Supplementary Materials for "Residual Balancing: A Method of Constructing Weights for Marginal Structural Models"

Xiang Zhou Geoffrey T. Wodtke

Harvard University University of Chicago

December 13, 2019

A. Minimization of Relative Entropy

Following Hainmueller (2012), we use the method of Lagrange multipliers to find a set of weights rbw_i that minimize their relative entropy with the base weights q_i subject to the balancing constraints. Substituting $\hat{\delta}(g_j(l_{it}))$ for $\delta(g_j(l_{it}))$ in equation (14) in the main text, the balancing constraints can be written as

$$\sum_{i=1}^{n} rbw_i \hat{c}_{ir} = 0, \quad 1 \leq r \leq n_c,$$

where $\hat{c}_{ir}$ is the rth element of $\hat{\boldsymbol{c}}_i = \{\hat{\delta}(g_j(l_{it}))h_k(\bar{l}_{i,t-1}, \bar{a}_i); 1 \leq j \leq J_t, 1 \leq k \leq K_t, 1 \leq t \leq T\}$. In addition, we impose a normalization constraint $\sum_i rbw_i = n$ such that the residual balancing weights sum to the sample size. Thus, the primal optimization problem is

$$\min_{rw_i} L^p = \sum_{i=1}^{n} rbw_i \log \frac{rbw_i}{q_i} + \sum_{r=1}^{n_c} \lambda_r \sum_{i=1}^{n} rbw_i c_{ir} + \lambda_0 (\sum_{i=1}^{n} rbw_i - n), \tag{1}$$

where $\{\lambda_1, \ldots, \lambda_{n_c}\}$ are the Lagrange multipliers for the balancing constraints and λ_0 is the Lagrange multiplier for the normalization constraint. Since the loss function L^p is strictly convex, the first order condition of equation (1) implies that the solution for each weight is

$$rbw_i^* = \frac{nq_i \exp(-\sum_{r=1}^{n_c} \lambda_r c_{ir})}{\sum_{i=1}^{N} q_i \exp(-\sum_{r=1}^{n_c} \lambda_r c_{ir})}. \tag{2}$$

Inserting equation (2) into L^p leads to the dual problem given by

$$\max_{\lambda_r} L^d = -\log\Big(\sum_{i=1}^{n} q_i \exp\big(-\sum_{r=1}^{n_c}\lambda_r c_{ir}\big)\Big),$$

or equivalently,

$$\min_{Z} L^d = \log\Big(Q' \exp\big(CZ\big)\Big),$$

where $Q = [q_1, q_2, \ldots, q_n]'$, $C = [\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_n]'$, and $Z = -[\lambda_1, \lambda_2, \ldots, \lambda_{n_c}]'$. Since both the gradient and the Hessian have closed-form expressions, this problem can be solved using Newton's method. Inserting the solutions for λ_r into equation (2) yields the residual balancing weights.

B. Performance of the Robust ("Sandwich") Variance Estimator

In most applications of marginal structural models (MSMs), standard errors are computed with the robust ("sandwich") variance estimator. In this section, we present a simulation study that evaluates the performance of the robust variance estimator for MSM coefficients estimated via IPW-GLM, IPW-GLM-Censored, IPW-CBPS, and residual balancing (under the same setup described in Section 4 of the main text). The results are shown in Figures S1-S4, where the box plots display the sampling distributions of the robust standard errors divided by the true standard errors estimated from the 2,500 random samples. Across nearly all scenarios, and especially when the confounder models are correctly specified, the robust variance estimator is conservative for residual balancing, that is, it tends to overestimate the true sampling variance. Consequently, as Tables S1-S2 show, when the confounder models are correctly specified, confidence intervals constructed with these standard errors typically ensure true coverage rates that are at least equal to, and often exceed, the nominal coverage rate. By contrast, results from this simulation study suggest that the robust variance estimator may underestimate the true sampling variance under IPW-GLM in many different situations, even though it is expected to be conservative in large samples (Robins 1999; Robins, Hernan and Brumback 2000). As a result, confidence intervals constructed with these standard errors often fall short of the nominal coverage rate, even when the propensity score models are correctly specified.

