Impact Evaluation, Causal Effects, Labour Market and Education Policies

Markus Froelich

University of St. Gall

My research activities as a PhD student and Post-doc at the University of St.Gallen mainly covered three fields: Nonparametric methods for *impact evaluation*, evaluation of active *labour market policies* and public employment services, and evaluation of *educational policies*, mainly primary education. These will also serve as the foundation for my future research at the University of Mannheim, where I will be assuming the Chair of Econometrics in 2008.

Impact evaluation and treatment effects

Policy evaluation and questions about causality have received substantial attention in economics in the recent decade. Causality aims to distinguish cause and effect, whereas in reality we only observe correlations. Clearly these questions have been of interest in philosophy for at least a few thousands of years. Nevertheless, recent advances in the computing power and the increasing abundance of microeconomic data permit us to develop and apply more robust nonparametric estimators for empirical applications. In particular, due to the increase in computational power new statistical methods become feasible, which rest on weaker assumptions about the real world. In the early days of econometrics, linear models have been assumed throughout. However, the assumption of linear relationships is often not very realistic in economic applications, nor in other sciences. Much of my research has been devoted to developing new nonparametric estimators, which do not require any parametric assumptions. In other words, for using these methods we do not have to assume that the world is linear.

The fundamental problem of impact evaluation can most simply be explained through the use of counterfactual outcomes. Consider the simplest case of a single binary treatment denoted by D. An individual i may either receive the treatment (D_i=1) or may not receive the treatment (D_i=0). As an example, we may think of a medical study where a new medicament is tested against a placebo. Each participant either receives the new drug or the placebo. After administering the treatment we observe some outcome some time later, e.g. health or survival status of that person. The potential outcomes are Y_i^1 and Y_i^0, where the former is the outcome that this particular individual would realize if it were

administered the new drug, whereas the outcome Y_i^0 would be realized if taking the placebo. The difference $Y_i^1 - Y_i^0$ is the effect of treatment for this particular individual and $E[Y_i^1 - Y_i^0]$ is the average effect in the population. The fundamental problem of evaluation is that for each individual only either Y_i^1 or only Y_i^0 can be observed. In other words, either individual i takes the placebo or he takes the new drug, and the other counterfactual outcome is then unobservable by definition. One could simply compare the average outcome of those who took the new drug with the average outcome of those who took the placebo. However, this difference $E[Y_i^1|D_i=1] - E[Y_i^0|D_i=0]$ would usually be a mixture between the treatment and the selection effect. Suppose that people simply choose whether they take the new drug or the placebo. It may well be that those with a more severe health status choose the new drug (with unknown side effects), while those not being acutely sick may be willing to take the placebo. From comparing the outcomes between these two groups, we could never ascertain the true causal effect of the new drug since we would be comparing different types of people. In the medical sciences double-blind randomised trials are therefore widely used. Individuals are randomly assigned to treatment or placebo (and, in addition, do not know their status), which guarantees that those who take the drug have on average the same observed and unobserved characteristics as those who received the placebo. In principle this implies that any observed differences in later outcomes between these two groups can be attributed to the effect of the new drug vis-à-vis the placebo. (Of course, issues of non-compliance might also become relevant here and the estimators discussed below might then also be useful.)

In the social and economic sciences, double-blind randomised trials are often not possible, or have not been implemented. The 'treatment' D_i might be here whether individual i received university education or did not, whether individual i attended a well equipped primary school with small classes and well trained teachers or did not, whether an unemployed person i attended a seminar or course in effective curriculum writing and job search training or did not, etc. Of course, D_i might not necessarily binary here anymore, but for convenience of exposition, I mostly focus on that case.

