Supplement to “Does affirmative action lead to mismatch? A new test and evidence”  

Peter Arcidiacono  
Department of Economics, Duke University and NBER

Esteban M. Aucejo  
Department of Economics, Duke University

Hanming Fang  
Department of Economics, University of Pennsylvania and NBER

Kenneth I. Spenner  
Department of Sociology, Duke University

Appendix B: Details about the implementation of the nonparametric estimation in Section 6

We propose an empirical strategy that consists of the following steps:

Step 1. Invoking Kotlarski’s (1967) theorem, we separately recover the marginal distributions of $X_C, X_U$, and $X_S$ from the observed joint distribution of $(W_U, W_S)$.

Step 2. We draw random samples of $\{X_{Ci}, X_{Ui}, X_{Si}\}$ from the marginal distributions of $X_C, X_U$, and $X_S$ recovered in Step 1.

Step 3. We obtain samples of $\{W_{Ui}, W_{Si}\}$ from the random samples of $\{X_{Ci}, X_{Ui}, X_{Si}\}$ generated in Step 2 and then recover a sample of $Y_i$ conditional on $\{W_{Ui}, W_{Si}\}$ using multiple imputation methods.29

Step 4. We run regressions of $Y$ on $X_C, X_U$, and $X_S$ using the pseudo-sample $\{Y_i, X_{Ci}, X_{Ui}, X_{Si}\}$ simulated above to estimate $\gamma_C, \gamma_U$, and $\gamma_S$, and to perform variance decomposition.

We now provide more details about each of the steps, beginning with recovering the marginal distributions of $X_C, X_U$, and $X_S$. Let

$$\Psi(t_1, t_2) = \exp(it_1W_U + it_2W_S) \quad (B1)$$

29See Rubin (1987) for an extensive description of this methodology.

DOI: 10.3982/QE83
denote the characteristics function for the observed joint random vector \((W_U, W_S)\) and let

\[
\Psi_1(t_1, t_2) \equiv \frac{\partial \Psi(t_1, t_2)}{\partial t_1}
\]

\[
= E[iW_U \exp(it_1W_U + it_2W_S)]
\]

(B2)

denote the derivative of \(\Psi(\cdot, \cdot)\) with respect to its first argument. Then the Kotlarski theorem shows that the characteristic functions for random variables \(X_C, X_U, \) and \(X_C\) are, respectively, given by

\[
\Psi_{X_C}(t) = \exp\left(\int_0^t \frac{\Psi(0, t_2)}{\Psi(0, t_2)} dt_2\right),
\]

\[
\Psi_{X_U}(t) = \frac{\Psi(t, 0)}{\Psi_{X_C}(t)},
\]

\[
\Psi_{X_S}(t) = \frac{\Psi(0, t)}{\Psi_{X_C}(t)}.
\]

Finally the characteristic functions of these three random variables uniquely determines the probability density function via an inversion formula. Let \(f_{X_C}, f_{X_U}, \) and \(f_{X_S}\), respectively, denote the marginal probability density function for random variables \(X_C, X_U, \) and \(X_S\). Following the inversion formula described in Horowitz (1998, p. 104), we have

\[
f_{X_K}(x_K) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp(-itx_K)\Psi_{X_K}(t) dt \quad \text{for } K \in \{C, U, S\}.
\]

We are now in a position to describe the somewhat standard estimation procedure needed to carry out Step 1. The key is to estimate \(\Psi(\cdot, \cdot)\) and \(\Psi_1(\cdot, \cdot)\) by their sample analogs: given a sample \(\{W_U^j, W_S^j\}_{j=1}^n\),

\[
\hat{\Psi}(t_1, t_2) = \frac{1}{n} \sum_{j=1}^n \exp(it_1W_U^j + it_2W_S^j),
\]

\[
\hat{\Psi}_1(t_1, t_2) = \frac{1}{n} \sum_{j=1}^n iW_U^j \exp(it_1W_U^j + it_2W_S^j).
\]

The characteristic functions \(\Psi_{X_K}(t)\) for \(K \in \{C, U, S\}\) can in turn be estimated by replacing \(\Psi(\cdot, \cdot)\) and \(\Psi_1(\cdot, \cdot)\) with their estimates above. Applying Kotlarski’s decomposition to \(\{W_U, W_S\}\) allows to generate data on \(\{X_{Ci}, X_{Ui}, X_{Si}\}\) and, therefore, \(\{W_{Ui}, W_{Si}\}\) (Steps 2 and 3) by simply drawing from the marginal distributions.

The next step, Step 4, is to obtain a sample of grades (i.e., \(Y_i\)) conditional on \(W_{Ui}\) and \(W_{Si}\) by multiple imputation. Here we follow Rubin (1987). The basic steps of Rubin multiple imputation are as follows:

\(^{30}\)See Krasnokutskaya (2011) for similar estimation procedure. Horowitz (1998, Chapter 4) described some useful suggestions for issues related to smoothing.
Supplementary Material

(i) Calculate $V = (W'W)^{-1}$, $\hat{\beta} = VW'Y$, and $\hat{Y} = W'\hat{\beta}$, where $W = \{W_U, W_S\}$.

(ii) Draw a random $g$ from $\chi^2$ distribution with degree of freedom $n_{\text{obs}} - r$.

(iii) Calculate $\sigma^2_* = (Y - \hat{Y})'(Y - \hat{Y})/g$.

(iv) Draw an $r$-dimensional Normal random vector $D \sim N(0, I_r)$, where $I_r$ is the identity matrix of dimension $r$.

(v) Calculate $\tilde{\beta}_* = \hat{\beta} + \sigma V^{1/2} D$, where $V^{1/2}$ is the triangular square root of $V$ obtained by the Cholesky decomposition.

(vi) Calculate predicted values $\hat{Y}_i = W'_i \hat{\beta}_*$.

(vii) For each missing value, find the respondent whose $\hat{Y}$ is closest to $\hat{Y}_i$ and take $Y$ of this respondent as the imputed value (predictive mean matching).\footnote{To test for robustness of the results, we also implemented a nonparametric approach to recover $Y_i$. Basically, we draw a sample of $Z_i$ conditional on $\{W_{Ui}, W_{Si}\}$ from the observed conditional distribution $G(Y|W_U, W_S)$, which was obtained using the Epanechnikov kernel ($K(u) = \frac{3}{4}(1 - u^2)1_{(|u| \leq 1)}$). The smoothing parameter was selected by following a refined plug-in method, which tries to find the bandwidth that minimizes the mean integrated square error. Results obtained using this strategy did not differ significantly from those using the multiple imputation technique.}

We then regress the generated outcomes on the generated regressors.

References


