Bounding average treatment effects: a linear programming approach

Thomas Demuynck*

Abstract

We show how to obtain bounds on the mean treatment effects by solving a simple linear programming problem. The use of a linear programme is convenient from a practical point of view because it avoids the need to derive closed form solutions. Imposing or omitting monotonicity or concavity restrictions is done by simply adding or removing sets of linear restrictions to the linear programme.
Keywords: treatment effect, linear programming, partial identification
JEL-code: C14, C61, J24

1 Introduction

Mean treatment effects are not point identified when treatments are not (mean)-independently assigned among all subjects or if no valid instruments are available. Although exact identification fails in such cases, it is nevertheless possible to place bounds on the treatment effects (Manski, 1990, 2003). In order to narrow down these bounds, additional monotonicity, concavity or support conditions are frequently imposed. This has led to a growing literature that derives closed form solutions for bounds under varying subsets of conditions (Manski, 1997; Manski and Pepper, 2000, 2009; Boes, 2010; Okumura and Usui, 2014)). These closed form solutions become increasingly complicated when several conditions are imposed simultaneously.

Linear programming methods have previously been applied in analysis of treatment response by Balke and Pearl (1997), Manski (2007) and Manski

*Maastricht University, Tongerstraat 53, 6370 Maastricht, Netherlands. Email: t.demuynck@maastrichtuniversity.nl.

and Pepper (2013) and it has been used by Honoré and Tamer (2006) for partial identification of dynamic nonlinear panel data models and by Molinari (2008) for dealing with the problem of data errors in discrete variables.

In this note, we demonstrate that it is possible to formalize monotonicity, concavity and support conditions by by means of linear inequalities. From this, it follows that bounds on the average treatment effects can be obtained by solving a simple linear programming problem. Adding or omitting conditions simply amounts to including or removing subsets of linear inequalities from the programme. We provide a short illustration using data from the U.S. National Longitudinal Survey of Youth.

2 The linear programming approach

We use the same framework as in Manski (1997). There is a probability space (J, Ω, P) of individuals. Each member j of the population J has an individual specific outcome function $y_j(.) : T \to Y$, mapping treatments $t \in T$ to outcomes $y_j(t) \in Y$. We assume that the set of possible treatments T is finite. Each individual $j \in J$ has a realized treatment $t_j \in T$ and a realized outcome $y_j = y_j(t_j)$. We assume that we observe the joint distribution of (y_j, t_j). The counterfactual outcomes $y_j(t)$ for $t \neq t_j$ are not observed. The objective is to derive bounds on the value of the mean treatment response $E(y(t)) = \int_J y_j(t)P(j)$ for particular treatment t and the average treatment effects $E(y(t)) - E(y(t'))$ between two distinct treatments t and t'.

By conditioning on the treatments, we can write the mean treatment response as a sum of conditional moments,

$$E(y(t)) = \sum_{z \in T} E(y(t)|z)P(z) \equiv \sum_{z \in T} \pi_{t,z} P(z).$$

Here, we introduced the notation $\pi_{t,z} = E(y(t)|z)$. The value of $\pi_{t,z}$ gives the expected outcome of $y_j(t)$ among all households j who face treatment $t_j = z$. In reality, we only observe (or estimate) $\pi_{z,z} = E(y(z)|z)$ and $P(z)$ for all treatments $z \in T$. This imposes the condition that for all $z \in T$, $\pi_{z,z} = E(y(z)|z)$. Computing an upper and lower bound on the treatment response $E(y(t))$ therefore corresponds to finding solutions to the following linear programming problems.

$$\max_{\pi_{s,z};s,z\in T}(\min) \sum_{z\in T} \pi_{t,z} P(z)$$
$$\text{s.t. } \pi_{z,z} = E(y(z)|z) \qquad \forall z \in T.$$

Bounds on the average treatment effect $E(y(t)) - E(y(t'))$ can be obtained by replacing the objective function by $\sum_{z\in T}(\pi_{t,z} - \pi_{t',z})P(z)$.

In the remaining part of this section, we show that imposing additional conditions on the treatment response and treatment selection amounts to adding restrictions which are linear in the unknowns $\pi_{t,z}$. We will focus on three kinds of restrictions: support restrictions, monotonicity restrictions and concavity restrictions.

