Archive number:

SCHOOL OF ECONOMICS AND MANAGEMENT
Lund University

Department of Economics
Lund University
Spring 2004
Bachelor thesis in Health Economics, 10 points

Review of models for evaluation of treatment effects based on nonexperimental data

Student: Aida Spahović

Mentors: Carl Hampus Lyttkens, Ulf-G Gerdtham and Curt Wells

Contents ***Page***

Abstract

In order to determine the true effects of new medicines, new medical approaches and health care policy, one has to resolve problem of selection faced in many econometric studies. This paper investigates the models that take account of this problem in nonexperimental data. This data, contra to experimental data sets, exhibit both some special characteristics and possibilities. Starting from the data and possible problems the scientist are meeting, I illustrate the feasible econometric solution to selection problems. Heckman's, matching, instrumental-variables and regression discontinuity estimators are discussed as cross-sectional estimators group. Before-after and difference-in-differences estimators are examined as part of longitudinal methods. Results show that these models are capable of addressing the selection problem.

Key words: selection problem, nonexperimental data, treatment effects and econometrics.

Introduction

Evaluations of health economics' issues are the assessment of efficacy, safety and cost-effectiveness for both new and established medical therapies. Health economics is commonly incorporated in the public policy making decisions process where attention on concerns about efficiency in health care has sharply increased lately. Health care is predominantly financed with public money. As there will always be more treatment options than the resources will allow, economists need right results from evaluation of available data in order to make appropriate decisions. New improvements in medicine whose goal is improvement of the welfare are invented every day. As we have to deal with scarce economic resources, beside evaluation of the effects, cost-benefits analysis plays an important role. In an ideal situation where resources were unlimited we could concentrate our consideration on only efficiency of the new inventions.

The range of models for evaluation of treatment effects in health economics has increased largely over the last years since the evaluation of the new medicine became a hot issue. Health economic evaluation is a tool to assess the benefits and the costs of different uses of scarce sources. It provides a comparative analysis of alternative courses of action in terms of their costs and consequences. One pharmaceutical product can be compared with itself, another product, another treatment as a surgery or with a watchful waiting approach whereby a patient receives no form of medication but any change in health status is checked (control group). As an example of a health economic validation I assume to have an access to data set of repeated observation on the same persons before and after treatment that comes from ill individuals with disease X. The cure for these patients is medicine A. Patients that receive this medicine belong to treatment group and those who do not receive any medicine belong to control group. With this data it is possible to calculate mean survival rate. If it exists a significant difference between patients who have taken the medicine and these who have not taken the medicine it becomes possible to valuate the efficiency of the drug. Still there are some problems that economist face.

The results that I observe come from the data on sick people when they choose to take the drug and become a part of the treatment group. The scientific strategy used for estimation of outcomes from this medical treatment is ordinary least square estimates (OLS) that estimates a general drug effect. This strategy fails because the chosen sample is not random and

achieved results overestimate drug effect. Differences in characteristics between the patients that have taken the medicine are not accounted for in evaluation of the problem. The data one observes is selected because individuals by themselves make the choice of proceed the treatment or nor. The question is why certain people chose the treatment. The answer may be grade of illness, education level, or something else by systematic process for which is not account. (Hill et al. (2001)) This phenomenon in economics leads to biased results and is called **selection problem**.

The general conclusion about cost-benefits of this treatment is even more difficult to make. Before treatment with pharmaceuticals was applied, this illness (e.g. Stomach Ulcer) was cured with surgery that caused additional cost for surgical approach itself, cost for sick leave before and after the operation as well as cost for special treatment for a post surgical treatment of the person. All of these are costly measurements for society. Beside that, one such surgery is always high-risk performance. That is why it can be wrong to say that today's pharmaceutical treatment is too expensive and not cost-efficient if we totally ignore the surgical treatment that was applied in the past and do not take its cost as a comparing reference even though the bills for anti ulcer pharmaceuticals are taking the greater part of total cost of pharmaceuticals. Further problem may arise if we consider unknown part of all (ulcer)ill people that, for some reason, do not seek medical help and are not included in any data. There is possibility that backgrounds factors influence people to (not)consume health and these factors may even be correlated and affect individual decision over amount of health consumption. Some of the reasons for ignorance of help seeking may be influenced by (in)appropriated results from statistics that come to individual knowledge.

Data type

Generally economics is the science where data on all variables are observed and not obtain from a controlled experiment. Using observational data on real-world population, economists analyse the relationship among provider characteristics, treatment and health outcomes. (Gould (1994)) In order to reach desired results they make some assumptions for different evaluations that might give biases for final decisions. With use of econometric methods it is possible to calculate the effects of some of these and similar social measures but because of assumptions made we still cannot know for sure what health improvements actually can be achieved, if any. The problem in principle is to find "the right" way of evaluation of these issues. The econometric valuations are based on two main groups of data type –

nonexperimental and experimental. For this review I also make distinction between cross-section and longitudinal data. The first is collected over discrete intervals of time while the last follow individuals over time. These types will be further studied in chapters where the respective econometrics models are analysed.

Experimental data

Experimental data comes from clinical trials and is randomised. Patients are selected for the study according to strict criteria and randomised in two groups. Patients from the first group get a pharmaceutical A (treatment group) while patients from the second group do not get any pharmaceutical (control group). Overall members of these two groups are the same. In this way the potential bias of patient selection on the outcomes will be eliminated but there will still be some unsolved problems. Factors other than the treatment may affect patient outcomes. One is the condition that is not really identical from one trial to another. There will be the errors in measurement and some factors that include preferences of patient, patient health status etc will be uncontrolled. This type of variation is called sampling variation or sample selection but may be accounted for in regression. (Maddala (1986))

The results on e.g. survival rate from the experiment above show a difference among the patients of the same group and indicate the benefits or lack of benefits from a specific, well-defined treatment's intervention. If the medicine is taken according to prescription from the physician, the results are better than those that come from a patient who did not follow the physician's advice. The results taken in economic valuations are only those that come from the patients who followed the rules. In that way the measured effect of the medicine in this controlled experiment will be the results from a perfect world.

In all evaluation models it is assumed that mean values from the treatment and control groups generated by random assignment give desired effect. This assumption does not take account of changes in the impact of participation due to the presence of random assignment or changes in the process of program participation. (Heckman et al. (1999)) Further problems will be that selection of patients for the treatment intervention may not be the same as the strict patient selection criteria in the randomised trials examining that intervention. Instead it is often modified or used with other treatments that were not tested in the randomised trial. (Gould (1994))

Using the comparisons of the means from the treatment and the control group cannot identify attendance of the dropping out from experiment. In order to obtain an estimate of the impact of treatment on those who actually receive it, an additional latent variable for the control group that indicates whether or not they receive the treatment will be required. This and other possible solutions to the problem with experimental data are beyond the scope of the paper and will not be further discussed here.

