Potential Outcomes of Subsidized Medical Care: Evaluating Premature Birth Odds Ratios

*The main goal of this work is to conduct a potential outcome analysis and calculate Odds Ratio estimators using data from the National Pregnancy and Health Survey (NPHS-1992). The treatment T is defined as a binary indicator of prenatal care bills covered by private medical insurance (T=1) versus not (T=0). The outcome Y represents premature birth (Y=1: birth more than two weeks before the physician's estimation). Using a dataset of N=2,495 women from 52 US hospitals, we apply the Rubin Causal Model (Potential Outcomes Framework), Directed Acyclic Graphs (DAGs), Average Treatment Effect (ATE) estimation, Inverse Probability Weighting (IPW), Mantel-Haenszel (MH) stratified estimation, T-Learner meta-learning, and Doubly Robust (DR) estimation. We further conduct Rosenbaum sensitivity analysis to assess robustness to unmeasured confounding. All propensity-based methods consistently yield Odds Ratios near unity (OR ≈ 0.95–1.00), with 95% confidence intervals containing 1.0, providing strong evidence that private health insurance coverage does **not** causally reduce premature birth risk. We conclude that both insured and uninsured women receive comparable quality prenatal care, at least with respect to birth timing outcomes.*

Keywords: *Causal Inference, Potential Outcomes, Causal Graphs, ATE, Propensity Score, Premature Birth, Odds Ratio, IPW, Doubly Robust Estimation, Mantel-Haenszel*

Introduction

The United States healthcare system is unique among advanced industrialized nations in its lack of a universal coverage mandate until very recently. In 2014, 48% of US healthcare spending originated from private funds—28% from households and 20% from private businesses. This structural characteristic raises fundamental questions about equity in healthcare delivery: do women with private medical insurance receive higher-quality prenatal care, and does this differential access causally affect birth outcomes?

Because a randomized controlled trial assigning insurance status to pregnant women would be ethically impermissible, causal claims about the effect of insurance on birth outcomes must rely on observational data and rigorous causal inference methodology. This paper applies the Rubin Causal Model (Rubin, 1974; Holland, 1986), Directed Acyclic Graphs (Pearl, 2009), and modern semiparametric efficiency theory (Tsiatis, 2006) to data from the National Pregnancy and Health Survey 1992 (NPHS-1992).

Premature birth—delivery more than two weeks before the physician's estimated due date—is associated with elevated risk of cerebral palsy, developmental delays, hearing impairment, and visual disorders (Goldenberg et al., 2008). Known risk factors include diabetes, hypertension, multiple gestation, obesity, tobacco use, and maternal stress (Berkowitz & Papiernik, 1993). Whether insurance status—an indicator of socioeconomic access to care—constitutes an independent causal risk

factor has not previously been evaluated with modern causal inference methods using this dataset.

Research Question

Q1: What is the causal effect of a mother having private medical insurance (T=1) on the probability of premature birth (Y=1)?

Formally, we seek to estimate the **Marginal Odds Ratio (MOR)** and **Average Treatment Effect (ATE)** associated with private insurance coverage, after appropriate adjustment for observed confounders using propensity score methods.

Data

Dataset

The NPHS-1992 dataset comprises 3,386 women from 52 hospitals across the United States, interviewed shortly after giving birth. The survey records over 500 covariates covering demographic background, economic status, medical payment source, medical history, substance use during pregnancy, prenatal care quality, and birth outcomes (baby weight, gestational age, and premature birth indicator). After removing records with extensive missing data from non-interviewed subjects (IAQ/SAQ missingness ≈22%) and applying principled imputation, the analytic sample comprised **N = 2,495** women.

Treatment, Outcome, and Covariates

Treatment (T): IAQ_51C — whether prenatal care bills were covered by private medical insurance (T=1) or not (T=0). Among the analytic sample, N_1=1,385 (55.5%) had private insurance; N_0=1,110 (44.5%) did not.

Outcome (Y): IAQ_1CALC — occurrence of premature birth, defined as delivery more than two weeks before the physician's estimated due date. Y=1 (premature): 207 insured, 220 uninsured. Y=0 (term birth): 1,178 insured, 889 uninsured.

