ONLINE APPENDIX: NONPARAMETRIC INSTRUMENTAL VARIABLE METHODS FOR DYNAMIC TREATMENT EVALUATION

GERARD J. VAN DEN BERG*, PETYO BONEV†, ENNO MAMMEN‡

This appendix contains supplementary materials for our paper "Nonparametric Instrumental Variable Methods for Dynamic Treatment". In particular, in section A.1 we provide proofs of our propositions. In section A.2, we explain how our variables were constructed from the data. In section A.3, we provide additional figures for the analysis of endogeneity section.

Appendix A. Appendix

A.1. **Proofs of propositions.**

Proof of proposition 2.1. First we show that from the no anticipation assumption the following result holds:

$$P(T(t) \geq t \mid X, S(t) = t) = P(T(t') \geq t \mid X, S(t) = t). \tag{A.1}$$

This is so because

$$\begin{aligned} & P(T(t) \geq t \mid X, S(t) = t, V) = \exp(-\Theta_{T(t)}(t \mid X, S(t) = t, V)) \\ \overset{\text{No anticipation}}{=} \; & \exp(-\Theta_{T(t')}(t \mid X, S(t) = t, V)) = P(T(t') \geq t \mid X, S(t) = t, V), \end{aligned}$$

so that we obtain

$$\begin{aligned} & P(T(t) \geq t \mid X, S(t) = t) = \mathbb{E}\left[I_{\{T(t)\geq t\}} \mid X, S(t) = t\right] \\ = \; & \mathbb{E}\left[\mathbb{E}\left[I_{\{T(t)\geq t\}} \mid X, S(t) = t, V\right] \mid X, S(t) = t\right] \\ = \; & \mathbb{E}\left[P(T(t) \geq t \mid X, S(t) = t, V) \mid X, S(t) = t\right] \\ = \; & \mathbb{E}\left[P(T(t') \geq t \mid X, S(t) = t, V) \mid X, S(t) = t\right] \\ = \; & \mathbb{E}\left[\mathbb{E}\left[I_{\{T(t')\geq t\}} \mid X, S(t) = t, V\right] \mid X, S(t) = t\right] = P(T(t') \geq t \mid X, S(t) = t) \end{aligned}$$

where $I_{\{T(s)\in B\}}$ is an indicator function equal to 1 when $T(s) \in B$ (of course from these steps we also see that $P(T(t) \geq t \mid X, S(t) = t, V) = P(T(t') \geq t \mid X, S(t) = t, V)$).

Date: March 3, 2016.

Keywords: hazard rate, duration variable, treatment effects, survival function, noncompliance, regression discontinuity design, unemployment, labor market policy reform, active labor market policy, unemployment benefits.

JEL codes: C14, C41, J64, J65.

* University of Bristol, IFAU-Uppsala, IZA, ZEW.

†MINES ParisTech, PSL Research University, CERNA - Centre for industrial economics, i3, CNRS UMR 9217. Email: petyo.bonev@mines-paristech.fr.

‡Heidelberg University and National Research University Higher School of Economics, Russian Federation. Email: mammen@math.uni-heidelberg.de.

Next, using result (A.1), we show $F_{V|T(t)\geq t,X,S(t)=t} = F_{V|T(t')\geq t,X,S(t)=t}$. Let B be a Borel set. With result (A.1), it holds

$$P(V \in B \mid T(t') \geq t, X, S(t) = t) = P(V \in B \mid T(t) \geq t, X, S(t) = t).$$

Now we show $F_{V|T(t)\geq t,X,S(t)=t} = F_{V|T\geq t,X,S=t,Z=t}$. First we observe that $Z \perp\!\!\!\perp \{T(s), S(z)\} \mid X, V$ and $Z \perp\!\!\!\perp V \mid X$ together imply $Z \perp\!\!\!\perp \{T(s), S(z)\} \mid X$ (Weak Union, see Pearl (2000)). Then, we have

$$P(V \in B \mid T(t) \geq t, X, S(t) = t) = \frac{P(V \in B \mid X, S(t) = t) P(T(t) \geq t \mid X, S(t) = t, V \in B)}{P(T(t) \geq t \mid X, S(t) = t)}.$$

We study the separate components of the right-hand side of the last expression.

