenico Depalo

July 2019

http://www.york.ac.uk/economics/postgrad/herc/hedg/wps/

Explaining the causal effect of adherence to medication on cholesterol through the marginal patient

Domenico Depalo *

Abstract

This paper investigates the relation between adherence to prescribed medication and reduction of cholesterol in Italy, by taking into account the possible sorting of patients into treatment and the heterogeneity of the effect. As predicted by a theoretical model, I find that patients who benefit most from medication are more likely to adhere to prescribed regime than those who benefit least. These results are used to study the effects of three hypothetical policies that aim at increasing the share of patients adherent to prescribed medication: one is directed towards patients, one towards physicians, and one is a policy mix. For each policy, the observable characteristics of patients switched into treatment are described. Although the most effective policy is directed towards patients, the policies differ substantially with respect to the population affected. Therefore, a less effective policy that targets best the desired population may be preferred to the most effective policy. Back of the envelope calculations suggest that even the most expensive policy would be cost-effective.

JEL classification: I12, I18, C21

Keywords: Cholesterol, Marginal Treatment Effect, Policy Evaluation

*I would like to thank Joshua Angrist, Anirban Basu, Stephen Bonhomme, Ylenia Brilli, David Card, Federico Cingano, Lucia Corno, Manisha Desai, Alessandra Fenizia, David Glynn, Jeremy D. Goldhaber-Fiebert, Alberto Holly, Andrew Jones, Pat Kline, Giuseppe Ilardi, Claudio Lucifora, Magne Mogstad, Cheti Nicoletti, Edward Norton, Alessandro Palma, Silvana Robone, Chris Walters, Jeffrey Wooldridge .... I would like to thank seminar participants at University of California (Berkeley), University of York, Stanford University, University of Washington (Seattle). Special thanks to Vincenzo Atella, who gave me the permission to use the data. I am grateful to Pierluigi Russo who kindly explained how to derive relevant information from AIFA. Replication files and additional results will be available at the webpage: http://sites.google.com/site/domdepalo/ The views expressed in this paper are those of the author and do not imply any responsibility of the Bank of Italy. Corresponding address: Domenico Depalo, Banca d'Italia, Economics and Statistics Department, Via Nazionale, 91 - 00184 Rome (Italy), Tel.: +39-06-4792 5989, e-mail: `domenico.depalo@bancaditalia.it`

1 Introduction

This paper investigates the relation between hypercholesterolemia, which is the high level of Low Density Lipoprotein (LDL) cholesterol in the blood stream, and adherence to prescribed medication, using register data from Italy. Hypercholesterolemia is a condition of both clinical and policy relevance. From a clinical perspective, it is an asymptomatic condition that may have serious consequences, like atherosclerosis and the related cardio-vascular diseases (CVD). These diseases constitute the largest causes of death for the population 65+ in the European Union (Niederlaender, 2006; Buchow et al., 2012). From a policy perspective, the prevalence of hypercholesterolemia in the population is very high: during the period 2009-12 almost 70% of the Italian population between 35–75 years suffered from this condition (Progetto Cuore, 2018); in the USA the percentage was about 75% for the population 35+ (American Heart Association, 2015); in UK and Scotland the percentage was about 60% of individuals (British Heart Foundation, 2015).

Hypercolesterolemia is typically treated with statin medications, which in randomized clinical trials have been shown to be safe and effective to reduce cholesterol (Athyros et al., 2010), provided that patients follow the prescribed treatment. However, due to the asymptomatic nature of the condition, adherence to prescribed medication is a critical issue out of the lab-experiments. Therefore, three questions are important: 1) why some patients are adherent and some are not? 2) Is it possible to implement policies that increase the share of patients adherent to prescribed medication (defined as "adherence rate" in what follows)? 3) If so, what is the best policy? The contribution of this paper to the existing literature is to address these issues.

In order to explain why some patients are adherent whilst some are not, I propose a simple theoretical model based on the Roy (1951) model, such that the behaviour of patients maximizing their utility function depends on the comparison between benefits (i.e., reduction of cholesterol) and opportunity costs from medication (e.g., side effects). Providing an evaluation of the net benefit from medication is one contribution of the paper. Standard quantities that are usually employed to this aim, like the Average Treatment Effect (ATE) for the entire population, or only for the treated (ATT) or untreated (ATU) population, may not be fully informative if benefits or

costs are heterogeneous in the population. In these situations, focusing on the treatment effect for the marginal individual who is induced to follow the prescribed treatment regime may be more appropriate. For this reason I employ the Marginal Treatment Effect (MTE) that identifies the reduction of cholesterol from adherence to treatment regime at any specific margin of indifference (i.e., the probability of being adherent). The MTE was introduced by Bjorklund and Moffitt (1987) and further developed in a series of contributions by Heckman and Vytlacil (2001, 2005, 2007a,b). Examples in health economics include: Basu et al. (2007); Doyle (2007); Maestas et al. (2013); French and Song (2014); Basu (2014); Kowalski (2016, 2018); Schoenberg et al. (2018). This study relates to this literature, but it is one of the very few papers that focuses specifically on health outcomes (Basu, 2014; Kowalski, 2018).

If the results from the empirical analysis are coherent with the predictions of the theoretical model, a second important question is how to increase the adherence rate, so to improve patients' welfare and save money (Lamiraud and Geoffard, 2007; CBO, 2018). Given the cost of non-adherence and the projected increase in the health care expenditures in several countries (New England Healthcare Institute, 2009; European Commission, 2015; CBO, 2018), how to increase the adherence rate is a concrete problem for policy makers around the world.[1] Several policies may be implemented to this aim. When different policies are available, two further questions arise. The first question concerns the magnitude of the treatment effect for individuals who are currently untreated but that would be treated under a different regime. This is the definition of the Policy Relevant Treatment Effect (PRTE) introduced by Heckman and Vytlacil (2005). The evaluation of PRTE for each policy is a second contribution of this paper. Looking at the magnitude of PRTEs suggests the most effective among the various alternatives. The second question concerns the individuals affected by each policy. This is important because in some cases looking only at the PRTE might hide important information. A prominent example is when two different policies have

[1] According to estimates by the New England Healthcare Institute (2009) the annual cost of non adherence in the United States is approximately $290 billion (2.3% in terms of GDP). These considerations will be increasingly important, overall in aging population societies with tight balance constraints and where the health care system is publicly funded (as in Italy). According to projections released by the European Commission (2015, p.128), in Italy the increase in public expenditure on health care over the period 2013-2060 with respect to GDP will be about 15%. Similar projection hold for the entire European Union. Given the forecast produced by the CBO (2018) of health care bill for the next decades in the US and the large incidence of hypercholesterolemia, similar budgetary considerations will be increasingly important even in countries where the health care system is privately funded (e.g., the USA).

the same returns but affect the behaviour of different individuals. Therefore, as a minor -albeit important- technical contribution to the existing literature, I describe the observable characteristics of the population that is actually affected by each policy. This exercise might help tailor the most appropriate policies to better target the desired population. Despite its relevance, I am not aware of previous papers addressing the question in a systematic way (the only paper that goes in a similar direction is Carneiro et al., 2003). Intuitively, this extension may be seen as the natural counterpart to what is nowadays standard for the Local Average Treatment Effect (LATE; Imbens and Angrist, 1994) when the characteristics of compliers are described (or characterization of compliers; Angrist, 2004).

In this analysis, the results obtained using MTE greatly enrich those based on standard techniques. Using Ordinary Least Square (OLS), I estimate that adherence to prescribed treatment regime reduces cholesterol by 1.3 %: this estimate is biased if the sample of patients who follow the therapy is not random, for example because they self-select on the basis of their expected gain from treatment. The standard cure for this issue is an Instrumental Variable (IV) technique that estimates a reduction of cholesterol caused by drugs by 3.5 %. Even this consistent estimate addresses the original question only partially: if the effect of the drug is heterogeneous in the population it is unlikely that a single number summarizes all the relevant features of the distribution of the effects of the drugs. The theoretical model proposed in the paper predicts that patients who benefit most from drugs 'need less arguments' to be convinced to be adherent to the treatment, and therefore it is more likely that they are treated; in contrast, those who benefit the least are harder to convince, and therefore it is less likely that they are observed into treatment. This is clearly reflected in the MTE estimates: the reduction of cholesterol for the former group of patients is the largest (about 5%), while for the latter group is remarkably smaller (negligible for some).

These results are further exploited to estimate the effect of three hypothetical policies aiming at improving the share of patients adherent to medication. The first policy affects the adherence attitude of the patients through TV advertisement; the second policy affects the communication skills of the physicians through on-line training courses; the third affects the overall probability of following the prescribed treatment (for unspecified reasons). The empirical results show that the

policy that targets directly the patients is the most effective (the reduction of cholesterol for those induced into treatment is about 3%). However, the other two policies may be preferred depending on which population the decision maker wishes to induce into treatment. For example, if the goal of the policy maker is to treat those relatively healthier in an attempt to reduce future possible occurrences of CVDs, then informing the patients more specifically through physicians might be preferable; if the goal is to treat a larger share of elderly patients that are in less healthy conditions, in an attempt to reduce current health care costs, then a mix of policies might be preferable. Therefore, a less effective policy that targets best the desired population may be preferred to the most effective policy.

A final important question for the Government is whether the implementation of the three hypothetical policies is cost-effective, i.e. economically convenient. To answer these questions I compare (monetary) costs and benefits attached to each policy. Back of the envelope calculations suggest that the most expensive policy would cost less than $30 mil.$ €; although with these data I can not estimate the total monetary benefits from the policies, avoiding 1,000 surgeries (that cost $25,000$€ each in the Italian health system) would make each policy cost effective. Because the medical literature suggests that this target is feasible (Monaldi et al., 2015), I conclude that even the most expensive policy is cost-effective.

The paper is organized as follows. In Section 2 I present a prototypical theoretical model that will be used to interpret the results. The identification of the parameters of interest is discussed in Section 3. Section 4 describes the data that are used in the empirical application (Section 5) and to evaluate hypothetical policies (Section 6). Some robustness checks are presented in Section 7. Section 8 offers some conclusions.

2 Theoretical framework

In this section I propose a model that explains why some patients are adherent to prescribed medication and why some are not. In order to predict the patients' behaviour I exploit a standard model of choice and associated outcomes (Roy, 1951). This model provides the background to

interpret the results of the empirical section (Heckman and Vytlacil, 2007a,b; Carneiro et al., 2003). Although the Roy model is standard in other fields of the economic literature, its application in health economics is rare (Diiro et al., 2016; Gibson et al., 2016; Deb et al., 2006). Furthermore, to the best of my knowledge this is the first paper in health economics that uses the Roy model to derive what quantities should be identified in the empirical analysis.

A patient i may be adherent to the therapy (or treated; subscript 1) or not (or untreated; subscript 0), but not simultaneously adherent *and* non adherent (i.e., the treatment states are mutually exclusive). The choice of the treatment status is made upon comparison between benefits and costs from adherence. Benefits are in terms of reduction of cholesterol under adherence vs non-adherence. Let $y_{i,d}$ be the outcome under treatment $d \in D$ for each patient i, the benefit amounts to $[y_{i,0} - y_{i,1}]$. Opportunity costs (c_i) are in terms of health outcomes, and they may be represented as a multidimensional vector including for example side effects from medication. A rational patient is adherent to the prescribed treatment if the benefit from the reduction of cholesterol thanks to adherence is no smaller than the corresponding cost, or more formally:

$$[y_{i,0} - y_{i,1}] - c_i \geq 0 \qquad \forall i. \tag{1}$$

Define g_i the rate of reduction of cholesterol for each patient, then $y_{i,1} = y_{i,0}\,(1+g_i)^{-1}$, such that the higher g_i the lower the cholesterol level under adherence as compared to non adherence. It follows that the adherence to prescribed medication will be endogenously preferred to non-adherence until

$$\begin{aligned} y_{i,0} - y_{i,1} &\geq c_i & (2)\\ y_{i,0} - y_{i,0}\,(1+g_i)^{-1} &\geq c_i & (3)\\ g_i &\geq \frac{c_i}{y_{i,0} - c_i}. & (4) \end{aligned}$$

As simple as it is, this model provides several useful predictions. First of all, adherence to prescribed medication will be endogenously preferred to non adherence until the rate of reduction of cholesterol under treatment ($g_i \equiv LHS$) is equal to or larger than the associated cost ($\frac{c_i}{y_{i,0}-c_i} \equiv RHS$). This

explains why some patients are adherent to prescribed regime and some are not. Second, the higher the cost of adherence (i.e., high c_i), the larger the minimum reduction of cholesterol to induce the individual into adherence ($\frac{\partial RHS}{\partial c_i} > 0$) and the lower the likelihood that the patient is treated. Third, the worse the health conditions without adherence (i.e., high y_0), the smaller the minimum reduction of cholesterol to self-select into treatment (as $\frac{\partial RHS}{\partial y_{i,0}} < 0$) and the higher the likelihood that the patient is treated. Figure 1 provides a graphical representation of these predictions. Finally, and most important for the following analysis, both costs and health conditions under non-adherence may vary at patient level (as emphasized by subscript i in condition 2), in which case the incentives to switch into treatment would be individual specific: the higher the net benefit the higher the incentives to sort into treatment, and thus the higher the likelihood that the patient is actually treated. Models where responses to interventions are heterogeneous and patients opt for treatment with at least partial knowledge of their idiosyncratic response are termed models with essential heterogeneity (Heckman et al., 2006). In this paper I identify the effect of adherence under essential heterogeneity.