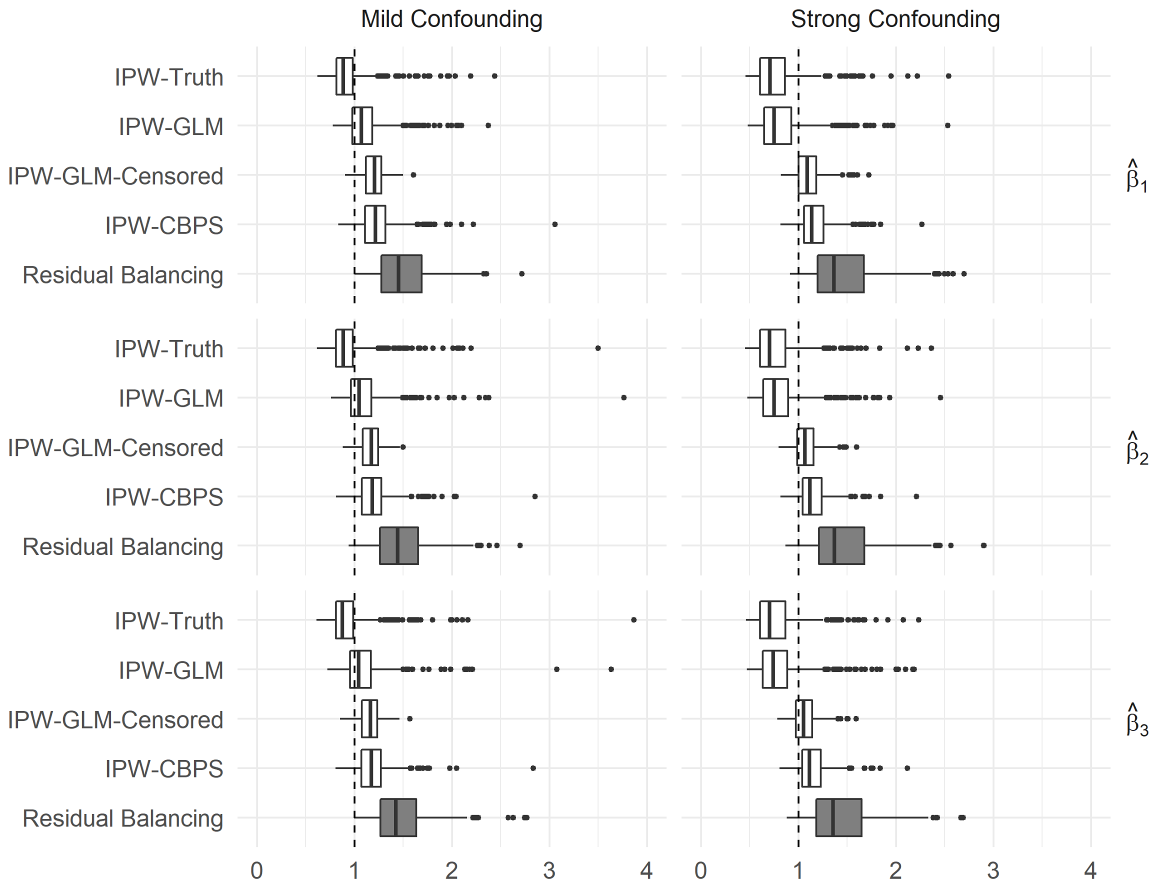

Figure S1: Performance of the robust ("sandwich") variance estimator for a binary treatment with correct model specification. The left and right panels correspond to the settings of "mild confounding" ($\alpha = 0.4$) and "strong confounding" ($\alpha = 0.8$) respectively. Four different methods are compared: IPW based on the standard logistic regression (IPW-GLM), IPW based on the standard logistic regression with weights censored at the 1st and 99th percentiles (IPW-GLM-Censored), IPW based on the CBPS (IPW-CBPS), and residual balancing. As a benchmark, results from IPW based on true treatment probabilities (IPW-Truth) are also reported. The box plots show the sampling distributions (from 2500 random samples) of the robust standard errors divided by the true standard errors (estimated via the 2500 random samples).

Table S1: Coverage of 95% confidence intervals constructed with robust ("sandwich") standard errors for a binary treatment with correct model specification.

