APPENDIX A: SUPPORTING INFORMATION - WEB APPENDIX

A.1 Simulation Design

In Scenario 1, both the true treatment assignment model and true outcome assignment model contain only linear terms as described in Equations A1 and A2 respectively.

$$log\left[\frac{Pr(Z_i \le k)}{Pr(Z_i > k)}\right] = \boldsymbol{\theta}_k + \frac{k}{g}\beta_1 x_{i,1} + \frac{k}{g}\beta_2 x_{i,2} + \frac{k}{g}\beta_{4,k} x_{i,4} + \frac{k}{g}\beta_{5,k} x_{i,5} + \beta_7 x_{i,7} + \beta_8 x_{i,8} \tag{A1}$$
$$\text{for } k = 1, 2, 3, \quad g = 2, \quad \boldsymbol{\theta} = (-1.5,\ 0.25,\ 2), \quad \beta_1 = \beta_4 = \beta_7 = 0.7,\ \boldsymbol{\beta}_2 = \boldsymbol{\beta}_5 = \beta_8 = 0.4$$

$$log\left[\frac{Pr(Y_i = 1)}{1 - Pr(Y_i = 1)}\right] = \alpha + \beta_Z Z_i + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \beta_3 x_{i,3} + \beta_4 x_{i,4} + \beta_5 x_{i,5} + \beta_6 x_{i,6} \tag{A2}$$
$$\alpha = -1.5, \quad \beta_Z = 0.7, \quad \beta_1 = \beta_2 = \beta_3 = 0.7, \quad \text{and} \quad \beta_4 = \beta_5 = \beta_6 = 0.4$$

In Scenario 2 the treatment assignment model is misspecified by introducing a non-linear and slightly mismeasured variable, $(x_{i,1}+0.5)^2$, into the treatment assignment model. The outcome model is not misspecified under this scenario. The misspecified treatment assignment model is shown in Equation A3.

$$log\left[\frac{Pr(Z_i \le k)}{Pr(Z_i > k)}\right] = \boldsymbol{\theta}_k + \frac{k}{g}\beta_1 x_{i,1} + \frac{k}{g}\beta_1 (x_{i,1} + 0.5)^2 + \frac{k}{g}\beta_2 x_{i,2} + \frac{k}{g}\beta_{4,k} x_{i,4} + \frac{k}{g}\beta_{5,k} x_{i,5} + \beta_7 x_{i,7} + \beta_8 x_{i,8} \tag{A3}$$
$$\text{for } k = 1, 2, 3, \quad g = 2, \quad \boldsymbol{\theta} = (-1.5,\ 0.25,\ 2), \quad \beta_1 = \beta_4 = \beta_7 = 0.7\ \boldsymbol{\beta}_2 = \boldsymbol{\beta}_5 = \beta_8 = 0.4$$

In Scenario 3 the outcome assignment model is misspecified by introducing a non-linear, and slightly mismeasured, variable, $(x_{i,1}+0.5)^2$, into the outcome model. However, this variable is not present in the treatment assignment model. This misspecified outcome assignment model is shown in Equation A4.

$$log\left[\frac{Pr(Y_i = 1)}{1 - Pr(Y_i = 1)}\right] = \alpha + \beta_Z Z_i + \beta_1 x_{i,1} + \beta_1 (x_{i,1} + 0.5)^2 + \beta_2 x_{i,2} + \beta_3 x_{i,3} + \beta_4 x_{i,4} + \beta_5 x_{i,5} + \beta_6 x_{i,6} \tag{A4}$$
$$\alpha = -2.5, \quad \beta_Z = 0.7, \quad \beta_1 = \beta_2 = \beta_3 = 0.7, \quad \text{and} \quad \beta_4 = \beta_5 = \beta_6 = 0.4$$

Finally, Scenario 4 generates treatment assignment from the misspecified treatment assignment shown in Equation A3 and generates outcome from the misspecified outcome assignment model shown in Equation A4. This is the only scenario in which the non-linear term is present in both the treatment and outcome assignment models.

Note: For the simulations comparing differing number of possible treatment levels (described in Section 4.1), the value of g (in equations A1 and A3) was adjusted to 4, 6, and 8 for the simulations with 6, 8, and 10 treatment groups respectively in order to prevent the treatment term from disproportionately increasing the logit.