3 Empirical Methods

In this section I review methods that identify the treatment effect under essential heterogeneity. Throughout the section I will focus only on the main issues for the present analysis; the reader interested in further technical details should refer to the original papers.

Define the potential outcomes under $D \in \{0, 1\}$ for each patient i (subscript omitted from now on for notational convenience) as:

$$\begin{aligned} \ln y_0 &= \alpha + U_0 \qquad \text{if D=0} \\ \ln y_1 &= \alpha - \overline{\beta} + U_1 \text{ if D=1} \end{aligned} \tag{5}$$

where α is an intercept, $-\overline{\beta}$ is a parameter equal to $(\ln y_1 - \ln y_0)$ and U_0 and U_1 are mean zero random errors under treatment states equal to 0 and 1. Since the two states of the world are mutually exclusive, the observed outcome is $\ln y = D \ln y_1 + (1 - D) \ln y_0 = D(\alpha - \overline{\beta} + U_1) + (1 -$

$D)\,(\alpha + U_0) = \alpha + \{-\overline{\beta} + (U_1 - U_0)\}\,D + U_0$. Upon the substitutions of $\eta = (U_1 - U_0)$, $\beta = -\overline{\beta} + \eta$ and $\epsilon = U_0$, the model for the observed outcome can be written as $\ln y = \alpha + \beta\, D + \epsilon$. This specification fits perfectly into the economic framework of Section 2, namely with the definition of essential heterogeneity. First, β is a *distribution* of parameters that therefore is able to capture the heterogeneity in the treatment effect.[2] Second, the model for the observed outcome shows that D may be correlated with ϵ, for example because of unobservable individual genetic endowment, and with $(U_1 - U_0)$, for example because adherent patients anticipate their (expected) gain from treatment. Third, the Roy model predicts that $D = 1$ if $\ln(y_1) \leq \ln(y_0)$, or equivalently $\beta \leq 0$, i.e. patients are adherent to prescribed treatment regime because they expect a positive net benefit from medication. More than this, under essential heterogeneity the model predicts that the higher the private net return to adherence (i.e., the larger the absolute value of β) the more likely that the patient will be adherent to the prescribed treatment regime. It follows that the reduction of cholesterol for patients adherent to treatment should be higher than for patients chosen at random from the population, or $ATT < ATE$ (viceversa, for non-treated individuals $ATE < ATU$). The relation between ATT, ATE, and ATU is a testable necessary condition for concordance between the model of Section 2 and the data under essential heterogeneity. In contrast, if the effect of adherence on cholesterol is homogeneous in the population, $ATE = ATT = ATU$. This clarification is important because if we knew all these quantities, they may be used as a null hypothesis for a test of non-essential heterogeneity (Section 5). An 'ideal estimator' should be able to jointly test all the relations between the various treatment effect parameters.

The fact that the sample of adherent patients is not a random sample of the population (Little, 1995), implies that the OLS is inconsistent (Heckman, 1979, 2010). To formalize the mechanism of selection into treatment, I define

$$D^* = \gamma\, Z - V \tag{6}$$

$$D = 1 \text{ if } \mathrm{D}^* \geq 0, \tag{7}$$

[2] Furthermore, from eq. 5 the expected gain from treatment is $\ln y_1 - \ln y_0 = -\overline{\beta}$. By definition of $y_1 = y_0(1+g)^{-1}$, it follows that $\ln y_1 = \ln y_0 - \ln(1+g)$. Therefore, $-\overline{\beta} = \ln y_1 - \ln y_0 = \ln y_0 - \ln(1+g) - \ln y_0 = -\ln(1+g)$.

where Z includes an indicator, called instrument, that is relevant to explain the adherence mechanism but not the outcome. For concreteness, in this paper Z is the 'pure' communication skill of the doctor, which aims at capturing the relation between patients and physisicians (Fichera et al., 2018), and has no impact on the patient's cholesterol as better discussed in Section 5. For later reference, using $F_V(\cdot)$ to denote the cumulative distribution function (CDF) of V, eq. 6-7 imply that $D = 1$ if $F_V(\gamma Z) \geq F_V(V)$, where $F_V(\gamma Z) \equiv P(Z)$ is a propensity score and $F_V(V) \equiv U_D \sim [0, 1]$ are quantiles of the distribution of the unobserved component.[3]

A standard approach to cure for the non-random selection of the sample is based on Instrumental Variable (IV). If the effect of the treatment is heterogeneous, IV identifies a *local* average treatment effect (LATE; Imbens and Angrist, 1994):[4]

$$\frac{E[y|Z = z'] - E[y|Z = z]}{E[D|Z = z'] - E[D|Z = z]} = E[y_{z'} - y_z | P(z) < U_D < P(z')]. \tag{8}$$

This is the treatment effect for *compliers* (Angrist et al., 1996), i.e. patients that are induced to be adherent to prescribed treatment regime by a change in the value of the instrument. It is well known that different instruments target two different subpopulations of compliers (i.e., estimate two different, consistent, parameters; Card, 1999); similarly, when the same instrument takes more than 2 values, e.g. $Z = \{z_1, z_2, z_3, z_4, \ldots\}$, LATE using (z_2, z_1) is different from LATE using (z_4, z_3) because different "marginal individuals" are induced into treatment (Heckman, 2010; for this reason, sometime I prefer the notation LATE(z', z) instead of the more compact LATE).

At this point, if Z is continuous one may wonder what is the treatment effect for infinitesimal changes of Z, i.e. as $z' \to z$. This effect is identified by the Marginal Treatment Effect (MTE), introduced by Bjorklund and Moffitt (1987) and further developed in various papers by Heckman

[3] For a proof, see Vytlacil (2002); Imbens and Newey (2009). The same result would follow by invoking results in Pollard (2001, ch.2.9).

[4] The following assumptions are needed: *1)* the potential outcomes for each patient are unrelated to the treatment status of other patient (Stable Unit Treatment Value Assumption); *2)* the instrument is randomly assigned; *3)* exclusion restriction (i.e. $y_d \equiv Y(0, d) = Y(1, d)$); *4)* nonzero average causal effect of Z on D (i.e. $E[D_1 - D_0] \neq 0$); *5)* monotonicity (i.e. $\forall z' > z$, $D_{z'} \geq D_z$ for all patients, such that an increase in the level of the instrument does not decrease the level of the treatment, or vice-versa). See Angrist et al. (1996) for an in-depth discussion.

and Vytlacil (2000, 2001, 2005, 2007a,b).[5] Formally, $p \equiv P(Z)$ (only for notational convenience),

$$\begin{aligned} \lim_{z' \to z} \text{LATE}(z, z') &= E[y_{z'} - y_z | p < U_D < p'] \\ &= \beta + \lim_{z' \to z} E[U_1 - U_0 | p < U_D < p'] \\ MTE(U_D = p) &= E[\Delta | U_D = p], \end{aligned} \tag{9}$$

i.e. the average reduction of cholesterol due to adherence for patients just indifferent between treatment and non treatment, at level of unobservables $U_D = p$.[6] Patients with low realizations of U_D are patients whose unobservable characteristics make them more likely to be adherent (because by definition $F_V(V) \equiv U_D$ and thus according to equation 6 these patients would prefer treatment even if γZ is low); at the opposite side, for realizations of U_D approaching 1, patients need high values of γZ to be induced to participate. Together this reasoning and the model of Section 2 provide two additional testable necessary conditions for coherency between the model and the data, namely that the MTE should be 1) negative, if adherence is effective to reduce LDL cholesterol; 2) a *non-decreasing* function of U_D if patients who expect a large reduction of cholesterol are more likely to be adherent.[7]

A particularly attractive feature of the MTE is the unifying approach to the treatment effects, as one can summarize the distribution of β that best addresses the economic question of interest. Heckman and Vytlacil (2005, Tab.1A) provide a set of weights that allow to go from MTE to other relevant quantities, like ATE, ATT, ATU and LATE. Since it is possible to identify the MTE over the entire support of U_D and it is possible to recover all the relevant parameters to test the necessary conditions for coherency between the data and the model of Section 2, MTE has all the properties to be an 'ideal estimator' in this analysis.

[5] Other approaches are available. Among them Imbens and Newey (2009); Chernozhukov et al. (2007). See Heckman and Vytlacil (2007b) for a comparison.

[6] See Carneiro et al. (2003, 2017); Cornelissen et al. (2016) for alternative admissible interpretations.

[7] The intuition is that patients who expect to enjoy a *large reduction* of cholesterol need less arguments to be convinced to be adherent, therefore even a physician with poor communication skill can succeed (a small value of Z and correspondingly a small value of U_D; hence $MTE(U_D \approx 0) << 0$). On the other hand, patients who enjoy a *small reduction* of cholesterol need more convincing arguments to be treated, i.e. the communication skills of the doctor must be high (i.e., high value of Z and thus of U_D; hence, $MTE(U_D \approx 1) \to 0$); see Carneiro et al. (2003, 2010); Basu et al. (2007); Cornelissen et al. (2016); Carneiro et al. (2017) for similar arguments.

Applications of MTE are increasing over time. Cornelissen et al. (2016) provide several references, mostly in labour economics, where a much explored research question involves the return to education. The MTE was introduced in the health economics literature by Basu et al. (2007) who investigate the effects on costs of breast conserving surgery as opposed to mastectomy after 5-years. Other examples in the field include Doyle (2007); Maestas et al. (2013); French and Song (2014); Brinch et al. (2012); Kowalski (2016); Schoenberg et al. (2018), but none of them is specifically interested in clinical outcomes. Among the very few papers that employ MTE for clinical outcomes are Kowalski (2018) who identifies the relationship between mortality and mammography receipt, finding that women are more likely to receive mammography are healthier, but are also more likely to be harmed by the receipt; Basu (2014) who compares survival rates for elderly men treated with surgery vs active surveillance in the case of prostate cancer, finding little difference between the two treatments.

It is important to mention that other estimators are available to identify an heterogeneous treatment effect with multivalued instruments. The closest alternative to MTE when a continuous instrument is available is the person centered treatment effect (PeT), introduced by Heckman and Vytlacil (1999) and made operational by Basu (2014). These effects are conditioned on the person's observed characteristics and averaged over the potential conditional distribution of observed characteristics that lead them to their observed treatment choice. In this application I prefer to work with MTE rather than PeT because it has a more direct connection to the population level policies, that are introduced next (nonetheless, I checked that the two approaches would lead to qualitatively identical conclusions; results are available on my website). More similar to IV-LATE setup is the approach in Card et al. (2018), who investigate the causal health effects of cesarean delivery for low-risk first births: their approach is well suited for discrete instruments, whereas MTE is point identified with a continuous instrument (however, see Brinch et al., 2012; Kowalski, 2016).

3.1 Policy Relevant Treatment Effect

If patients were adherent to prescribed treatment regime their health condition would improve (Monaldi et al., 2015), and the Government would save money, e.g. thanks to foregone hospitalization (New England Healthcare Institute, 2009). Hence, an important question for the policy maker is whether it is possible to change the state of the world from the current D to a new D^* that increases the treatment probability from p to $p*$. In this case, how large would the gain be?

To answer this question Heckman and Vytlacil (2005) introduce the Policy Relevant Treatment Effect (PRTE), which is the average effect of going from a baseline policy to an alternative policy per net person shifted, or $\frac{E[y|D^*]-E[y|D]}{E[D^*]-E[D]}$. Necessary for the *interpretation* of this quantity is the assumptions that MTE will be policy invariant, so that an external manipulation of the policy does not affect anything in the model apart from the selection into treatment. Necessary for the *estimation* of this quantity is that $E[D^*] - E[D] \neq 0$, so that the policy has the power to shift some individuals previously untreated into treatment. This definition allows for a large spectrum of potential policies. In order to increase the share of adherent patients, one can implement policies that affect the behaviour of the patients (externally influencing their attitude to adherence; this is a 'direct policy'), or policies that improve the communication skill of the physicians (that then affects the behaviour of the patients and ultimately their adherence; this is a 'mediated policy'). In both cases, specific components of the probability of treatment would be manipulated. With the direct policy, the *baseline* probability of being adherent to the prescribed policy regime would go from $p = [1, Z][\gamma_1, \gamma_Z]'$ to $p* = [1, Z][\gamma_1 + \alpha, \gamma_Z]'$; with the mediated policy, the instrument that affects the probability of being adherent would go from $p = [Z_j, Z_{\neq j}][\gamma_j, \gamma_{\neq j}]'$ to $p* = [Z_j, Z_{\neq j}][\gamma_j + \alpha, \gamma_{\neq j}]'$. Carneiro et al. (2010, 2011) consider also the case of policies that increase the probability of being adherent by an amount α (from p to $p* = p + \alpha$) or that shifts each person's probability of being adherent by a proportion $(1 + \alpha)$ (from p to $p* = p\,(1 + \alpha)$). The mechanism behind the last two policies is left unspecified, so I refer to them as a policy mix.

