

Estimation Method	Model Type	Inputs	Assumptions	Description
Anderson et al. (2019)	GMM	Linked dataset with multiple matches $(x_i, \{y_{i\ell}\}_{\ell=1}^{L_i})_{i=1}^n$	Datasets are random samples conditional on the matching variables; each $x_i$ has exactly one true link among its matched outcomes $\{y_{i\ell}\}_{\ell=1}^{L_i}$ , and each $y_{i\ell}$ is equally likely to be the correct match	Rewrite the moments as a difference between the sum of the moments for all possible links and a conditional expectation of the moment function based on $i$ 's identifying variables; estimation and variance estimation proceed as standard 2-step semiparametric GMM problem
Lahiri and Larsen (2005)	GLM	Files A and B, linked via Fellegi-Sunter/EM algorithm method	Each record in File A has exactly one true link in File B; distribution of matching variables is independent of the response variable	Apply OLS to a transformed regression model constructed using parameter estimates from the Fellegi-Sunter/EM algorithm method; estimate the variance using parametric bootstrap
Scheuren and Winkler (1991, 1993, 1997)	GLM	tbd		Propose a bias correction for OLS, using a weighted average of responses from all observations and the actual response $y_i$
Nix and Qian (2015)				
Bleakley and Hoyt (2016)				
Hirukawa and Prokhorov (2018)	GLM	Nearest-neighbor matching (Unique matches only)	The samples jointly identify the regression models	Bias correction for nearest-neighbor matched samples; could be used if we define an appropriate metric for matching variables and use a nearest neighbor matching rule
Goldstein et al., 2012	GLM			
Anderson (2019)	GMM	Files A and B, linked via Fellegi-Sunter EM method	Want to relax the assumption where matching variables are independent of the response variable	