Working Paper No. 16/02

A Low-key Social Insurance Reform - Treatment Effects for Back Pain Patients in Norway

by

Arild Aakvik, Tor Helge Holmås and Egil Kjerstad

SNF-project No. 2435: Poliklinikktakst

The project is financed by The Ministry of Health

INSTITUTE FOR RESEARCH IN ECONOMICS AND BUSINESS ADMINISTRATION

BERGEN, APRIL 2002
ISSN 0803-4028

A Low-key Social Insurance Reform - Treatment Effects for Back Pain Patients in Norway

Arild Aakvik*

Department of Economics, University of Bergen

Tor Helge Holmås

Department of Economics, University of Bergen

Egil Kjerstad

Institute for Research in Economics and Business Administration (SNF) and

Programme for Health Economics in Bergen (HEB)

March 15, 2002

*We thank the Norwegian Ministry of Health for financial support.

Abstract

This paper estimates treatment effects for back pain patients using observational data from a low-key social insurance reform in Norway. Using a latent variable model we estimate the average treatment effects (ATE), the average effect of treatment on the treated (TT), and the distribution of treatment effects for outpatient treatment at three different locations. To estimate these parameters and the distribution of treatment effects we use a discrete choice model with unobservables generated by a factor structure model. Distance to nearest hospital (in kilometers) is used as an instrument in estimating the different treatment effects. We find a positive effect of treatment of 6 percentage points on the probability of leaving sickness benefits after allowing for selection effects and full heterogeneity in treatment effects. We also find that there are sound arguments for increasing the outpatient program of treating back pain patients.

1 Introduction

In many countries, the social insurance system is under pressure from an aging population and an increased number of people on disability pension. Sickness absence represents the third major type of financial transfers from the social insurance system to individuals. There is also a worrying connection between long-term sickness absence and recruitment into disability pensions. Thus, reduction in sickness absence is high on the political agenda since a reduction will contribute to lessen the burden facing the social insurance system.

One of the main reasons for sickness absence in Norway is related to back pains. Traditionally, treatment of patients with back pain has not been prioritized at Norwegian hospitals, resulting in long waiting lists for inpatient treatment. As a consequence, people with back pain are on sick leave for relatively long periods of time. However, clinical evidence emerging in the last five to ten years shows that multidisciplinary *outpatient* care (medical doctors, physiotherapists or psychologists working in teams) give promising results regarding the transition back to work for people with back pains.[1]

Partly based on these results, the Norwegian government decided that one possible way forward in order to lessen the burden on the social insurance system is to boost the number of treated patients at outpatient clinics. To this end, the Norwegian government settled for financial incentives by introducing a new and higher outpatient tariff for multidisciplinary treatment of patients with back pains. The aim is two-fold: First, create incentives for hospitals to establish designated outpatient clinics for back patients or to increase capacity in already established clinics. Second, reward multidisciplinary outpatient treatment. The rationale for the reform is that increased

[1] See for instance Haldorsen et al. (2002) and Indahl et al. (1995).

capacity and increased utilization of treatment techniques with an expected high success rate should result in shorter sickness spells, finally leading to reduction in sickness payment from the social insurance system.

The major question asked in this paper is whether this low-key reform can be termed a success. We have taken the following approach: The main part of this paper is the estimation of multidisciplinary treatment effects for patients with back pains (mainly disk herniation and non-specific low back pain) using observational data from three different locations. We estimate an econometric model for evaluating treatment effects when outcomes are discrete and estimate a flexible model were responses to treatment vary among observationally identical persons. The outcome variable is a dichotomous variable indicating if the patient leaves the sickness benefit scheme after nine months. Our structural model can be used to generate a variety of mean treatment effects (the average treatment effect (ATE) and treatment effect on the treated (TT)) from a common set of parameters as well as distributions of treatment effects. The estimates produced from our model are economically interpretable and can be used to conduct out-of-sample forecasts and to pool evidence across studies - the usual benefits of a structural econometric approach.

We address four questions in this paper: 1) What type of patients are being treated in an outpatient hospital of those who have been examined at the outpatient hospital? 2) What is the overall effect of treatment on the probability of the leaving sickness benefit scheme after 9 months? 3) Which groups of individuals benefit most from treatment? 4) How important is it to control for observables and unobservables in understanding the selection and outcome processes?

We find that hospitals select patients well in terms of observable charac-

teristics. Without adjusting for observed selection into treatment we find a treatment effect of 7.3 percentage points.[2] After allowing for observed characteristics we find a treatment effect of those who are treated (TT) of 9.3 percentage points. However, after running a very flexible selection model where we account for heterogeneous treatment effects, we find an effect of treatment on the treated (TT) of 6.3 percentage points.

In addition we find that the average treatment effect (ATE), i.e. the effect of randomly picking a person in the pool of eligible patients, are higher than TT. Adjusting for observed characteristics gives a average treatment effect (ATE) of 12.3 percentage points, while after adjusting for unobserved selection gives a ATE of 9.5 percentage points. This indicates that expanding treatment slots may increase the overall benefits of treating back pain patients, since ATE is greater than TT.