4 Data and descriptive statistics

The data for the empirical analysis are taken from Health Search Database (HSD), a longitudinal observational database collected by the Italian Society of General Medicine (SIMG), an association of the Italian College of General Practitioners (GP). The availability of real world register data may be preferable to questionnaire in this paper, because the former reveal preferences in real situations rather than in hypothetical scenarios as the latter does (Lamiraud and Geoffard, 2007). Important features of these register data are common to administrative sources, namely the information provided is affected neither by measurement errors nor by systematic unit-/item- non-response. It is not surprising then that the observable characteristics of the GPs and patients observed in the HSD overlap to those from the official figures collected by the Italian Ministry of Health and National statistical Institute with respect to the entire National Health System (Bianchini et al., 2016). Fabiani et al. (2004) show that these data are suitable to carry out pharmacoepidemiological studies generalised to the whole Italian population. The information recorded by GPs is rich and detailed with respect to health indicators; in contrast, information on individual socio-economic conditions are limited to an indicator for exemption of payment due to low income. However, all the information required by the economic model of Section 2 is available in the data.

In this paper, for the year 2004-08 I consider patients born between 1925 and 1975, diagnosed of hypercholesterolemia and with at least a prescription of statin medications. The initial sample is then made of 704,857 observations, pertaining to 71,668 patients observed for 2.5 years on average (9.8 quarters).

As indicator of cholesterol I use the (log-)ratio between serum levels of LDL and total cholesterol, which is considered a better predictor of future heart attack risk than total cholesterol itself (Baigent et al., 2005; Barter et al., 2007).

The model of Section 2 predicts that the likelihood to follow the treatment regime is a function of health conditions under non-adherence and the associated cost. Therefore three aspects are of interest: 1) how I define the treatment regime, 2) what observed characteristics affect the cholesterol condition without adherence to medication (y_0 in the model), and 3) what the associated cost (c)

is.

A crucial aspect for policy makers is the extent to which patients follow the medical prescriptions, or adherence. Although in principle one may have access to this information, with real world data almost always only proxies are available (Hughes et al., 2001). As proxy for adherence I use the medication possession ratio (MPR) at patient level, defined as the number of days of medication supplied within the refill interval over the number of days in the refill interval. The smaller the value, the smaller the adherence to prescribed regime. This proxy is widely used in the literature, thanks to its simplicity of calculation and interpretation (Cramer et al., 2008). Fairman and Motheral (2000) discuss pros and cons of various measures of medication-adherence and conclude that MPR is an economic, valid, and timely indicator. Monaldi et al. (2015) conducted a randomized clinical trial specific to the Italian population and conclude that an MPR at least equal to 0.75 is crucial for effectiveness of statin medication; Atella et al. (2017) run several robustness checks using data similar to this study and confirmed the finding of Monaldi et al. (2015). These results suggest to employ a binary indicator of adherence, as standard in the health economics literature (Atella et al., 2017; Monaldi et al., 2015; Lamiraud and Geoffard, 2007): if MPR is at least 0.75 the patient is defined as adherent.[8]

The second important aspect to test the theoretical model is related to the conditions that affect the cholesterol level under non-adherence. The data provide several details that will be exploited as covariates in the empirical application. A comprehensive measure of general health conditions is the Charlson et al. (1987) index, a composite measure for the seriousness of the disease that increases as health conditions worsen (the largest value is equal to 6 and refers to HIV/AIDS and metastatic tumor; I collapsed together the conditions above 4 for reasons related to sample frequencies above this threshold). I also have access to clinical information related to specific pathologies that are associated to high cholesterol levels, like diabetes and hypertension (British Heart Foundation, 2014), or that may affect the intake and/or the effectiveness of statins, like ischemia, ictus, heart attack, and other cardiovascular diseases (SPARCL, 2006; Adams et al., 2008); occurrences of hospitalization are recorder. Finally, the health conditions under non-adherence may be influenced

[8] All the conclusions reached in this paper are valid if I change the threshold to be defined as adherent patient to $\{0.70, 0.72, 0.75, 0.77, 0.80\}$.

by some demographic characteristics, like gender and age, that are registered in the data.

It is widely known that the molecular components of some statins are more powerful than others for the management of cholesterol. This property is summarized by an equipotency index (Maron et al., 2000; Table 1): for a given quantity of statins, the higher the index the larger the reduction of cholesterol and, in terms of the stylized model, the lower the opportunity cost of adherence (e.g., the shoe-leather costs). In addition, the data provide the specific molecular composition of the statin (i.e., Simvastatin, Atorvastatin, Rosuvastatin, Fluvastatin, Lovastatin, Pravastatin).

In Table 2 I report descriptive statistics for selected characteristics by treatment status, for the pooled sample. The cholesterol is lower for adherent patients than for non-adherent patients by 1 percentage points and the heterogeneity is slightly larger for adherent patients than for non adherent. The difference in cholesterol across the two groups is slightly decreasing in quantiles: it is about 2 percentage points at 10th quantile, 1 at the median and about zero at 90th quantile. This suggests that there might be heterogeneity in the treatment effect (although -at this stage- not necessarily 'essential' in the sense of Heckman et al., 2006). The differences for the other observable characteristics are usually negligible.

5 Empirical analysis

In this section I present the empirical analysis. First, I check that the required assumptions for the identification of MTE are verified (Section 5.1). Then I estimate the net benefit from statin medication to reduce LDL cholesterol (Section 5.2). If the empirical results are coherent with the predictions of the theoretical model it is meaningful to implement policies that increase the adherence rate.

5.1 Checking the assumptions

As a preliminary step for the analysis I tackle three issues: 1) the existence of a valid instrument; 2) the common support assumption (required for a meaningful comparison); 3) whether 'essential heterogeneity' is a feature of the data. The first two aspects are common to any treatment evaluation

with self-selection; the last is somewhat specific to the approach taken in this paper.

5.1.1 The existence of a valid instrument

A valid instrument should be relevant to explain the sorting mechanism into adherence without impact on the cholesterol indicator (i.e., exogenous). Increasingly relevant strands of the health economics and medical literatures emphasize the role of the physicians for the adherence of patients to prescribed medications (Fichera et al., 2018; Murtagh, 2007).[9] Based on these growing literatures as instrument I exploit the communication skill of the GPs. For each patient i at time t, the instrument is constructed as the average adherence of the other patients cured by the same GP (excluding the observation $\{i,t\}$), i.e. $Z_{it} = \sum_{j \neq i} \text{adherence}_{jt}$, for each patient j cured by the same doctor of patient i. A similar instrument was introduced by Kling (2006) under the name of 'judge design', that uses the differences in the harshness/leniency of judges, and it has been used to address several questions in other fields of economics (Doyle, 2007; Maestas et al., 2013); in this paper the expression 'gatekeeper design' was suggested (by David Card, personal communication). The need to break any dependence (even due to unobservable characteristics) affecting both the treatment status and the outcome to satisfy the exogeneity condition is the reason why I exclude the $\{i,t\}$-th observation from the construction of Z.

Two main concerns would threaten the exogeneity of the instrument and eventually its validity: first, that the communication skills of the GPs may be related to the technical ability of GPs (i.e., a 'knowledge channel'); second, that there might be some unobservable components of the instrument that are related to cholesterol, e.g. a preference for diet (i.e., a 'preference channel'). With respect to the knowledge channel, the concern is that patients of GP 'A' are more adherent to prescribed treatment regime than patients of GP 'B' because GP 'A' has a better clinical knowledge than GP 'B', and thus is more successful in making patients adherent because provides more and better information on hypercholesterolemia. In this case the instrument would capture an

[9] For example, in the health economics literature the prediction of a formal theoretical model that doctors' effort exerts a positive effect on patients behaviour could not be rejected by Fichera et al. (2018) using data from UK. In the medical literature, the Royal Australian College of General Practitioners identifies the communication skills of the physicians and the doctor-patient relationship as main requirements for GPs practice, at the same level of importance as the clinical knowledge and skills (Murtagh, 2007).

ability endowment of GP with direct, although unobservable, effect on cholesterol. This mechanism would induce a non random sorting of patients across physicians. In the Italian system, several institutional and clinical reasons make this argument largely irrelevant. As for institutional reasons, a sorting of patients across GPs due to the management of hypercholesterolemia is very unlikely because in Italy all the individuals are enrolled in the National Health System (NHS). The NHS is financed through general taxation and provides universal coverage free of charge at the point of service (Lo Scalzo et al., 2009). Thanks to the enrollment in the NHS patients are entitled to a GP who is responsible for providing a list of services (among which are prescription of drugs, requests for specialist visits and diagnostic tests). When an individual turns 14, she/he chooses a GP from a list of available doctors that is administered by the Local Health Unit (LHU). The choice is made on the basis of geographical proximity (Atella et al., 2018). Each GP can manage at most 1,500 patients and the salary is proportional to the actual number of patients. The relationship between GP and an individual can be broken at any time for two reasons: the GP has the power to refuse the assistance to an individual by communicating the motivations to LHU that afterwards selects a new GP;[10] the individual may choose to change a physician without motivating the decision. In practice, splits of patients from physicians are rare (about 1% of obervations in these data), which rules out the possibility of non-random sorting in this analyis: GP would likely face a trial; for patients it would be unlikely to find a GP in their neighborhood (the cap to the number of patients within each physician is reached relatively fast because children are typically cured by GPs rather than by pediatricians, as described by Ministry of Health, 2013, p.10). To support the argument that splits of patients from GPs are random I implement a strategy introduced by Altonji et al. (2005). This strategy is based on the idea that the amount of selection on the observable characteristics is informative about the amount of selection on the unobservable characteristics. It follows that if the sorting of patients within GPs is at random the observable characteristics should be similar between patients who split from their GPs and patients who do not split. In Table 3 I report key observable characteristics conditioning on the splits of patients from GPs. The characteristics of the two

[10]Most common causes are related to the mutual trust between patients and GP. Among the causes is excluded the 'complexity of the case'. The Italian Constitution (art.32) guarantees access to health services for everybody, therefore several decisions of the judicial courts were against physicians refusing care to patients. Hence, incentives to break the relation are virtually absent for GPs.

groups are remarkably similar, thus supporting the random allocation of patients across physicians thanks to the Italian institutional framework. As for the clinical motivations the management of hypercholesterolemia is not an hard task for GPs: 1) the condition is widespread (the incidence of high LDL values in the Italian population is 70% for basically the same ages considered in this analysis); 2) the information flow from medical research to GPs is spread in a short time period (Atella et al., 2018); 3) official guidelines are available to assist the GP in each possible situation (Adams et al., 2008); 4) if guidelines are still not enough, GP can prescribe a specialist visit, and the outcome would still be recorded in HSD database. Continuing with the strategy of Altonji et al. (2005), in Table 3 I build subsamples based on the idea that richest/poorest patients may select the best/worst GPs in which case there would be an indirect sorting of patients within GPs based on the technical ability of GPs. Even in this case, the observable characteristics across the two groups are remarkably similar. These falsification tests support the claim that the allocation of patients across physicians is indeed as good as randomly assigned and that the knowledge channel is not a concerning falsification channel.

A second reason of concern against the validity of the instrument is that the instrument is not exogenous because of unobservable components related to cholesterol, namely a preference for diet that may be a substitute for statin medications.[11] Among the available data, as a proxy for diet it may be used the Body Mass Index (BMI), that is recorded for about 7.5% of the sample (hence the indicator is not used in the analysis) To check the relevance of this falsification channel, in Table 4 I report the average BMI across level of the instrument, by gender and treatment status: there is no systematic difference of the index along these dimensions, thus even this source of potential instrument falsification is dismissed.

More formally, I tested the assumptions related to monotonicity and exogeneity of the instrument using a statistical test proposed by Mourifié and Wan (2016) (see also Kitagawa, 2015; Laffers and Mellace, 2016): the joint null hypotheses are never rejected at standard confidence level (Table 5), thus the validity of the instrument is not rejected.[12] As for the relevance of the instruments,

[11] Notice that the individual preference for diet over medication would bias downward the estimated benefit, thus providing a lower bound of the true treatment effect: the indicator of cholesterol of a patients not adherent to medication improves thanks to a better diet, thus decreasing the estimated benefit from medication.