Nonexperimental data

Nonexperimental data values are, like for the experimental data unknown until we observe them. The values of its random variables are however not under complete control of any person but are affected by the factors that cannot be controlled. This means that nonexperimental data is not random and continuous non-observable factors that cause selection problem. Selection problem leads to issue of questioning the results that come from evaluation of scientific experiments, which vary often continuo positive results. These are some examples of situations where selections problem is important. The experimental results may for instance suggest that subvention of the preventive drugs and subvention of an effective break on AIDS are efficient, or that the treatment effects of new chemical entities have successful results that assure effective improvement of welfare. The question is if we can accept evaluation results as reliable. The original task lies in data type that can lead to different problems where selection is one of them. The best summary for above said is words of Gisela Kobelt, managing director of Health Dynamics International Ltd: *"An economic evaluation can only be as good as the underlying effectiveness data."* (Kobelt 2002, page 27)

There are couple of econometric methods that deal with evaluation of nonexperimental data. Selection problems that bias results from study of treatment effects can be solved with two main groups of econometric methods – cross-sectional and longitudinal model. Heckman's, matching, instrumental variables (IVs) and regression discontinuity estimators belongs to group of cross-sectional estimators. Before-after and difference-in-differences are longitudinal estimators.

Summary of the introduction

The aim of this paper is to examine methods to eliminate selection problems in the sampling data that cause selection bias. A possible solution can be to use randomised trials rather than observational studies to estimate treatment effects but these trials are not always feasible or it

can be extremely costly. The use of observational data is often an exception and most applied studies rely on data from secondary sources. This research attempts to assess the economic discussion of different economic approach to the selection problem in observational studies of treatment effects.

This paper is organized as follows. First I describe methods applied and material used. Then I formalize the evaluation problem and the parameter of interest in evaluating nonexperimental data. Here I describe selection models as well as treatment effects and give needed limitation of this paper. Afterwards I report results from reviewed material. In two following sections I provide an analysis of cross-section and longitudinal methods. I examine matching, Heckman's, IVs and regression discontinuity estimators as well as before-after and difference-in-differences estimators. I focus on their main application the problem that they deal with. Finally I give conclusion and suggestions for further research.

Material and method

The central goal of many researchers in the health economics is to identify the impact of some treatment by analysing its outcome. The problem that very often appears is the identification of the effects of the treatment. Almost all available data is contaminated with some kind of selection bias. In order to review econometrics techniques that take account of selection problem I studied two groups of literature. The first was handbooks of econometrics and health economics where different econometric models were studied in detail. This was initially intended to be the main source of literature reviewed. To gain the further inside into the applied methods the second group of additional papers from available literature about application of econometric models were also included.

The papers were obtained from a reading list for post graduate course "Econometric evaluate of labour programmes" given by Professors M. Lechner and J. Smith in Uppsala, Sweden, August 11-15, 2003. As a complement to literature list Electronic Library Information Navigator (ELIN) database was searched using economic terms as a selection bias, treatment effect or the name of authors (e.g. Heckman) whose articles I read before. A total of approximately fifteen economic studies were examined. All articles in general have some relationship with sample selection and / or treatment effects. They were read with caution. The literature studied led to deeper examination of methods that fulfil criteria for sample selection or treatment effects. Models as Heckman and IVs were discussed actively. Matching method is investigated as well as regression discontinuity estimators and longitudinal methods.

Method used for this work is simple reviewing of found literature. No date or language limits were applied though the fact that the search terms were in English implies some limitation. The whole study is meant to be a handbook for the econometric study in future of different data types in health economics.

Formalisation of the evaluation problem and parameter of interest in evaluating nonexperimental data

This paper examines the evidence on the effectiveness of health care improvements such as treatment effects with new medical drugs, and the methods used to obtain these effects. An econometric evaluation makes a comparison between treated and untreated persons. The data on potential outcome from treated person is available and paired with potential outcomes from comparable person who has not participated in the treatment. The outcome of this non-participant is treated as proxy for potential outcome from untreated state of the treated individual. The population mean difference between treated and untreated person is then calculated by averaging over all pairs. For that problem, the challenge is to find mean value of potential outcome from treatment for non-participant. (Heckman et al. (1999))

Treatment effects in nonexperimental data

The model of selectivity has been applied to measurement of program effectiveness. In this model the data on total population is available but there are one or more regressors of interest that take different values as a result of some type of selection process. (Heckman et al. (1999), page 1930) Value of program effectiveness may be measured as value of drug strength that participant administrate in the program which is then called treatment effect. An effectiveness function that account for the value is:

$$Y = \beta X + \delta C + \varepsilon \qquad (1)$$

where C is a dummy variable indicated whether or not the individual is taking the drug (i.e. individual is under treatment). Observations are available from G different persons that both take and do not take the medicine. Individuals that take a drug belong to group 1 while these who do not take a drug belong to group 0. The goal is to determine the effect of membership in the group 1 versus group 0 on a continuous outcome variable Y. Membership in the takers group implies treatment and the aim is to measure the treatment effect. The question is if δ measures the value of drug strength (assuming that the rest of the regression model is correctly specified). The answer is no if the typical individual who chooses to take a medicine would become healthy whether or not he or she take the drug. This is an example of the problem of self-selection. In this case any estimation by using the OLS estimates of δ will overestimate the treatment effect. (Greene (1999), page 981-982)

Heckman (1979) was first to point that there are variables that affect selection of the individuals into the treatment but do not affect expected behaviour for each individual once that person is a group member. The effects of shifting a person with individual characteristics X from group 0 to 1 gives ΔY. Heckman claims that increase in selection variables never decreases the expected outcome for an individual within a given group. An increase in selection variables may have an effect on the outcome variable Y that is defined as an endogenous interaction effect. (Brock et al. (2001), page 3343-3345)

First to examine is the general selection model. After giving the brief insight into this problem, I continuo by explaining two different selection models, traditional and continuous, as two basic examples of methods used in field of health economics.

Selection models

Constructing counterfactuals is the central problem in evaluation strategies. Persons involved in study are imagined as being able to occupy either treated "1" state or untreated "0" state. D = 1 denotes treatment (participation in the program being evaluated) and D = 0 denotes non-treatment. Potential outcomes from Y_1 and Y_0 correspondent to treated and untreated states and they are a function of conditioning variable X. (Heckman et al. (1999), page 1877)

General for this paper is that outcomes from each state are:

$$Y_1 = \mu_1(X, D) + U_1 \qquad (2)$$

$$Y_0 = \mu_0(X, D) + U_0 \qquad (3)$$

where it is assumed that $E(U_1 \mid X) = 0$ and $E(U_0 \mid X) = 0$. However it is not said that $E(U_1\text{-}U_0 \mid X, D=1) = 0$ because D may be dependent on U_1, U_0 or $U_1 - U_0$ and X.