This paper is organized in the following way. In Section 2, we present a class of latent variable models that can be used to generate and produce structure on the classical model of potential outcomes. The specification can be used to estimate structural econometric models. We define commonly used treatment effect parameters in terms of the latent variables. We consider both means and distributions of treatment effects. Section 3 presents background information on the program and the data used in the empirical section. In Section 4 we discuss the selection process into treatment using a probit model. In section 5 we present the main estimation results from the model. The paper concludes in Section 6.

[2] 48.1 percent of those under treatment left sickness benefits after 9 months, while 40.8 percent of those not under treatment left sickness benefits after 9 months.

2 A Latent Variable Model

For each person i we have two potential outcomes (Y_{0i}, Y_{1i}) corresponding, respectively, to the potential outcomes in the untreated and treated state.

Let $D_i = 1$ denote the receipt of treatment and $D_i = 0$ denote non-receipt. Let Y_i be the measured outcome variable so that

$$Y_i = D_i Y_{1i} + (1 - D_i) Y_{0i}.$$

This is the classical model of potential outcomes[3] that can be used to estimate structural econometric models. The model has two potential outcome states of which only one is observed for each individual.

We specify a discrete-choice framework where the unobserved heterogeneity is assumed to follow a factor structure. The decision rule for outpatient treatment is given by

$$\begin{aligned} D_i^* &= Z_i \beta_D + U_{Di} \\ D_i &= 1 \text{ if } D_i^* \geq 0, \ D_i = 0 \text{ otherwise,} \end{aligned} \tag{1}$$

where D_i^* is a latent index that determines treatment or not, Z_i is the vector of background variables, γ is a set of parameters that reflect the effect of changes in background variables on the treatment index, and U_D is the unobservables.

We specify an outcome equation that depends on the whether the individual is in the treated or non-treated state. We have the following outcome equation for the treatment state

$$\begin{aligned} Y_{1i}^* &= X_i \beta_1 + U_{1i} \\ Y_{1i} &= 1 \text{ if } Y_{1i}^* \geq 0, \ Y_{1i} = 0 \text{ otherwise,} \end{aligned} \tag{2}$$

[3] See Neyman (1923), Fisher (1935), Roy (1951), Cox (1958), Quandt (1972), Rubin (1978) and Heckman and Honoré (1990).

where Y_{1i}^* is the latent index of leaving the sickness benefits scheme after 9 months, and X_i is a vector of background variables that affect the outcome. X_i and Z_i are not necessarily the same vectors. In particular, we have included a variable in Z_i that is not in X_i. The identifying exclusion restriction we are using is the distance in kilometers to the nearest hospital treating back pain patients. The outcome in the non-treatment state is

$$\begin{aligned} Y_{0i}^* &= X_i\beta_0 + U_{0i} \\ Y_{0i} &= 1 \text{ if } Y_{0i}^* \geq 0,\ Y_{0i} = 0 \text{ otherwise.} \end{aligned} \tag{3}$$

The effects of the unobservables is the same in both states if $U_{1i} = U_{0i}$. In this case individuals with the same observed x will have the same treatment effect. However, the model allows for treatment effects to vary by observed individual charateristics. The model is termed the common coefficient model, see Heckman (1978). In this paper we assume $U_{1i} \neq U_{0i}$ and thus allow for a idiosyncratic gain of treatment for each individual. This is a random coefficient model if patients act on U_{1i} and U_{0i}, see Heckman (1997). To build a structural random coefficient model we assume the following factor structure for the error terms

$$U_{Di} = \alpha_D\theta_i + \epsilon_{Di} \tag{4}$$

$$U_{1i} = \alpha_1\theta_i + \epsilon_{1i} \tag{5}$$

$$U_{0i} = \alpha_0\theta_i + \epsilon_{0i} \tag{6}$$

where ϵ_D, ϵ_1, ϵ_0, and θ have mean zero, are mutually independent, and are independent of the exogenous variables in the model. The parameter α_D is the factor loading for the selection outcome, and α_1 and α_0 is the factor loading for the outcome equation with and without treatment, respectively. The interpretation of this specification considers θ, which are common to

all states, to be an unobserved covariate that affects the outcomes, and the α's to be regression coefficients. From the model we can formulate several interesting treatment effects parameters within the framework of flexible but parsimonious specification, see for instance Aakvik et al. (2000).

To identify the model we assume $\alpha_D = 1$ and that θ follows the standard normal distribution. The standard normality assumption of θ is not needed, see Aakvik et al. (2000). We assume access to an i.i.d. sample and suppress the i subscript. We focus on three correlations, derived from equations (4)-(6), for the unobservables in the model

$$\text{Corr}(U_0, U_1) = \frac{\text{Cov}(U_0, U_1)}{\sqrt{\text{Var}(U_0)}\sqrt{\text{Var}(U_1)}} = \frac{\alpha_1 \alpha_0}{\sqrt{1+\alpha_0^2}\sqrt{1+\alpha_1^2}} \tag{7}$$

$$\text{Corr}(U_D, U_0) = \frac{\text{Cov}(U_D, U_0)}{\sqrt{\text{Var}(U_D)}\sqrt{\text{Var}(U_0)}} = \frac{\alpha_0}{\sqrt{2}\sqrt{1+\alpha_0^2}} \tag{8}$$

$$\text{Corr}(U_D, U_1) = \frac{\text{Cov}(U_D, U_1)}{\sqrt{\text{Var}(U_D)}\sqrt{\text{Var}(U_1)}} = \frac{\alpha_1}{\sqrt{2}\sqrt{1+\alpha_1^2}} \tag{9}$$

which are easy to verify given our assumptions. We also have

$$\text{Cov}(U_0, \theta) = \alpha_0, \text{Cov}(U_1, \theta) = \alpha_1, \text{Cov}(U_D, \theta) = \alpha_D. \tag{10}$$

since $\text{Var}(\theta) = 1$.