[12] The test takes the form of two inequalities that are necessary to identify a Local Average Treatment Effect

the F-statistics from the first stage are much higher than 20, so finite sample bias is not a concern (Bound et al., 1995; Stock and Yogo, 2002).

Finally I am in the unique position to run a *validation* test of the interpretation of the instrument in terms of communication skills by considering conditions where communication is the only relevant channel, and technical knowledge plays a minor role. This is the case of smoking behaviour and sport attitudes. In a regression of these indicators on the instrument and other covariates, I find that a GP with better communication skills can successfully increase the likelihood that a patient suffering from hypercholesterolemia quits smoking or practice sport activities (Table 6).[13] This supports my interpretation of the instrument in terms of communication skills.

5.1.2 The common support assumption

With a valid instrument, a comparison between treated and untreated patients is meaningful only if $0 < Pr(D = 1|X = x) < 1$, i.e. for treated individuals there is a comparison group of untreated patients. I check the probability of adherence spanned by the instrument (Figure 2). There is huge overlap between the distributions of adherence for treated and untreated patients over a large support of the probability. Only in the tails (for very large or very small estimated probabilities of adherence to prescribed treatment regime) there are few observations and the common support may turn out to be restrictive. For this reason, the discussion of results will not pay much attention to most extreme estimated probabilities (say less than 0.05 and more than 0.95). (For a similar approach, see for example Carneiro et al., 2003.)

5.1.3 Essential heterogeneity

The final important check is whether in the data there is essential heterogeneity in the sense of Heckman et al. (2006), namely that responses to interventions are heterogeneous and patients opt

(Imbens and Angrist, 1994): $P[y, D = 1|Z = 0] \leq P[y, D = 1|Z = 1]$ and $P[y, D = 0|Z = 1] \leq P[y, D = 0|Z = 0]$. If any of the two inequalities is violated, the validity of the instrument is falsified. In the latter case, the test is not informative about which of the two assumptions fails.

[13] The information is not used in the analysis because it is recorded for about 30% of the observations. For the indicator of smoking attitude the estimated coefficient is negative and large for men and women- although strongly significant only for women; for the indicator of sport attitudes the estimated coefficient is positive, but not significant at standard confidence level as a consequence that sport attitudes are self-reported and that the share of individuals that practice sport is very high (Istat, 2017).

for treatment with at least partial knowledge of their idiosyncratic response. I exploit two different methods. The first method uses the fact that $MTE(U_D = p)$ is a derivative of $E[\ln y|p]$ with respect to p (Heckman and Vytlacil, 1999), so that if the IV estimate is linear in the propensity score the treatment effect is constant and the empirical content of LATE can be extended to the whole population (Carneiro et al., 2003). In the regression $y = x\beta + \gamma_1 p + \sum_{k=2}^{K} \gamma_k p^k$ I test the joint null hypothesis that $\{p^k\}_{k=2}^{K} = 0$, necessary for linearity of the propensity score. The null hypothesis is strongly rejected even with with $K = 3$ (Table 7), and therefore MTE is needed.

A second method is proposed by Kowalski (2016) and is based on the representation

$$y = \lambda_D D + \lambda_Z Z + \lambda_{DZ} DZ,$$

where the joint null hypothesis $\lambda_D = \lambda_{DZ} = 0$ tests whether the treatment effect is globally externally valid and equal to zero. Coherent with the previous test, these null hypotheses are strongly rejected, either jointly or separately. As a consequence, a constant IV does not jointly identify the relevant economic effects (LATE/ATE/ATT/ATU). This result empirically confirms that in this analysis the IV, although consistent, identify a local treatment effect and therefore it is not appropriate to test the economic model of Section 2. For completeness, the Hausman (1978) test strongly rejects the null hypothesis of non systematic difference between IV and OLS: although under heterogeneous treatment effect the empirical content of the Hausman (1978) test is less clear than under homogeneity, I interpret this result as pointing towards the need for methods robust to endogeneity *and* heterogeneous effect.

This preliminary analysis provides large evidence in favour of a MTE approach rather than an IV generalized to the entire population, using as instrument the communication skills of the physicians, constructed on the base of the gatekeeper design. The results delivered by this approach will be valid for a very large population (as identified by the common support). Through the discussion of the results, it shall be clear that the information provided by MTE greatly enriches what we might know about the effects of medication on cholesterol using standard techniques.

5.2 Results

There are three (families of) estimators for MTE: fully parametric (usually, but not necessarily, based on normality assumptions as in Bjorklund and Moffitt, 1987), semi-parametric (Heckman and Vytlacil, 2007b; Carneiro and Lee, 2009; Carneiro et al., 2017), and fully non-parametric (Heckman and Vytlacil, 2005). I estimate all of them, but prefer the semi-parametric method, on which the empirical analysis presented in this section is based. The results from semi-parametric method are indistinguishable from those of the non-parametric approach (and, if anything, are more conservative; see Figure 4), but the computational burden is much lower. To see how semi-parametric method works, write the observed outcome as $E[y|X,p] = E[\alpha + U_0 + D\,X\,\beta + \underbrace{D\,E[U_1 - U_0|p]}_{K(p)}]$, and use the method of Local Instrumental Variable (Heckman and Vytlacil, 2000) to estimate $MTE = \frac{\partial E[y|P(Z)=p]}{\partial p} = X\,\beta + K'(p)$. The selection rule is based on a logit and $K'(p)$ is based on a polynomial of degree 3.[14]

The model specification is suggested directly by the economic model of Section 2. The set of covariates includes clinical and socio-demographic characteristics described in Section 4.[15] I also control for the initial level of cholesterol to take into account that the required reduction of cholesterol is different across different patients. In addition, dummies for years, quarters and geographical areas are added to adjust for possible systematic differences along these dimensions.

The three necessary conditions for coherency between the theoretical predictions of the model

[14] The fully non parametric approach requires that the support of the distribution U_D conditional on the design matrix to span the full unit interval (Carneiro et al., 2011), and is computationally slower than the semi-parametric approach (in this application, the time of execution of non-parametric estimator is 4 times that of semi-parametric estimator). The semi-parametric approach does not impose distributional assumptions (in contrast to the fully parametric approach) and has weaker requirement on U_D with respect to the non-parametric approach. See Carneiro et al. (2003) for further technical details. The fully parametric approach has the disadvantage of fully specifying the Data Generating Process (DGP), thus being consistent and more efficient than the other methods when the assumptions are indeed verified, but at the cost that even minimal departures from them leads to inconsistency (Arabmazar and Schmidt, 1981, 1982). Under this setup I can implement an Hausman (1978) test. The results under normality assumption are not shown because the Hausman test strongly rejects the null hypothesis of non-systematic difference between the fully parametric and semi-/non-parametric estimators. The results from a parametric estimator are remarkably different from those presented here (available upon request).

[15] The list includes: an indicator of general health condition (Charlson et al., 1987), dummies for occurrences of hospitalization, diabetes, hypertension, ischemia, ictus, heart attack, and other cardiovascular diseases; equipotency index across different statins, dummies for consumption of specific statins (namely, simvastatin -possibly paired with other molecular compositions-, lovastatin, pravastatin, fluvastatin, atorvastatin, rosuvastatin). As for socio-demographic characteristics, I include an indicator for gender and age.

of Section 2 and the data are satisfied.

1. Adherence to prescribed medication causes a reduction of cholesterol.

The average reduction of the cholesterol indicator thanks to medication is about 2.5% (Table 8), which is comparable to, but slightly lower than the effect estimated in Randomized Clinical Trials (4S, 1994).[16]

2. There exists heterogeneity in return to adherence.

The MTE parameter increases monotonically over the support of U_D (Figure 3), so that patients whose reduction of cholesterol is larger (i.e., have a more negative treatment effect) enter first into treatment (i.e., at low values of U_D).

3. $ATT < ATE$ and $ATE < ATU$.

The sorting gain (i.e., $ATT - ATE = -0.020$) is negative and significant, confirming that treated patients enjoy a larger benefit than a patient taken at random from the population. At the same time, $ATE - ATU = -0.013$, because non treated patients enjoy a lower benefit than a patient taken at random from the population. This 'full decomposition' of the sorting gain suggests yet another test of the (non-)essential heterogeneity, based on the joint null hypotheses that $(ATE = ATT)\&(ATE = ATU)$. With respect to existing tests run in Section 5.1 this test requires each quantity to be known in advance (which is clearly a disadvantage; although notice that $(ATE = ATT) \Leftrightarrow (ATE = ATU)$).[17] However, when we reject the null hypothesis we obtain more information than can be gained from the existing tests, namely which component of the heterogeneity is most important. In this analysis the most important component of the heterogeneity is the large gain from medication (a positive sorting) because -with respect to the average patient- the surplus of the benefit from medication for patients that are actually treated (i.e., ATT-ATE) is

[16] Deaton and Cartwright (2018) discuss at length the reasons behind the differences of effectiveness estimated in the real world and in randomized clinical trials. One prominent reason is the adherence to medication, that in lab-experiment is not an issue. Accordingly, if I focus on adherent patients, the estimated effect is close to that expected on the base of randomized clinical trials (4S, 1994). In this respect, this paper provides an empirical support to the arguments in Deaton and Cartwright (2018).

[17] Write $ATE = \alpha\, ATT + (1-\alpha)ATU$, if $ATE = ATT$, then $ATE = \alpha\, ATE + (1-\alpha)ATU$ and $(1-\alpha)\, ATE = (1-\alpha)ATU$. By the same token, $(1-\alpha)\, ATT = (1-\alpha)ATU$.

almost twice the benefit that an untreated patients would enjoy if treated (i.e., ATE-ATU). This is coherent with the unconditional descriptive statistics of Table 2. Finally, the selection effect is positive and significant (i.e., $OLS - ATT = 0.031$), as one would expect from the Roy (1951) model if patients who enjoy a larger benefit self-select into treatment.

The comparison with standard techniques

The results obtained using MTE greatly enrich those based on standard techniques. The treatment effect estimated with the OLS (-0.01252) is smaller than that from MTE because it suffers from a bias due to the omission of the unobservable components related to the sorting mechanism (Section 2). The treatment effect from IV (-0.03512) is larger than from MTE reflecting the local identification power of the estimator: the weights that map the MTE to the IV (Heckman and Vytlacil, 2005, Tab 1A) are maximized between $p = (0.4, 0.5)$, where the treatment effect of MTE is larger (graph available upon request). Moreover, neither of the standard estimators is informative about the huge heterogeneity of the treatment effect that MTE shows. In particular, nor the OLS nor the IV informs that drugs benefits more some patients than others, being not very effective for part of the population (Heckman et al., 1997).

Since the predictions of the economic model and the data are coherent with each other, it is meaningful to make a step forward and understand whether it is possible to implement policies that increase the adherence rate, i.e. the share of patients who are adherent to prescribed medication.

6 Hypothetical policies to increase the adherence rate

Using the results from a Randomized Clinical Trial on cholesterol in Italy Monaldi et al. (2015) conclude that physicians need to increase the awareness of patients regarding the consequences from underuse of prescribed medication. Monaldi et al. (2015) recommend that physicians encourage adherence above 0.75. Motivated by this reccommendation in this section I propose three alternative hypothetical policies that aim at increasing the adherence rate. One policy affects the baseline probability of adherence of the patients (e.g., through TV and radio advertisements, that have

been shown to be important to affect this behaviour; see Oberlander et al., 2017); one policy affects the communication style of the physicians (e.g., through courses that improve their communication skill); one policy affects the overall probability of adherence (e.g., through a mix of interventions that are left unspecified). For each policy, I shock the current state of the world by 10%. The policy maker may wish to know: 1) how large would the average gain be for patients that are currently non adherent, but that would be in a different scenario; 2) which policy maximizes the return in terms of lower cholesterol; 3) what are the characteristics of the individuals that change behaviour.

The first question is the definition of policy relevant treatment effect (PRTE; Table 9). The most effective policies are the advertisements directed to patients and (slightly less effective) the courses that improve the communication skill of the physicians: the average reduction of cholesterol for patients currently untreated but that would opt for treatment if these policies were implemented would be about 3%, compared to 4.3% for patients already treated under the current scenario; the smallest return is obtained when a policy mix is implemented: the average reduction of cholesterol would be about 2%.