The gain to individual of moving from untreated to treated state is $\Delta = Y_1 - Y_0$. It is required that realization of D does not determine X given the potential outcomes. Otherwise the parameter $E(Y_1 - Y_0 \mid X, D = 1)$ does not capture full effect of treatment on treated. (Heckman et al. (1999), page 1879)

Fundamental evaluation problem is that neither Y_1 nor Y_0 are possible to observe from the same person and hence Δ for anybody is not known. All approaches for solving this problem attempt to estimate the missing data, i.e. to make assumption about unknown factors and how those are related to the available data.

According to Lewis (1963), quoted by Heckman et al. (1999), impact of the program is not same for all participants. There are two different effects: the direct that affects these explicitly enrolled in the program, and the indirect that does not follow from direct participation. Further different persons may value the same state differently which shows that preferences may be state dependent.

In order to find out the impact of the program especially for those exhibiting the problems above, one has to know distribution of program gains. That is why it becomes of interest to know the proportion (Pr) of people taking the program who benefits from it:

$$\Pr (Y_1 > Y_0 \mid D = 1) \qquad (4)$$

Consequently the major interest of this problem is evaluation of the available data on persons who are in the treated or the untreated state available at any time. For certain time periods there are some persons in both states but there is no information on any single person who is in both states at the same time. If indirect effect are negligible, the no treatment state might be approximated with the no program state. (Heckman et al. (1999), page 1882)

Evaluation of the treatment effect is based on estimation of the mean of the group who takes the treatment. Estimation of the treatment effect $E(\Delta \mid X, D = 1)$ measures the average gain in the outcome for person who chose to participate in a program compared to what they would experienced in the base state. As long as $var(U_1) = var(U_0) = \sigma^2$ among people with the same X, there is no heterogeneity in gains moving from "0" to "1". This assumption is strong and almost always rejected because there the agents selected into state "1" from "0" that do act on $Y_1 - Y_0$ in making their decision to participate in the program. (Heckman et al. (1999), page 1885)

Selection problems that bias research with nonexperimental data come from differences in outcomes among patients who receive the treatments. These differences are based on

correlation between D and U. Then factors are not observed by the analyst and gives biased results in estimation of the treatment effects. I go further with giving the main characteristics of a traditional selection model.

Traditional selection model

The traditional econometric method for selection problem allows for selection into the program on the basis of unobserved components of outcomes. With this method it becomes possible to estimate the effects of the treatment on the treated. (Heckman et al. (1999), page 1956) Selection bias that arises from using a nonexperimental comparison group is defined as follows:

$$B(X) = \text{var}(U_0 \mid X, D = 1) - \text{var}(U_0 \mid X, D = 0) \quad (5)$$

The standard evolution goal is to eliminate this bias based on differences between unobservables among participants (D=1) and not participants (D=0) given the same X, and not correlation between (U_0, U_1) and X and D.

Sample selection bias equation applied to non-participants and participants respectively are:

$$E(Y_0 \mid X, D = 0) \neq E(Y_0), \quad (6)$$

$$E(Y_1 \mid X, D = 1) \neq E(Y_1). \quad (7)$$

Selection bias strategies estimate mean values of difference between potential outcomes from treated and untreated states. (Heckman et al. (1999), page 1958-1959)

Next to examine is continuous selection model where outcomes, in contrast to traditional selection model with discrete variables in outcome, are continuous variables.

Continuous selection model

Continuous selection model consists of two equations. The first equation is an outcome equation with a continuously distributed dependent variable (drug effect). The estimated outcome gives the effects of the treatment. The second equation is a dichotomous choice equation (choice of taking medicine). Here probability that chosen people take medicine is

estimated. The dependent variable is observed or not depending of the choice decision. (Geweke et al. (2001), page 3530)

Equation of that determines sample selection is:

$$I^* = Z\delta + \upsilon \qquad (8)$$

The equation of primary interested is following:

$$Y = \beta X + \varepsilon \qquad (9)$$

The sample rule is observed only when I^* is greater than zero. Further it is assumed that ε and υ have bivariate normal distribution with zero means and correlation ρ. In the most cases selection variable I^* is not observed but one can observe its sign. The model formulated is described deeper in Heckman's method section. (Dhrymes (1986), page 1604 and Greene (1993), page 977-978) These two selection models give a fundamental concept of econometric theory that I will use through the paper.

Econometric is used in health economics for the evaluation of available data that requires making some comparison between treated and untreated persons in order to find out why certain persons are in a certain state. It may be done in many ways. For this work I chose to study methods that deal with comparison data at a point in time as in the cross-section, between same persons before and after treatment as in the before-after estimator or as a mixture of two principles as in the difference-in-differences estimator. Two last estimators belong to longitudinal method group. This group of estimators is more robust than cross-section selection bias correction methods. However these estimators depend on functional form assumptions. (Heckman et al. (1999), page 1942)

These are only a part of all existing methods and a limitation for this thesis is needed. Next section describes undertaken restrictions.

Limitations

Working with nonexperimental data implies selection problems. There are several methods and models that can properly fit the available data and can be useful in this case. Some of

them are Multiple indicators-multiple causes (MIMIC) model that deals with unobservable (latent) variables, maximum likelihood (ML) estimation that is used for nonlinear models involving qualitative or limited dependent variables and pseudo maximum likelihood (PML) methods. Often used to purge the bias is control function in form of Heckit estimator. Sample selection and hurdle models deal with the problem of limited dependent variables. (Griliches (1986), page 1486)

Analysis of all of them is beyond the scope of this paper. I will restrict my studies to examination of matching, Heckman's model, IVs method and regression discontinuity estimators as good representatives of cross-sectional models. Further longitudinal models represented with before-after and difference-in-differences estimators will be estimated. In the next chapter I will analyse these chosen models and give further insight of their possible applications in economics.

Econometric approach to treatment effects in nonexperimental data

The topic on sample selection has been a popular subject of recent literature. Selection problems are caused by the existence of unobservable individual effects that are often likely to be correlated with observed explanatory variables. (Jones et al. (2002)) In some cases the selection problem arises from the dependence between unobservables in non-treated state and the treatment. Selection may appear also when data is missing for some reason of self-selection. This "behavioural missing" (Griliches 1986) can give efficiency losses and bias in the estimated coefficients of models that do not take this behavioural missing into account. Same data set will allow modelling difference in behaviour across individuals and there are two frameworks for analysing them - the fixed effect approach and the random effects approach. The fixed model is analysis of the effects present on the observed sample. The random effects model assumes that individual effects are not correlated with other regressors and therefore may be inconsistent with omitted variables. More about these frameworks can be found elsewhere. (Greene (1993), page 632-633)

Different method for nonexperimental data

This section will explain different methods for nonexperimental data in more detail. First part of the section considers cross-sectional methods with method of matching, Heckman's method, instrumental variables and regression discontinuity estimator in focus. The second part deal with longitudinal methods where before-after and difference-in differences estimators are the most important. I will start with brief introduction to cross-section methods.