In the following we approximate the distribution of θ with a finite number of support points. This is a common estimation strategy see Butler and Moffitt (1982).

In a three-equation model with dichotomous outcomes we can form the following equations. First, the probability of leaving the sickness benefit scheme in the treated state is given by

$$\Pr(Y_1 = 1|X) = \sum_{j=1}^{m} \pi_j \Phi(X\beta_1 + \alpha_1 \theta_j), \tag{11}$$

where m is the number of support points, π is the mass probabilities which sums to 1, and Φ is the standard normal cumulative density function. The probability of leaving the sickness benefit scheme in the non-treated state is given by

$$\Pr(Y_0 = 1|X) = \sum_{j=1}^{m} \pi_j \Phi(X\beta_0 + \alpha_0\theta_j). \tag{12}$$

This set up is very flexible since we allow $\beta_1 \neq \beta_0$ and $\alpha_1 \neq \alpha_0$. The probability of being treated in an outpatient hospital is given by

$$\Pr(D = 1|Z) = \sum_{j=1}^{m} \pi_j \Phi(Z\beta_D + \alpha_D\theta_j). \tag{13}$$

Equation (11)-(13) is a structural model in the sense that we can predict the outcome in the treated and non-treated state for each individual even if we do not observe each individual in both states. Using Bayes' rule we get

$$\Pr(Y_1 = 1|D = 1, X) = \frac{1}{\Phi(Z\beta_D)} \sum_{j=1}^{m} \pi_j \Phi(X\beta_1 + \alpha_1\theta_j)\Phi(Z\beta_D + \alpha_D\theta_j), \tag{14}$$

and

$$\Pr(Y_0 = 1|D = 1, X) = \frac{1}{\Phi(Z\beta_D)} = \sum_{j=1}^{m} \pi_j \Phi(X\beta_0 + \alpha_0\theta_j)\Phi(Z\beta_D + \alpha_D\theta_j). \tag{15}$$

The average treatment effect (ATE) and the effect of treatment on the treated (TT) is given by

$$\text{ATE}(X) = \Pr(Y_1 = 1|X) - \Pr(Y_0 = 1|X) \tag{16}$$

and

$$\text{TT}(X) = \Pr(Y_1 = 1|D = 1, X) - \Pr(Y_0 = 1|D = 1, X) \tag{17}$$

see Heckman and Robb (1984) and Heckman (1997). To find the average treatment effect we insert equation (12) and (13) into equation (16), and to

find the effect of treatment on the treated we insert equation (14) and (15) into equation (17). To find ATE we average ATE(X) for the full sample. To find TT we average TT(X) over the sample of treated patients ($D = 1$).

3 Data and Institutional Settings

In this study we use a data set drawn from the Norwegian Patient Register (NPR). NPR is a large database containing patient data from all public general hospitals in Norway, as well as from some private clinics. The register provides detailed information on variables like age, gender, medical diagnosis, treatments and the date of hospitalization. Since the register does not contain social security information, we have merged the data from NPR with data from the Norwegian National Insurance Administration (NIA). Among other information, these data provides us with exact dates for each patient's sick leave spell during the period 01.01.00 until 31.12.00.

Three conditions have to apply for a patient to be part of the sample: 1) The patient has been examined (but not necessarily treated) at an outpatient spine clinic in year 2000. 2) The patient has a M-diagnosis according to the ICD-10 classification, which comprises patients suffering from musculoskeletal pain. 3) The patient is eligible for sickness benefits from the Norwegian mandatory sickness insurance system, and has started a sick leave spell during the three first months of year 2000.

This gives us a sample of 656 individuals. Since the sampling period is the three first months in 2000, and since we have social security data until 31.12.00, we can track the individuals in our sample for a period of minimum 9 months. Obviously, there is a trade-off between the length of the sampling and follow-up periods. A longer sampling period gives a larger sample but a

shorter observation period. We have experimented with different lengths of the sampling period up to six months.[4]

The treatment offered to the patients is a multidisciplinary program were for instance neurologists, psychologist, physiotherapists and nurses are involved. Over several years, clinical research has aimed at establishing new cause-and-effect mechanism behind back pain illnesses and to evaluate the effect of different treatment programs. Researchers have succeeded in the sense that advises given by the medical profession to people with back pain is slowly changing. From a period dominated by what can be termed passive treatment and a belief that rest and minimal physical activity would eventually free the patient from pain, now the strategy is early intervention and light physical activity. Advice and instruction concerning how to cope with different diagnosis is an important part of the treatment programs. Patients are given information concerning the reason for their pains and why it hurts, and thereby also motivating the importance of light exercises even if the pains are relatively strong. Most back pain related illnesses do in fact disappear in a relatively short period of time and surgical interventions should in many cases be avoided. The best treatment is to motivate the patients to exercise and to lessen their anxiety through information.