In many circumstances this is all the information the decision maker needs, and indeed this is the point where the current literature stops. However, in other circumstances more is required. Suppose that the advertisement to patients and the course to GPs, two policies that have the same return, switch into treatment two different individuals. If we knew which patient is induced into adherence the two policies would no longer be equivalent and the policy maker may prefer one policy over the other. As a technical contribution of this paper to the existing literature, I describe the observable characteristics of the patients induced into treatment by each policy (Appendix A). The description of observable characteristics for patients switched into adherence is reported in the lower panel of Table 9 (for selected covariates). A ratio $\frac{E[X|\Delta P, u_d]}{E[X]} > 1$ implies that the share of patients with the characteristic X that are induced into treatment thanks to the policy is larger than the share of patients with the characteristic X in the overall population. The results are highly informative. A training course that improves the communication skill of the GP induces into treatment healthier patients than the advertisement to patients would do (i.e., the share of patients

switched into treatment and suffering from bad health conditions -as summarized by the Charlson index- is always lower with the policy towards physicians than towards patients). In contrast, a policy mix induces into treatment less healthy patients (i.e., the share of patients switched into treatment after a policy mix and suffering from bad health conditions is the highest among the proposed policies). Therefore, a policy that successfully increases the communication of the GP may be preferred in an attempt to reduce avoidable, *future*, occurrences of worse health conditions through prevention. In contrast, if the aim of the policy is to switch into adherence patients with worst current conditions, in an attempt to reduce *current* health care costs, then a mix of policies may work better. Therefore which of the three policies is actually preferred depends on the shape of the utility function of the policy maker. The preference for a less effective policy may be perfectly rational to the extent that it targets better the desired population.

6.1 Is the policy cost-effective?

These policies might be implemented provided that they are 'convenient'. From the previous discussion it is clear that the definition of convenience depends on the preferences of the decision maker. As the preferences of the decision maker are not observable to the researcher, in this section I consider only the cost-effectiveness of each policy. In particular, I focus on monetary components of costs (C) and benefits (B) in order to derive the Total Net Benefit (TNB), i.e. $TNB = B - C$. The quantities of interest are incremental with respect to current state of the world D and refer to the entire population, which is coherent with the PRTE parameter estimated above.

With the data at hand, I can estimate precisely only the total cost of each policy (all the details are reported in Appendix B). It shall be clear that even though the information on associated benefits is scant I am still in the position to address the issue of cost-effectiveness. Indeed, I find that a subset of the total benefits are larger than the total cost of implementation of the policy, so that I can already draw the conclusion that the policy is cost-effective. Furthermore, the partial information on benefits provides conservative conclusions in the sense that I find a *lower bound* of the true Total Net Benefit.

I implement three different hypothetical policies. For policy directed towards patients, I consider

the advertisement by TV and radio; for policy directed towards physicians, I consider on-line communication courses; for policy mix, I consider as if the policies directed to patients and physicians are jointly implemented. A very conservative estimate of the cost of the implementation of the policies ($C_{TOT.} = C_{DRUGS} + C_{POLICY}$) are about $C_{TOT.} < 10\,mil.€$ when the target is the physician, $C_{TOT.} \approx 25\,mil.€$ when the target is the patient, $C_{TOT.} \approx 30\,mil.€$ when a policy mix is implemented (Table 10).

The array of potential benefits from these policies is large and difficult to estimate with the data at hand. Benefits may be 'direct' or 'indirect'. Among direct benefits the most important is hospitalization. For the Government, the cost of hospitalization is about $2,500€$ for CVDs *without* surgery and between $15,000$ and $25,000€$ for CVDs *with* surgery (Ialian Official Journal n. 23, 28 Jan. 2013). Therefore, avoiding 1,000 CVDs surgeries makes any hypothetical policy discussed in this paper cost effective. The data do not allow to estimate the odds of these hospitalization. However, according to the results from randomized clinical trials this target is feasible (Athyros et al., 2010; 4S, 1994; Monaldi et al., 2015). As for indirect benefits they are relevant because, for example, a lower cholesterol level might reduce the occurrences of other pathologies that are typically associated with hypercholesterolemia (Kasteridis et al., 2014). To the extent that avoiding these other pathologies reduces health expenditures, the saving should be considered as a benefit from the policy. I am not be able to evaluate these indirect benefits, hence the TNB estimated here may be understood as a lower bound of the true total net benefit. Even considering only the direct benefit of the policies is enough to draw conclusions in favour of the cost-effectiveness of any of the proposed policy.

7 Working with cohort subsamples

To shed more light on the effectiveness of adherence to reduce cholesterol I investigate whether relevant differences are estimated by age. I split the sample in 3 categories (of approximately same sample size): patients born before 1935 (the 'old' patients), patients born between 1935 and 1944 (the 'middle age' patients), and patients born after 1944 (the 'young' patients). Other

age splits would be legitimate, but I view these as a fair compromise between ages relevant for hypercolesterolemia and (sub)sample sizes.

For all birth cohorts, the conclusions are coherent with those from the whole sample, namely the three necessary conditions for coherency between the model of Section 2 and the data are satisfied. Interesting differences emerge when I compare the magnitudes of the estimated effects. Patients who on average benefit most from adherence to prescribed treatment are the middle-age and young patients (an ATE larger than 3%), whose benefit is more than three times larger than that of the patients who benefit the least (the old patients, whose ATE is 1.0%). More insights can be gained from the full decomposition of ATE. The positive sorting gain (ATT-ATE) is largest for the old patients, thus implying that for these patients the gain from medication with respect to an individual taken at random from the population of the same age is the largest when compared to younger patients. However, for elder patients the ATU is positive (although not very precisely estimated), so that untreated patients would experience a *negative* benefit if they were treated. This happens because old patients are in worst general health conditions on average (i.e., have an higher Charlson index), and therefore side effects from medication would more than compensate the benefit (Heckman et al., 1997). Differently, for middle age and young patients the ATT is the largest in absolute value but the sorting gain (ATT-ATE) is smaller than for elder patients because also the ATU is negative reflecting the better general health conditions of younger patients.

8 Concluding remarks

This paper investigates the relation between adherence to prescribed medication and cholesterol, by jointly taking into account the possible sorting of patients into treatment and the heterogeneity of the treatment effect. Marginal Treatment Effect is most convenient in this setup, because it identifies the average reduction of cholesterol from adherence for patients just indifferent between treatment and non treatment, at a given level of unobservables. To the extent that there exists substantial heterogeneity (Heckman et al., 2006), MTE will be different across levels of unobservables.

Confirming the predictions of a stylized economic model I find that drugs successfully reduce the level of cholesterol, and patients who benefit most from the treatment are more likely to adhere to prescribed regime than patients who benefit the least. The heterogeneity of the treatment effect is substantial: patients who benefit most from medication enjoy a reduction of cholesterol indicator as large as 5%, whilst patients who benefit least enjoy a remarkably small, perhaps even negligible, reduction (Heckman et al., 1997). The average reduction of cholesterol is comparable to, but slightly lower than the effect estimated in Randomized Clinical Trials (4S, 1994). This is an empirical contribution to a recent but growing literature that emphasizes the differences of drug effectiveness between Randomized Clinical Trials and evryday practice (Deaton and Cartwright, 2018).

These results are used to study the effects of alternative hypothetical policies aiming at increasing the share of patients adherent to medication, so to improve the individual health conditions (Monaldi et al., 2015). The proposed policies affect either the patients or the GPs. As a technical contribution to the existing literature the paper provides the description of the observable characteristics of patients switched into treatment thanks to the policies. This is an important improvement over the current literature because the policies differ much both in terms of reduction of cholesterol per patient shifted (about 3% when the patients or the physicians are targeted, as opposed to 2% when a policy mix is implemented), and in terms of the population affected. As a consequence, less effective policies may be rationally preferred to the most effective policy depending on desired target population. The possibility to inform the policy maker with respect to the population induced into treatment may be very helpful to tailor the most appropriate policies, and to evaluate the trade-off between effectiveness of the policy and targeted population. Back of the envelope calculations suggest that any of the proposed policy is convenient when evaluated from a purely monetary perspective.

References

4S, S. (1994, nov). Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). *The Lancet 344*(8934), 1383–1389.

Adams, R. J., G. Albers, M. J. Alberts, O. Benavente, K. Furie, L. B. Goldstein, P. Gorelick, J. Halperin, R. Harbaugh, S. C. Johnston, I. Katzan, M. Kelly-Hayes, E. J. Kenton, M. Marks, R. L. Sacco, and L. H. Schwamm (2008). Update to the aha/asa recommendations for the prevention of stroke in patients with stroke and transient ischemic attack. *Stroke 39*(5), 1647–1652.

AIFA (2017). *Elenco confezioni per Principio Attivo*. Agenzia Italiana del Farmaco. In Italian.

Altonji, J. G., T. E. Elder, and C. R. Taber (2005, feb). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy 113*(1), 151–184.

American Heart Association (2015). Heart disease and stroke statistics2015 update. Technical report.

Angrist, J. and J. Pischke (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Angrist, J. D. (2004). Treatment effect heterogeneity in theory and practice*. *The Economic Journal 114*(494), C52–C83.

Angrist, J. D., G. W. Imbens, and D. B. Rubin (1996). Identification of causal effects using instrumental variables (Disc: p456-472). *Journal of the American Statistical Association 91*, 444–455.

Arabmazar, A. and P. Schmidt (1981). Further evidence on the robustness of the tobit estimator to heteroskedasticity. *Journal of Econometrics 17*(2), 253 – 258.

Arabmazar, A. and P. Schmidt (1982). An investigation of the robustness of the tobit estimator to non-normality. *Econometrica 50*(4), 1055–1063.

Atella, V., F. Belotti, J. Bhattcharya, and D. Depalo (2018). Technological Advance in Cholesterol Medication Meets Physician Learning: A Non-Parametric Bounding Approach. Technical report. Mimeo.

Atella, V., F. Belotti, and D. Depalo (2017, September). Drug therapy adherence and health outcomes in the presence of physician and patient unobserved heterogeneity. *Health Economics 26*, 106–126.

Atella, V., F. Belotti, J. Kopinska, A. Palma, and A. P. Mortari (2018, February). Economic Crisis, Mortality and Health Status. A New Perspective. CEIS Research Paper 425, Tor Vergata University, CEIS.

Athyros, V. G., K. Tziomalos, T. D. Gossios, T. Griva, P. Anagnostis, K. Kargiotis, E. D. Pagourelias, E. Theocharidou, A. Karagiannis, and D. P. Mikhailidis (2010). Safety and efficacy of long-term statin treatment for cardiovascular events in patients with coronary heart disease and abnormal liver tests in the Greek Atorvastatin and Coronary Heart Disease Evaluation (GREACE) Study: a post-hoc analysis . *The Lancet 376*(9756), 1916–1922.

Baigent, C., A. Keech, P. Kearney, and L. Blackwell (2005). Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90,056 participants in 14 randomised trials of statins. *Lancet 366*, 1267–1278.

Barter, P., A. M. Gotto, J. C. LaRosa, J. Maroni, M. Szarek, S. M. Grundy, J. J. Kastelein, V. Bittner, and J.-C. Fruchart (2007). Hdl cholesterol, very low levels of ldl cholesterol, and cardiovascular events. *New England Journal of Medicine 357*(13), 1301–1310. PMID: 17898099.

Basu, A. (2014). Estimating person-centered treatment (pet) effects using instrumental variables: An application to evaluating prostate cancer treatments. *Journal of Applied Econometrics 29*(4), 671–691.

Basu, A., J. J. Heckman, S. Navarro-Lozano, and S. Urzua (2007). Use of instrumental variables in the presence of heterogeneity and self-selection: an application to treatments of breast cancer patients. *Health Economics 16*(11), 1133–1157.

Basu, A., A. B. Jena, D. P. Goldman, T. J. Philipson, and R. Dubois (2014, mar). HETEROGENEITY IN ACTION: THE ROLE OF PASSIVE PERSONALIZATION IN COMPARATIVE EFFECTIVENESS RESEARCH. *Health Economics 23*(3), 359–373.

Bianchini, E., O. Brignoli, C. Cricelli, I. Cricelli, F. Lapi, G. Medea, A. Pasqua, S. Pecchioli, C. Piccinni, and M. Simonetti (2016). *IX Report Health Search*. SIMG.

Bjorklund, A. and R. Moffitt (1987). The estimation of wage gains and welfare gains in self-selection models. *The Review of Economics and Statistics 69*(1), pp. 42–49.

Bound, J., D. Jaeger, and R. Baker (1995). Problems with instrumental variable estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association 90*, 443–450.

Brinch, C. N., M. Mogstad, and M. Wiswall (2012, September). Beyond LATE with a discrete instrument. Heterogeneity in the quantity-quality interaction of children. Discussion Papers 703, Statistics Norway, Research Department.

British Heart Foundation (2014). *Reducing your blood cholesterol*, Volume 146.

British Heart Foundation (2015). *CARDIOVASCULAR DISEASE STATISTICS 2015*, Volume 146.

Buchow, H., E. Cayotte, and L. Agafitei (2012). Circulatory diseases - main causes of death for persons aged 65 and more in europe, 2009. *Statistics in focus 7*, 1–12.

Card, D. (1999). The causal effect of education on earnings. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3 of *Handbook of Labor Economics*, Chapter 30, pp. 1801–1863. Elsevier.

Card, D., A. Fenizia, and D. Silver (2018, April). The health effects of cesarean delivery for low-risk first births. Working Paper 24493, National Bureau of Economic Research.

