Supplementary Materials for
"How Should We Estimate Inverse Probability Weights with Possibly Misspecified Propensity Score Models?"

A Literature Review of the Weighting and Matching Methodology in Political Science

To investigate the prevalence of the IPW in political science, I survey all the articles published in the *American Journal of Political Science*, the *American Political Science Review*, and the *Journal of Politics* from 2000 and 2022. I employ the following procedure:

1. Search the following words in the three journals via Google Scholar on March 15, 2022.
2. Carefully read all the articles found by the search to code whether they use each of the popular weighting and matching methods (the IPW, entropy balancing, propensity score matching, genetic matching).

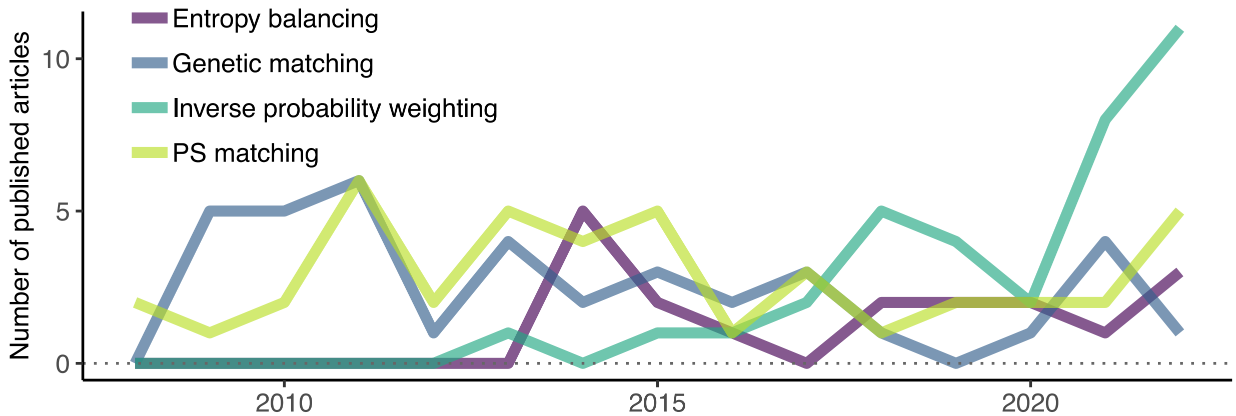
The search words are following:

- IPW
- inverse probability weighting
- inverse probability weights
- inverse probability weighted
- propensity score weighting
- propensity score weights
- propensity score weighted

- CBPS
- covariate balancing propensity scores
- covariate balancing propensity score
- entropy balancing
- entropy balance
- (matching OR matched OR match) AND ("propensity score" OR "propensity scores")
- "genetic" AND "matching"

The identified articles are presented in Table A.1, where I code the journal, author(s), and title of the article (Journal, Author, and Title), the year, volume, and issue of the publication (Year, Vol., and Issue), and whether the article uses the IPW, CBPS, entropy balancing, propensity score matching, and genetic matching (IPW, CBPS, EB, PSM, GM). Note that the publication year of the forthcoming articles is coded as 999 and treated as publication in 2022 in the following analysis.

To summarise the results, Figure A.1 shows the number of published articles that use each of the IPW (green), entropy balancing (purple), propensity score matching (yellow), and genetic matching (blue) in the three journals each year from 2008 and 2022. The IPW is increasingly utilized in the past five years, making it the most widely used method in each of the five years. The decline of the propensity score matching and the rise of the IPW after 2015 may be partly attributable to (King and Nielsen 2019), an influential work (cited 1106 times as of March 14, 2022) presented at Polmeth 2015 that recommends not to use the propensity score matching.



Notes: This figure shows the number of published articles that use each of the IPW (green), entropy balancing (purple), propensity score matching (yellow), and genetic matching (blue) in the *American Journal of Political Science*, the *American Political Science Review*, and the *Journal of Politics* in each year from 2008 and 2022. The IPW is increasingly utilized in the past five years, making it the most widely used method in each of the five years.