This is the kind of treatment programs the Norwegian government wants to encourage by the new outpatient tariff. By differentiating the tariff, so that clinics get paid twice as much for treating patients on sick leave compared to other clients, the government shows that they want clinics to be selective concerning the patient they use their scarce resources on.

In Table 1 we define the variables used in this paper. The outcome variable (Sickness) is a dummy variable, which takes the value 1 if the individual

[4]Results from this type of sensitivity analysis are available from the authors.

leaves the sickness benefit scheme after 9 months, and 0 otherwise. Whether the patients receive treatment or not are also measured by a dummy variable. The treatment variable equals 1 if the patient receives one ore more treatments, 0 otherwise. Further, we have information about age and gender, where age is a continuous variable. We also have detailed information on medical diagnoses. The largest and most important groups of medical diagnoses are disk herniation[5] and low back pain[6]. They constitute more than 80 percent of the total sample and we have generated dummy variables for these two groups. In addition we have information about the yearly income and the number of sickness days prior to the sickness period, i.e. in 1999.

Lastly, we have information about where people live. From this variable we have constructed a new variable that measures the distance in kilometers to the nearest hospital that offers a treatment program for these kinds of patients. As will be nearer explained in the next section, this variable is being used as the excluded variable in a model where we control for selection into treatment. Our hypothesis is that there is a higher probability of getting treatment if the patient lives close to a hospital that offers treatment. This variable should not affect the transition out of sickness benefits, except indirectly through treatment.

[5] A disk may herniate because of sudden trauma, anything from a fall on an icy sidewalk to an athletic injury to simply lifting the wrong bag of groceries in the wrong way at the wrong time. They may also be caused by the cumulative long term effects of what doctors like to call poor body mechanics - a lifetime of too much bending and twisting in too many awkward positions. Disks herniate most commonly in the lower back, although they also occur frequently in the lower neck and more uncommonly may occur anywhere.

[6] Low back pain or non-specific low back pain is a symptom that can arise from many causes. Many cases of back pain are caused by stresses on the muscles and ligaments that support the spine. Both increased weight on the spine and increased pressure on the discs can cause low back pain. A low back problem may come on suddenly or gradually.

Descriptive statistics for the full sample are reported in Table 2a, while Table 2b and Table 2c describes the samples for the treated and non-treated patients, respectively. Interestingly, we find that the proportion leaving sickness benefit after nine months is higher among the treated than in the non-treatment group (0.481 vs. 0.408). The unconditional mean difference is 7.3 and this would have been a consistent measure of the treatment effect if our data were truly experimental. However, our data are observational, and the two samples are therefore potential unbalanced both in observables and unobservables. This could be illustrated by looking at for example the diagnosis variables. We see that the proportion of patients with the diagnosis disk herniation is much higher in the non-treatment group than in the treatment group, while the proportion low back pain is highest in the treatment group. As expected we find that the distance to the nearest hospital that offers treatment (Distance) is highest among individuals in the non-treatment group. For the variables male, age, income and sm99 there are only small differences between the two samples. To take the non-experimental nature of our data into consideration we analyse the selection process into treatment formally using the econometric model outlined in section 2.

4 Selection into treatment

We first discuss the parameters related to selection into treatment within the framework of a regression model. The selection parameters reported in Table 3 offer a straightforward way to examine the presence of non-random selection into treatment. Table 3 presents the estimated coefficients of the probit model. Several of the estimated coefficients are statistically different from zero, as can be seen from the z-values. This indicates that individuals under

treatment differ significantly from eligible non-participants with respect to observable characteristics. The last column in Table 3 shows the marginal effects in percent.

From Table 3 we see that distance to nearest hospital (our instrument) is significantly different from zero. Thus we pass the first test of having a valid instrument: The instrument should be correlated with the treatment decision. The instrument should however not affect the outcome directly, only indirectly through the treatment variable.

Age and gender are not significant in the selection equation. However, both the medical dummies are different from zero. Income is also significant. Higher income in 1999 reduces the probability of being treated at an outpatient hospital. Table 4 shows the number of correct number of predictions in the probit model. The model fit is relatively high. The Pseudo R2 reported in Table 3 is 0.24.

Figure 1 shows the support and the distribution of the propensity score for participants $D = 1$), while Figure 2 shows the support and the distribution of the propensity score for non-participants ($D = 0$). Both figures are drawn using a kernel density estimate. The support is very good given the relatively low number of observations we have in the data set. The support of the propensity score for non-participants is mostly concentrated near zero, as can be seen from Figure 1. However, we have almost full support for the propensity score for non-participants, which in theory should be in the interval [0, 1]. The support for participants is similar to that of non-participants. However, the shape of the distribution is different, as we would expect. The probability of entering the treatment program is clearly higher for participants than non-participants. Mean propensity score for participants and non-participants is 0,34 and 0.12, respectively. The support region for par-

ticipants is [.024, .820] and [.0001, .850] for non-participants.[7] The support region is larger for non-participants, which could be explained by the relatively large number of non-participants (552) compared to participants (104).