Robust statistical nonparametric methods

Traditionally, linear regression models have been used as a primary approach in empirical economics and the social sciences, essentially assuming that the world is linear, by supposing that Y_i = a D_i + , X_i + $\hat{A}_i$, where X_i are other observed characteristics of individual i. Clearly, the linearity assumption is not very credible, but slow computers and limited amount of data made other approaches less appealing. With the increases in computing power and also the availability of large datasets, often being made available from administrative registers, more robust *nonparametric* methods for *impact evaluation* have received considerable interest. Nonparametric methods do not assume a linear relationship and permit any statistical relationship between the variables of interest. This nonparametric approach starts from a structural causal model, as in the following two graphs, which entails our assumptions about the causal relationships between certain variables. These minimal-identifying assumptions cannot be tested from the data and depend on our understanding of the causal ordering and potential confounding variables. These assumptions do not impose any linearity or other functional form, only the causal ordering is supposed. Based on such a causal model, complex statistical estimators can be calculated to estimate the average causal treatment effect $E[Y_i^1 - Y_i^0]$.

Most of my research in this field evolved around the development of new robust nonparametric estimators for treatment evaluation and the exploration of their asymptotic and finite-sample properties. My interest focussed a lot on nonparametric instrumental variable estimators, with potentially endogenous instrumental variable, as depicted in the previous graph. In this situation, there exists an instrumental variable Z that has an impact on treatment choice but not directly on the potential outcome. A simple example refers to the impact of university education on wages later in life. Clearly, those who studied at university have different observed and unobserved characteristics than those who decided not to pursue tertiary education. Therefore, a simple comparison between these two groups cannot tell us if and by how much university education raises individual wages. However, there might be certain factors that affect attendance but not wages directly. For example, it is quite likely that an individual i who grew up in a city with a

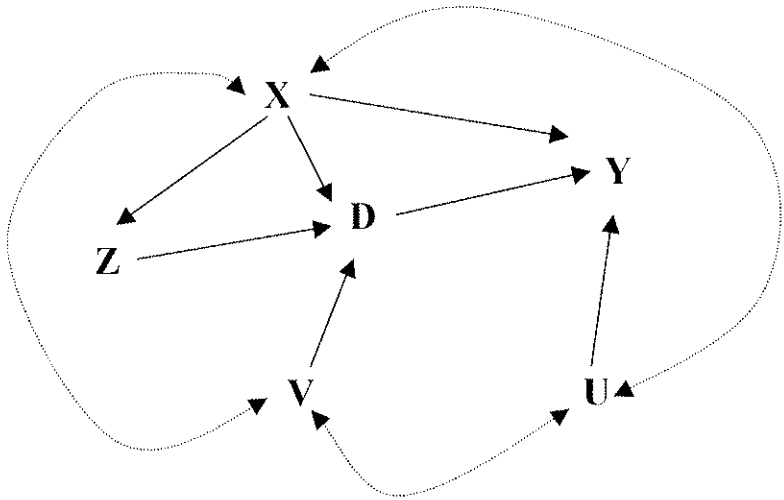

Fig. 1. Nonparametric identification with instrumental variables

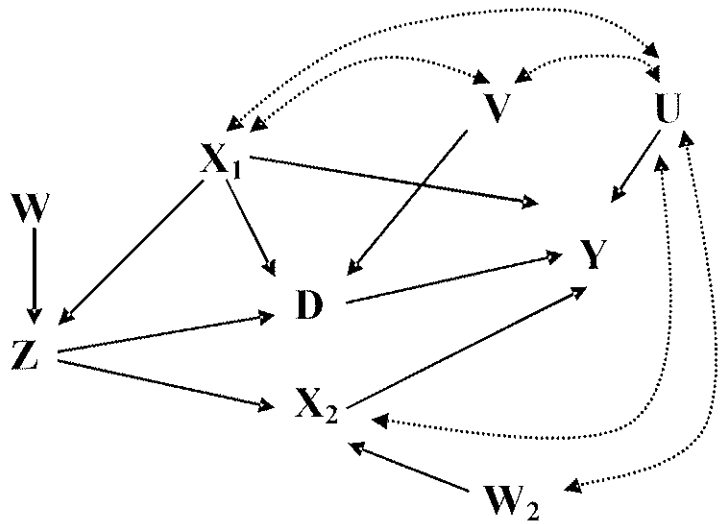

Fig. 2. Nonparametric identification with instrumental and control variables

university close by may be (at least a little bit) more likely to attend university later than an individual j who grew up in a city or region without a university. In this situation, the distance to university at the age 16 may act as such an instrumental variable Z. However, since location of residence was an active choice by their parents we need to control also for many other characteristics X, to make a causal chain as in the preceding or following graph plausible.