Cross-section methods

Cross-section methods have to do with the programs where participation is voluntary. For this case distribution of outcomes between participants and non-participants can be construct. The approximation of distribution of potential outcomes from an untreated state for untreated persons with the distribution of potential outcomes from the untreated state for treated persons gives a risk for selection bias:

$$B(X) = E(Y_0 \mid X, D = 1) - E(Y_0 \mid X, D = 0) \qquad (10)$$

This bias arises from selective differences in uncontrolled variables. Many methods have been proposed to avoid the bias. Differential outcomes due to selection bias can be eliminated by

method of matching, Heckman's method, IVs estimation and regression discontinuity estimators. These estimators are frequently used because they are quite straightforward and easy to understand. (Heckman et al. (1999), page 1941-1956)

General for the cross-section estimator is that it compares mean outcomes of participants and non-participants at time t. Its disadvantage is that it does not compare the same person in both states and cannot estimate the distributions of gains. In order to eliminate the bias (equation (10)) some assumption must be taken. The key assumption is that the person who does not participate in the program has the same no-treatment outcomes as those who do participate. Then the cross-section estimator is given by difference in following means:

$$(\overline{Y}_{1t})_1 - (\overline{Y}_{0t})_0 \qquad (11)$$

where $\overline{Y}_{1t}$ stands for post-program earnings of a person who participate in the program, $\overline{Y}_{0t}$ stands for post-program earnings of a person who do not participate in the program, the subscript "1" denotes conditioning on D=1 and bar denotes sample means.

This estimator is valid if participation in the program is unrelated to outcomes in the no-program state in the post-program period. Moreover it is robust as long as underlying macro economic and aging factors of involved persons operate identically on participant and non-participants. (Heckman et al. (1999), page 1898)

There are four different cross-sectional methods: matching, Heckman's, IVs method and regression discontinuity estimators. All of them are based on comparison between treated and untreated persons at a point of time. Following sections are describing them separately.

Method of matching

Matching method is applied on cross-sectional data. It is based on the idea of having a group of comparable persons for each person *i* in treatment group. The same individual may be in both groups if she is treated at one time and untreated at another. Outcomes from the treatment group are matched to the outcomes from the comparison group in order to estimate a treatment effect. Here one makes the assumption that selection into a program does not occur on the basis of unobservable gains from program. Instead X in the selection bias

(equation (10)) is set of conditioning variables and it does not influence the potential outcome from treated and untreated states. This allows for measure of the effect of treatment among participants if they have not participated by using potential outcomes of non-participants. (Heckman et al. (1999), page 1950)

Further one has to assume that

$$0 < \Pr(D = 1 \mid X) < 1 \qquad (12)$$

e.g. there are participants and non-participants for each X which allows for making comparison. Failure to satisfy this condition is major reason why the matching method may produce biased estimates. (Heckman et al. (1999), page 1951)

A comparison group in the matching method is conditioned on X. The distribution of the counterfactual outcome, Y_1, for the participants is the same as the observed distribution of outcome Y_0 for the comparison group. This implies that means of potential outcome from an untreated state for a treated person is same as the mean of the potential outcome from an untreated state for a non-treated person. At the same time the mean of the potential outcome from a treated state for a treated person is same as the mean of the potential outcome from a treated state for a non-treated person. To summarise, individuals from the same state who do not receive treatment would respond in the same way as those who do. (Jones (2000), page 292) Therefore the bias B(X) is zero but selection bias that is caused because the participant and the comparison group may not completely overlap still exists. Consequently matching method balances following bias:

$$\mathrm{var}(U_0 \mid X, D = 1) = \mathrm{var}(U_0 \mid X, D = 0). \qquad (13)$$

Instead of obtaining the comparison group via randomisation among participants as in an ideal experiment, matching methods aver conditioned on a set of X variables for randomly selected participants. There is no selective difference in Y_0 outcomes between participants and non-participants given X. The matching generated group is often smaller than randomised comparison group. (Heckman et al. (1999), page 1952)

Further "treatment on the treated" parameter and "non-treatment on the treated" are constructed and persons can be selected into the program on the basis of unobservables for treated but not on the basis of unobservables for untreated. This allows for identification of the mean effect of non-treatment on the non-treated. (Heckman et al. (1999), page 1952)

For demonstration of the method I use an example from Heckman et al. (1999) where two samples are analysed – a sample "t" as a treatment and a sample "c" as a comparison group. These observations are statistically independent. Y_i^t denotes outcomes from the treatment group and they are matched with outcomes from comparison group denoted Y_i^c. In this example the matches are constructed on the basis of a neighbourhood $C(X_i)$, where X_i is a vector of characteristics for person *i*. There are N_c persons in the comparison sample and N_t in the treatment sample. Further it is stated that person *j* from comparison group is a neighbour to the person *i* from treatment group. Weight W(i,j) is stated and is placed on observation *j* in order to make comparison with observation *i* and sums to one. Then a weighted comparison group mean for person *i* is formed $\overline{Y}_i^c = \sum_{j=1}^{N} W(i,j) Y_i^c$ and the estimated treatment effect for person *i* is $Y_i^t - \overline{Y}_i^c$. (Heckman et al. (1999), page 1953)

Example above is one of the several matching methods that are used in econometrics. Matching is most practical in cases where the causing variable takes on two values. (Angrist et al. (1999)) Compared with method of randomised trials, the method of matching has the same uncertainty about X. Randomisation method works for any choice of X. The general rule for matching method is to include in X only variables that are not caused by treatment given the unobservables. This leads to elimination of selection bias by making condition on X. (Heckman et al. (1999), page 1955-1956)

Heckman's method

Second cross-sectional method that can be used in evaluation of non-experimental data is Heckman's method. This method consists of two-stage estimation and is used for a test for selectivity bias. It is done by estimation the covariance between U_0 and U_1. When there is no selection bias in the data this covariance is zero. (Maddala (1986)) Sample selection bias may be written as following:

$$Y = X\beta + \varepsilon \quad (14)$$

$$I^* = Z\delta + \upsilon \quad (15)$$

$$I = 1 \text{ if } I^* > 0, \quad (16)$$

$$I = 0 \text{ if } I^* \leq 0. \quad (17)$$

Z is variable that vary across individuals e.g. characteristics of their occupation. X and I are exogenous variables. I^* is selected variable that is never observed. Model estimated is:

$$E(Y \mid X, I = 1) = X\beta + \theta\lambda(Z\delta) \quad (18)$$

where $\theta = \rho\sigma$ and $\lambda(Z\delta)$ is inverse Mill's ratio. I is monotone decreasing function of the probability that an observation is selected into sample. (Greene (1993), page 977-978)

Since the selection variable I^* is not observed, the disturbance variance in the selection equation cannot be estimated by usual means. Thus the model is reformulated as follows:

Selection mechanism: $I^* = Z\delta + \upsilon$, $I = 1$ if $I^* \geq 0$ and 0 otherwise; Prob $(I = 1)$ and $\text{Prob}(I = 1) = \Phi(Z\delta)$ and $\text{Prob}(I = 1) = 1 - \Phi(Z\delta)$.