	Mild Confounding			Strong Confounding		
	β_1	β_2	β_3	β_1	β_2	β_3
IPW-Truth	0.94	0.92	0.93	0.90	0.85	0.88
IPW-GLM	0.95	0.97	0.97	0.92	0.90	0.90
IPW-GLM-Censored	0.94	0.95	0.97	0.93	0.79	0.82
IPW-CBPS	0.90	0.83	0.95	0.87	0.40	0.67
Residual Balancing	0.98	1.00	0.99	0.98	1.00	0.98

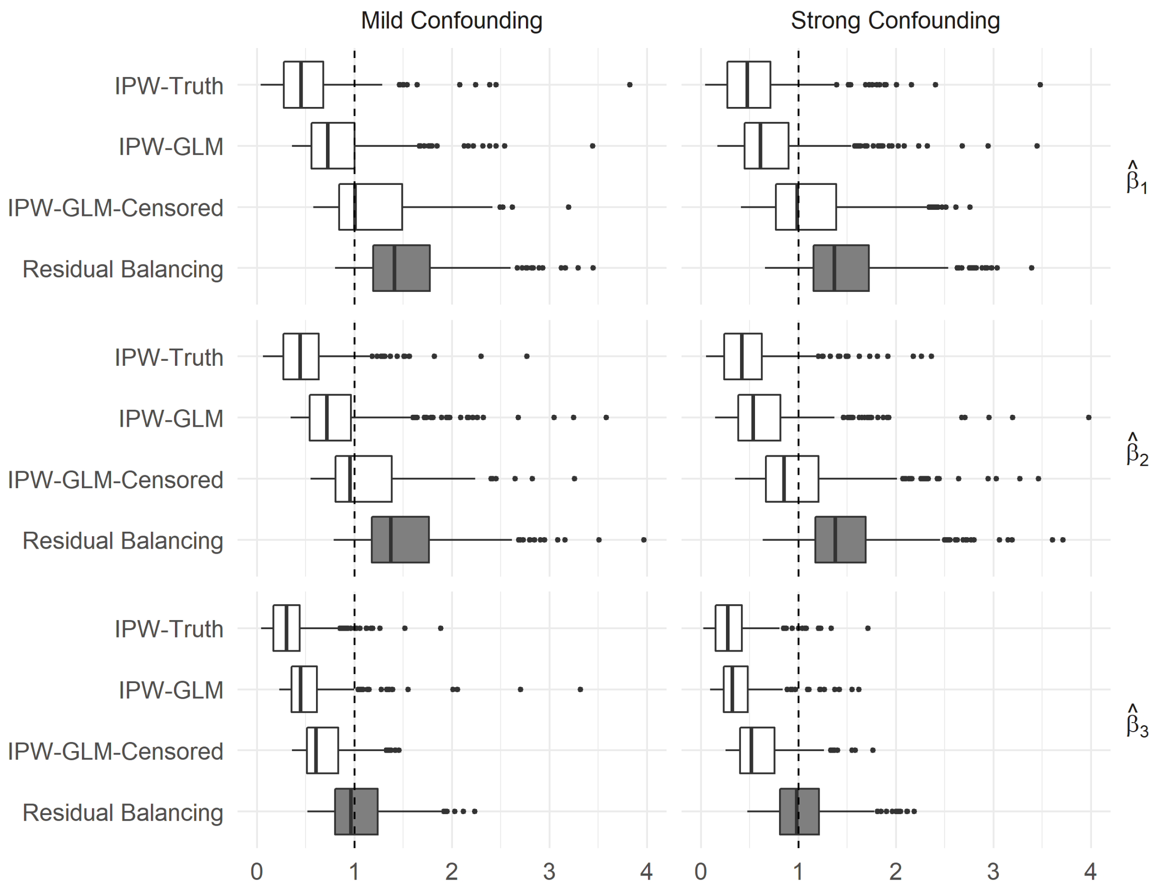

Figure S2: Performance of the robust ("sandwich") variance estimator for a continuous treatment with correct model specification. The left and right panels correspond to the settings of "mild confounding" ($\alpha = 0.4$) and "strong confounding" ($\alpha = 0.8$) respectively. Three different methods are compared: IPW based on the standard logistic regression (IPW-GLM), IPW based on the standard logistic regression with weights censored at the 1st and 99th percentiles (IPW-GLM-Censored), and residual balancing. As a benchmark, results from IPW based on true treatment probabilities (IPW-Truth) are also reported. The box plots show the sampling distributions (from 2500 random samples) of the robust standard errors divided by the true standard errors (estimated via the 2500 random samples).