(1) With assumptions A3 and A4, it holds

$$P(V \in B \mid X, S(t) = t) = P(V \in B \mid X, S = t, Z = t).$$

(2) Further,

$$P(T(t) \geq t \mid X, S(t) = t, V \in B) = P(T \geq t \mid X, S = t, V \in B, Z = t).$$

(3) Using $Z \perp\!\!\!\perp \{T(s), S(z)\} \mid X$ instead of $Z \perp\!\!\!\perp \{T(s), S(z)\} \mid X, V$, we obtain

$$P(T(t) \geq t \mid X, S(t) = t) = P(T \geq t \mid X, S = t, Z = t)$$

So finally we get the equality

$$\begin{aligned} & P(V \in B \mid T(t) \geq t, X, S(t) = t) \\ = & \frac{P(V \in B \mid X, S = t, Z = t) P(T \geq t \mid X, S = t, V \in B, Z = t)}{P(T \geq t \mid X, S = t, Z = t)} \\ = & P(V \in B \mid T \geq t, X, S = t, Z = t) \end{aligned}$$

□

Proof of corollary 2.1. With proposition 2.1,

$$\begin{aligned} TE(t,t',a) &= \mathbb{E}\Big[P(T(t) \in [t, t+a) \mid T(t) \geq t, X, V, S(t) = t) \mid T(t) \geq t, X, S(t) = t\Big] \\ &- \mathbb{E}\Big[P(T(t') \in [t, t+a) \mid T(t') \geq t, X, V, S(t) = t) \mid T(t') \geq t, X, S(t) = t\Big] \\ &= P(T(t) \in [t, t+a) \mid T(t) \geq t, X, S(t) = t) - P(T(t') \in [t, t+a) \mid T(t') \geq t, X, S(t) = t). \end{aligned}$$

□

Lemma A.1. *Set* $B = [t, t+a)$ *where* $a \leq t' - t$. *Under Assumptions A1-A4, it holds for all* $\infty \geq t' \geq t \geq 0$ *that*

(A.2)

$$P(T(t) \in B \mid T(t) \geq t, X, S(t) = t) = P(T \in B \mid T \geq t, X, S = t, Z = t),$$

(A.3)

$$P(T(t') \in B \mid T(t') \geq t, X, S(t) = \infty) = P(T \in B \mid T \geq t, X, S = \infty, Z = t) \qquad \textit{and}$$

(A.4)

$$P(T(t') \in B \mid T(t') \geq t, X) = P(T \in B \mid T \geq t, X, Z = t').$$

Proof of Lemma A.1. First, observe that with randomization and consistency, it holds

$$P(T(t) \in B \mid X, S(t) = t) = P(T \in B \mid X, S = t, Z = t),$$
$$P(T(t) \geq t \mid X, S(t) = t) = P(T \geq t \mid X, S = t, Z = t),$$

so that

$$P(T(t) \in B \mid T(t) \geq t, X, S(t) = t) = P(T \in B \mid T \geq t, X, S = t, Z = t)$$

where the r.h.s of the equality consists only of observables.

Next, we have

$$\begin{aligned} & P(T \in B \mid X, S = \infty, Z = t) = P(T(\infty) \in B \mid X, S = \infty, Z = t) \\ = \; & P(T(\infty) \in B \mid X, S(t) = \infty, Z = t) = P(T(\infty) \in B \mid X, S(t) = \infty) \\ = \; & P(T(t') \in B \mid X, S(t) = \infty), \end{aligned}$$

where the first and the second equalities follow due to consistency, the third due to randomisation and the fourth due to no anticipation. Equality (A.4) follows analogically. □

Lemma A.2. *Under Assumptions A1-A4, it holds for all* $\infty \geq t' \geq t \geq 0$ *that*

$$P(S(t) = t \mid T(t) \geq t, X) = P(S = t \mid T \geq t, X, Z = t), \tag{A.5}$$
$$P(S(t) = t \mid T(t') \geq t, X) = P(S(t) = t \mid T(t) \geq t, X). \tag{A.6}$$

Proof of Lemma A.2. First, it holds

$$\begin{aligned} & P(S = t \mid T \geq t, X, Z = t) = \frac{P(T \geq t \mid S = t, X, Z = t) P(S = t \mid X, Z = t)}{P(T \geq t \mid X, Z = t)} \\ = \; & \frac{P(T(t) \geq t \mid S(t) = t, X) P(S(t) = t \mid X)}{P(T(t) \geq t \mid X)} = P(S(t) = t \mid T(t) \geq t, X), \end{aligned}$$

where the second equality follows with assumptions A1-A4.