5 Transition out of sickness benefits

We are interested in the effect of treatment on leaving the sickness benefit scheme after 9 months. We first look at the outcome model that does not include unobserved heterogeneity. We report the estimated outcome regression coefficients in Table 4 where the β_0-vector without selection is reported in Column 1 of Table 4, and the β_1-vector without selection is reported in Column 3 of Table 4. For both outcome equations, all the estimated coefficients have reasonable signs but it is surprising that variables like gender, age and diagnoses are not statistically significant from zero on the probability of leaving the sickness benefit scheme. However, the income variable and the number of sickness days in 1999 are statistically significant. Higher income increases the probability of leaving the sickness benefit scheme. Higher number of sickness days in 1999 also increases the probability of leaving the scheme.

For the model with no unobserved heterogeneity, if we condition on a given X-value, the average treatment effect and the effect of treatment on the treated parameters are equivalent. However, if we average over different distributions of X to get the unconditional average treatment effect and the effect of treatment on the treated (average over the unconditional distribution

[7] Any non-experimental evaluation can non-parametrically estimate treatment effects only over the common support region, see Heckman et al. (1998). Due to the relatively low number of observations we do not pursuite non-parametric estimation of treatment effects. For a non-parametric matching strategy see Aakvik (2001).

of X for the average treatment effect and average over the distribution of X conditional on $D = 1$ for the effect of treatment on the treated), the resulting averaged version of these parameters will be different. Using the results of our model without unobserved heterogeneity, the estimated average effect of training (averaging over the unconditional empirical distribution of X) is 12.3 percentage points. The estimated average effect of training on those treated (averaging over the empirical distribution of X conditional on $D = 1$) is 9.3 percentage points.

The trainees have observable characteristics that are associated with a slightly lower effect of training, so that on average their treatment effect is lower than it would be for a person drawn at random from the pool of patients. The unconditional mean difference of leaving the sickness benefit scheme between treated and non-treated when we do not control for observed background variables is 7.3 percentage points. Thus the training effect increases once we adjust for observed variables, and it increases more for a random patient in the sample.

The distributional effects are plotted in Figure 3 for the average treatment effect (ATE) and Figure 4 for the effect of treatment on the treated (TT). In Figures 3 and 4 we have plotted the treatment effects in terms of the probability of treatment. We could also have plotted the treatment effects against different background variables, like age and income, to show the heterogeneous treatment the structural model allows us to estimate. Such plots are available from the authors.

The model with unobserved heterogeneity allows selection both on observables and unobservables. Column 2 in Table 4 shows the estimated parameter vector in the sickness outcome for non-treated, while Column 4 in

Table 4 shows the results for treated patients.[8] The effect of treatment on the treated drops to 5.9 percentage points. Although the factor structure model estimates insignificant factor loadings, the selection specification still affects the estimated effect of treatment. In fact it reduces the treatment effect slightly. This suggests that the unobserved elements of selection into treatment reduce the treatment effect. This is caused by the fact that α_1 is negative and α_0 is positive, although both are not significantly different from zero.

Figure 5 plots each treated person's estimated treatment effect (TT) based on the random coefficient model. The treatment effect is plotted against the propensity score. Different plots of the treatment effect against individual background characteristics are available from the authors. Most of the patients have a positive estimated treatment effect. The maximum estimated treatment effect is 50 percentage points. Figure 6 shows the distribution of TT within the random coefficient framework. The distribution of TT is slightly right skewed. The majority of the treated patients has a treatment effect in the interval [-0.1, 0.3] with a mean of 0.059.

[8] We have used distance to nearest hospital treating patients as our identifying exclusion restriction. This variable was close to zero and with a z-value of 0.1 when we included this variable in the outcome equations, while it was highly significant in the selection equation with a z-value of almost 10. We have no indication that distance to hospital should affect the transition out of the sickness benefits scheme. Graphic variation is also a highly utilized instrument in the literatur on the returns to education, see for instance Card (1995), where it affects the probability of entering college but not subsequent earnings.

6 Conclusions

In this paper we have estimated both mean and distributional treatment effects for back pain patients using observational data. We find a positive effect of treatment of 6 percentage points on the probability of leaving sickness benefits after allowing for selection effects and full heterogeneity in treatment effects. We also find that there are sound arguments for increasing the outpatient program of treating back pain patients. Usually, the average treatment effect is expected to be lower than the effect of treatment on the treated if individuals are rational and can act on the unobserved element of selection on the transition out of sickness benefits. We do not find any significant unobserved selection effects in this data set. This may be explained by the fact that it is difficult for individuals to predict the outcome of treatment of back pain diagnoses.

References

[1] Aakvik, A. (2001), "Bounding a Matching Estimator: The Case of a Norwegian Training Program," *Oxford Bulletin of Economics and Statistics*, 63(1), 115-143.

[2] Aakvik, A., J. Heckman, and E. Vytlacil (2000), "Treatment Effects For Discrete Outcomes When Responses To Treatment Vary Among Observationally Identical Persons: An Application to Norwegian Vocational Rehabilitation Programs," forthcoming *Journal of Econometrics.*

[3] Butler, J. and R. Moffitt (1982), "A Computationally Efficient Quadrature Procedure for the One-Factor Multinomial Probit Model," *Econometrica*, 50(3), pp. 761-764.