Regression model: $Y = X\beta + \varepsilon$ observed only if $I = 1$, $(\upsilon, \varepsilon) \sim$ bivariate normal $[0, 0, 1, \sigma, \rho]$.

Again I and Z are observed for a random sample of individuals but Y is observed only when I = 1. (Greene (1993), page 978)

The strategy for estimating this model is to use the two-step estimator. Heckman's two-step estimation procedure is following:

1. Estimate the probit equation by maximum likelihood to obtain estimates of δ. For each observation in the selected sample compute $*\lambda$[1] = $\varphi(*\delta Z) / \Phi(*\delta Z)$ and $*\delta = *\lambda(*\lambda + *\delta Z)$. In the context of the problem of estimating e.g. effect of the treatment for a group member, the probit model explains why a curtain person is member of examined group.

2. Estimate β and $\theta = \rho\sigma$ by least squares regression of Y an X and $*\lambda$. In this stage relation between effect of treatment and variable $\lambda(Z\delta)$ is created and accounts for the fact that observed sample of group members is not random.

The difference in expected effects of e.g. the medicine (expected value of the treatment for an individual) between participants and non-participants is:

$$E(Y \mid X = 1) - E(Y \mid X = 0) = \delta + \rho\sigma(\varphi / \Phi(1 - \Phi)) \qquad (19)$$

where φ and Φ are respectively the density and distribution function for a standard normal variable. (Greene (1993), page 978, 982)

To illustrate this method I use an example from Greene (1993) where a model of migration analysed by Nackosteen and Simmer (1980) is described. The model includes following three equations: net benefit of moving equation, equation for income if moving and equation for income if staying. One of the benefits individuals could get if they move is the market wage they achieve compared with that if they stay. That is why even factors that affect the income received in either places are determinants of the net benefit equation. In this case these determinants for equation for net benefit of moving are dummy for self-employment, rate of growth of state employment, growth of state per capita, dummies for age, race, sex and individual change of industry. The earnings equations included dummies for self-employment and the change of industry. This model implies an income after moving for all (9223 individuals) but it is only observed for those who actually moved (1078 individuals). Reported results show that selectivity ratio (step 1) is higher for non-migrant's than for migrant's earnings and estimated earnings (step 2) are higher for migrants than for non-migrants. (Greene (1993), page 979)

[1] * in front of symbol means that this value is estimated.

Empirical practice of this method in labor economics has slightly declined (Moffitt (1997)). The possible reasons are that $\lambda(Z\delta)$ is often highly collinear with X and estimates of β tend to be unstable, non-robust and sensitive to changes in specification of X. Further the inverse Mill's ratio is close to unity over some ranges of selection probabilities. For these reasons estimation by maximum likelihood might be preferred by economist. (Moffitt (1997), page 1390-1391) However ρ may then fall outside of the permitted range of ± 1.

Instrumental variables method

The third cross-sectional estimator is based on the method of instrumental variable (IV). This method is well known in econometrics and widely used to estimate treatment effects. (Heckman (1995)) It is a good predictor of the treatment, but it is not independently related to outcome. It is invoked when persons are sorted into programs on a basis of unobservable factors that influence outcomes from treatment but that are not due to treatment being evaluated. This randomises patient to different likelihoods of receiving alternative treatments and gives the estimates of treatment effect that are not affected by selection bias. The method differs however from randomised trials in two aspects – it assumes that the proposed IVs are not correlated with unobserved differences in characteristics that directly affected outcomes, and it estimates an incremental effect of treatment only across IV groups. Further IVs methods compare outcomes from different treatment groups that appear identical except for their values of the IVs, which are associated with differences in treatment. Thus IVs methods isolate the effect of treatment variation in the observational data that is independent of the unobserved patient characteristics. (McClellan et al. (1994), page 860)

The method of IV is a variant of the method of matching. (Hill et al. (2001), page 294) If *z* is called IV, it satisfies following condition:

$$E(z, \varepsilon_i) = 0 \rightarrow E[z(Y - \beta X)] = 0. \qquad (20)$$

In large sample the IVs estimators have approximate normal distributions. This approach can eliminate for instance the bias of patient selection on mortality among these who receive varying treatments. It is developed on the basis of the general regression model $Y = \beta X + \varepsilon$ where K variables may be correlated with ε. The method exhibits a set of L variables *z* where

L is at least as large as K and where *z* is correlated with X but not with ε. With this in mind one can now estimate β using the assumed relationships among *z*, X and ε. (Greene (1993), page 288)

As above-mentioned, strategies based on IV approximate randomised trials where an underlying causal relationship exists and leads to biasing results. The conventional randomised trial tests treatment of well-defined individuals while IVs analysis tests application of that treatment to a large population in clinical practice. An IV, *z* that is correlated with X but otherwise independent of potential outcomes can eliminate selection bias. (Angrist et al. (1999), page 1300) The IV method estimates the mean of differences between potential outcomes from treated and untreated states and it augments the X variables in matching with instruments *z* as follows:

$$E(Y_1 - Y_0 \mid X, z, D = 1) = E(Y_1 - Y_0 \mid X, D = 1) \qquad (21)$$

that is, there is no dependence between $U_1 - U_0$ given X and D

$$E(U_0, z) = 0 \qquad (22)$$

which means that U_0 may depend on X but not on *z*.

If one also assumes that $E(U_0 \mid X) = 0$, $E(U_1 \mid X) = 0$ and $(U_0, U_1) \perp\!\!\!\perp$[2] $D \mid X, z$, i.e. treatment being evaluated has the same effect for everyone among persons with a given value of the regressor X, then IV also identifies the effect of treatment on randomly chosen persons with characteristics X. (Heckman et al. (1999), page 1960-1962)

If individuals are selected into the treatment on the basis of the unobservables or the variables that are dependent on them, U_1 must equal U_0 or their difference has no effect on participation decision. For the instrument *z* to be valid it has to be independent of error term. Generally it is assumed that the IV estimation can only provide a relative scaling or "incremental" effect of treatment, not the likelihood of benefits (or lack of benefit) of a specific treatment in an individual patient. Fundamental assumption on which IV estimation rests is that the proposed

[2] $\perp\!\!\!\perp$ denotes stochastic independence.