Next,

$$\begin{aligned} & P(S(t) = t \mid T(t') \geq t, X) = \frac{P(S(t) = t, T(t') \geq t \mid X)}{P(T(t') \geq t \mid X)} \\ = \; & \frac{P(T(t') \geq t \mid S(t) = t, X) P(S(t) = t \mid X)}{P(T(t) \geq t \mid X)} = \frac{P(T(t) \geq t \mid S(t) = t, X) P(S(t) = t \mid X)}{P(T(t) \geq t \mid X)} \\ = \; & P(S(t) = t \mid T(t) \geq t, X), \end{aligned}$$

where the second equality holds due to no anticipation. □

Proof of proposition 2.2. First, write

$$\begin{aligned} & P(T(t') \in [t, t+a) \mid T(t') \geq t, X) \\ = \; & P(T(t') \in [t, t+a) \mid T(t') \geq t, X, S(t) = t) P(S(t) = t \mid T(t') \geq t, X) \\ + \; & P(T(t') \in [t, t+a) \mid T(t') \geq t, X, S(t) = \infty) P(S(t) = \infty \mid T(t') \geq t, X), \end{aligned} \tag{A.7}$$

and then express $P(T(t') \in [t, t+a) \mid T(t') \geq t, X, S(t) = t)$ in terms of the other three components of equality (A.7). Plugging in the results of lemma A.1 and lemma A.2, we obtain for $F_{C,0} := P(T(t') \in B \mid T(t') \geq t, X, S(t) = t)$

$$\begin{aligned} & P(T(t') \in B \mid T(t') \geq t, X, S(t) = t) \\ = \; & \frac{P(T \in B \mid T \geq t, X, Z = t') - P(T \in B \mid T \geq t, X, Z = t, S = \infty) P(S = \infty \mid T \geq t, X, Z = t)}{P(S = t \mid T \geq t, X, Z = t)}. \end{aligned}$$

Finally, with $F_{C,1} := P(T(t) \in B \mid T(t) \geq t, X, S(t) = t)$, the treatment effect is equal to $F_{C,1} - F_{C,0}$ which after simplification is equal to

$$\frac{P(T \in B \mid T \geq t, X, Z = t) - P(T \in B \mid T \geq t, X, Z = t')}{P(S = t \mid T \geq t, X, Z = t)}.$$

□

Proof of proposition 2.7. For notational simplicity we drop the dependence on 0 and x_0. First note, that the results of Theorem 1 Nielsen and Linton (1995) remain valid at the boundary when we replace the symmetric kernel k with its boundary counterpart k_+ and adapt the constants. The validity of proposition 2.7 i) follows from $\sqrt{nb^{q+1}}((\widehat{\Psi} - \Psi^*) = \frac{\sqrt{nb^{q+1}}}{\widehat{p}_1}((\widehat{\theta}_1 - \theta_1^*) - (\widehat{\theta}_2 - \theta_2^*))$, the independence of $(\widehat{\theta}_1 - \theta_1^*)$ and $(\widehat{\theta}_2 - \theta_2^*)$, and the adapted proof of Theorem 1 i) in Nielsen and Linton (1995). Next, it holds

$$b^{-2}(\Psi^* - \Psi) = \frac{b^{-2}}{\widehat{p}_1}((\theta_1^* - \theta_1) - (\theta_2^* - \theta_2)) + b^{-2}(\theta_1 - \theta_2)(\frac{1}{\widehat{p}_1} - \frac{1}{p_1}). \tag{A.8}$$

The second term on the right-hand side of (A.8) is equal to $o_p(1)$ when b is of order $O(n^{-1/(q+5)})$ or $o(n^{-1/(q+5)})$. Proposition 2.7 ii) follows with Theorem 1 b) in Nielsen and Linton (1995). Finally, proposition 2.7 iii) follows directly from the adapted proof of Theorem 1 c) Nielsen and Linton (1995) and the continuous mapping theorem. □

A.2. **Description of variables.** The variables used in our empirical application have been constructed in the following way:

- The variable **age** gives the age at the begin of the unemployment spell and is defined as the year in which the spells begins minus the year of birth.
- **Marital status** consists of four categories: single, married, divorced and widowed.
- the variable for **educational level** summarizes the 31 categories used in the administrative data set into 6 categories according to the highest degree attained. The correspondence is roughly as follows: value 1 if the degree is in niveau I and II (university degree, maîtrise and licence), value 2 if the degree is in niveau III - BTS and DUT (brevet de technicier supérieur and diplôme univeritaire de technologie, respectively, both technical degrees obtained in 2 years after high school), value 3 for all Baccalauréat (high school degree, the general part of lycée) diplomas and for all dropouts from niveau III, 4 for all BEP ,CEP (professional Baccalauréat, specialised part of lycée) and all dropouts from Baccalauréat, 5 for BEPC (brevet d'études du premier cycle, junior high school), and 6 for below.
- The variable **experience** states the number of years of experience in the job (type and position), which the individual is looking for. The types of jobs are specified in an administrative nomenclature table (ROME table). There are several hundred different types.