Carneiro, P., J. J. Heckman, and E. Vytlacil (2003). Understanding what instrumental variables estimate: Estimating marginal and average returns to education. *processed, University of Chicago, The American Bar Foundation and Stanford University, July 19*.

Carneiro, P., J. J. Heckman, and E. Vytlacil (2010, 01). Evaluating Marginal Policy Changes and the Average Effect of Treatment for Individuals at the Margin. *Econometrica 78*(1), 377–394.

Carneiro, P., J. J. Heckman, and E. J. Vytlacil (2011, October). Estimating marginal returns to education. *American Economic Review 101*(6), 2754–81.

Carneiro, P. and S. S. Lee (2009). Estimating distributions of potential outcomes using local instrumental variables with an application to changes in college enrollment and wage inequality. *Journal of Econometrics 149*(2), 191–208.

Carneiro, P., M. Lokshin, and N. Umapathi (2017). Average and marginal returns to upper secondary schooling in indonesia. *Journal of Applied Econometrics 32*(1), 16–36.

CBO (2018). *The Budget and Economic Outlook: 2018 to 2028*.

Charlson, M., P. Pompei, K. Ales, and C. R. MacKenzie (1987). A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chron Dis 40*(5), 373–383.

Chernozhukov, V., G. W. Imbens, and W. K. Newey (2007). Instrumental variable estimation of nonseparable models. *Journal of Econometrics 139*(1), 4 – 14. Endogeneity, instruments and identification.

Cornelissen, T., C. Dustmann, A. Raute, and U. Schonberg (2016). From LATE to MTE: Alternative methods for the evaluation of policy interventions. *Labour Economics 41*, 47 – 60. SOLE/EALE conference issue 2015.

Coyle, D., M. J. Buxton, and B. J. O'Brien (2003). Stratified cost-effectiveness analysis: a framework for establishing efficient limited use criteria. *Health Economics 12*(5), 421–427.

Cramer, J., A. Roy, and A. Burrell (2008). Medication compliance and persistence: Terminology and definitions. *Value in Health 11*.

Deaton, A. and N. Cartwright (2018, aug). Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine 210*, 2–21.

Deb, P., M. K. Munkin, and P. K. Trivedi (2006, oct). Private Insurance, Selection, and Health Care Use. *Journal of Business & Economic Statistics 24*(4), 403–415.

Diiro, G. M., H. D. Affognon, B. W. Muriithi, S. K. Wanja, C. Mbogo, and C. Mutero (2016, dec). The role of gender on malaria preventive behaviour among rural households in Kenya. *Malaria Journal 15*(1), 14.

Doyle, J. J. (2007, December). Child protection and child outcomes: Measuring the effects of foster care. *American Economic Review 97*(5), 1583–1610.

European Commission (2015). *The 2015 Ageing Report*. Belgium: European Commission.

Fabiani, L., M. Scatigna, K. Panopoulou, A. Sabatini, E. Sessa, F. Donato, M. Marchi, R. Nardi, C. Niccolai, F. Samani, and G. Ventriglia (2004). Health search: istituto di ricerca della societá italiana di medicina generale: la realizzazione di un database per la ricerca in medicina generale. *Epidemiol and Prev 28*, 156–162.

Fairman, K. A. and B. Motheral (2000). Evaluating medication adherence: Which measure is right for your program? *Journal of Managed Care Pharmacy 6*(6), 499–506.

Fichera, E., J. Banks, L. Siciliani, and M. Sutton (2018, nov). Does patient health behaviour respond to doctor effort? *Journal of Economic Behavior & Organization* (62).

French, E. and J. Song (2014, May). The effect of disability insurance receipt on labor supply. *American Economic Journal: Economic Policy 6*(2), 291–337.

Gibson, L. A., M. Hernández Alava, M. P. Kelly, and M. J. Campbell (2016, sep). The effects of breastfeeding on childhood BMI: a propensity score matching approach. *Journal of Public Health*.

Hausman, J. A. (1978, November). Specification Tests in Econometrics. *Econometrica 46*(6), 1251–71.

Heckman, J. (2010, June). Building bridges between structural and program evaluation approaches to evaluating policy. *Journal of economic literature 48*(2), 356–398.

Heckman, J. J. (1979, January). Sample Selection Bias as a Specification Error. *Econometrica, Econometric Society 47*(1), 153–61.

Heckman, J. J., J. Smith, and N. Clements (1997). Making the most out of programme evaluations and social experiments: Accounting for heterogeneity in programme impacts. *The Review of Economic Studies 64*(4), 487–535.

Heckman, J. J., S. Urzua, and E. Vytlacil (2006, August). Understanding Instrumental Variables in Models with Essential Heterogeneity. *The Review of Economics and Statistics 88*(3), 389–432.

Heckman, J. J. and E. Vytlacil (2005, 05). Structural Equations, Treatment Effects, and Econometric Policy Evaluation. *Econometrica 73*(3), 669–738.

Heckman, J. J. and E. J. Vytlacil (1999). Local instrumental variables and latent variable models for identifying and bounding treatment effects. *Proceedings of the National Academy of Sciences of the United States of America 96*(8), 4730–4734.

Heckman, J. J. and E. J. Vytlacil (2000). The relationship between treatment parameters within a latent variable framework. *Economics Letters 66*(1), 33 – 39.

Heckman, J. J. and E. J. Vytlacil (2001). Instrumental variables, selection models, and tight bounds on the average treatment effect. In *Econometric Evaluations of Active Labor Market Policies in Europe. Physica-Verlag.*

Heckman, J. J. and E. J. Vytlacil (2007a, January). Econometric Evaluation of Social Programs, Part I: Causal Models, Structural Models and Econometric Policy Evaluation. In J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics*, Volume 6 of *Handbook of Econometrics*, Chapter 70. Elsevier.

Heckman, J. J. and E. J. Vytlacil (2007b, January). Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast their Effects in New. In J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics*, Volume 6 of *Handbook of Econometrics*, Chapter 71. Elsevier.

Hughes, D., A. Bagust, A. Haycox, and T. Walley (2001). The impact of non-compliance on the cost-effectiveness of pharmaceuticals: a review of the literature. *Health Economics 10*, 601–615.

Imbens, G. W. and J. D. Angrist (1994, March). Identification and Estimation of Local Average Treatment Effects. *Econometrica 62*(2), 467–75.

Imbens, G. W. and W. K. Newey (2009, 9). Identification and estimation of triangular simultaneous equations models without additivity. *Econometrica 77*(5), 1481–1512.

Institute of Medicine (2009). *Initial National Priorities for Comparative Effectiveness Research.* Washington, DC: The National Academies Press.

Istat (2017). La pratica sportiva in italia. `https://www.istat.it/it/files/2015/10/Slide-CONI_Alleva_2017.pdf`. In Italian; Accessed: 2019-07-10.

Kasteridis, P., A. D. Street, M. Dolman, L. Gallier, K. Hudson, J. Martin, and I. Wyer (2014). The importance of multimorbidity in explaining utilisation and costs across health and social care settings : evidence from south somerset?s symphony project. Working paper, ARRAY(0x7fb7acc94b90), York, UK.

Kitagawa, T. (2015). A test for instrument validity. *Econometrica 83*(5), 2043–2063.

Kling, J. R. (2006). Incarceration length, employment, and earnings. *The American Economic Review 96*(3), 863–876.

Kowalski, A. (2018, sep). Biology Meets Behavior in a Clinical Trial: Two Relationships Between Mortality and Mammogram Receipt. Technical report, National Bureau of Economic Research, Cambridge, MA.

Kowalski, A. E. (2016, June). Doing more when you're running late: Applying marginal treatment effect methods to examine treatment effect heterogeneity in experiments. Working Paper 22363, National Bureau of Economic Research.

Laffers, L. and G. Mellace (2016). A note on testing instrument validity for the identification of late. *Empirical Economics*, 1–6.

Lamiraud, K. and P. Geoffard (2007, 11). Therapeutic non adherence: a rational behavior revealing patient preferences? *Health Economics 16*(11), 1185–1204.

Little, R. J. A. (1995). Modeling the drop-out mechanism in repeated-measures studies. *Journal of the American Statistical Association 90*(431), 1112–1121.

Lo Scalzo, A., A. Donatini, L. Orzella, A. Cicchetti, S. Profili, and A. Maresso (2009). Italy: Health system review. *Health Systems in Transition 11*(6), 1–216.

Maestas, N., K. J. Mullen, and A. Strand (2013, August). Does disability insurance receipt discourage work? using examiner assignment to estimate causal effects of ssdi receipt. *American Economic Review 103*(5), 1797–1829.

Maron, D., S. Fazio, and M. Linton (2000). Current perspectives on statins. *Circulation 101*, 207–213.

Ministry of Health (2013). Annuario statistico del servizio sanitario nazionale. Report, Ministero della Salute - Ufficio statistica.

Monaldi, B., G. Bologna, G. G. Costa, C. D'Agostino, F. Ferrante, M. Filice, A. M. Grion, A. Mingarelli, L. Paloscia, R. Tettamanti, C. Veronesi, and L. Degli Esposti (2015). Adherence to statin treatment following a myocardial infarction: an Italian population-based survey. *ClinicoEconomics and outcomes research : CEOR 7*, 273–80.

Mourifié, I. and Y. Wan (2016). Testing local average treatment effect assumptions. *Review of Economics and Statistics na*, n/a–n/a.

Murtagh, J. (2007). *John Murtagh's general practice*. North Ryde, N.S.W 2007: McGraw-Hill Australia.

New England Healthcare Institute (2009). Thinking outside the pillbox a system-wide approach to improving patient medication adherence for chronic disease.

Niederlaender, E. (2006). Causes of death in the eu. *Statistics in focus 10*, 1–12.

Oberlander, L., A.-C. Disdier, and F. Etilé (2017, sep). Globalisation and national trends in nutrition and health: A grouped fixed-effects approach to intercountry heterogeneity. *Health Economics 26*(9), 1146–1161.

Pollard, D. (2001). *A User's Guide to Measure Theoretic Probability*. Cambridge: Cambridge University Press.

Progetto Cuore (2018). Assetto lipidico, italia: Confronto tra 1998-02 e 2008-12. eta' 35-74 anni. `http://www.cuore.iss.it/fattori/colesterolemia.asp`. In Italian; Accessed: 2018-02-08.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers 3*(2), 135–146.

Schoenberg, U., T. Cornelissen, C. Dustmann, and A. Raute (2018). Who Benefits from Universal Child Care? Estimating Marginal Returns to Early Child Care Attendance. *Journal of Political Economy 126*(6), 699979.

SPARCL, T. (2006). High-dose atorvastatin after stroke or transient ischemic attack. *New England Journal of Medicine 355*(6), 549–559. PMID: 16899775.

Stock, J. and M. Yogo (2002). Testing for weak instruments in linear iv regression. In D. Andrews and J. Stock (Eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg*. Cambridge University Press.

Vytlacil, E. (2002). Independence, monotonicity, and latent index models: An equivalence result. *Econometrica 70*(1), pp. 331–341.

Table 1: Statin conversion table: the required mg required to reduce LDL by a a given percentage

Class	% LDL reduction	Type of statin					
		Simva.	Ator.	Rosu.	Fluva.	Lova.	Prava.
1	<24	5	-	-	20	10	10
2	25–32	10	-	-	40	20	20
3	31–39	20	10	-	80	40	40
4	37–45	40	20	5	-	80	80
5	48–52	80	40	10	-	-	-
6	55–60	-	80	20	-	-	-
7	60–63	-	-	40	-	-	-

Table 2: Descriptive statistics

	Non adherent					Adherent				
Obs:	411,695 (58.4%)					293,162 (41.6%)				
Stat.	Chol.	Female	Charlson	Age	Equipot.	Chol.	Female	Charlson	Age	Equipot.
Mean	0.592	0.561	0.920	66.2	4.167	0.583	0.494	1.039	65.5	4.555
SD	0.092	0.496	1.096	9.4	0.629	0.099	0.500	1.144	9.5	0.666
10th quant.	0.482	0.000	0.000	53.0	3.000	0.465	0.000	0.000	53.0	4.000
Median	0.594	1.000	1.000	67.0	4.000	0.585	0.000	1.000	67.0	5.000
90th quant.	0.694	1.000	2.000	78.0	5.000	0.692	1.000	3.000	77.0	5.000

Table 3: Comparison of the observable characteristics by potential falsification channel of the instrument: if the instrument is randomly assigned the distribution of observable characteristics should be similar (Altonji et al., 2005).