Support conditions: Support restrictions simply add bounds on the values of $\pi_{t,z}$. For example if $\pi_{t,z}$ is bounded between $\ell_{t,z}$ and $h_{t,z}$, we add the following set of inequalities to the linear programme,

$$\begin{aligned} \pi_{s,z} &\leq h_{s,z} & \forall s, z \in T \\ \pi_{s,z} &\geq \ell_{s,z} & \forall s, z \in T. \end{aligned}$$

Monotonicity restrictions: The outcome functions $y_j(.)$ satisfies the **monotone treatment response condition (MTR)** if it is increasing in t (Manski, 1997):

$$y_j(t) \leq y_j(t') \text{ if } t \leq t'.$$

Given that $\pi_{t,z}$ is the mean of the values $y_j(t)$ over all individuals j for which $z_j = z$, MTR implies that $\pi_{t,z}$ is also monotone in t. This imposes the following conditions on the variables $\pi_{t,z}$.

$$\pi_{t,z} \leq \pi_{t',z} \qquad \forall t, t', z \in T \text{ with } t \leq t'. \tag{1}$$

MTR imposes monotonicity in the first argument of $\pi_{t,z}$. Monotonicity in the second argument is related to the **monotone treatment selection (MTS)** condition (Manski and Pepper, 2000). MTS requires that subjects that receive a higher treatment have a higher (mean) outcome function. This adds the following set of linear constraints to the linear programme.

$$\pi_{s,z} \leq \pi_{s,z'} \qquad \forall s, z, z' \in T \text{ with } z \leq z'. \tag{2}$$

Concavity conditions: If the outcome function $y_j(t)$ is concave in t, then $\pi_{t,z}$ is also concave in t. This is known as the **concave treatment response condition (CTR)** (Manski, 1997). Concavity can be imposed by adding a sub-linearity condition: there should exist numbers $\lambda_{t,z} \in \mathbb{R}$ such that,

$$\pi_{t',z} - \pi_{t,z} \leq \lambda_{t,z}(t - t') \qquad \forall t, t', z \in T. \tag{3}$$

On the other hand, any set of values $\pi_{t,z}$ that satisfy this set of linear inequalities must be concave in t. In order to see this, let $\tilde{t} \in T$ where $\tilde{t} = \alpha t + (1-\alpha)t'$. Then (3) gives,

$$\begin{aligned}
\pi_{t,z} - \pi_{\tilde{t},z} &\leq \lambda_{\tilde{t},z}(t - \tilde{t}), \\
\pi_{t',z} - \pi_{\tilde{t},z} &\leq \lambda_{\tilde{t},z}(t' - \tilde{t}),
\end{aligned}$$

Adding α times the first inequality with $(1-\alpha)$ times the second gives the following inequality,

$$\alpha \pi_{t,z} + (1-\alpha)\pi_{t',z} \leq \pi_{\tilde{t},z}.$$

This shows that the function $\pi_{t,z}$ is concave in t.

Convexity can be modelled by simply reversing the inequalities (3). Also, the MTR condition can be combined with the CTR condition by requiring that $\lambda_{t,z}, \lambda_{t',z} \geq 0$. Indeed, in this case we obtain that $t \geq t'$ if and only if $\pi_{t,z} \geq \pi_{t',z}$ for all $z \in T$, which is equivalent to MTR.

The model satisfies **concavity in the treatment selection condition (CTS)** (Boes, 2010) if $\pi_{t,z}$ is concave in z. This implies that there exist numbers $\delta_{t,z} \in \mathbb{R}$ such that,

$$\pi_{t,z'} - \pi_{t,z} \leq \delta_{t,z}(z' - z). \qquad \forall t, z, z' \in T \tag{4}$$

On the other hand, any set of values $\pi_{t,z}$ that satisfies these inequalities is concave in z. The CTS condition can be combined with MTS by imposing $\delta_{t,z} \geq 0$.