A.2 Supporting Figures

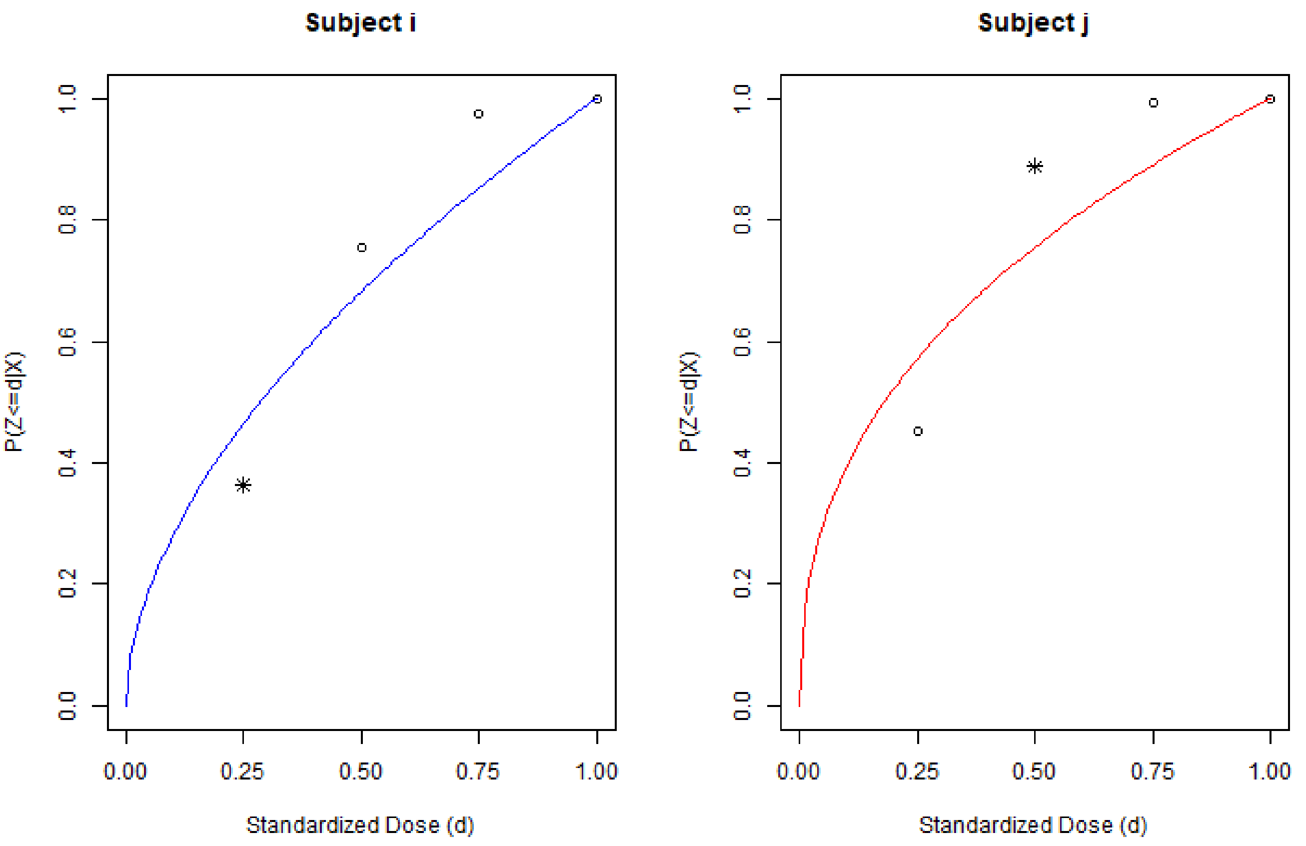

FIGURE A1 Comparison of the CDF of the estimated $\hat{\mathbf{P}}_{i,GPS}$ and $\hat{\mathbf{P}}_{j,GPS}$ vectors and fitted power functions of matched subjects, i and j, $\tilde{a}_i = -0.9400$, $\tilde{a}_j = -0.5678$, with different levels of observed treatment (indicated by the * symbol), $Z_i = 2$ (d=0.50), $Z_j = 3$ (d=0.75).