Evaluation of active labour market programmes

These estimators served also as the basis for my empirical research on active labour market policies. In Switzerland, as in other countries, various different types of training programmes, e.g. language training, computer training, further training, job search training, employment programmes, are offered to unemployed people to increase their chances to find a job (earlier). On the other hand, these active labour market policies might also generate incentives for unemployed individuals to extend their dependence on public income support. Active labour market programmes may not only benefit but could also harm. Several research projects have been pursued to assist in the improvement of the effectiveness of the public employment services in Switzerland. In one of these, the variation introduced by a minimum quota of year places to be provided by the different cantons was used as an instrumental variable to evaluate the effects of participation on later employment outcomes.

In several other projects for the State Secretariat for Economic Affairs (seco), other aspects of active labour market policies and of the public employment services in Switzerland were examined. One of these projects evaluated the strategies and organizations of employment offices and provided advice on possibilities to enhance their effectiveness. In another project a computer programme was developed to support case workers in choosing appropriate assistance programmes for their unemployed clients. The motivation for this latter project is based on the possibility of treatment effect heterogeneity. It may well be that some people do benefit from certain training programmes, whereas others may be harmed. If it were possible to identify which persons benefit

and which persons do not, one could increase efficiency by sending the former to training but the latter not. If these groups could be classified by a low dimensional set of covariates, e.g. young women benefit, whereas older men do not, this allocation mechanism would be very simple. If the heterogeneity in the effect of treatment is more complex, though, a more involved estimation strategy is necessary and it will be necessary to provide this information to the caseworker in the employment office on a case by case basis. For each individual client, predictions of individual expected programme impacts should be made by taking the characteristics of this particular individual into account. Such kind of targeting approach would thus acknowledge to the fullest extent possible that people may be different such that evaluation should be done on an individual level.

Evaluation of educational policies

Educational policies are another field of interest where the proposed nonparametric estimators have been applied. Most of this research focussed on primary education in Francophone Africa, where quality and quantity are least developed compared to other regions of the world. In addition, high population growth strains the public resources for providing free (or almost free) primary education. In a first research project I examined the impact of textbooks on students cognitive achievement and in particular their potential externalities on classmates. It appeared that textbooks seem to have large positive effects also on classmates in the sense that if one additional child receives a textbook also her peers benefit from this. The channels for this impact may be several fold, but it appears that teacher's changes in instructional methods may be an important part of this. One of the conclusions from this study is that the externalities may lead to private under provision of textbooks and therefore call for public support for textbook provision, e.g. in the form of subsidized books.

Apart from textbooks and similar items, teachers are the most important input for schooling and at the same time the most expensive one. Due to high population pressure and high relative teacher salaries, several Francophone African countries changed in the late 1990s the terms of employment for newly hired public school teachers. Traditionally teachers

entered as civil servants with almost absolute employment security. In the new regime, many were now hired on short-term contracts with much lower pay, lower job security and often lower educational requirements. Whereas this clearly reduced the financial public burden and thereby permitted the hiring of many additional teachers, there is a concern that education quality might suffer from teachers with low job satisfaction, less qualifications and less job security. On the other hand, less job security may also impose incentives to provide (at least) a minimum level of effort as a teacher since the threat of being fired for poor performance, e.g. not attending at the workplace etc., is real. The empirical results point out that in some countries these new forms of teacher contracts may have reduced educational quality a little, whereas in others it even partly increased. Contrasting this with the large increase of newly hired teachers made possible by their lower costs, the overall assessment of this policy change is positive.

Further research

From early 2008 I will be assuming the Chair of Econometrics at the Universität Mannheim, where I will be continuing this line of research, on the one hand to develop robust methods for *impact evaluation* and, on the other hand, applying these to real world problems, mostly in Africa where the need for empirical research seems to be largest and at the same the potential for capacity building is greatest.

I would like to thank the Latsis Foundation for this encouraging award and also the Universität St.Gallen, and in particular my Doktorvater Professor Dr. Michael Lechner, for their support and for providing such a stimulating and productive research environment.