IVs are not correlated with unobserved differences on population subgroups that may lead to outcome differences. (Gould (1994))

To show an application of the IV method, an example of a macroeconomic version of the consumption function is described.[3] The National income is the sum of consumption (X), investment, government spending and net exports. The precise relation between consumption, X, and income, Y, i.e. $X = f(Y, \varepsilon)$ is ambiguous and covariance between Y and ε is not zero, but in the same time ε is uncorrelated with past values of consumption and income. Therefore the suitable instrumental variables are lagged values of Y and X, i.e. Y_{t-1} and X_{t-1}. Then the IV estimates of consumption function can be computed using these lagged values of Y and X. (Greene (1993), page 288-289, 294)

The reviewed literature goes further in examining instrumental variable method and states that the standard model of the evaluation problem can be based on three parameters. First, the mean effects of treatment on treated, answers the question about the change in outcome from the program if participants are compared with non-participants, i.e. $E(Y_1 - Y_0 \mid D = 1, X)$. Problems arise because $E(Y_0 \mid D = 1, X)$ cannot be observed. The second parameter is the effect of randomly assigning a persons selected in the population at large to the program. Finally third parameter is the effect of treatment on person at the margin of being treated. The details about these different IV methods are beyond the scope of this paper and can be found elsewhere (Heckman (1995)). Worth to mention is that in the most general case, when responses to treatment vary among persons with same X, these three parameters are different.

Regression discontinuity estimator

The last estimator applied on cross-sectional data is regression discontinuity estimator. It constitutes a special case of selection on observables. In this model treatment depends on n observed variable here named S so that if S is less than its mean value then there are participants in the treatment (D=1), otherwise there is non-participant in the treatment (D=0). The difference from the earlier methods is that there is no common support for participants and non-participants. If outcome from untreated state also depends on S then a discontinuity between the state that depends of the untreated state and S at the point equal its mean is introduced. (Heckman et al. (1999), page1969-1970)

[3] In this example the income is a variable that is possible to observe.

A relevant example for this study can be a program for subsidising of the medicine that is based on price of the treatment there outcome variable of interest is the effect of the treatment. Another simple example of this estimator from other economic field is given by Barnow et al. quoted by Heckman et al. (1999) where a hypothetical enrichment program for disadvantaged children based on family income is considered. Children with family income below a cutoff level receive the program, other do not. The children's test scores present the outcome variable. Figure 1 shows linear relationship between the test scores and the family income.

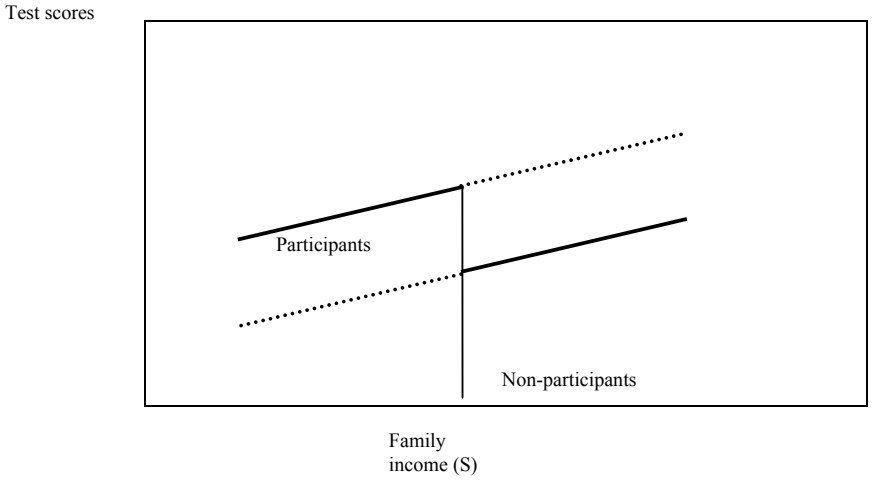

Figure 1. Regression discontinuity estimator

The broken line segment above the cutoff line represents this relationship that could continue to lower levels of family income in the absence of the subsidy program. The discontinuity in the regression line at the cutoff point reflects the effect of the program that is assumed to be α. It can be estimated without bias by OLS estimation[4] under the assumptions of the common effect model and of linear relationship between the effect and the price of the treatment.

If this model considers the random coefficient case where α varies among the persons and with S then the main impact of the treatment on treated may differ from the main impact of the treatment on the randomly selected person in population. Because there is no common support, there is no information about the impact of treatment on the untreated either except at

[4] Estimated equation is: $Y = \beta_0 + \alpha D + \beta_1 S + U$.

the point of discontinuity where extrapolation via functional form assumptions can give upper broken line in Figure 1. (Heckman et al. (1999), page 1971)

A possible application of the model is policy that can be interested to change the cutoff line so the treatments that are presently excluded become be allowed in the program. In order to do that one have to assume that potential outcome of untreated that depends on S is known. Then the impact of the treatment for person in the program can be estimated as the difference between the extrapolated potential outcomes of untreated that depends on S and the observed outcomes of participants at each value of S. (Heckman et al. (1999), page 1971)

One important issue that has to be solved in this method is assignment rule where assignment to treatment is not completely determined by S. This leads to that the problem changes from being a problem of selection on observables to being the problem of selection on unobservables conditional on S. Finally one problem can be that some therapies have "wrong" values of S that make persons being involved in the program be non-participants. Then the model no longer identify the mean impact of the treatment satisfying cutoff condition even if the functional form of potential outcome of untreated that depends on S is known. However unbiased estimates of the impact of the treatment on the treated are still provided. (Heckman et al. (1999), page 1972)

This regression discontinuity estimator together with the method of matching, Heckman's method and instrumental variable method represent cross-sectional estimators. They are only one group of all estimators that solve selection problem. In next chapter I will describe another group of estimators that deal with same problem but in slightly different way. These two groups are chosen to illustrate two strategies that can be used when dealing with selection problem.

Longitudinal methods

The evaluation strategy based on longitudinal data is called longitudinal methods. This data is, contra to cross-section data that is sampled in a particular tiem period, follow individual micro-units over time. There are two basic principals that can be used for the evaluation of this data- before-after estimator and difference-in-differences estimator. The next two sections are discussing these estimators in more detail.