Finally, one could also assume that $\pi_{t,z}$ is jointly concave in both t and z. This would require the existence of numbers $\lambda_{t,z}$ and $\delta_{t,z} \in \mathbb{R}$ such that for all $t, t', z, z' \in T$,

$$\pi_{t',z'} - \pi_{t,z} \leq \lambda_{t,z}(t' - t) + \delta_{t,z}(z' - z).$$

The linear programme: By combining the different linear constraints, we can obtain bounds on the mean treatment response which imposes various combinations of conditions. For example, if we require MTR, MTS, CTR and CTS with some support conditions, then we can compute the bounds on the treatment response by solving the following linear programme,

$$\max_{\pi_{s,z};s,z\in T}(\min) \sum_{z\in T} \pi_{t,z} P(z)$$

$$\begin{aligned}
\text{s.t. } & \pi_{z,z} = E(y(z)|z) & & \forall z \in T, & \\
& \pi_{s,z} \leq h_{s,z} & & \forall s, z \in T, & \\
& \pi_{s,z} \geq \ell_{s,z} & & \forall s, z \in T, & \\
& \pi_{s,z} - \pi_{s',z} \leq \lambda_{s',z}(s - s') & & \forall s, s', z \in T, & \text{(CTR)} \\
& \pi_{s,z} - \pi_{s,z'} \leq \delta_{s,z'}(z - z'). & & \forall s, z, z' \in T & \text{(CTS)} \\
& \lambda_{s,z} \geq 0 & & \forall s, z \in T, & \text{(MTR)} \\
& \delta_{s,z} \geq 0 & & \forall s, z \in T. & \text{(MTS)}
\end{aligned}$$

3 Illustration

We illustrate our results by studying the returns to schooling. We use the 1996 wave of the U.S. National Longitudinal Survey of Youth 1979 (NLSY79). We use the years of schooling as our treatment variable and log of hourly earnings as the outcome. We follow Okumura and Usui (2014) and divide the treatment into four groups (see table 1). Also, following Manski and Pepper (2000) and Okumura and Usui (2014)), we restrict ourselves to a random sample of white men who reported that they were full-time, year-round workers and not self-employed. This gives a sample of 1210 observations. Table 1 contains the sample estimates of the variables $\pi_{z,z}$ and the probabilities $P(z)$. We construct bounds for various combinations of monotonicity and concavity restrictions and we impose the lower bounds $y_j(t) \geq 0$. The linear programme is easy to compute. Setting up and computing the bounds for all treatment levels for a certain collection of conditions takes on average 1.27 seconds on a standard laptop computer using the MATLAB *linprog* or R library *lpSolve*. The code for the programs are available from http://www.revealedpreferences.org/codes/treatment.zip. Table 1 gives the results.

Given that the bounds are obtained as the solution from a linear maxi-

Table 1: Sample statistics

treatment	schooling	$P(z)$	$\pi_{z,z}$
1	<12	0.0876	2.2932
2	12	0.4165	2.5443
3	>12; $\leq$ 15	0.1818	2.7413
4	16	0.1876	2.9737
5	> 16	0.1264	3.0490

mization or minimization programming problem, the usual bootstrap procedure can not be used to construct confidence intervals (see Andrews (2000) for a formal argument). Given this, we construct 95% confidence intervals for both lower and upper bound by using a subsampling procedure which is asymptotically valid under very weak conditions (see Politis, Romano, and Wolf (1999)). The subsampling procedure is similar to the bootstrap procedure but instead of drawing samples (with replacement) of size n equal to the sample size, subsampling uses samples of size m ($\ll n$) without replacement, where $m/n \to 0$.[1] We focus on the construction of Bonferroni type intervals. We refer to Imbens and Manski (2004) and Stoye (2009) for more details on the intricacies that are involved when constructing confidence intervals for partially identified treatment effects.

4 Conclusion

We have shown how to model monotonicity, concavity and support restrictions on the mean treatment effect as a set of linear inequalities. Given this, we show that it is possible to compute bounds on the treatment effect by solving a simple linear programming model. This is especially convenient from a practical point of view because it avoids the need to compute closed form solutions for every subset of assumptions.

[1] We choose $m = n^{0.6}$.