Characteristic	Full sample	Potential Falsification Channels			
		Split between GP and Patient		Drug payment Exemption	
		Yes	No	Yes	No
Cholesterol	0.588	0.593	0.588	0.587	0.588
Adherence	0.416	0.420	0.416	0.452	0.415
GP Communication	0.430	0.419	0.430	0.421	0.430
Initial chol.	0.621	0.625	0.621	0.616	0.621
Equipotency	4.328	4.303	4.328	4.338	4.328
Charlson 2	0.319	0.324	0.319	0.348	0.318
Charlson 3	0.158	0.127	0.159	0.170	0.158
Charlson 4	0.096	0.096	0.096	0.110	0.096
Simvastatin	0.314	0.285	0.315	0.306	0.315
Lovastatin	0.020	0.012	0.020	0.022	0.020
Pravastatin	0.101	0.154	0.101	0.104	0.101
Fluvastatin	0.065	0.071	0.065	0.068	0.065
Atorvastatin	0.310	0.339	0.310	0.299	0.310
Rosuvastatin	0.171	0.133	0.171	0.185	0.170
Simva eze.	0.019	0.006	0.019	0.016	0.019

Table 4: Distribution of Body Mass Index across instrument and treatment status, by gender

Instrument Level	Men		Women	
	Adherent			
	No	Yes	No	Yes
[0.05,0.30]	28.394	28.444	28.113	28.491
[0.30,0.45]	28.396	28.672	28.311	28.707
[0.45,0.60]	28.819	29.253	28.666	29.106
[0.60,0.75]	29.111	29.616	29.221	30.132
[0.75,0.95]	28.333	28.351	28.067	28.823

Table 5: Test of falsification for the instrument validity (Mourifié and Wan, 2016)

	Men			Women		
CI level	10	5	1	10	5	1
Year						
2004	NR	NR	NR	NR	NR	NR
2005	NR	NR	NR	NR	NR	NR
2006	NR	NR	NR	NR	NR	NR
2007	NR	NR	NR	NR	NR	NR
2008	NR	NR	NR	NR	NR	NR

Notes: For the parameter θ defined in Mourifié and Wan (2016), the hypotheses are $H_0 : \theta_0 \leq 0$, $H_1 : \theta_0 > 0$. 'R' stands for reject and 'NR' for NON reject. For the test, I set $Z = 1$ if $Z_i > \frac{1}{N-1} \sum_1^N Z_{j \neq i}$.

Table 6: Validation of the interpretation of the instrument; effect & S.E.

	Men		Women	
	Smoke	Sport	Smoke	Sport
Women	-0.008	0.008	-0.030 ***	0.010
	0.007	0.010	0.005	0.010

Notes: ***[**](*) denotes significance at the 1[5](10)% confidence level.

Table 7: Various tests of non-heterogeneity of the treatment effect.

H_0 : **NON**-hetero.	Covariates			
	Yes		No	
Carneiro et al. (2003): $y = x\beta + \gamma_1 p + \sum_{k=2}^K \gamma_k p^k$				
$\sum_{k=2}^K \gamma_k = 0$	34.072	***	26.247	***
Kowalski (2016): $y = \lambda_D D + \lambda_Z Z + \lambda_{DZ} DZ$				
$\lambda_D = 0$	197.623	***	1471.194	***
$\lambda_Z = 0$	52.647	***	326.545	***
$\lambda_{DZ} = 0$	64.36	***	28.564	***
$\lambda_D = \lambda_{DZ} = 0$	123.993	***	1051.402	***
Hausman (1978)	76.113	***	295.79	***

Notes: ***[**](*) denotes significance at the 1[5](10)% confidence level.

Table 8: Treatment effects; effect & S.E.

Sample	ATE	ATT	ATU	Sorting (ATT-ATE)	Sorting* (ATE-ATU)	Selection (OLS-ATT)
All	-0.02303	-0.04346	-0.01007	-0.02043	-0.01296	0.03094
	0.00376	0.00492	0.00606	0.00529	0.00335	0.00494

Note: Sorting is the difference between the ATT and the ATE, i.e. between $E(Y_1 - Y_0|D = 1)$ and $E(Y_1 - Y_0)$, a measure of the 'surplus' of the treatment effect enjoyed by patients who sort into adherence with respect to the treatment effect enjoyed by a randomly drawn patient. Sorting* is the same kind of difference, taken between the ATE and the ATU. Selection is the difference between the OLS and the ATT, i.e. between $E(Y_1|D = 1) - E(Y_0|D = 0)$ and $E(Y_1 - Y_0|D = 1)$: in this application, a positive difference suggests a comparative advantage from adherence to prescribed regime (i.e., positive selection). All these quantities are calculated using the procedure introduced by Carneiro et al. (2017).

Table 9: Policy relevant treatment effect (PRTE)

	Policy directed towards:		
Parameter	Physicians	Patients	Mix
PRTE	-0.02996	-0.03063	-0.02091
SE	0.00264	0.00272	0.00356

Description of characteristics of patients who change behaviour thanks to the policy

Chracteristic	Overall	Affected	Ratio	Affected	Ratio	Affected	Ratio
Charlson 2	0.31988	0.32480	1.01540	0.32594	1.01895	0.33526	1.04809
Charlson 3	0.16197	0.16595	1.02460	0.16652	1.02812	0.17616	1.08765
Charlson 4	0.09684	0.09911	1.02341	0.09976	1.03011	0.10973	1.13306
Hosp	0.03777	0.03876	1.02618	0.03847	1.01868	0.03911	1.03567
Hyper	0.68503	0.68979	1.00695	0.68897	1.00575	0.70451	1.02844
Diabetes	0.26281	0.27152	1.03314	0.27221	1.03578	0.29193	1.11082
Ictus	0.11171	0.11576	1.03623	0.11540	1.03305	0.12650	1.13242
Hearth fail	0.02764	0.02864	1.03613	0.02874	1.04007	0.03250	1.17591
By pass	0.00322	0.00339	1.05270	0.00340	1.05595	0.00427	1.32730
Ischemia	0.01317	0.01352	1.02623	0.01339	1.01664	0.01376	1.04426
CVD	0.01729	0.01727	0.99875	0.01742	1.00718	0.01755	1.01485
Atrial fibril	0.04562	0.04707	1.03170	0.04652	1.01970	0.04895	1.07296
Age 45 50	0.05212	0.05351	1.02650	0.05346	1.02571	0.05360	1.02830
Age 50 55	0.10185	0.10302	1.01148	0.10362	1.01741	0.10342	1.01538
Age 55 60	0.15486	0.15589	1.00667	0.15690	1.01317	0.15981	1.03195
Age 60 65	0.19969	0.19934	0.99827	0.19923	0.99769	0.19981	1.00062
Age 65 70	0.20957	0.20846	0.99467	0.20834	0.99410	0.20978	1.00101
Age 70 75	0.16920	0.16521	0.97639	0.16525	0.97667	0.16285	0.96245
Age 75 80	0.06891	0.06928	1.00543	0.06824	0.99033	0.06652	0.96534

Table 10: Cost effectiveness analysis

Component	**Policy**		**€**
$C_{TOT.}$ =	Physician	=	8.2 (=5.2+3.0)
C_{DRUGS}+C_{POLICY}	Patient	=	23.5 (= 2.5+21.0)
	Mix	=	29.9 (=5.9+3.0+21.0)
	Definition		**Relevant numbs.**
C_{DRUGS}	$\underbrace{\Delta P}_{switched} \times \underbrace{\overline{HC}}_{AVG.Hyper.cost}$	=	$2.5mil$.per year − $5.9mil$.per year
	Physician$^{\triangle}$	=	4.5%
ΔP	Patient$^{\triangle}$	=	2.2%
	Mix$^{\triangle}$	=	5.1%
$\overline{HC}$	$N_{PAT.} \times C_{GEN.DRUGS}$	=	115.9mil.per year
$N_{PAT.}$	$70\% \times POP[40, 80]$	=	22.5mil.
$C_{GEN.DRUGS}$	Official Journal of Italy	=	
C_{POLICY}	$C_{\{\text{Broadcast TV,On line courses}\}}$	=	{21,3.0} mil. per year

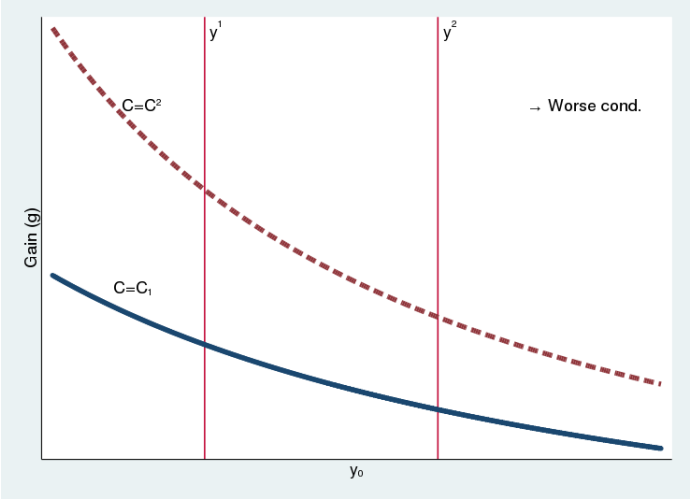
Notes: All the details to derive these amounts are in Appendix B. $^{\triangle}$: the share of patients induced into treatment is derived from Table 9.

Table 11: Treatment effects by age; effect & S.E.

Sample	ATE	ATT	ATU	Sorting (ATT-ATE)	Sorting* (ATE-ATU)	Selection (OLS-ATT)
Elders	-0.01025	-0.03355	0.00341	-0.02330	-0.01366	0.02096
	0.00758	0.00824	0.01099	0.00864	0.00505	0.00809
Middle	-0.03153	-0.05015	-0.01948	-0.01863	-0.01205	0.03781
	0.00699	0.00759	0.01123	0.00883	0.00570	0.00757
Young	-0.03102	-0.05129	-0.01725	-0.02028	-0.01377	0.03858
	0.00674	0.00921	0.01105	0.00957	0.00646	0.00913

Note: See Table 8.

Figure 1: A graphical representation of the economic model



The minimum required gain to sort into adherence is plotted as a function of $\frac{c}{y_0-c}$ (subscript i is omitted for notational convenience). In order to have a 2-dimensions plot, I fixed $c \in \{C_1, C_2\}$, with $C_1 < C_2$, and let health conditions under non-adherence (y_0) free to vary. For any fixed health condition under non-adherence (e.g., $y_0 = y^1$), the likelihood to be treated decreases as cost increases (the curve defined by C_2 is above the curve defined by C_1, hence the minimum required gain to self-select into treatment increases). For any fixed cost, as going from y^1 to y^2 (i.e., from better to worse conditions under non adherence), the likelihood to be treated increases for both levels of costs (the curves are decreasing in the level of cholesterol, hence the minimum required gain decreases).

Figure 2: Histogram of the predicted probability of compliance

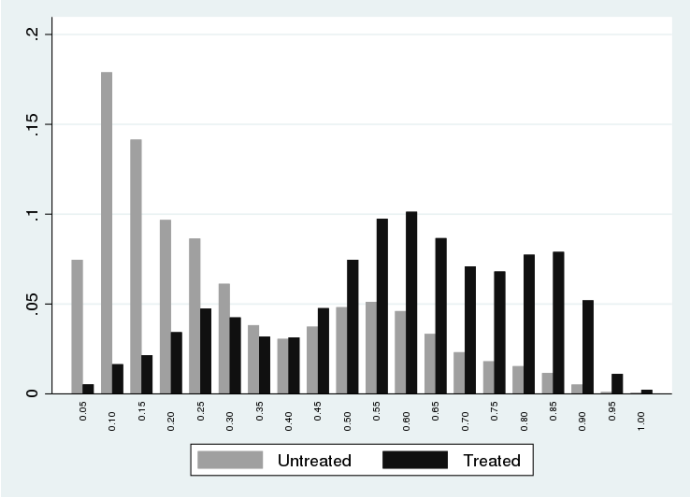
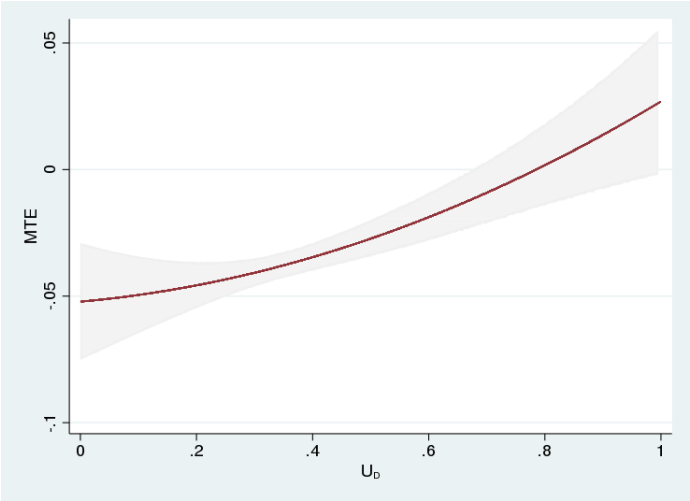


Figure 3: MTE of pills to cholesterol



The inference is based on bootstrap approximation as suggested by Carneiro et al. (2017) and 250 draws.