Table S2: Coverage of 95% confidence intervals constructed with robust ("sandwich") standard errors for a continuous treatment with correct model specification.

	Mild Confounding			Strong Confounding		
	β_1	β_2	β_3	β_1	β_2	β_3
IPW-Truth	0.72	0.63	0.56	0.68	0.39	0.34
IPW-GLM	0.91	0.85	0.88	0.8	0.58	0.66
IPW-GLM-Censored	0.91	0.72	0.77	0.83	0.27	0.40
Residual Balancing	0.98	0.99	1.00	0.97	0.99	0.99

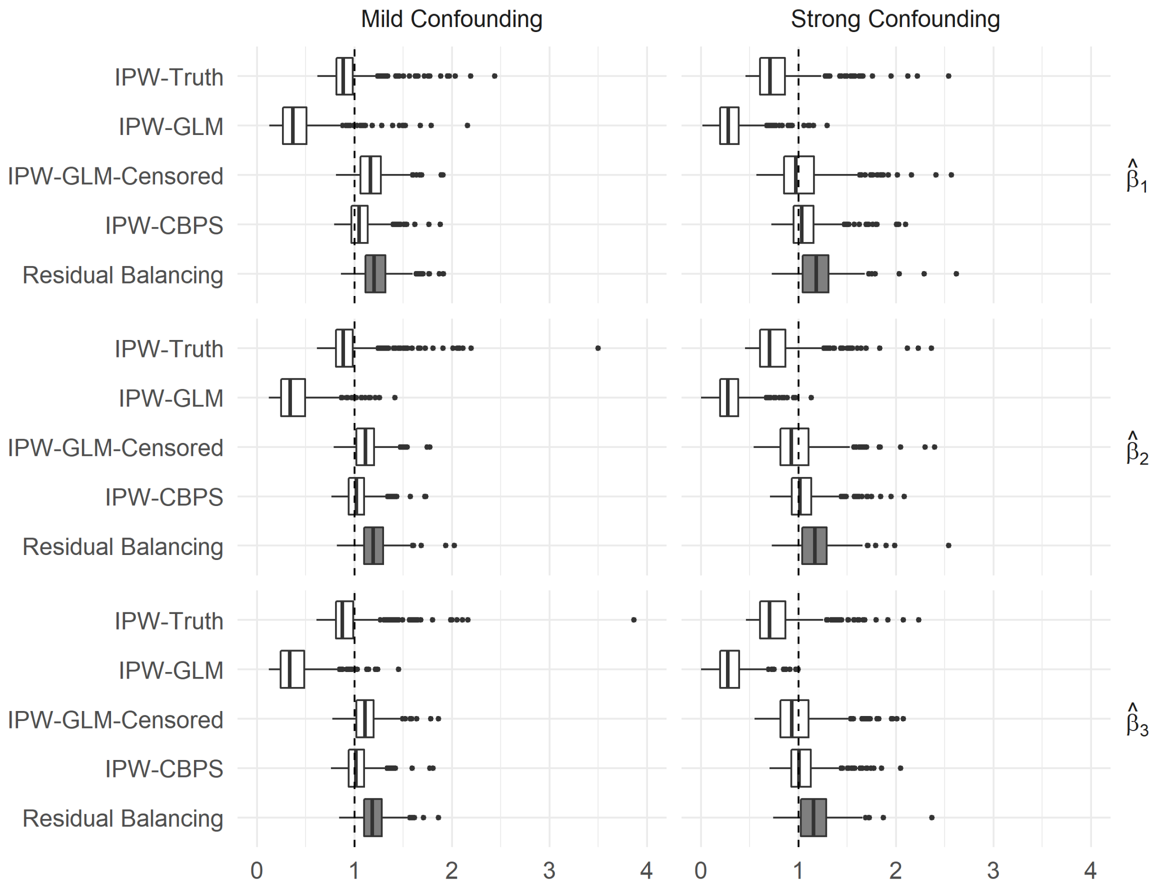

Figure S3: Performance of the robust ("sandwich") variance estimator for a binary treatment with incorrect model specification. The left and right panels correspond to the settings of "mild confounding" ($\alpha = 0.4$) and "strong confounding" ($\alpha = 0.8$) respectively. Four different methods are compared: IPW based on the standard logistic regression (IPW-GLM), IPW based on the standard logistic regression with weights censored at the 1st and 99th percentiles (IPW-GLM-Censored), IPW based on the CBPS (IPW-CBPS), and residual balancing. As a benchmark, results from IPW based on true treatment probabilities (IPW-Truth) are also reported. The box plots show the sampling distributions (from 2500 random samples) of the robust standard errors divided by the true standard errors (estimated via the 2500 random samples).