Table 2: Treatment responses

Assumptions		Treatments 1	2	3	4	5
MTR	lower conf	0.1709	1.1932	1.6985	2.2713	2.677
	lb	0.2009	1.2607	1.7591	2.317	2.7025
	ub	2.7025	∞	∞	∞	∞
	upper conf	2.7254	∞	∞	∞	∞
MTS	lower conf	2.2221	2.2753	1.2876	0.8644	0.3371
	lb	2.2932	2.3214	1.3593	0.9339	0.3855
	ub	∞	∞	∞	∞	3.0490
	upper conf	∞	∞	∞	∞	3.1197
MTR/MTS	lower conf	2.2279	2.4920	2.5926	2.6678	2.6804
	lb	2.2932	2.5223	2.6200	2.6930	2.7025
	ub	2.7025	2.7245	2.8238	2.9833	3.0490
	upper conf	2.7276	2.75191	2.8579	3.0204	3.1347
MTR/CTR	lower conf	1.0918	1.998	2.3856	2.6028	2.6777
	lb	1.1135	2.0261	2.4088	2.6254	2.7025
	ub	2.7025	2.9034	3.6342	4.5311	5.5675
	upper conf	2.7270	2.9388	3.6982	4.6256	5.6891
MTS/CTS	lower conf	2.2281	2.3948	2.1502	1.9777	1.6852
	lb	2.2932	2.4329	2.2006	2.0239	1.7372
	ub	6.5329	3.6241	2.7413	2.9738	3.0490
	upper conf	6.7481	3.7072	2.7985	3.0230	3.12048
MTR/CTR/MTS	lower conf	2.2249	2.4983	2.6037	2.6667	2.6787
	lb	2.2932	2.5223	2.6256	2.6930	2.7025
	ub	2.7025	2.7245	2.8238	2.9589	3.0490
	upper conf	2.7269	2.7491	2.8521	2.9823	3.0875
MTR/CTR/MTS/CTS	lower conf	2.2185	2.4960	2.6055	2.6747	2.6864
	lb	2.2932	2.5223	2.6256	2.6962	2.7057
	ub	2.6736	2.6892	2.7386	2.9329	3.0490
	upper conf	2.6917	2.7093	2.7709	2.9582	3.0794

References

Andrews, D. W. K., 2000. Inconsistency of the bootstrap when a parameter is on the bound of the parameter space. Econometrica 68, 399–405.

Balke, A., Pearl, J., 1997. Bounds on treatment effects from studies with imperfect compliance. Journal of the American Statistical Association 92, 1171–1176.

Boes, S., 2010. Convex treatment response and treatment selection. Working paper 1001, University of Zurch.

Honoré, B. E., Tamer, E., 2006. Bounds on parameters in panel dynamic discrete choice models. Econometrica 74, 611–629.

Imbens, G. W., Manski, C. F., 2004. Confidence interval for partially identified parameters. Econometrica 72, 1845–1857.

Manski, C. E., 2003. Partial Identification of Probability Distributions. Springer-Verlag, New York.

Manski, C. F., 1990. Nonparametric bounds on treatment effects. American Economic Review 80, 319–323.

Manski, C. F., 1997. Monotone treatment response. Econometrica 65, 1311–1334.

Manski, C. F., 2007. Partial identification of counterfactual choice probabilities. International Economic Review 48, 1393–1410.

Manski, C. F., Pepper, J. V., 2000. Monotone instrumental variables: With an application to the returns to schooling. Econometrica 68, 997–1010.

Manski, C. F., Pepper, J. V., 2009. More on monotone instrumental variables. The Econometrics Journal 12, 200–216.

Manski, C. F., Pepper, J. V., 2013. Deterrence and the death penalty. partial identification analysis using repeated cross sections. Journal of Quantitative Criminology 29, 123–141.

Molinari, F., 2008. Partial identification of probability distributions with misclassified data. Journal of Econometrics 144, 81–117.

Okumura, T., Usui, E., 2014. Concave-monotone treatment response and monotone treatment selection: With an application ot the returns to schooling. Quantitative Economics 5, 175–194.

Politis, D. N., Romano, J. P., Wolf, M., 1999. Subsampling. Springer.

Stoye, J., 2009. More on confidence intervals for partially identified parameters. Econometrica 77, 1299–1315.