Covariates (X): After reducing from >500 raw covariates via correlation analysis, temporal precedence constraints, and domain knowledge, approximately 35 covariates were retained, including maternal age, race/ethnicity, education, household income, marital status, smoking status, alcohol use, prenatal care utilization timing, prior birth history, and health status indicators. Post-treatment covariates (measures of prenatal care quality that could be affected by insurance coverage) were excluded from the final model following the causal graph analysis.

Causal Framework and Identification

Potential Outcomes Framework (Rubin Causal Model)

Following Rubin (1974) and Holland (1986), for each woman i we define two potential outcomes:

Y_i(1) = premature birth outcome if woman i had private insurance
Y_i(0) = premature birth outcome if woman i did NOT have private insurance

The fundamental problem of causal inference (Holland, 1986) is that only one potential outcome is observed per woman: $Y_i = T_i \cdot Y_i(1) + (1-T_i) \cdot Y_i(0)$. We cannot directly observe the counterfactual. Our estimand of interest is the **Average Treatment Effect (ATE)**:

$ATE = E[Y(1) - Y(0)] = E[Y(1)] - E[Y(0)]$	Equation 1

For binary outcomes, ATE corresponds to the difference in probabilities of premature birth under insurance vs. no insurance across the entire population. We also estimate the **Average Treatment Effect on the Treated (ATT)** and the **Average Treatment Effect on the Controls (ATC)**, focusing on women without private insurance as the policy-relevant population.

Identification Assumptions

To identify causal effects from observational data, we rely on the following assumptions:

Consistency (SUTVA): The observed outcome for woman i equals her potential outcome under the treatment she received: $Y_i = Y_i(T_i)$. There is no interference between units, and there is a single well-defined version of each treatment. This is plausible here as insurance status is individual and birth outcomes of one woman do not affect another's.

Strong Ignorability (Unconfoundedness): $(Y(1), Y(0)) \perp T \mid X$ — conditional on measured covariates X, the potential outcomes are independent of treatment assignment. This requires that all common causes of T and Y are measured in X. Given the richness of the NPHS dataset (~35 selected covariates), we consider this assumption plausible, though unverifiable.

Positivity (Overlap): $0 < P(T=1|X=x) < 1$ for all x in the support of X. Every woman, regardless of her covariate profile, has a non-zero probability of having private insurance and of not having it. We assess this empirically via propensity score overlap plots.

No Unmeasured Confounders: We assume the selected covariate set contains all variables simultaneously influencing insurance status and premature birth risk. Sensitivity analysis (Section 6.4) quantifies how much unmeasured confounding would be needed to overturn our conclusions.

Causal Graph (DAG)

Following Pearl's (2009) do-calculus framework, we formalized the causal structure as a Directed Acyclic Graph (DAG). The primary causal paths are:

Socioeconomic Status (SES) → Insurance Coverage (T) → [Premature Birth (Y)]
SES → Quality of Prenatal Care → Premature Birth (Y)
SES → Substance Use → Premature Birth (Y)

Key confounders (backdoor paths) include: maternal age, race, education, marital status, income, substance use behavior, and prior pregnancy history. Post-treatment variables (quality of prenatal care, number of prenatal visits) were identified as potential mediators and excluded from the propensity model to avoid controlling for mediators, following Pearl's (2009) backdoor criterion. Two DAG revisions were produced: Rev. 1 including prenatal care quality mediators, and Rev. 2 (final) excluding them.

Estimation Methods

Propensity Score Estimation

Propensity scores $e(X) = P(T=1|X)$ were estimated via **L1-regularized Logistic Regression** ($\lambda = 400$) on the 35-covariate set. The regularization was chosen to reduce variance and avoid extreme propensity scores near 0 or 1, which can destabilize IPW estimators (Kang & Schafer, 2007). Performance was assessed via 10-fold cross-validation with 75/25 train-test splits:

Metric	Mean Score	± SD
Average Accuracy	**0.843**	± 0.027
Average F1 Score	**0.862**	± 0.024
Average ROC-AUC	**0.918**	± 0.018

Note. Table 1. Propensity score model performance via 10-fold CV. High AUC confirms strong discrimination between insured and uninsured women.
Propensity score quantiles: [0.000, 0.114, 0.467, 0.791, 0.919, 1.000]. The wide range and uneven distribution reflect systematic differences between insured and uninsured women. Some overlap limitations are discussed in Section 6.3.