Table S3: Coverage of 95% confidence intervals constructed with robust ("sandwich") standard errors for a binary treatment with incorrect model specification.

	Mild Confounding			Strong Confounding		
	β_1	β_2	β_3	β_1	β_2	β_3
IPW-Truth	0.94	0.92	0.93	0.90	0.85	0.88
IPW-GLM	0.64	0.69	0.69	0.44	0.49	0.44
IPW-GLM-Censored	0.84	0.39	0.57	0.88	0.60	0.74
IPW-CBPS	0.69	0.03	0.13	0.47	0.00	0.01
Residual Balancing	0.94	0.80	0.82	0.90	0.75	0.76

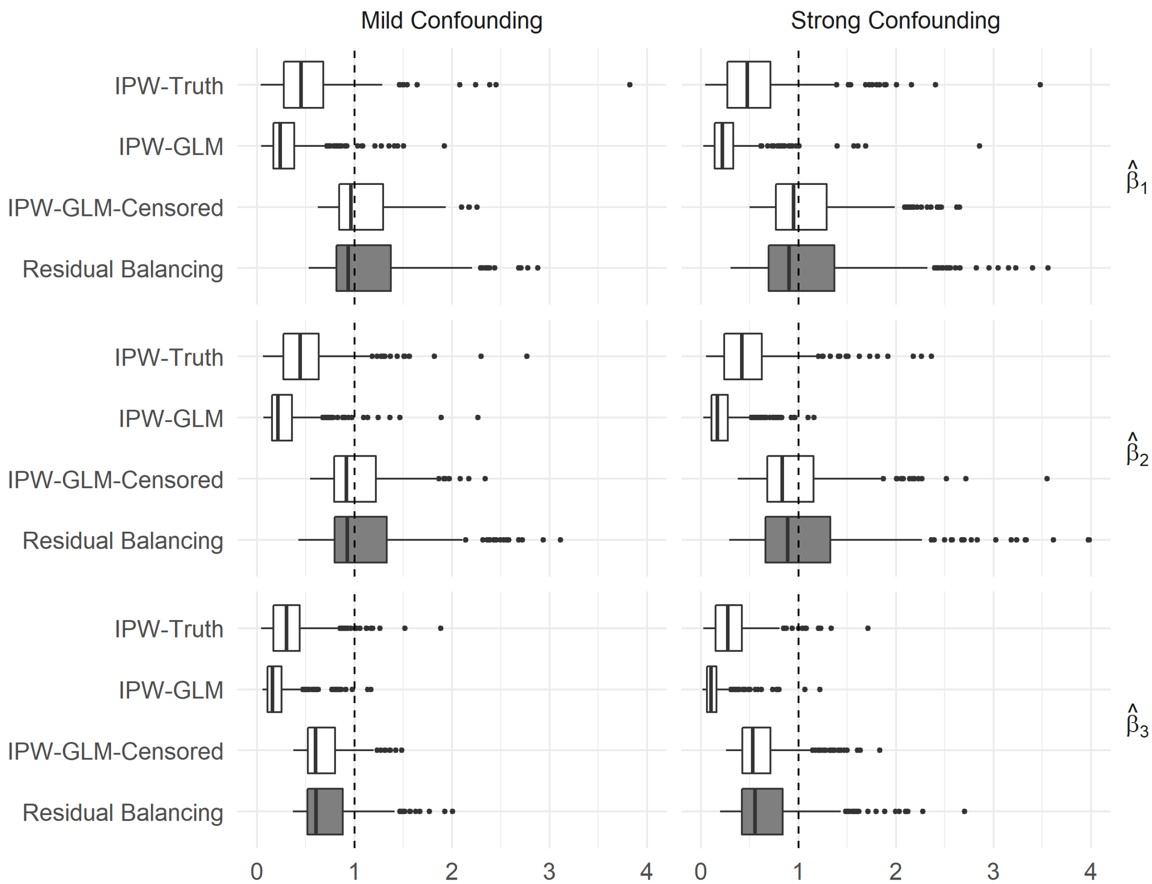

Figure S4: Performance of the robust ("sandwich") variance estimator for a continuous treatment with incorrect model specification. The left and right panels correspond to the settings of "mild confounding" ($\alpha = 0.4$) and "strong confounding" ($\alpha = 0.8$) respectively. Three different methods are compared: IPW based on the standard logistic regression (IPW-GLM), IPW based on the standard logistic regression with weights censored at the 1st and 99th percentiles (IPW-GLM-Censored), and residual balancing. As a benchmark, results from IPW based on true treatment probabilities (IPW-Truth) are also reported. The box plots show the sampling distributions (from 2500 random samples) of the robust standard errors divided by the true standard errors (estimated via the 2500 random samples).

Table S4: Coverage of 95% confidence intervals constructed with robust ("sandwich") standard errors for a continuous treatment with incorrect model specification.

	Mild Confounding			Strong Confounding		
	β_1	β_2	β_3	β_1	β_2	β_3
IPW-Truth	0.72	0.63	0.56	0.68	0.39	0.34
IPW-GLM	0.48	0.11	0.06	0.29	0.02	0.02
IPW-GLM-Censored	0.33	0.00	0.00	0.10	0.00	0.00
Residual Balancing	0.89	0.69	0.72	0.87	0.64	0.66

C. Illustrative R Code

In this appendix, we illustrate the implementation of residual balancing using the R package `rbw` for the two empirical examples.

```
devtools::install_github("xiangzhou09/rbw")
library(rbw); library(survey)
## Example 1: The Cumulative Effect of Negative Advertising on Candidate's Voteshare ##
# models for time-varying confounders
m1 <- lm(dem.polls ~ (d.gone.neg.l1 + dem.polls.l1 + undother.l1) * factor(week),
