Peter Hull September 28, 2018

Subtracting the Propensity Score in Linear Models

Often in quasi-experimental designs, a binary treatment or instrument is assumed to be as-good-as-randomly assigned conditional on a set of observed controls. Since Rosenbaum and Rubin (1981) researchers have leveraged this assumption by matching or weighting observations via the propensity score, which gives the conditional probability of treatment assignment. This note motivates an alternative use of the propensity score, as an additive correction term in ordinary least squares (OLS) or instrumental variables (IV) regressions. In particular I show how subtracting the propensity score off the treatment or instrument identifies a convex average of heterogeneous treatment effects, with identified weights. This extends an earlier OLS result of Angrist (1998), in which one controls for a set of mutually-exclusive group indicators and implicitly subtracts a linear propensity score estimate off the treatment variable. Here the result is shown with no restrictions on the propensity score, which may even be zero or one for some values of the controls.

Formally, consider the linear IV regression of

$$Y_i = \alpha + \beta D_i + \epsilon_i \tag{1}$$

$$D_i = \gamma + \pi(Z_i - P(X_i)) + \eta_i, \tag{2}$$

for an outcome Y_i, treatment D_i, binary instrument Z_i, vector of controls X_i, and propensity score $P(x) = Pr(Z_i = 1 \mid X_i = x)$. For simplicity suppose D_i is binary, though the result extends to non-binary treatment by the results of Angrist, Imbens, and Rubin (1996) and Angrist, Graddy, and Imbens (2000). Write $Y_i = Y_{0i}(1-D_i)+Y_{1i}D_i$ and $D_i = D_{0i}(1-Z_i)+D_{1i}Z_i$, where $(Y_{0i}, Y_{1i}, D_{0i}, D_{1i})$ is a vector of potential outcomes and treatments.[1] Furthermore assume:

A1 *(Conditional independence)*: $(Y_{0i}, Y_{1i}, D_{0i}, D_{1i}) \perp\!\!\!\perp Z_i \mid X_i$

A2 (*Monotonicity*): $Pr(D_{1i} \geq D_{0i}) = 1$

Here A1 makes the instrument as-good-as-randomly assigned given the controls, while A2 imposes a monotone effect of the instrument on treatment. Note that this setup accommodates the case of conditional random assignment of a binary treatment, with $D_i = Z_i$ and A2 satisfied trivially. In this case $\pi = 1$ and β is equivalent to the coefficient from regressing Y_i on $D_i - P(X_i)$.[2]

In the case of constant X_i (and thus constant $P(X_i)$), Imbens and Angrist (1994) show that A1 and A2 ensure β captures the local average treatment effect (LATE), $E[Y_{1i} - Y_{0i} \mid D_{1i} > D_{0i}]$. Here I show that in general the regression identifies a weighted average of conditional LATEs, $\beta(x) = E[Y_{1i} - Y_{0i} \mid D_{1i} > D_{0i}, X_i = x]$, with identified weights. In the case of conditionally random treatment, this shows the OLS regression of Y_i on $D_i - P(X_i)$ recovers a weighted average of conditional average treatment effects, $E[Y_{1i} - Y_{0i} \mid X_i = x]$.

The proof starts from observing that $\beta = \rho/\pi$, where ρ comes from the reduced form regression

$$Y_i = \mu + \rho(Z_i - P(X_i)) + \nu_i. \tag{3}$$

Note that

$$\begin{aligned}
\rho &= \frac{Cov(Y_i, Z_i - P(X_i))}{Var(Z_i - P(X_i))} \\
&= \frac{E[Cov(Y_i, Z_i \mid X_i)]}{E[Var(Z_i \mid X_i)]} \\
&= E\left[\frac{\sigma_Z^2(X_i)}{E[\sigma_Z^2(X_i)]}\left(E[Y_i \mid Z_i = 1, X_i] - E[Y_i \mid Z_i = 0, X_i]\right)\right],
\end{aligned} \tag{4}$$

[1] As usual writing (Y_{0i}, Y_{1i}) without instrument subscripts imposes an exclusion restriction, that Z_i only affects outcomes through its effect on D_i. I also implicitly assume the vectors $(Y_{0i}, Y_{1i}, D_{0i}, D_{1i}, Z_i, X_i)$ are independently and identically distributed, satisfying a stable unit treatment value assumption. Finally, I assume a nonzero first stage, $\pi \neq 0$, so that the regression is well-defined. Note that we do not require a bounded propensity score, unlike with typical approaches. That is, $P(X_i)$ may equal zero or one with positive probability.