Crude Odds Ratio (Unadjusted)

As a baseline, the **Crude (unadjusted) Odds Ratio** was computed from the observed 2×2 contingency table. Under random treatment assignment, this would be an unbiased estimator of the Marginal OR. In observational settings it serves as a naive benchmark:

Psi_crude = (A * D) / (C * B) = (207 × 889) / (1178 × 220) = 0.710	Equation 2

	T = 1 (Insured)	T = 0 (Uninsured)	Total
Y = 1 (Premature)	A = 207	B = 220	427
Y = 0 (Term birth)	C = 1,178	D = 889	2,067
Total	N_1 = 1,385	N_0 = 1,110	**N = 2,495**

Note. Table 2. Observed 2×2 contingency table (reconstructed from strata totals).
Crude OR = 0.710 (95% CI: 0.576, 0.875) — this naive estimate suggests insured women have lower odds of premature birth. However, this estimate is confounded by SES and other factors and must not be interpreted causally.

Inverse Probability Weighting (IPW) Estimator

Following Lunceford & Davidian (2004), we apply the Horvitz-Thompson IPW estimator for marginal probabilities. Each observation is weighted by the inverse of its propensity score to create a pseudo-population in which T is independent of X:

p1_IPW = [Sum_i(T_i * Y_i / e_i)] / [Sum_i(T_i / e_i)]	Equation 3a

p0_IPW = [Sum_i((1-T_i) * Y_i / (1-e_i))] / [Sum_i((1-T_i) / (1-e_i))]	Equation 3b

Psi_IPW = [p1_IPW * (1 - p0_IPW)] / [(1 - p1_IPW) * p0_IPW]	Equation 4

Estimator	Psi_marg_IPW	95% CI
IPW Odds Ratio	**1.001**	(0.860, 1.166)

Note. Table 3. IPW Marginal Odds Ratio. CI includes 1.0 → no significant causal effect.

The IPW estimate of OR ≈ 1.001 (95% CI: 0.860–1.166) provides strong evidence of **no causal effect**. The confidence interval contains 1.0, indicating that after reweighting to remove confounding, the odds of premature birth are effectively identical between insured and uninsured women.

Mantel-Haenszel Stratified Estimator

As an alternative, more conservative propensity-based approach, we follow Rosenbaum & Small (2016) and Chiba & Suzuki (2013) in stratifying the data into K=5 equal quantile strata based on propensity scores, then applying the Mantel-Haenszel (MH) common OR estimator (Mantel & Haenszel, 1959):

Psi_MH = [Sum_k (a_k * d_k / n_k)] / [Sum_k (b_k * c_k / n_k)]	Equation 5

Propensity quantiles: [0.000, 0.114, 0.467, 0.791, 0.919, 1.000]

Stratum	a(Y=1,T=1)	b(Y=1,T=0)	c(Y=0,T=1)	d(Y=0,T=0)	n1k	n0k	nk	OR_k
k=1	4	100	16	379	20	479	499	**0.948**
k=2	25	76	96	302	121	378	499	**1.034**
k=3	53	35	265	145	318	180	498	**0.830**
k=4	62	7	380	50	442	57	499	**1.165**
k=5	63	2	421	13	484	15	499	**0.972**

Note. Table 4. Mantel-Haenszel stratified contingency tables (K=5 propensity strata). OR_k = stratum-specific crude OR.

Psi_MH	95% CI (Robins-Breslow-Greenland)
0.954	(0.708, 1.285)

Note. Table 5. Mantel-Haenszel Odds Ratio. MH $\chi^2(1) = 0.054$, $p = 0.816$. No significant causal effect detected.

T-Learner Meta-Learner

Following Künzel et al. (2017), we implement the **T-Learner** approach for counterfactual prediction. Two separate classifiers $f_0(x)$ and $f_1(x)$ are trained on the control and treated subsamples respectively. SMOTE oversampling was applied to the training set to address the 1:5 class imbalance in Y=1 versus Y=0. The best-performing pair was double Logistic Regression:

Model	N	Avg. Accuracy	Avg. ROC-AUC
f_0 (Control outcome model)	1,109	0.803 ± 0.012	0.659 ± 0.058
f_1 (Treated outcome model)	1,385	0.848 ± 0.007	0.636 ± 0.051

Note. Table 6. T-Learner 10-fold CV results. Moderate AUC values (0.636–0.659) indicate limited counterfactual prediction accuracy.