Before-after estimator

Comparison that is based on the idea that persons can be in both treated and untreated state at different times is used for evaluation with before-after estimator. Here a person is compared with himself / herself. Outcomes from one state at one time are assumed to be good approximates for outcomes from the same state at the other time. (Heckman et al. (1999), page 1891)

The post-program outcome for a person who participate in the program is Y_{1t} and the pre-program outcome of the same person is $Y_{0t'}$. Program participation occurs between period t and t' there $t > t'$. The before-after estimator uses $Y_{0t'}$ to proxy the no-treatment state in the post-program period, i.e. $Y_{0t'} = Y_{0t}$. The test for the before-after estimator is then given by:

$$E(Y_{0t} - Y_{0t'} \mid D = 1) = 0. \qquad (23)$$

In other words, estimator is defined by the sample mean value for participants, i.e. $(\overline{Y}_{1t} - \overline{Y}_{0t'})_1$. The gain from the program for each individual is difference between Y_{1t} and $Y_{0t'}$, i.e. the impact of participation on those who participate in program is remaining value after subtracting participants' mean pro-program health status from the mean of their post-program health status. (Heckman et al. (1999), page 1892)

An example that practically shows both advantages and weaknesses of this estimator is the work of Ashenfelter quoted in Heckman et al. (1999), page 1892. Here Ashenfelter studies person in a training program. A decline in their earnings is observed prior to persons' training experience. If this decline is transitory and eventually restored even in the absence of participation in the program, the before-after estimator overstates the average effects of treatment on treated. This is the major disadvantage of this estimator and it comes from the assumption that the mean outcome in the no-treatment state is the same in post- and pre-program period. Another potential defect of this estimator is that it will attribute to the program any trend in health status due to macro or lifecycle factors. (Heckman et al. (1999), page 1893)

One possible approach to solve these problems is to extrapolate post-program outcomes Y_{1t} with data on pre-program outcomes prior to data t'. This method is appropriate if there are no

errors of extrapolation or if it is safe to assume that such errors average out to zero across persons in period t. In that way missing data on counterfactual state in period t is replaced with extrapolated values. (Heckman et al. (1999), page 1894)

Difference-in-differences estimator

When the mean change in the no-program outcomes is the same for participants and non-participants difference-in-differences estimator is used to evaluate the longitudinal but even panel data. It is given by

$$(\overline{Y}_{1t} - \overline{Y}_{0t'})_1 - (\overline{Y}_{0t} - \overline{Y}_{0t'})_0 \qquad (24)$$

and is based on assumption that expected values of difference between Y_{0t} and $Y_{0t'}$ for individuals in treated and untreated state are the same. If this assumption is valid the change in outcome in the comparison group represents common year or age effects among participants. The change in outcomes between the treated and untreated states is however not detectable. In order to identify gains from the program it is assumed that $(Y_{1t} - Y_{0t'})_1$ is independent of $(Y_{0t} - Y_{0t'})_1$ and that the distribution of $(Y_{1t} - Y_{0t})_1$ is the same as the distribution of $(Y_{0t} - Y_{0t'})_0$. (Heckman et al. (1999), page 1894)

Madrian's experiment quoted by Folland et al. (2001), page 235-236, provides an example of a case where difference-in-differences method is used. Madrian studies private health insurance in the USA, which are mostly obtained through employment and may inhibit worker mobility (job lock). In the example following matrix is used to consider the impact of the job lock:

	Employer-provided Health Insurance	
	No	***Yes***
No Other Health Insurance	a	B
Other Health Insurance	c	D

A test derived is whether having other health insurance increases mobility more for those who have employment-based insurance than for those who do not:

$$(d - b) > (c - a) \rightarrow (d - b) - (c - a) > 0$$

This test refers to differences between differences in d and b and differences in c and a, i.e. difference-in-differences.

The disadvantage of this estimator becomes prominent when participants and non-participants do not experience the same health change without the treatment. In that case the main assumption for this estimator is violated because the time path of no-treatment health status between pre-program time, t‘, and post-program time, t, will be different between participants and non-participants. The difference-in-differences estimator overstates then the average impact of treatment on treated. (Heckman et al. (1999), page 1895, 1896)

The methods that I have discussed in this paper are only a limited sample if all possible estimators that can be used for solving the selection problem. I chose those that I found often quoted and applied by economists in studied literature. This paper should be considered as a brief introduction into the field of models for evaluation of treatment effects based on nonexperimental data. More details about all of the studied models can be found in literature given in reference list. The interested reader may extend these studies with newer of different sources.

Summary

Empirical work in economics is going towards less restrictive, more robust and simpler methods that attempt to isolate key sources for identification of different phenomena in the economy. As the scientists are still looking for an appropriate answer to different questions about efficiency in improvements of welfare given scarce sources, economic evaluation is becoming more important in social-economic arena. Econometrics as a set of research tools facilitates evaluation of this development.

The purpose of this paper is to illustrate the diversity of applied econometric work over the last decades in the fields with a selection problem. Most of the scientific studies use observed, nonexperimental data and it has led to particular way of dealing with the selection problem bias. This work attempts to provide an overview of the empirical strategies in modern health economics used to deal with the issue of selection. It examines the methods used to evaluate effectiveness of new inventions like new pharmaceutical products. When these medicines are effective they modestly increase health status. But the gains form existing treatments have some times not been sufficiently large to completely lift one out of sickness or to scientifically reduce unhealthiness. That is the reason why the gains that actually occur are likely to overstate the medical improvement of the pharmaceuticals. Another important point of view in the discussion is the cost of these new products and their relation to actual benefits. That is why an appropriate cost-benefit analysis may be desirable.

Based on the problems discussed above this paper analyses the bias that come up when using nonexperimental data. These biases give rise to sample selection and errors in final conclusions. One should be aware that available nonexperimental evaluations of treatment effects might contain large and unknown biases resulting from specification errors. No methods are effective for bad and incomplete data. Good nonexperimental data is very often not available and likely to be expensive if it is available. Existing data can contain either too few (non)participants or contain too little information about characteristics that indirectly might affect the outcomes and that is important for conducting acceptable evaluations. Analysts have to rely on available data and to make the best possible interpretations of results from these evaluations. (Heckman et al. (1999))

I focused on identifying methods that illuminate mean outcome and in particular the mean impact of treatment on the treated. I found a great multiplicity of programs for evaluation and several parameters of interest that take account of treatment effects. They are based on heterogeneity in response among participants and the possibility that they participate in the program because of other factors than those being controlled for in available data, so called unobservable factors. This alters econometric estimators previously used and requires different assumptions for different parameters.

The choice of an evaluation method depends on the question being asked and on the economic model generating participation and outcomes because both questions and models vary among economic environments. There is no universally correct method for evaluation that applies in all contexts. This statement is contradicting a large part of current literature that treat matching, Heckman's, difference-in-differences and IV methods as cure-all-selection-problems-method. It is however clear that all methods are based on identifying assumptions that are difficult to test unless additional data about unobservables are collected. Different estimators solve different selection problems under different assumptions. Common for all of them is the importance of measuring outcome variables in the same way for participant and non-participant.

In the paper I have discussed various solutions to the selection problems which all have been undertaken to evaluate treatment effects based on nonexperimental data. The studied models are separate into two main sections - cross-section and longitudinal methods. In the first group I examine matching, Heckman's, IVs and regression discontinuity estimators. To the second group I count before-after and difference-in-differences estimators. This categorisation is done according to Heckman et al. (1999). I focus on their main application to the problem that they deal with and illustrate them with one economic study done by other scientists. All models above, summarised in Figure 2 serve more or less to the same aim, i.e. to solve the selection problem in nonexperimental data and make evaluations reliable.