A Characterization of patients induced to change behaviour thanks to the hypothetical policy

In this section I show how to describe the observable characteristics of the patients induced into treatment in response to the policy . To the best of my knowledge this is a technical contribution of my paper to the existing literature. The only example that goes in this direction and that I am aware of is Carneiro et al., 2003. The same question may be relevant to explain the results from MTE: examples are Carneiro et al. (2003); Basu (2014). Thanks to the results in Vytlacil (2002) the description of the characteristics of the patients who change behaviour can be accomplished exploiting existing results available for the characterization of compliers in the LATE context (Angrist, 2004; Angrist and Pischke, 2008). For a characteristic X that can be described by dummy variables (i.e., X is Bernoulli distributed), the theoretical justification for the exercise is based on the equivalence (Angrist and Pischke, 2008):

$$F[X, \Delta P] = F[X|\Delta P]\, F[\Delta P] = F[\Delta P|X]\, F[X] \tag{10}$$

$$E[X|\Delta P] = \frac{E[\Delta P|X]\, E[X]}{E[\Delta P]}, \forall u_D. \tag{11}$$

Although one may be interested in measuring the observable characteristics at specific u_D, I summarize it over all the support of the selection probability, i.e. $\int E[X|\Delta P, u_d] d\, u_d$. Similar to the case of LATE, a limitation of this exercise is that I can describe the average characteristics of the switchers, but I am not able to flag who is actually switched from non-treatment to treatment.

To characterize the patients using the procedure by Carneiro et al. (2017), I only need to derive the sample of patients that are switched into treatment by the simulated policy. Once I have this 'new' sample of individuals, I simply evaluate the sample averages of their characteristics obtaining $E[X|\Delta P, u_d]$. In order to compare the characteristics of the shifted patients to those of the entire population, I also evaluate $E[X]$: for example, if X is diabetes then a ratio $\frac{E[X|\Delta P, u_d]}{E[X]} < 1$ implies that the share of patients that suffers from diabetes and are induced into treatment by the policy is lower than in the overall population (and viceversa if the same ratio is greater than 1).

B Deriving Cost-Effectiveness

In this appendix I describe how I derive the Total Net Benefit (TNB) from each policy, defined as the difference between benefits (B) and costs (C), i.e. $TNB = B - C$. Here I describe step-by-step how I calculated the total costs. The benefits are described directly in the main text.

The exercise may be conducted at different levels of detail. Coyle et al. (2003) show that the finer the stratification the greater the opportunity for positive net benefits, because it makes it possible to exclude from treatment patients whose benefits would be lower than costs. In this paper the costs are evaluated at individual level. This is an important feature of the calculation for two reasons. First, on the empirical side this guarantees a unifying approach with the previous empirical analysis. Second, it better serves the scope of the Comparative Effectiveness Research (CER) to improve the *individualized* patient care (Institute of Medicine, 2009; Basu et al., 2014).

Total costs ($C_{TOT.}$) are the sum of two components: the additional cost of drugs for patients non-adherent under the current policy, but otherwise adherent (C_{DRUGS}); the cost of actual implementation of the policy (C_{POLICY}), either advertisement or courses of communication. Cost C_{DRUGS} is relevant under the very conservative assumption that drugs are given directly to the patients by the Government. This is the relevant cost if all the switchers are entitled to exemption from payment of drugs. This choice is made so to obtain the upper bound of the cost of the policy. The component C_{POLICY} is the cost due to the actual implementation of the policy.

The additional cost of drugs is obtained as $C_{DRUGS} = \Delta P \times \overline{HC}$, where ΔP is the share of patients switched into treatment thanks to the policy, derived directly from the PRTE of Section 6; $\overline{HC}$ is the total cost of drugs for hypecholesterolemia, defined as $\overline{HC} = N_{PAT.} \times C_{GEN.DRUGS}$. In this notation, $N_{PAT.}$ is the number of patients suffering from hypercholesterolemia and is obtained applying the percentage of 70% (Progetto Cuore, 2018) to the Italian population between 40 and 80 years old ($32.2mil.$ individuals), so that $N_{PAT.} \approx 22.5mil.$; $C_{GEN.DRUGS}$ is the actual expenditure of the Government for generic drugs, that may be derived using the official prices set by the Italian authority for drugs (AIFA, 2017). Given the observed composition of statin intake, it follows that $\overline{HC} = N_{PAT.} \times C_{GEN.DRUGS} = 115.9mil.$per year.[18] The policy directed towards

[18] This is the step where the advantages from using a personalized version of CER are the largest with respect to a

patients would cost $C_{DRUGS} = 2.5mil.$per year €; the policy directed towards physicians would cost $C_{DRUGS} = 5.2mil.$per year €; the policy mix would cost $C_{DRUGS} = 5.9mil.$per year €.

The cost of advertisement (C_{POLICY}) depends on the specific policy. For policy directed towards patients, I consider the advertisement by TV and radio; for policy directed towards physicians, I consider on-line communication courses. In both cases, there are only fix costs, whilst the marginal cost per patient is zero. It is likely that both policies exert positive externalities (e.g., there is no reason that a better communication of the GP should be confined to a specific disease). With the data at hand, these externalities can not be considered. In this respect, the costs estimated are an upper bound of the true costs.

As for TV, I downloaded (26th sept 2018) the prices on the main Italian public broadcast (RAI): on average, the cost is 30,000 €, each 15 seconds. As for the radio the prices are much cheaper: on average, the cost is at most 5,000 euro, each 30 seconds, during peak time. I implement a policy where advertisements run once per hours, 10 hours a day, for 60 days, for a total cost of about $C_{POLICY} = 21\, mil.$ €. Concerning on-line communication courses, I found courses between 15 and 60 €(data as for 26th sept 2018) per person. According to official statistics the number of GP during the period of this analysis were approximately 46,500 (44,279 in 2016, the latest available figure according to data from Eurostat), the total cost for the policy directed towards physicians is between $C_{POLICY} = 0.7$ $(= 46,000 \times 15)$ and 3.0 mil. €.

Therefore, the total cost for the implementation of the policies ($C_{TOT.} = C_{DRUGS} + C_{POLICY}$) are about $C_{TOT.} \approx 25\, mil.$€ when the target is the patient, $C_{TOT.} < 10\, mil.$€ when the target is the physician, or $C_{TOT.} \approx 30\, mil.$€ when a policy mix is implemented (under the assumption that it implements both a course to GPs and advertisement).

C Further results

standard CER based on average effectiveness. The age is a typical risk–factor used to build subgroups in a standard CER analysis. Whilst one can establish the drug composition -and the attached costs- for patients in a given age range, there may be many other factors that determine individual choices within the same age range. Therefore, in this paper I do not condition on specific risk factors, selected a priori, and I run the analysis at individual level. See Basu et al. (2014, p.363) for a more thorough discussion on this issue.

Table 12: Detail of the estimates

Variables	OLS	IV	MTE	
			Selection	Main eq.
Init	0.536 ***	0.530 ***	-1.236 ***	0.602 ***
Init sq	-0.129 ***	-0.125 ***	0.591 ***	-0.179 ***
Equip	-0.005 **	0.015 ***	5.472 ***	-0.037 ***
Equip sq	0.001 ***	-0.000	-0.351 ***	0.006 ***
Charlson 2	-0.003 ***	-0.002 ***	0.143 ***	-0.001 **
Charlson 3	-0.003 ***	-0.002 ***	0.189 ***	-0.002 **
Charlson 4	-0.001 **	-0.001	0.216 ***	0.002 **
Hosp.	0.001 **	0.001 *	-0.099 ***	0.001
Hyper	-0.007 ***	-0.007 ***	0.122 ***	-0.006 ***
Diab	-0.018 ***	-0.017 ***	0.048 ***	-0.016 ***
Ictus	-0.004 ***	-0.003 ***	0.141 ***	-0.003 ***
BMI	-0.006 ***	-0.006 ***	0.040 ***	-0.003 ***
Hearth disfunct.	-0.000	0.000	0.127 ***	0.000
ByPass	0.001	0.003	0.271 ***	-0.009 *
Ischemia	0.001	0.000	-0.125 ***	-0.004 **
Vascular Dis.	0.002 **	0.002 **	-0.052 **	0.004 **
Atrial Fibr.	-0.004 ***	-0.004 ***	0.073 ***	-0.003 ***
Simvastatin	-0.008	-0.024 **	-3.865 ***	-0.032
Lovastatin	0.010	0.001	-3.183 ***	-0.019
Pravastatin	0.006	0.000	-1.808 ***	-0.021
Fluvastatin	0.007	0.004	-1.119 ***	-0.015
Atorvastatin	-0.015	-0.035 ***	-4.927 ***	-0.043
Rosuvastatin	-0.020 *	-0.039 ***	-4.785 ***	-0.041
Simva eze	-0.010	-0.007	0.454	-0.040
Female	0.002 ***	0.001 ***	-0.229 ***	-0.001
Eta 2	-0.007 ***	-0.006 ***	0.060 ***	-0.003 **
Eta 3	-0.009 ***	-0.009 ***	0.072 ***	-0.010 ***
Eta 4	-0.013 ***	-0.013 ***	0.141 ***	-0.017 ***
Eta 5	-0.015 ***	-0.015 ***	0.123 ***	-0.019 ***
Eta 6	-0.019 ***	-0.019 ***	0.111 ***	-0.022 ***
Eta 7	-0.019 ***	-0.019 ***	0.073 ***	-0.023 ***
Eta 8	-0.021 ***	-0.020 ***	0.082 ***	-0.025 ***
Q 1	0.003 ***	0.003 ***	-0.085 ***	0.002 ***
Q 2	0.002 ***	0.002 ***	-0.086 ***	0.002 ***
Q 3	0.000	0.001 *	-0.001	0.000
Y 2005	-0.006 ***	-0.004 ***	0.419 ***	-0.004 ***
Y 2006	-0.008 ***	-0.005 ***	0.641 ***	-0.004 ***
Y 2007	-0.009 ***	-0.006 ***	0.618 ***	-0.004 ***
Y 2008	-0.011 ***	-0.008 ***	0.616 ***	-0.004 ***
NW	-0.005 ***	-0.006 ***	-0.136 ***	-0.004 ***
NE	-0.009 ***	-0.009 ***	-0.059 ***	-0.004 ***

C	-0.008 ***	-0.009 ***	-0.192 ***	-0.006 ***
S	-0.007 ***	-0.008 ***	-0.052 ***	-0.005 ***
Compliance	-0.013 ***	-0.035 ***		
Others' compliance			2.692 ***	
P				-0.093
P^2				0.010
P^3				0.020
Init × P				-0.159 ***
Init sq × P				0.117 ***
Equip × P				0.037 ***
Equip sq × P				-0.005 ***
Charlson 2 × P				-0.002 **
Charlson 3 × P				-0.002
Charlson 4 × P				-0.007 ***
Hosp. × P				0.000
Hyper × P				-0.003 ***
Diab × P				-0.004 ***
Ictus × P				-0.000
BMI × P				-0.006 ***
Hearth disfunct. × P				-0.000
ByPass × P				0.021 ***
Ischemia × P				0.010 **
Vascular Dis. × P				-0.005
Atrial Fibr. × P				-0.002
Simvastatin × P				0.031
Lovastatin × P				0.062
Pravastatin × P				0.051
Fluvastatin × P				0.038
Atorvastatin × P				0.032
Rosuvastatin × P				0.022
Simva eze × P				0.061
Female × P				0.005 ***
Eta 2 × P				-0.007 **
Eta 3 × P				0.004 *
Eta 4 × P				0.011 ***
Eta 5 × P				0.012 ***
Eta 6 × P				0.009 ***
Eta 7 × P				0.011 ***
Eta 8 × P				0.011 ***
Q 1 × P				0.001
Q 2 × P				0.001
Q 3 × P				0.001
Y 2005 × P				-0.003 *
Y 2006 × P				-0.006 ***
Y 2007 × P				-0.009 ***

Y 2008 × P				-0.014***
NW × P				-0.003
NE × P				-0.011***
C × P				-0.005***
S × P				-0.004**
Intercept	0.360***	0.321***	-14.922***	0.417***

Notes: ***[**](*) denotes significance at the 1[5](10)% confidence level.

Table 13: Distribution of weights

U_D	Doctor	Patient	Mix
(0,.1]	0.00760	0.00747	0.00762
(0.1,0.2]	0.02379	0.02113	0.02372
(0.2,0.3]	0.02597	0.02134	0.02808
(0.3,0.4]	0.01788	0.01630	0.02473
(0.5,0.6]	0.03309	0.03066	0.05310
(0.6,0.7]	0.03838	0.03104	0.07207
(0.7,0.8]	0.02132	0.01661	0.05595
(0.8,0.9]	0.01276	0.00984	0.05742
(0.9,1)	0.00101	0.00062	0.02679

Figure 4: Comparison of semi-parametric and non-parametric estimators.