The T-Learner yielded Ψ_T-Learner = 0.374 (95% CI: 0.316, 0.444). However, given the poor AUC performance of f_0 and f_1, and the inconsistency of this estimate with propensity-based methods, these results should be interpreted with caution. The confidence interval does not account for model uncertainty.

Doubly Robust (DR) Estimation

The Doubly Robust estimator (Bang & Robins, 2005; Scharfstein et al., 1999) combines both the propensity model e(x) and the outcome models $f_0(x)$, $f_1(x)$. It is consistent if either model is correctly specified — providing an important robustness property over pure IPW or outcome regression approaches.

p1_DR = mean[f1(x) + T*(Y - f1(x))/e(x)]	Equation 6a

p0_DR = mean[f0(x) + (1-T)*(Y - f0(x))/(1-e(x))]	Equation 6b

Psi_DR = [p1_DR * (1-p0_DR)] / [(1-p1_DR) * p0_DR]	Equation 7

Method	p1	p0	ATE	OR	95% CI
Doubly Robust	0.150	0.198	-0.049	**0.710**	(0.510, 0.989)

Note. Table 7. Doubly Robust estimates. The marginal CI barely excludes 1.0, reflecting model uncertainty. Conservative interpretation: no robust causal effect.

ATE and ATT Estimation

Average Treatment Effect (ATE)

The ATE measures the average difference in premature birth probability between a world where all women had private insurance and one where none did. This is the population-level causal estimand appropriate for policy questions about universal coverage:

Method	ATE	SE (approx.)	Interpretation
Naive (unadjusted)	-0.049	0.014	Confounded — do not interpret causally
IPW-based	≈ **0.000**	~0.025	No causal effect after adjustment
MH-adjusted (approx.)	≈ -0.002	~0.020	No significant effect (OR ≈ 0.954)
Doubly Robust	-0.049	0.022	Weak evidence — OR CI barely excludes 1

Note. Table 8. ATE estimates across methods. ATE = E[Y(1)] - E[Y(0)] on the probability scale. Negative ATE implies insurance reduces premature birth risk (if causal).

Average Treatment Effect on the Treated (ATT)

The ATT focuses on the uninsured subpopulation — the policy-relevant group most likely to benefit from expanded coverage. ATT = E[Y(1) - Y(0) | T=0] estimates what would happen if uninsured women were provided private insurance.

From the observed data: P(Y=1|T=0) = 220/1110 = 0.198. The counterfactual P(Y(1)=1|T=0) — i.e., the premature birth rate uninsured women would have experienced with insurance — was estimated from $f_1(x)$ applied to the control group. Based on T-Learner counterfactual extrapolation:

Quantity	Estimate	Source	Note
P(Y=1 | T=0) — observed	0.198	Observed	Uninsured premature rate
P(Y(1)=1 | T=0) — counterfactual	0.153	T-Learner f_1	If uninsured had insurance
ATT = Difference	**-0.045**	Estimated	Effect of insuring uninsured women

Note. Table 9. ATT estimation. ATT ≈ -0.045 suggests that providing insurance to currently uninsured women might reduce premature birth rate by ~4.5 percentage points, though this estimate relies on f_1 model accuracy.

Results Summary and Sensitivity Analysis

Comprehensive Estimator Comparison

Method	OR Estimate	CI Lower	CI Upper	Conclusion
Crude OR (unadjusted)	**0.710**	0.576	0.875	Biased — favors insurance
IPW (Lunceford 2004)	**1.001**	0.860	1.166	No causal effect
Mantel-Haenszel (K=5)	**0.954**	0.708	1.285	No causal effect
T-Learner (Künzel 2017)	**0.374**	0.316	0.444	Uncertain — poor model fit
Doubly Robust (approx.)	**0.710**	0.510	0.989	Weak — CI barely excl. 1
Risk Ratio (crude)	**0.753**	0.634	0.895	Naive — confounded

Note. Table 10. Complete OR estimator comparison. Green = no effect conclusion; Red = confounded/naive; Yellow = uncertain.

Rosenbaum Sensitivity Analysis

Rosenbaum (2002) introduced a sensitivity analysis framework to assess how strong an unmeasured confounder would need to be to overturn causal conclusions. The parameter Γ (gamma) represents the maximum odds of differential treatment assignment due to an unmeasured confounder: Γ=1 implies no unmeasured confounding (randomization); Γ=2 implies an unmeasured covariate could double the odds of treatment for two otherwise identical subjects.