Figure A.1: The rise of inverse probability weighting

Table A.1: Literature review

Journal	Year	Vol.	Issue	Author	Title	IPW	CBPS	EB	PSM	GM
AJPS	999	0	0	Brandon L. Bartels, Jeremy Horowitz, Eric Kramon	Can Democratic Principles Protect High Courts from Partisan Backlash? Public Reactions to the Kenyan Supreme Court's Role in the 2017 Election Crisis	1	0	0	0	0
AJPS	999	0	0	Adriana Molina-Garzón, Boulder Tara Grillos, Alan Zarychta, Krister P. Andersson	Decentralization Can Increase Cooperation among Public Officials	0	0	0	1	0
AJPS	999	0	0	Saad Gulzar, Miguel R. Rueda, Nelson A. Ruiz	Do Campaign Contribution Limits Curb the Influence of Money in Politics?	1	0	0	0	0
AJPS	999	0	0	Jennifer L. Tobin, Christina J. Schneider, David Leblang	Framing Unpopular Foreign Policies	0	0	1	0	0
AJPS	999	0	0	Kosuke Imai, In Song Kim, Erik H. Wang	Matching Methods for Causal Inference with Time-Series Cross-Sectional Data	1	1	0	1	0
AJPS	999	0	0	Mitchell Kilborn and Arjun Vishwanath	Public Money Talks Too: How Public Campaign Financing Degrades Representation	1	0	0	0	0
AJPS	999	0	0	Soledad Artiz Prillaman	Strength in Numbers: How Women's Groups Close India's Political Gender Gap	0	0	0	0	1
AJPS	999	0	0	Adam Zelizer	Talking Shops: The Effects of Caucus Discussion on Policy Coalitions	1	0	0	0	0
AJPS	999	0	0	Vincenzo Bove, Jessica Di Salvatore, Leandro Elia	UN Peacekeeping and Households' Well-Being in Civil Wars	1	0	0	1	0

Table A.1: *(continued)*

Journal	Year	Vol.	Issue	Author	Title	IPW	CBPS	EB	PSM	GM
AJPS	2009	53	2	Jonathan McDonald Ladd, Gabriel S. Lenz	Exploiting a Rare Communication Shift to Document the Persuasive Power of the News Media	0	0	0	0	1
AJPS	2010	54	1	Melissa J. Marschall, Anirudh V. S. Ruhil, Paru R. Shah	The New Racial Calculus: Electoral Institutions and Black Representation in Local Legislatures	0	0	0	1	0
AJPS	2010	54	2	Michael MacKuen, Jennifer Wolak, Luke Keele, George E. Marcus	Civic Engagements: Resolute Partisanship or Reflective Deliberation	0	0	0	0	1
AJPS	2010	54	2	Christina L. Boyd, Lee Epstein, Andrew D. Martin	Untangling the Causal Effects of Sex on Judging	0	0	0	1	0
AJPS	2011	55	2	Matthew Adam Kocher, Thomas B. Pepinsky, Stathis N. Kalyvas	Aerial Bombing and Counterinsurgency in the Vietnam War	0	0	0	1	1
AJPS	2011	55	2	Steven E. Finkel, Amy Erica Smith	Civic Education, Political Discussion, and the Social Transmission of Democratic Knowledge and Values in a New Democracy: Kenya 2002	0	0	0	1	0
AJPS	2011	55	2	Richard A. Nielsen, Michael G. Findley, Zachary S. Davis, Tara Candland, Daniel L. Nielson	Foreign Aid Shocks as a Cause of Violent Armed Conflict	0	0	0	1	1
AJPS	2011	55	4	Taylor C. Boas, F. Daniel Hidalgo	Controlling the Airwaves: Incumbency Advantage and Community Radio in Brazil	0	0	0	0	1
AJPS	2012	56	3	Paul S. Herrnson, Michael J. Hanmer, Richard G. Niemi	The Impact of Ballot Type on Voter Errors	0	0	0	1	0

Table A.1: *(continued)*

Journal	Year	Vol.	Issue	Author	Title	IPW	CBPS	EB	PSM	GM
AJPS	2013	57	1	Casey A. Klofstad, Anand Edward Sokhey, Scott D. McClurg	Disagreeing about Disagreement: How Conflict in Social Networks Affects Political Behavior	0	0	0	1	0
AJPS	2013	57	1	Jonathan P. Kastellec	Racial Diversity and Judicial Influence on Appellate Courts	0	0	0	1	0
AJPS	2013	57	2	Matthew Blackwell	A Framework for Dynamic Causal Inference in Political Science	1	0	0	0	0
AJPS	2013	57	2	Nahomi Ichino, Noah L. Nathan	Do Primaries Improve Electoral Performance? Clientelism and Intra-Party Conflict in Ghana	0	0	0	0	1
AJPS	2013	57	4	Yonatan Lupu	The Informative Power of Treaty Commitment: Using the Spatial Model to Address Selection Effects	0	0	0	1	0
AJPS	2013	57	4	Lisa Hultman, Jacob Kathman, Megan Shannon	United Nations Peacekeeping and Civilian Protection in Civil War	0	0	0	0	1
AJPS	2013	57	4	Cesar Zucco Jr.	When Payouts Pay Off: Conditional Cash Transfers and Voting Behavior in Brazil 2002–10	0	0	0	1	1
AJPS	2014	58	2	Carly Urban, Sarah Niebler	Dollars on the Sidewalk: Should U.S. Presidential Candidates Advertise in Uncontested States?	0	0	0	1	1
AJPS	2014	58	2	Jon C. Rogowski	Electoral Choice, Ideological Conflict, and Political Participation	0	0	0	0	1
AJPS	2014	58	4	Michael M. Bechtel, Jens Hainmueller, Yotam Margalit	Preferences for International Redistribution: The Divide over the Eurozone Bailouts	0	0	1	0	0