- The **job type** variable contains general information about the type of the activity in the job preceding the current unemployment spell. It summarizes the 9 administrative categories into 6 categories: white collar skilled, white collar unskilled, technical, supervisor (a production team leader) and manager. This summarized categorization is in line with existing literature, see for example Crépon et al. (2010). The initial administrative variable is contained in the FH data set. This holds also for the variable, which states which job is the unemployed looking for, while the following employment type and position is contained in the DADS data set. Unfortunately, there is no clear matching between the variables from the two different data sets, which leads to some unclarity regarding the question whether the unemployed actually found the job he/she was looking for. This restricts our definition of censoring. Therefore, in this application each observation with known job destination is considered uncensored.
- **Censoring indicator**: there are several possibilities, when an observation is considered as censored. These are:
  - when the unemployment spell in the data set is not finished at the time of the data collection, or
  - when the individual exits the labor market. This includes exits to maternity, accident, illness or invalidity, invalidity pension, military service, administrative change of insurance status, attrition because of insufficient administrative control, dropout because of irregular notifications, and other, unspecified reasons. While reasons such as maternity, military services and invalidity pension are normally known well in advance by the unemployed and can therefore be related to search activity (as well as to compliance behavior), they represent a small fraction of the observations.
- **Unemployment history**: it is constructed as a binary variable which equals 1 if the individual had been already unemployed before the last employment spell. There are various ways to define unemployment history. One example is the total length of previous unemployment spells. Alternatively, one could take the number of unemployment spells, or both. All possibilities suffer from disadvantages. The last possibility seems to provide the most complete information, but it also demands more data, since it provides many different categories. The total length of previous unemployment lacks any information about the lengths of the separate spells, and the number of spells alone doesn't give any information about the length of unemployment. The binary indicator also does not provide any information at all about the dispersion of previous unemployment, but it is easy to understand and requires only two categories, which makes it computationally attractive. Additional, more serious drawback for the other two indicators is, that the data set is left censored: the earliest information about employment is from 1993. This problem is less severe, if one only looks at the indicator of having been unemployed.

A.3. **Analysis of endogeneity.** This section contains additional figures for the analysis of endogeneity section of our paper.

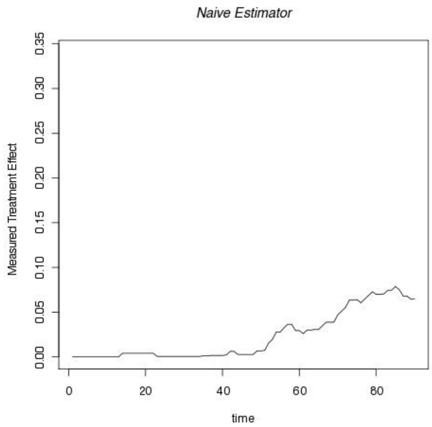
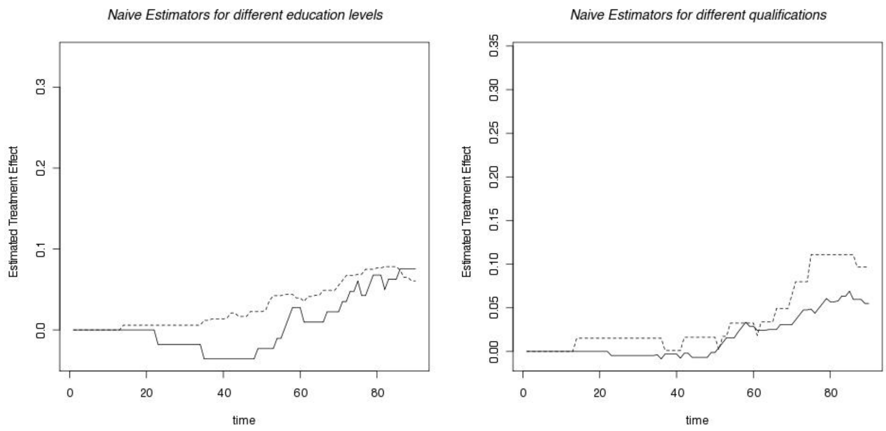


FIGURE 1. A naive estimator. Male vs. Female. Time measured in days.

FIGURE 2. A naive estimator for subgroups



(A) Education. Low educated dashed line

(B) Qualification. Blue collars dashed line

References

Crépon, B., Ferracci, M., Jolivet, G., and Van den Berg, G. J. (2010). Analyzing the anticipation of treatments using data on notification dates. IZA Discussion paper 5265, Institute for the Study of Labor (IZA), Bonn, Germany.

Nielsen, J. P. and Linton, O. B. (1995). Kernel estimation in a nonparametric marker dependent hazard model. *The Annals of Statistics*, 23(5):1735–1748.

Pearl, J. (2000). *Causality. Models, reasoning, and inference*. Cambridge University Press.