[4] Card, D. (1995), "Using Geographic Variation in College Proximity to Estimate the Return to Schooling," in *Aspects of Labour Market Behaviour: Essays in honour of John Vanderkamp*, ed. by Christofides, L., Grant, E., and R. Swidinsky. Toronto: University of Toronto Press, 201-222.

[5] Cox, D. R. (1958) *The Planning of Experiments*, (New York: Wiley).

[6] Fisher, R. A. (1935) *Design of Experiments*, (London: Oliver and Boyd).

[7] Haldorsen, E.M.H, A.L. Grasdal, J.S. Skouen, A.E. Risa, K. Kronholm og H. Ursin (2002). "Is there a right treatment for a particular patient group? Comparison of ordinary treatment, light multidisciplinary treatment, and extensive multidisciplinary treatment for long-term sick-listed employees with musculoskeletal pain," *Pain*, 95, 49-63.

[8] Heckman, J. (1978) "Endogenous Variables in a Simultaneous Equation System," *Econometrica*, 6(4), 931-959.

[9] Heckman, J. (1997) "Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations," *Journal of Human Resources*, 32, 441-462.

[10] Heckman, J. and B. Honoré (1990) "The Empirical Content of the Roy Model," *Econometrica*, 58, 1121-1149.

[11] Heckman, J, H. Ichimura, J. Smith, and P. Todd (1998) "Characterizing Selection Bias Using Experimental Data," *Econometrica*, 66, 1017-1098.

[12] Heckman, J, and R. Robb (1985) "Alternative Methods for Evaluating The Impact of Interventions," in *Longitudinal Analysis of Labor Market Data*, ed. by J. Heckman and B. Singer. New York: Cambridge University Press, 1985, pp. 156-245

[13] Indahl, A., L. Velund, and J. Reikerås (1995). "Good Prognosis for Low Back Pain when left Untreated," *Spine*, 4, 473-477.

[14] Neyman, J. (1923) "Statistical Problems in Agricultural Experiments," Supplement to the *Journal of the Royal Statistical Society*, 2(2), 107-180.

[15] Quandt, R. (1972) "Methods for Estimating Switching Regressions," *Journal of The American Statistical Association*, 67(338), 306-310.

[16] Roy, A. (1951), "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3, 135-146.

[17] Rubin, D. (1978) "Bayesian Inference for Causal Effects: The Role of Randomization," *Annals of Statistics*, 6, 34-58.

Table 1. Variables used in the regressions

Variable name	Definitions
Treatment	Dummy variable for treatment (1=treated, 0=untreated)
Sickness	Dummy variable for leaving sickness benefits (1 = out of sickness, 0 = not out of sickness)
Distance	Distance in kilometers (in logarithms) to nearest hospital treating patients
Male	Dummy variable (1=male, 0=female)
Age	Age in years
Age_sq	Age squared
Disk herniation	Medical diagnoses (1=Disk herination, 0=otherwise)
Low back pain	Medical diagnoses (1=Low back pain, 0=otherwise)
Lumbago	Medical diagnoses (1=Lumbago, 0=otherwise)
Income	Income in 1999, NOK 1000
SM99	Number of sickness days in 1999

Table 2a. Descriptive statistics. Full sample, n=656.

Variable	Mean	Std. Dev.	Min	Max
Distance	4.092251	1.554143	0	7.58
Male	0.586890	0.492767	0	1
Age	41.37957	10.96173	19	67
Age sq.	1832.245	916.3665	361	4489
Disk herniation	0.304878	0.460707	0	1
Low back pain	0.495426	0.500360	0	1
Income	221.9415	111.2831	9.7	1328
SM99	41.80030	62.18449	0	309
Treatment	0.158536	0.365522	0	1
Sickness	0.419207	0.493805	0	1

Table 2b. Descriptive statistics. Treated, n=104.

Variable	Mean	Std. Dev.	Min	Max
Distance	2.523911	1.564216	0	5.55
Male	0.567307	0.497848	0	1
Age	41.06731	10.31001	20	63
Age sq.	1791.798	833.657	400	3969
Disk herniation	0.134615	0.342965	0	1
Low back pain	0.567307	0.497848	0	1
Income	203.9994	80.14556	9.7	500
SM99	42.90385	69.13579	0	283
Treatment	1	0	1	1
Sickness	0.480769	0.502049	0	1

Table 2c. Descriptive statistics. Non-treated, n=552.

Variable	Mean	Std. Dev.	Min	Max
Distance	4.387735	1.364604	0	7.58
Male	0.590579	0.492172	0	1
Age	41.43841	11.08816	19	67
Age sq.	1839.866	931.634	361	4489
Disk herniation	0.336956	0.473098	0	1
Low back pain	0.481884	0.500124	0	1
Income	225.3219	115.9674	10.67	1328
SM99	41.59239	60.8524	0	309
Treatment	0	0	0	0
Sickness	0.407608	0.491835	0	1

Table 3. Probit model of treatment decisions.