Table A.1: *(continued)*

Journal	Year	Vol.	Issue	Author	Title	IPW	CBPS	EB	PSM	GM
AJPS	2014	58	4	Jens Blom-Hansen, Kurt Houlberg, Søren Serritzlew	Size, Democracy, and the Economic Costs of Running the Political System	0	0	0	1	0
AJPS	2015	59	3	Yonatan Lupu	Legislative Veto Players and the Effects of International Human Rights Agreements	0	0	0	1	0
AJPS	2016	60	1	Guy Grossman, Oren Gazal-Ayal, Samuel D. Pimentel, Jeremy M. Weinstein	Descriptive Representation and Judicial Outcomes in Multiethnic Societies	1	0	0	0	0
AJPS	2016	60	3	Adam S. Chilton, Mila Versteeg	Do Constitutional Rights Make a Difference?	0	0	0	1	0
AJPS	2016	60	4	Leah C. Stokes	Electoral Backlash against Climate Policy: A Natural Experiment on Retrospective Voting and Local Resistance to Public Policy	0	0	1	0	0
AJPS	2016	60	4	David E. Broockman, Timothy J. Ryan	Preaching to the Choir: Americans Prefer Communicating to Copartisan Elected Officials	0	0	0	0	1
AJPS	2017	61	2	Florian Foos, Eline A. de Rooij	All in the Family: Partisan Disagreement and Electoral Mobilization in Intimate Networks-A Spillover Experiment	1	0	0	0	0
AJPS	2017	61	3	Tali Mendelberg, Katherine T. McCabe, Adam Thal	College Socialization and the Economic Views of Affluent Americans	0	0	0	0	1
AJPS	2017	61	3	Sophie Schuit, Jon C. Rogowski	Race, Representation, and the Voting Rights Act	0	0	0	0	1
AJPS	2018	62	1	Connor Huff, Joshua D. Kertzer	How the Public Defines Terrorism	0	0	1	0	0

Table A.1: *(continued)*

Journal	Year	Vol.	Issue	Author	Title	IPW	CBPS	EB	PSM	GM
AJPS	2018	62	1	Charles J. Finocchiaro, Scott A. MacKenzie	Making Washington Work: Legislative Entrepreneurship and the Personal Vote from the Gilded Age to the Great Depression	0	0	0	0	1
AJPS	2018	62	1	Clayton Nall, Benjamin Schneer, Daniel Carpenter	Paths of Recruitment: Rational Social Prospecting in Petition Canvassing	1	0	0	0	0
AJPS	2018	62	3	Greg Distelhorst. Richard M. Locke	Does Compliance Pay? Social Standards and Firm-Level Trade	0	0	1	0	0
AJPS	2019	63	4	Bradford Jones, Kristina Flores Victor, David Vannette	Alien Citizens and the Canonical Immigrant: Do Stigmatized Attributes Affect Latina/o Judgment about Discrimination?	0	0	1	0	0
AJPS	2019	63	4	Daniel W. Gingerich	Ballot Reform as Suffrage Restriction: Evidence from Brazil's Second Republic	1	0	0	0	0
AJPS	2019	63	4	Daniel Bischof, Markus Wagner	Do Voters Polarize When Radical Parties Enter Parliament?	0	0	0	1	0
AJPS	2020	64	2	Kevin Arceneaux, Johanna Dunaway, Martin Johnson, Ryan J. Vander Wielen	Strategic Candidate Entry and Congressional Elections in the Era of Fox News	0	0	0	0	1
AJPS	2020	64	4	Margaret E. Roberts, Brandon M. Stewart, Richard A. Nielsen	Adjusting for Confounding with Text Matching	0	0	0	1	0
AJPS	2021	65	1	Paul Frymer, Jacob M. Grumbach	Labor Unions and White Racial Politics	0	0	0	0	1