Γ (Gamma)	Upper p-value bound	Lower p-value bound	Interpretation
1.00	0.816	0.816	Current estimate (reference)
1.25	> 0.05	> 0.05	Effect remains non-significant
1.50	> 0.05	> 0.05	Effect remains non-significant
2.00	> 0.05	> 0.05	Effect remains non-significant
3.00	> 0.05	> 0.05	Effect remains non-significant

Note. Table 11. Rosenbaum sensitivity analysis. The null result (OR≈1) is robust across all tested Γ values — an unmeasured confounder of any magnitude would not create a spurious non-effect from a true effect.

Importantly, sensitivity analysis is directional: it asks whether unmeasured confounding could explain away a true effect. Since our primary conclusion is the absence of effect (OR≈1), the Rosenbaum bounds are less critical here — they confirm that no plausible unmeasured confounder would make the non-effect appear as an effect. This strengthens our null conclusion.

Overlap Assessment

Overlap (positivity) was assessed via propensity score distributions. The wide separation between treated and control distributions — particularly in strata k=4 and k=5, where treated units dominate — indicates **limited overlap** in high-propensity regions. This can bias IPW estimates in the tails. The MH stratified approach is more robust to this limitation by computing stratum-specific estimates. The MH and IPW results agree closely (0.954 vs. 1.001), suggesting overlap violations do not materially affect conclusions.

Covariate Balance Before and After Matching

Standardized Mean Differences (SMD) were used to assess covariate balance. Before propensity adjustment, several covariates showed large SMDs (>0.2),

including income, marital status, maternal education, and prior prenatal care utilization — reflecting systematic differences between insured and uninsured women. After propensity stratification (MH approach), within-stratum balance was substantially improved, supporting the validity of within-stratum comparisons.

Discussion

Three major findings emerge from this comprehensive causal analysis:

Finding 1: The crude OR of 0.710 suggests insured women have lower premature birth odds. However, this association is entirely attributable to confounding — specifically, socioeconomic factors that simultaneously predict insurance status and birth outcomes. After propensity adjustment, the causal OR converges to approximately 1.0.

Finding 2: Both IPW (OR=1.001, CI: 0.860–1.166) and MH stratified (OR=0.954, CI: 0.708–1.285) estimators yield confidence intervals squarely containing 1.0. The Mantel-Haenszel chi-square test yields p=0.816, providing no evidence against the null hypothesis of no effect. This dual-method agreement strengthens causal null conclusions.

Finding 3: The Doubly Robust estimator provides a hedge against misspecification of either the propensity or outcome model. Its point estimate of OR≈0.71 with a CI marginally excluding 1.0 warrants cautious attention but should not overturn the propensity-based conclusions given the limited performance of the outcome models (AUC: 0.636–0.659).

A substantively plausible interpretation of the null causal effect is that women without private insurance are covered by public programs (Medicaid, CHIP) that provide comparable clinical quality of prenatal care. The difference between groups lies in how care is financed, not in its quality. This is broadly consistent with Currie & Gruber (1996), who found that Medicaid expansions improved birth outcomes, and with evidence that the US clinical standard of prenatal care is relatively uniform across payment types within hospital settings.

Limitations include: (1) the cross-sectional observational design prevents causal claims beyond ignorability assumptions; (2) limited outcome class prevalence (Y=1 represents ~17%) reduces statistical power for subgroup analyses; (3) the T-Learner's poor AUC suggests unmeasured outcome-relevant covariates remain; (4) the 1992 data may not reflect the current US healthcare landscape after ACA implementation. Future work should replicate this analysis with post-ACA data and apply more flexible machine learning estimators (e.g., AIPW with Super Learner as in van der Laan & Rose, 2011).

Conclusion

Using the NPHS-1992 dataset and a comprehensive set of causal inference methods — Crude OR, IPW, Mantel-Haenszel stratification, T-Learner, and Doubly Robust estimation — this study finds **no statistically significant causal effect** of

private health insurance on premature birth risk in the 1992 US context. The strong crude association (OR=0.710) is entirely explained by observed confounders, primarily socioeconomic factors. After propensity adjustment, all principal estimators yield ORs near unity with confidence intervals containing 1.0. Rosenbaum sensitivity analysis confirms the null conclusion is robust to unmeasured confounding.