[2] A straightforward extension shows that when a non-binary D_i is as-good-as-randomly assigned given X_i, the regression of Y_i on $D_i - M(X_i)$ identifies a weighted average causal effect, where $M(x) = E[D_i \mid X_i = x]$ generalizes the propensity score. Contrast this with the generalized propensity score of Hirano and Imbens (2004), which gives the full conditional distribution of D_i given X_i.

where $\sigma_Z^2(X_i) = Var(Z_i \mid X_i)$ denotes the conditional instrument variance. Here the second equality follows by the law of total covariance, and the third by the fact that $Cov(Y_i, Z_i \mid X_i)/Var(Z_i \mid X_i) = E[Y_i \mid Z_i = 1, X_i] - E[Y_i \mid Z_i = 0, X_i]$. A similar derivation holds for π; thus

$$\begin{aligned}
\beta &= \rho/\pi \\
&= \frac{E\left[\sigma_Z^2(X_i)\left(E[Y_i \mid Z_i = 1, X_i] - E[Y_i \mid Z_i = 0, X_i]\right)\right]}{E\left[\sigma_Z^2(X_i)\left(E[D_i \mid Z_i = 1, X_i] - E[D_i \mid Z_i = 0, X_i]\right)\right]} \\
&= E\left[\frac{\sigma_Z^2(X_i)\pi(X_i)}{E\left[\sigma_Z^2(X_i)\pi(X_i)\right]} \frac{E[Y_i \mid Z_i = 1, X_i] - E[Y_i \mid Z_i = 0, X_i]}{E[D_i \mid Z_i = 1, X_i] - E[D_i \mid Z_i = 0, X_i]}\right] \\
&= E\left[\frac{\sigma_Z^2(X_i)\pi(X_i)}{E\left[\sigma_Z^2(X_i)\pi(X_i)\right]}\beta(X_i)\right], \qquad (5)
\end{aligned}$$

where $\pi(X_i) = E[D_i \mid Z_i = 1, X_i] - E[D_i \mid Z_i = 0, X_i]$ and the last line follows by A1 and A2 via the Imbens and Angrist (1994) result.

Equation (5) shows that the IV coefficient in (1) captures a weighted average of conditional local average treatment effects, with weights that are proportional to the conditional variance of the instrument and the conditional first stage $\pi(X_i)$. Under A1 and A2, $\pi(X_i)$ captures the conditional share of instrument compliers, $Pr(D_{1i} > D_{0i} \mid X_i) \geq 0$, so the weighting scheme is convex. Note that in the $D_i = Z_i$ case, $\pi(X_i) = 1$ and the OLS coefficient β captures a variance-weighted average of conditional average treatment effects, as in Angrist (1998). Alternatively when propensity scores are constant $\sigma_Z^2(X_i) = E[\sigma_Z^2(X_i)]$; then the $\beta(X_i)$ are weighted only by the conditional complier shares, yielding the unconditional LATE.[3]

Rather than matching on or weighting by the propensity score, researchers may therefore wish to subtract it off the treatment or instrument in OLS or IV regressions. In some cases the score may be known, such as in a randomized control trial or when quasi-experimental variation is generated from a random mechanism that can be simulated with arbitrary precision (Abdulkadiroğlu et al., 2017; Aronow and Samii, 2017). While it is outside the scope of this note to study large-sample properties of regressions that subtract the *estimated* propensity score, they may prove favorable relative to some properties of matching or weighting estimators (King and Nielsen, 2016; Kahn and Tamer, 2010).

A clear drawback to the propensity score adjustment approach is it does not, in general, produce a population average causal effect but rather a convex average of conditional causal effects. An exception is when the conditional LATEs and conditional complier shares are mean-independent of X_i (or, more simply, when conditional LATEs are constant); in this case it can be shown from equation (6) that $\beta = E[Y_{1i} - Y_{0i} \mid D_{1i} > D_{0i}]$. Consequently in the OLS case where $D_i = Z_i$, mean-independence of treatment effects with respect to the controls ensures $\beta = E[Y_{1i} - Y_{0i}]$.

[3] It is straightforward to verify that the auxiliary regression of Z_i on $P(X_i)$ produces a coefficient of one and intercept of zero, and therefore a residual of $Z_i - P(X_i)$. By the Frisch-Waugh-Lovell theorem, (5) is thus also identified by the regression of Y_i on D_i that instruments with Z_i and controls for $P(X_i)$.

References

Abdulkadiroğlu, A., J. D. Angrist, Y. Narita, and P. A. Pathak. "Research Design Meets Market Design: Using Centralized Assignment for Impact Evaluation," *Econometrica*, 85 (2017), 1373-1432.

Angrist, J. D. "Estimating the Labor Market Impact of Voluntary Military Service Using Social Security Data on Military Applicants," *Econometrica*, 66 (1998), 249-288.

Angrist, J. D., K. Graddy., and G. W. Imbens. "The Interpretation of Instrumental Variables Estimators in Simultaneous Equations Models with an Application to the Demand for Fish," *Review of Economic Studies*, 67 (2000), 499-527.

Angrist, J. D. and G. W. Imbens. "Two-stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity," *Journal of the American Statistical Association*, 90 (1995), 431-442.

Aronow, P. M. and C. Samii. "Estimating Average Causal Effects Under General Interference, with an Application to a Social Network Experiment," *Annals of Applied Statistics*, 11 (2017), 1912-1947.

Hirano, K. and G.W. Imbens. "The Propensity Score with Continuous Treatments," in A. Gelman and X. Meng (eds.), *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*, Wiley.

Imbens, G. W. and J. D. Angrist. "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62 (1994), 467-475.

Khan, S. and E. Tamer. "Irregular Identification, Support Conditions, and Inverse Weight Estimation," *Econometrica*, 78, 2021-2042.

King, G. and R. Neilsen. "Why Propensity Scores Should Not Be Used for Matching," Working Paper (2016).

Rosenbaum and Rubin. "The Central Role of the Propensity Score in Observational Studies for Causal Effect," *Biometrika*, 70 (1981), 41-55.