Treatment	Coef.	Std.Err.	P>|z|	[95% Conf.	Inter.]	dF/dx
Distance	-0.43137	0.04453	0.000	-0.5186	-0.3440	-7.64215
Male	0.07167	0.14544	0.622	-0.2133	0.3567	1.25985
Age	0.06944	0.04789	0.147	-0.0244	0.1633	1.23025
Age_sq	-0.00080	0.00057	0.160	-0.0019	0.0003	-0.01433
Disk herniation	-0.59433	0.20457	0.004	-0.9952	-0.1933	-9.14283
Low back pain	-0.38125	0.16302	0.019	-0.7007	-0.0617	-6.76448
Income	-0.00192	0.00085	0.025	-0.0036	-0.0002	-0.03409
SM99	0.00123	0.00103	0.234	-0.0007	0.0032	0.02180
Constant	-0.19691	0.96214	0.838	-2.0826	1.6888	

Number of obs = 656, LR chi2(8) = 137.42, Pseudo R2 = 0.2396
Prob > chi2 = 0.0000, Log likelihood = -218.11563,

Table 4. Number of correct predictions in selection equation.

Treatment	Treatment predicted 0	1	Total
0	427	125	552
1	22	82	104
Total	449	207	656

Table 5. Sickness outcomes

Sickness	Participation state		Non-participation state	
	No selection (1)	Selection (2)	No selection (3)	Selection (4)
Male	.0193795	.0228920	.1706879	.1618567
	(.2794496)	(.2815361)	(.1196573)	(.1195558)
Age	-.1957280*	-.1957736*	.0303364	.0221880
	(.1014047)	(.0882371)	(.0377263)	(.0382571)
Age_sq	.0020683*	.0020775*	-.0005983	-.0005012
	(.0012315)	(.0010705)	(.0004499)	(.0004558)
Disk herniation	.1412425	.1791952	.1366012	.2046501
	(.4262911)	(.4333067)	(.1615563)	(.1675164)
Low back pain	-.3287608	-.3134319	-.0708599	-.0474808
	(.2990292)	(.2978321)	(.1533144)	(.1518404)
Income	.0034397*	.0033739*	.0014444**	.0015400**
	(.0018676)	(.0019191)	(.0005599)	(.0005198)
SM99	.0017996*	.0018794	.0025232**	.0024574**
	(.0019139)	(.0019063)	(.0009148)	(.0008875)
Constant	3.668909*	3.776815*	-.9446188	-.8966531
	(1.967996)	(1.758069)	(.7583738)	(.7629868)
Corr		-.1104519		.3069103
		(.2864983)		(.2548640)
Alpha		-.1581400		.4817800

We have used the Huber/White/sandwich estimator of the variance.

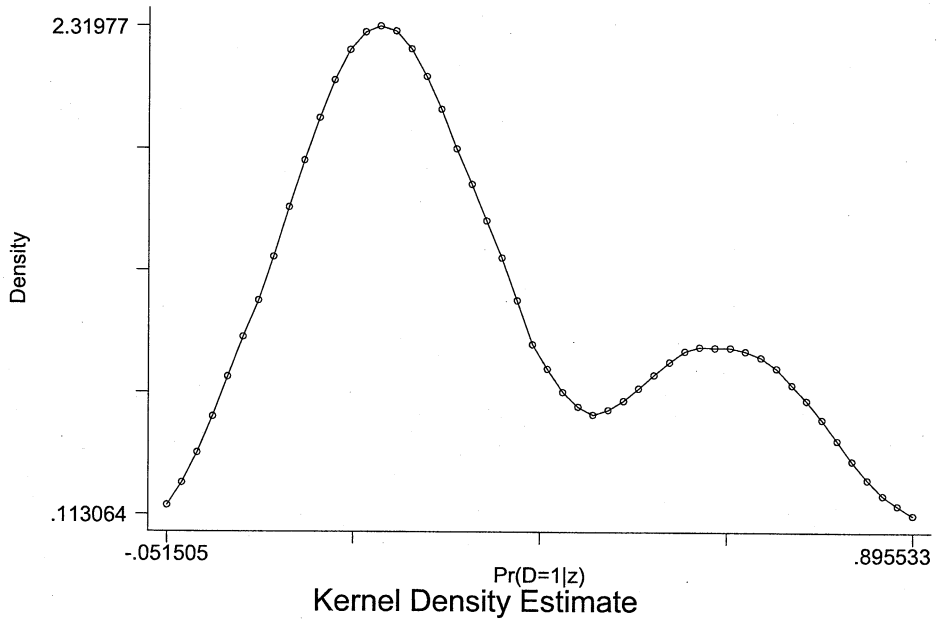

Figure 1. Kernel density estimate of the probability of treatment for participants.

