Written for Targeted Learning by Mark J. van der Laan and Sherri Rose, forthcoming 2011.

TECHNICAL REPORT
R-376
February 2011

Foreword

Targeted Learning, by Mark J. van der Laan and Sherri Rose, fills a much needed gap in statistical and causal inference. It protects us from wasting computational, analytical, and data resources on irrelevant aspects of a problem and teaches us how to focus on what is relevant – answering questions that researchers truly care about.

The idea of targeted learning has its roots in the early days of econometrics, when Jacob Marschak (1953) made an insightful observation regarding policy questions and structural equation modeling (SEM). While most of his colleagues on the Cowles Commission were busy estimating each and every parameter in their economic models, some using maximum likelihood and some least squares regression, Marschak noted that the answers to many policy questions did not require such detailed knowledge – a combination of parameters is all that is necessary and, moreover, it is often possible to identify the desired combination without identifying the individual components. Heckman (2000) called this observation "Marschak's Maxim" and has stressed its importance in the current debate between experimentalists and structural economists (Heckman 2010). Today we know that Marschak's Maxim goes even further – the desired quantity can often be identified without ever specifying the functional or distributional forms of these economic models.

Until quite recently, however, Marschak's idea has not attracted the attention it deserves. For statisticians, the very idea of defining a target quantity not as a property of a statistical model but by a policy question must have sounded mighty peculiar, if not heretical. Recall that policy questions, and in fact most questions of interest to empirical researchers, invoke causal vocabulary laden with notions such as "what if," "effect of," "why did," "control," "explain," "intervention," "confounding," and more. This vocabulary was purged from the grammar of statistics by Karl Pearson (1911), an act of painful consequences that has prevented most data-driven researchers from specifying mathematically the quantities they truly wish to be targeted. Understandably, seeing no point in estimating quantities they could not define, statisticians showed no interest in Marschak's Maxim.

Later on, in the period 1970–1980, when Donald Rubin (1974) popularized and expanded the potential-outcome notation of Neyman (1923) and others and causal vocabulary ascended to a semilegitimate status in statistics, Marschak's Maxim met with yet another, no less formidable, hurdle. Rubin's potential-outcome vocabulary, while powerful and flexible for capturing most policy questions of interest, turned out to be rather inept for capturing substantive knowledge of the kind carried by structural equation models. Yet this knowledge is absolutely necessary for turning targeted questions into estimable quantities. The opaque language of "ignorability," "treatment assignment," and "missing data" that has ruled (and still rules) the potential-outcome paradigm is not flexible enough to specify transparently even the most elementary models (say, a three-variable Markov chain) that experimenters wish to hypothesize. Naturally, this language could not offer Marschak's Maxim a fertile ground to develop because the target questions, though well formulated mathematically, could not be related to ordinary understanding of data-generating processes.

Econometricians, for their part, had their own reasons for keeping Marschak's Maxim at bay. Deeply entrenched in the quicksands of parametric thinking, econometricians found it extremely difficult to elevate targeted quantities such as policy effects, traditionally written as sums of products of coefficients, to a standalone status, totally independent of their component parts. It is only through nonparametric analysis, where targeted quantities are defined procedurally by transformational operations on a model (as in $P(y \mid do(x))$; Pearl 2009), and parameters literally disappear from existence, that Marschak's Maxim of focusing on the whole without its parts has achieved its full realization.

The departure from parametric thinking was particularly hard for researchers who did not deploy diagrams in their toolkit. Today, as shown in Chap. 2 of this book, students of graphical models can glance at a structural equation model and determine within seconds whether a given causal effect is identified while paying no attention to the individual parameters that make up that effect. Likewise, these students can write down an answer to a policy question (if identified) directly in terms of probability distributions, without ever mentioning the model parameters. Jacob Marschak, whom I had the great fortune of befriending a few years before his death (1977), would have welcomed this capability with open arms and his usual youthful enthusiasm, for it embodies the ultimate culmination of his maxim in algorithmic clarity.

