Three Training Parameters within a Semiparametric Factor Model

Arild Aakvik
Foundation for Research in Economics and Business Administration
Breiviken 2
5035 Bergen-Sandviken, Norway
Arild.Aakvik@snf.no

James J. Heckman
University of Chicago and American Bar Foundation, Department of Economics
1126 E. 59th Street
Chicago IL 60637, USA
j-heckman@uchicago.edu

Edward J. Vytlacil
University of Chicago, Department of Economics
1126 E. 59th Street
Chicago IL 60637, USA
klmedvyt@lily.src.uchicago.edu

1. Introduction

Most studies that have examined the effects of training programs on employment assume that the effect of training is constant for all potential trainees. We formulate an econometric framework for studying heterogeneous training effects on discrete outcomes. We allow the treatment effect to vary depending on the trainee's observable and unobservable characteristics, and allow selection into training to be determined in part by the trainee's idiosyncratic treatment effect. We examine the special case where the unobserved heterogeneity is given by a factor structure. Within this framework, we show how to semi-parametrically define and estimate the average treatment effect, treatment on the treated, and local average treatment effect.

We apply our methodology to estimate the effect of Norwegian vocational rehabilitation training programs on employment, where training are offered individuals with reduced productivity in the labor market due to medical conditions. The main purpose of training is to increase participant’s employment probabilities rather than earnings.

2. The Training Model and Parameters of Interest

Let \( D_i \) be the treatment (participation in training) variable and \( Y_i \) be the outcome (employment) variable. We specify a discrete-choice framework where the unobserved heterogeneity is assumed to follow a factor structure and we examine the forms of different treatment parameters within this

---


2 Several recent studies have also used the factor structure assumption, including Card and Sullivan (1988), Gritz (1993), and Ham and Lalonde (1996).
framework. We specify the following decision rule for training:

\[ D_i^* = Z \beta_D - \eta_{D_i} \]
\[ D_i = 1 \quad \text{if} \quad D_i^* \geq 0, = 0 \quad \text{otherwise} \]

We specify an employment outcome equation that depends on whether the individual is in the training or non-training state. We specify the following employment outcome equation for the training state:

\[ Y_{i1}^* = X \beta_1 - \eta_{1i} \]
\[ Y_{i1} = 1 \quad \text{if} \quad Y_{i1}^* \geq 0, = 0 \quad \text{otherwise} \]

and the following employment equation for the non-training state:

\[ Y_{0i}^* = X \beta_0 - \eta_{0i} \]
\[ Y_{0i} = 1 \quad \text{if} \quad Y_{0i}^* \geq 0, = 0 \quad \text{otherwise} \]

We assume the following factor structure:

\[ \eta_{Di} = -\theta_i + \epsilon_{Di} \]
\[ \eta_{1i} = -\alpha_1 \theta_i + \epsilon_{1i} \]
\[ \eta_{0i} = -\alpha_0 \theta_i + \epsilon_{0i} \]

where \( \theta_i \) is a common unobserved factor. Let \( F_{\theta} \), \( F_{\epsilon_D} \), \( F_{\epsilon_1} \), \( F_{\epsilon_0} \) denote the distribution of \( \theta \), \( \epsilon_D \), \( \epsilon_1 \), \( \epsilon_0 \). We assume that any subset of these shocks is jointly independent of the remaining shocks, and the shocks are jointly independent of the regressors. We thus have that

\[ F_{\eta_1}(t) = \int F_{\epsilon_1}(t + \alpha_1 \theta) dF_{\theta} \]
\[ F_{\eta_0}(t) = \int F_{\epsilon_0}(t + \alpha_0 \theta) dF_{\theta} \]

where \( j = 0, 1 \). We examine the form of three different parameters within this framework: the average treatment effect (ATE), effect of treatment on the treated (TT), and the local average treatment effect (LATE). The average treatment effect is given by

\[ \Delta^{ATE}(X,Z) = E(Y_1 - Y_0 | X,Z) \]
\[ = F_{\eta_1}(X \beta_1) - F_{\eta_0}(X \beta_0) \]

The expected effect of treatment on the treated is the most commonly estimated parameter, see Heckman and Robb(1985), and we define it as:

\[ \Delta^{TT}(X,Z) = E(Y_1 - Y_0 | X,Z,D = 1) \]
\[ = \int [E(Y_1 | X,0) - E(Y_0 | X,0)] dF_{\theta | D=1,X,Z} \]
\[ = \frac{1}{F_{\eta_0}(Z \beta_D)} \int [E(Y_1 | X,0) - E(Y_0 | X,0)] F_{\epsilon_D}(Z \beta_D + \theta) dF_{\theta} \]

LATE was introduced by Imbens and Angrist (1994). Following Heckman (1997) we define the LATE parameter as follows:
$$\Delta^{LATE}(X,Z) = \frac{\partial E(Y|X,Z=z)}{\partial z} \frac{\partial \Pr(D=1|X,Z=z)}{\partial z}$$

Furthermore, we can show that

$$\Delta^{LATE}(X,Z) = \frac{\int E(Y_1 - Y_0 | X, \theta) f_{e_0}(Z\beta_D + \theta) dF_\theta}{\int f_{e_0}(Z\beta_D + \theta) dF_\theta}$$

Further details about the different training parameters, how they are related, identification, and estimation results, see Aakvik, Heckman and Vytlacil (1998).

3. Empirical Results

The model is identified subject to appropriate normalizations and exclusion restrictions. Following Cameron and Heckman (1987), we estimate the model using nonparametric maximum likelihood estimation (NPMLE). The distribution of $\theta$ is approximated by a discrete distribution with a finite number of mass points. For details, see Aakvik, Heckman and Vytlacil (1998).

We apply our methodology to estimate the impact of Norwegian vocational rehabilitation (VR) programs on employment. The VR programs offer income maintenance payments and training programs for individuals with reduced productivity in the labor market due to medical conditions. Today, more than one percent of the Norwegian working-age population participates in a VR training program each day (around 50,000 individuals). Our random sample consist of 4,416 individuals that were eligible for VR training and had applied for training in 1989. Of these individuals, 2,908 were accepted and participated in training. Employment is recorded in 1993. For a description of the data and institutional settings, see Aakvik (1998).

We find that managers of VR programs select candidates for program participation in a manner consistent with a hypothesis of cream-skimming: that is, they select participants who would have higher employment rates even without the training program. The difference in the employment rates between participants and nonparticipants is eight percentage points. After adjusting for observed differences between participants and nonparticipants, the training effect reduces to around six percentage points. When the treatment effect is averaged over participants, it falls to 5.3 percentage points. The trainees have observable characteristics that are associated with a lower training effect, so that on average their treatment effect is lower than it would be for a person drawn randomly from the pool of applicants.

The model with unobserved heterogeneity allows selection both on observables and unobservables. Controlling for unobserved differences using our semiparametric factor model, the average training effect reduces to negative 1.8 percentage points, while the effect of treatment on the treated is negative 7.4 percentage points. This suggest that unobserved factors are important when estimating different training effects.

REFERENCES


**FRENCH RÉSUMÉ**