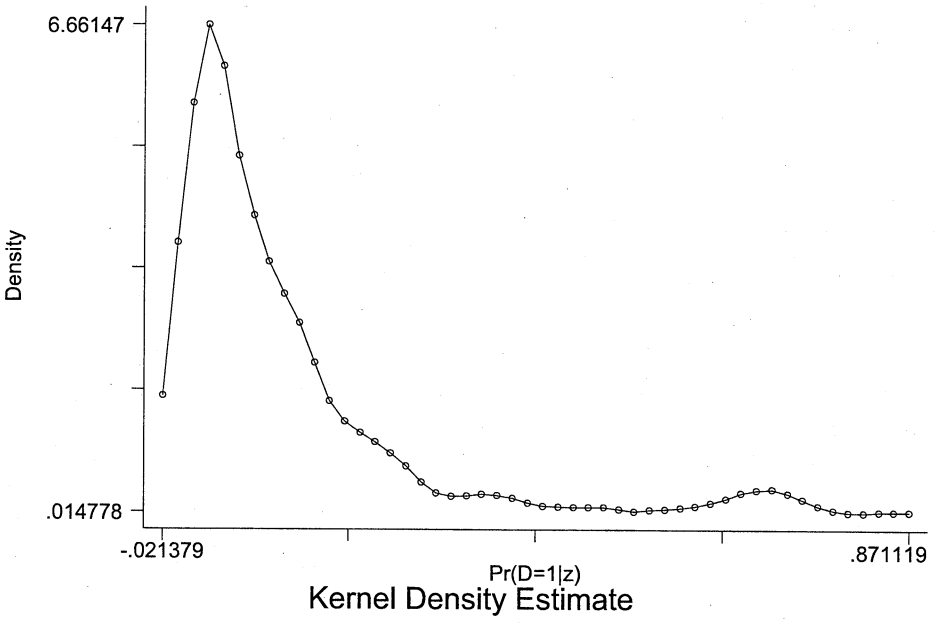

Figure 2. Kernel density estimate of the probability of treatment for non-participants.

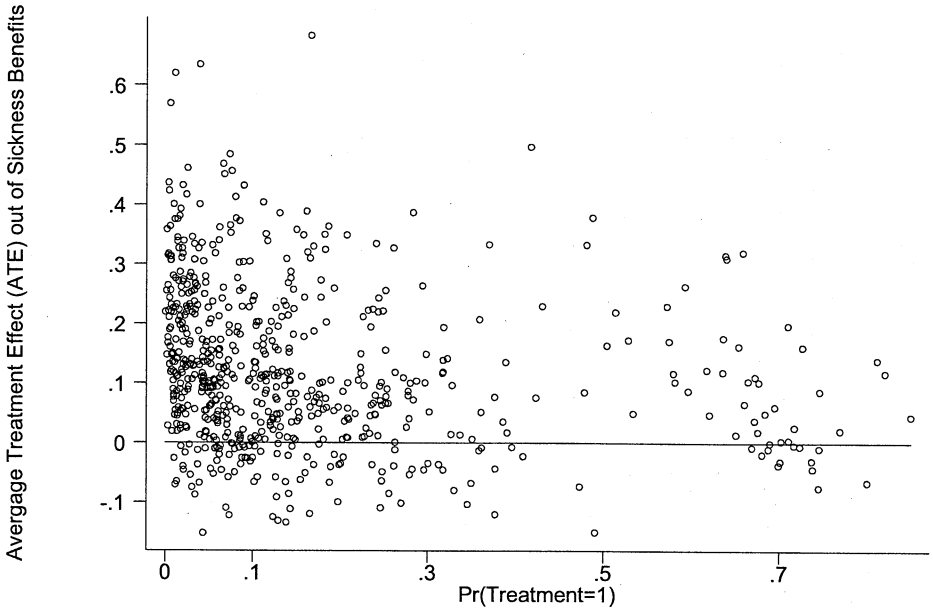

Figure 3. Average treatment effect (ATE) without unobserved selection (mean ATE = 12.3 percentage points).

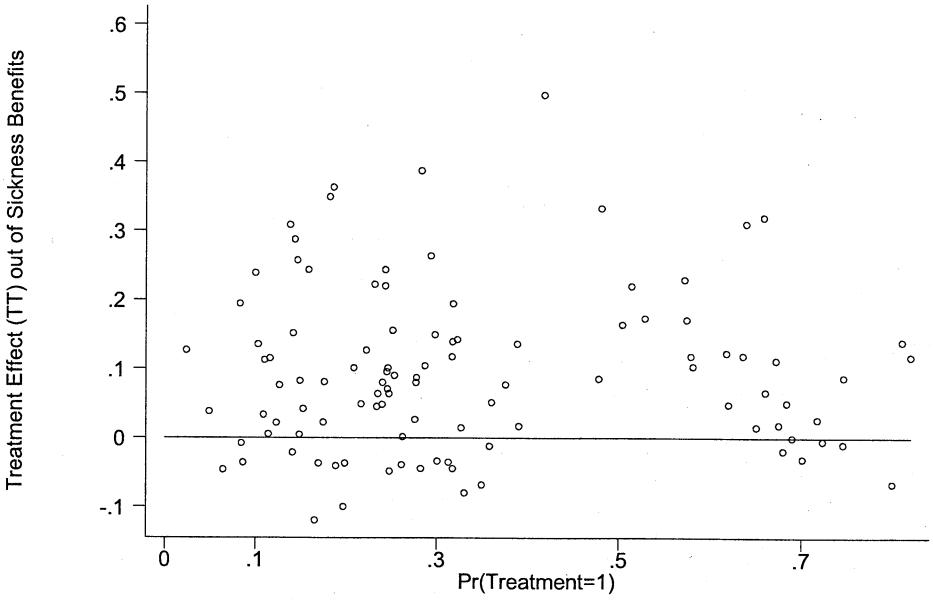

Figure 4. Treatment effect on the treated (TT) without unobserved selection (mean TT = 9.3 percentage points).