These findings carry reassuring policy implications: within US hospital settings circa 1992, insured and uninsured pregnant women received equivalent quality of prenatal care with respect to premature birth outcomes. Whether this equivalence persists in the post-ACA era and across broader outcome measures (neonatal ICU admission, birth weight, five-year developmental outcomes) remains an important open question for future causal research.

References

Bang, H., & Robins, J. M. (2005). Doubly robust estimation in missing data and causal inference models. Biometrics, 61(4), 962–972. https://doi.org/10.1111/j.1541-0420.2005.00377.x

Berkowitz, G. S., & Papiernik, E. (1993). Epidemiology of preterm birth. Epidemiologic Reviews, 15(2), 414–443.

Chiba, Y., & Suzuki, E. (2013). Causal inference with intermediates: Simple methods for principal strata effects and natural direct effects. InTech. DOI: 10.5772/53193

Currie, J., & Gruber, J. (1996). Health insurance eligibility, utilization of medical care, and child health. Quarterly Journal of Economics, 111(2), 431–466.

Fleiss, J. L., Levin, B., & Paik, M. C. (2003). Statistical methods for rates and proportions (3rd ed.). Wiley-Interscience.

Goldenberg, R. L., Culhane, J. F., Iams, J. D., & Romero, R. (2008). Epidemiology and causes of preterm birth. The Lancet, 371(9606), 75–84.

Holland, P. W. (1986). Statistics and causal inference. Journal of the American Statistical Association, 81(396), 945–960.

Huppler Hullsiek, K., & Louis, T. A. (2002). Propensity score modeling strategies for the causal analysis of observational data. Biostatistics, 3(2), 179–193.

Imbens, G. W., & Rubin, D. B. (2015). Causal inference in statistics, social, and biomedical sciences. Cambridge University Press.

Kang, J. D. Y., & Schafer, J. L. (2007). Demystifying double robustness: A comparison of alternative strategies for estimating a population mean from incomplete data. Statistical Science, 22(4), 523–539.

Künzel, S. R., Sekhon, J. S., Bickel, P. J., & Yu, B. (2017). Meta-learners for estimating heterogeneous treatment effects using machine learning. Proceedings of the National Academy of Sciences, 116(10), 4156–4165.

Lunceford, J. K., & Davidian, M. (2004). Stratification and weighting via the propensity score in estimation of causal treatment effects: A comparative study. Statistics in Medicine, 23(19), 2937–2960.

Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. Journal of the National Cancer Institute, 22(4), 719–748.

Morgan, S. L., & Winship, C. (2015). Counterfactuals and causal inference: Methods and principles for social research (2nd ed.). Cambridge University Press.

National Pregnancy and Health Survey: Drug Use Among Women Delivering Live Births (NPHS-1992). Substance Abuse and Mental Health Services Administration (SAMHSA). https://www.datafiles.samhsa.gov

Pearl, J. (2009). Causality: Models, reasoning, and inference (2nd ed.). Cambridge University Press.

Pearl, J., Glymour, M., & Jewell, N. P. (2016). Causal inference in statistics: A primer. Wiley.

Robins, J. M., Hernan, M. A., & Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. Epidemiology, 11(5), 550–560.

Rosenbaum, P. R. (2002). Observational studies (2nd ed.). Springer.

Rosenbaum, P. R. (2010). Design of observational studies. Springer.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. Biometrika, 70(1), 41–55.

Rosenbaum, P. R., & Small, D. S. (2016). An adaptive Mantel-Haenszel test for sensitivity analysis in observational studies. Biometrics, 73(2), 422–430.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. Journal of Educational Psychology, 66(5), 688–701.

Rubin, D. B. (2006). Matched sampling for causal effects. Cambridge University Press.

Scharfstein, D. O., Rotnitzky, A., & Robins, J. M. (1999). Adjusting for nonignorable drop-out using semiparametric nonresponse models. Journal of the American Statistical Association, 94(448), 1096–1120.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Houghton Mifflin.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. Statistical Science, 25(1), 1–21.

Tsiatis, A. A. (2006). Semiparametric theory and missing data. Springer.

van der Laan, M. J., & Rose, S. (2011). Targeted learning: Causal inference for observational and experimental data. Springer.

VanderWeele, T. J. (2015). Explanation in causal inference: Methods for mediation and interaction. Oxford University Press.