         data = campaign_long)
m2 <- lm(undother ~ (d.gone.neg.l1 + dem.polls.l1 + undother.l1) * factor(week),
         data = campaign_long)
xmodels <- list(m1, m2)
# residual balancing weights
fit <- rbwPanel(exposure = d.gone.neg, xmodels = xmodels, id = id, time = week,
                data = campaign_long)
campaign_wide <- merge(campaign_wide, fit$weights, by = "id")
# fitting a marginal structural model
rbw_design <- svydesign(ids = ~ 1, weights = ~ rbw, data = campaign_wide)
msm_rbw <- svyglm(demprcnt ~ cum_neg * deminc + camp.length + factor(year) + office,
                  design = rbw_design)
## Example 2: The Controlled Direct Effect of Shared Democracy on Public Support for War ##
haven::read_dta("peace.dta")
# models for post-treatment confounders
m1 <- lm(threatc ~ ally + trade + h1 + i1 + p1 + e1 + r1 + male + white + age + ed4 + democ,
         data = peace)
m2 <- lm(cost ~ ally + trade + h1 + i1 + p1 + e1 + r1 + male + white + age + ed4 + democ,
         data = peace)
m3 <- lm(successc ~ ally + trade + h1 + i1 + p1 + e1 + r1 + male + white + age + ed4 + democ,
         data = peace)
# residual balancing weights
fit <- rbwMed(treatment = democ, mediator = immoral, zmodels = list(m1, m2, m3),
              data = peace)
peace$rbw <- fit$weights
# fitting a marginal structural model
rbw_design <- svydesign(ids = ~ 1, weights = ~ rbw, data = peace)
msm_rbw <- svyglm(strike ~ ally + trade + h1 + i1 + p1 + e1 + r1 + male + white +
     age + ed4 + democ + democ * immoral, design = rbw_design)
```

References

Hainmueller, Jens. 2012. "Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies." *Political Analysis* 20(1):25–46.

Robins, James M. 1999. "Marginal Structural Models versus Structural Nested Models as Tools for Causal Inference." *Statistical Models in Epidemiology: The Environment and Clinical Trials* .

Robins, James M, Miguel Angel Hernan and Babette Brumback. 2000. "Marginal Structural Models and Causal Inference in Epidemiology." *Epidemiology* 11(5):550–560.