Unfortunately, many economists and SEM researchers today are still not versed in graphical tools, and, consequently, even authors who purport to be doing nonparametric analysis (e.g., Heckman 2010) are unable to fully exploit the potentials of Marschak's Maxim. Lacking the benefits of graphical models, nonparametric researchers have difficulties locating instrumental variables in a system of equations, recognizing the testable implications of such systems, deciding if two such systems are equivalent, if two counterfactuals are independent given another, whether a set of measurements will reduce bias, and, most importantly, reading the causal and counterfactual information that such systems convey (Pearl 2009, pp. 374–380).

Targeted learning aims to fill this gap. It is presented in this book as a natural extension to the theory of structural causal models (SCMs) that I introduced in Pearl (1995) and then in Chaps. 3 and 7 of my book *Causality* (Pearl 2009). It is a simple and friendly theory, truly nonparametric, yet it subsumes and unifies the potential outcome framework, graphical models, and structural equation modeling in one mathematical object. The match is perfect.

I will end this foreword with a description of a brief encounter I recently had with another area in dire need of targeted learning. I am referring to the analysis of mediation, also known as "effect decomposition" or "direct and indirect effects" (Robins and Greenland 1992; Pearl 2001).

The decomposition of effects into their direct and indirect components is of both theoretical and practical importance, the former because it tells us "how nature works" and the latter because it enables us to predict behavior under a rich variety of conditions and interventions. For example, an investigator may be interested in assessing the extent to which an effect of a given exposure can be reduced by weakening one specific intermediate process between exposure and outcome. The portion of the effect mediated by that specific process should then become the target question for mediation analysis.

Despite its ubiquity, the analysis of mediation has long been a thorny issue in the social and behavioral sciences (Baron and Kenny 1986; MacKinnon 2008) primarily because the distinction between causal parameters and their regressional surrogates have too often been conflated. The difficulties were amplified in nonlinear models, where interactions between pathways further obscure their distinction. As demands grew to tackle problems involving categorical variables and nonlinear interactions, researchers could no longer define direct and indirect effects in terms of sums or products of structural coefficients, and all attempts to extend the linear paradigms of effect decomposition to nonlinear systems, using logistic and probit regression, produced distorted results (MacKinnon et al. 2007). The problem was not one of estimating the large number of parameters involved but that of combining them correctly to capture what investigators mean by direct or indirect effect (forthcoming, Pearl 2011).

Fortunately, nonparametric analysis permits us to define the target quantity in a way that reflects its actual usage in decision-making applications. For example, if our interest lies in the fraction of cases for which mediation was *sufficient* for the response, we can pose that very fraction as our target question, whereas if our interest lies in the fraction of responses for which mediation was *necessary*, we would pose this fraction as our target question. In both cases we can dispose of parametric analysis altogether and ask under what conditions the target question can be identified/estimated from observational or experimental data.

Taking seriously this philosophy of "define first, identify second, estimate last" one can derive graphical conditions under which direct and indirect effects can be identified (Pearl 2001), and these conditions yield (in the case of no unmeasured confounders) simple probability estimands, called mediation formulas (Pearl 2010a), that capture the effects of interest. The mediation formulas are applicable to both continuous and categorical variables, linear as well as nonlinear interactions, and, moreover, they can consistently be estimated from the data.

The derivation of the mediation formulas teaches us two lessons in targeted learning. First, when questions are posed directly in terms of the actual causal relations of interest, simple probability estimands can be derived while skipping the painful exercise of estimating dozens of nonlinear parameters and then worrying about how to combine them to answer the original question.

Second, and this is where targeted learning comes back to parametric analysis, the expressions provided by the mediation formulas may demand a new parameterization, unrelated to the causal process underlying the mediation problem. It is this new set of parameters, then, that need to be optimized over while posing the estimation accuracy of the mediation formula itself as the objective function in the maximum likelihood optimization. Indeed, in many cases the structure and dimensionality of the mediation formula would dictate the proper shaping of this reparametrization, regardless of how intricate the multivariate nonlinear process is that actually generates the data.

I am very pleased to see the SCM serving as a language to demonstrate the workings of targeted learning, and I am hopeful that readers will appreciate both the transparency of the model and the power of the approach.

Los Angeles, January 2011 *Judea Pearl*