Name of the model	Type of the model	Example of possible applications
Method of matching	Cross-section method	Estimation of the effects on earnings after training of workers
Heckman's method	Cross-section method	Health care consumption: 1 step: How often the patient seeks the health? 2 step: How much health care the patient seek once visiting the physician
Instrumental variables method	Cross-section method	Estimation of the intensity of e.g. treatment of acute myocardial infarction in the elderly and its effect on mortality
Regression discontinuity estimator	Cross-section method	Estimation of the effects of subsidy of the medicine baser on the price of the treatment
Before-after estimator	Longitudinal method	Estimation of the effects on schooling
Difference-in-differences estimator	Longitudinal method	Estimation of the effects of having health insurance on mobility among the workers

Figure 2: Summary of the models for evaluation of treatment effects in nonexperimental data

The matching method is a comparative method that provides an estimation strategy that gives more control over the average effects. According to leading economists in this field this method reduces biases in nonexperimental data. It is emphasised that crucial parts in a program evaluation are the construction and the weighting of the comparison group. It suggests that the comparison group should be selected to balance the regressors in the comparison group and to make them comparable to the treatment group.

Heckman's method has been used in several empirical applications. The case of selectivity is based on several criteria and selectivity bias is due to error in measurement of explanatory variables. This suggests that all the usual problems in the regression models need to be

analysed in the presence of the selection problem. The method includes even weakness whereas other methods like maximum likelihood may be more appropriate for certain cases.

When responses to treatment from different participants in the program are heterogeneous, IV method to estimate the parameters becomes weaker. An obvious task is to find a suitable set of IVs. The best choice of instruments is variables that are highly correlated with independent (explanatory) variables. When responses to treatment vary among participants the use of IVs fails unless person-specific responses to treatment do not influence decisions to participate in the program being evaluated. This requires that gains from the program do not influence the decision of the persons being studied to participate in the program. IV methods are extremely sensitive to assumption about how participant process information. This method has the same effect as prospective randomisation in eliminating unrecognised biases of patient selection on outcomes of differing treatments types.

The regression-discontinuity method uses a model to control for smooth and gradually evolving trends, when the variable of interest changes abruptly for non-behavioural reason. This method is a hybrid version of the instrumental variables method. The model is based on idea of how to identify the causal effect of a treatment that is a function of an observed covariate, which is also related to the outcome of interest. The choice of controls is even more important in the regression discontinuity than conventional IV method since the instrument is a function of one of the control variables. It suggests that when compared regression discontinuity method depends more on the functional form assumption than IV procedure.

Longitudinal methods include two mean principles - before-after and difference-in-differences models. The before-after estimator is also based on making comparisons undertaking certain assumptions and making adjustment to the means. An advantage of this approach is that it only requires information on the participants and their pre-participation histories to evaluate the program as long as the approximation error average out. The last assumption is seen to be the major drawback of the method. The other group of longitudinal methods represented by difference-in-differences estimators work well for an average version of treatment on the treated parameter when a good comparison group is available.

One task for future studies will be to apply these different methods and to see which of them fits better to certain data and the problem set. Another task will be to give them the validation

degree that will illustrate the goodness and applicability of the specific model. Finally it also has to be mentioned that even other models exist and an evaluation of them would be desired in order to get a full picture of the econometric tools that can be used today. Another important issue for investigation is the microeconomic treatment effect that ignores the effects of programs in the interactions among agent and may be misleading. In the future an extension of these studies with social setting would be even more satisfactory.

References

1. J. D. Angrist and A. B. Krueger (1999): "Empirical Strategies in Labor Economics". In O. Ashenfelter and D. Card, eds., *Handbook of Labor economics, Volume 3A*. North-Holland. 1277-1366.

2. W. A. Brock and S N. Durlauf (2001): "Interactions-Based Models". In J. J. Heckman and E. Leander, eds., *Handbook of Econometrics 2, Volume 5*. North-Holland. 3297-3380.

3. P. J. Dhrymes (1986): "Limited Dependent Variables". In Z. Griliches and M. D. Intriligator, eds., *Handbook of Econometrics 2, Volume 3*. North-Holland. 1567-1632.

4. S. Folland, A. C. Goodman and M. Stano (2001): "The Economics of Health and Health Care", third edition. Practice Hall, Inc.

5. J. Geweke and M. Keane (2001): "Computationally Intensive Methods for Integration in Econometrics". In J. J. Heckman and E. Leander, eds., *Handbook of Econometrics 2, Volume 5*. North-Holland. 3463-3568.

6. K. L. Gould (1994): "Invasive Procedures in Acute Myocardial Infarction; Are They Beneficial?", eds., *The Journal of the American Medical Association, Volume 272, Number 11*. 891-893.

7. W. H. Greene (1993): Econometric analyses, third edition. Practice Hall, Inc.

8. Z. Griliches (1986): "Economic Data Issues ". In Z. Griliches and M. D. Intriligator, eds., *Handbook of Econometrics 2, Volume 3*. North-Holland. 1465-1514.

9. J. Heckman (1979): "Sample Selection Bias as a Specification Error". eds., Econometrica*, Volume 47, Number 1*. 153-161.

10. J. Heckman (1995): “Instrumental Variables; A study of Implicit Behavioral Assumptions Used In Making Program Evaluations” eds., *The Journal of Human Resources, Volume 32, Number 3*. 441-461.

11. J. J. Heckman, R. J. Lalonde and J. A. Smith (1999): “The Economics and Econometrics of Active Labor Market Programs”. In O. Ashenfelter and D. Card, eds., *Handbook of Labor economics, Volume 3A*. North-Holland. 1865-2097.

12. R. C. Hill, W. E. Griffiths and G. G. Judge (2001): Undergraduate econometrics, second edition. John Wiley & Sons, Inc.

13. A. M. Jones (2000): “Health Econometrics”. In A. J. Culyer and J. P. Newhouse, eds., *Handbook of Health Economics, Volume 1A*. North-Holland. 256-344.

14. A. M. Jones and O. O’Donnell (2002): “Introduction” *Econometric Analysis of Health Data* 1-12. John Wiley & Sons, Ltd.

15. G. Kobelt (2002): “Health economics: An introduction to economic evaluation”, second edition. Office of Health Economics.

16. G. S. Maddala (1986): “Disequilibrium, Self-selection and Switching Models”. In Z. Griliches and M. D. Intriligator, eds., *Handbook of Econometrics 2, Volume 3*. North-Holland. 1633-1655.

17. M. McClellan, B. J. McNeil and J. P. Newhouse (1994): “Does More Intensive Treatments of Acute Myocardial Infarction in the Elderly Reduce Mortality?”, eds., *The Journal of the American Medical Association, Volume 272, Number 11*. 859-866.

18. R. A. Moffitt (1997): “Econometric Methods for Labor Market Analysis”. In O. Ashenfelter and D. Card, eds., *Handbook of Labor economics, Volume 3A*. North-Holland. 1367-1397.