Errata to

“The performance of estimators based on the propensity score”

by Martin Huber, Michael Lechner, and Conny Wunsch,

Journal of Econometrics, Volume 175, Issue 1, July 2013, Pages 1‑21

Table 2.1: Descriptive statistics of the 'population'

Variable	Treated		Control		Standardized difference	Probit est. of selection equat.	
	mean	std.	mean	std.	in %	Marg.eff. in %	std. error
3 years since beginning of UE spell	0.63	0.48	0.56	0.50	9	-	-
some unsubsidized employ.							
av. monthly earnings (EUR)	1193	1115	1041	1152	9	-	-
Constant term	-	-	-	-	-	-	-
Age / 10	3.67	0.84	3.56	1.11	8	7.3	0.5
... squared / 1000	1.42	0.63	1.39	0.85	3	-9.1	0.6
20 - 25 years old	0.22	0.41	0.36	0.48	22	0.9	0.2
Women	0.57	0.50	0.46	0.50	15	-5.8	1.5
Not German	0.11	0.31	0.19	0.39	16	-0.5	0.1
Secondary degree	0.32	0.47	0.22	0.42	15	1.1	0.1
University entrance qualification	0.29	0.45	0.20	0.40	15	1.0	0.1
No vocational degree	0.18	0.39	0.34	0.47	26	-0.3	0.1
At least one child in household	0.42	0.49	0.28	0.45	22	-0.2	0.1
Last occupation: Non-skilled worker	0.14	0.35	0.21	0.41	13	0.3	0.1
Last occupation: Salaried worker	0.40	0.49	0.22	0.41	29	1.8	0.2
Last occupation: Part time	0.22	0.42	0.16	0.36	12	2.1	0.3
UI benefits: 0	0.33	0.47	0.44	0.50	16	-0.6	0.1
> 650 EUR per month	0.26	0.44	0.22	0.41	7	0.7	0.1
Last 10 years before UE: share empl.	0.49	0.34	0.46	0.35	8	-1.4	0.2
share unemployed	0.06	0.11	0.06	0.11	1	-2.5	0.5
share in programme	0.01	0.04	0.01	0.03	9	5.1	1.2
Last year before UE: share minor em.*	0.07	0.23	0.03	0.14	15	-1.0	0.7
share part time	0.16	0.33	0.11	0.29	10	-1.0	0.2
share out-of-the labour force (OLF)	0.28	0.40	0.37	0.44	14	-1.3	0.2
Entering UE in 2000	0.26	0.44	0.19	0.39	13	1.6	0.2
2001	0.29	0.46	0.26	0.44	5	0.9	0.1
2003	0.20	0.40	0.27	0.44	12	0.0	0.1
Share of pop. living in/ close to big city	0.76	0.35	0.73	0.37	6	0.4	0.1
Health restrictions	0.09	0.29	0.15	0.36	13	-0.6	0.1
Never out of labour force	0.14	0.34	0.11	0.31	6	0.6	0.2
Part time in last 10 years	0.35	0.48	0.29	0.45	9	-0.5	0.1
Never employed	0.11	0.31	0.20	0.40	17	-1.0	0.1
Duration of last employment > 1 year	0.41	0.49	0.43	0.50	4	-0.6	0.1
Av. earn. last 10 yrs when empl./1000	0.59	0.41	0.52	0.40	13	-0.4	0.2
Women x age / 10	2.13	1.95	1.65	1.94	17	2.6	0.6
x squared / 1000	0.83	0.85	0.65	0.90	15	-2.6	0.8
x no vocational degree	0.09	0.28	0.16	0.36	15	-0.9	0.1
x at least one child in household	0.32	0.47	0.17	0.37	25	0.9	0.2
x share minor employment last year	0.06	0.22	0.02	0.13	16	3.2	0.7
x share OLF last year	0.19	0.36	0.18	0.35	3	1.0	0.2
x average earnings last 10 y. if empl.	0.26	0.34	0.19	0.30	16	-1.0	0.2
x entering UE in 2003	0.10	0.30	0.13	0.33	6	-0.6	0.1
$x_i\hat{\beta}$	-1.7	0.42	-2.1	0.42	68	-	-
$\Phi(x_i\hat{\beta})$	0.06	0.03	0.05	0.03	59	-	-
Number of obs., Pseudo-R^2 in %	3266		114349			3.6	

Note: * Minor em. is minor employment with earnings of no more than 400 EUR per month, which are not or only partially subject to social insurance contributions. 'binary': indicates a binary variable (standard deviation can be directly deduced from mean). $\hat{\beta}$ is the estimated probit coefficients and $\Phi(a)$ is the c.d.f. of the standard normal distribution evaluated at *a. Pseudo-R^2* is the so-called Efron's R^2 $\left(1-\sum_{i=1}^{N}[d_i-\hat{p}(x_i)]\Big/\sum_{i=1}^{N}\left[d_i-\sum_{i=1}^{N}(d_i)/N\right]\right)$. The *Standardized difference* is defined as the difference of means normalized by the square root of the sum of estimated variances of the particular variables in both subsamples (see e.g. Imbens and Wooldridge, 2009, p. 24). Marg. effect: Average marginal effect based on discrete changes for binary variables and derivatives otherwise.

Table A.1: Matching protocol for the estimation of a counterfactual outcome and the effects

Step A-1	Choose one observation in the subsample defined by *d*=1 and delete it from that pool.
Step B-1	Find an observation in the subsample defined by *d*=0 that is as close as possible to the one chosen in step A-1) in terms of either (i) $p(x)$ (matching on the propensity score only), or (ii) $p(x)$ and additional predictors (matching on the propensity score and a subset of *X*). In the latter case, 'closeness' is based on the Mahalanobis distance, in which $p(x)$ and the additional predictors may or may not be weighted.
Step C-1	Repeat A-1) and B-1) until no observation with *d*=1 is left.
Step D-1	Compute the maximum distance (*maxdist*) obtained for any comparison between a member of the reference distribution and matched comparison observations. Alternatively, one may also compute the quantile at a particular rank in the distribution of distances (*quantdist*).
Step A-2	Repeat A-1).
Step B-2	Repeat B-1). If possible, find other observations in the subsample of *d*=0 that are at least as close as *R* * *maxdist* or *R** *quantdist*, respectively, to the one chosen in step A-2), where *R* denotes the radius multiplier. Do not remove these observations, so that they can be used again. Compute weights for all chosen comparisons observations that are proportional to their distance. Normalise the weights such that they add to one.
Step C-2	Repeat A-2) and B-2) until no participant in *d*=1 is left.
Step D-2	For any potential comparison observation, add the weights obtained in A-2) and B-2).
Step E	Using the weights $w(x_i)$ of the comparison observations obtained in D-2), run a weighted linear regression of the outcome variable on an intercept, the propensity score, its square, and any further variables used to define the distance.
Step F-1	Predict the potential outcome $y^0(x_i)$ of every observation using the coefficients of this regression: $\hat{y}^0(x_i)$.
Step F-2	Estimate the bias of the matching estimator for $E(Y^0 \mid D=1)$ as: $\sum_{i=1}^{N} \frac{(1-d_i)w_i\hat{y}^0(x_i)}{N_0} - \frac{d_i\hat{y}^0(x_i)}{N_1}$.
Step G	Using the weights obtained by weighted matching in D-2), compute a weighted mean of the outcome variables in *d=0*. Subtract the bias to this estimate to get $\widehat{E(Y^0 \mid D=1)}$.

Note: For estimation of the ATENT the counterfactual distribution can be obtained by replacing d by 1-d and repeating steps A to G.