Table A.1: *(continued)*

Journal	Year	Vol.	Issue	Author	Title	IPW	CBPS	EB	PSM	GM
AJPS	2021	65	1	Tara Slough, Christopher Fariss	Misgovernance and Human Rights: The Case of Illegal Detention without Intent	1	0	0	0	0
AJPS	2021	65	1	Richard Clark, Lindsay R. Dolan	Pleasing the Principal: U.S. Influence in World Bank Policymaking	1	0	0	0	0
AJPS	2021	65	2	Maria Jose Hierro, Didac Queralt	The Divide Over Independence: Explaining Preferences for Secession in an Advanced Open Economy	0	0	1	0	0
AJPS	2021	65	4	Molly Offer-Westort, Alexander Coppock, Donald P. Green	Adaptive Experimental Design: Prospects and Applications in Political Science	1	0	0	0	0
AJPS	2021	65	4	Rune Slothuus, Martin Bisgaard	How Political Parties Shape Public Opinion in the Real World	0	0	0	1	0
AJPS	2022	66	1	Gabriel López-Moctezuma, Leonard Wantchekon, Daniel Rubenson, Thomas Fujiwara, Cecilia Pe Lero	Policy Deliberation and Voter Persuasion: Experimental Evidence from an Election in the Philippines	1	0	0	0	0
APSR	999	0	0	COLE J. HARVEY	Can Courts in Nondemocracies Deter Election Fraud? De Jure Judicial Independence, Political Competition, and Election Integrity	1	1	1	0	0
APSR	999	0	0	DONGHYUN DANNY CHOI, J. ANDREW HARRIS, FIONA SHEN-BAYH	Ethnic Bias in Judicial Decision Making: Evidence from Criminal Appeals in Kenya	1	0	0	0	0

Table A.1: *(continued)*

Journal	Year	Vol.	Issue	Author	Title	IPW	CBPS	EB	PSM	GM
APSR	999	0	0	NINA MCMURRY	From Recognition to Integration: Indigenous Autonomy, State Authority, and National Identity in the Philippines	0	0	0	1	0
APSR	2004	98	4	JASON BARABAS	How Deliberation Affects Policy Opinions	0	0	0	1	0
APSR	2005	99	2	KOSUKE IMAI	Do Get-Out-the-Vote Calls Reduce Turnout? The Importance of Statistical Methods for Field Experiments	0	0	0	1	0
APSR	2005	99	4	BETH A. SIMMONS, DANIEL J. HOPKINS	The Constraining Power of International Treaties: Theory and Methods	0	0	0	1	0
APSR	2006	100	1	KASPAR RICHTER	Wage Arrears and Economic Voting in Russia	0	0	0	1	0
APSR	2009	103	1	MATTHEW KROENIG	Exporting the Bomb: Why States Provide Sensitive Nuclear Assistance	0	0	0	0	1
APSR	2009	103	3	ANDREW HEALY, NEIL MALHOTRA	Myopic Voters and Natural Disaster Policy	0	0	0	1	0
APSR	2009	103	4	SANFORD C. GORDON	Assessing Partisan Bias in Federal Public Corruption Prosecutions	0	0	0	0	1
APSR	2009	103	4	ANDREW C. EGGERS, JENS HAINMUELLER	MPs for Sale? Returns to Office in Postwar British Politics	0	0	0	0	1
APSR	2010	104	1	DANIEL J. HOPKINS	Politicized Places: Explaining Where and When Immigrants Provoke Local Opposition	0	0	0	0	1

Table A.1: *(continued)*

Journal	Year	Vol.	Issue	Author	Title	IPW	CBPS	EB	PSM	GM
APSR	2010	104	4	VESLA M. WEAVER, AMY E. LERMAN	Political Consequences of the Carceral State	0	0	0	0	1
APSR	2011	105	2	DAVID DREYER LASSEN, SØREN SERRITZLEW	Jurisdiction Size and Local Democracy: Evidence on Internal Political Efficacy from Large-scale Municipal Reform	0	0	0	1	0
APSR	2012	106	1	JASJEET S. SEKHON, ROCIO TITIUNIK	When Natural Experiments Are Neither Natural nor Experiments	0	0	0	0	1
APSR	2013	107	1	YOTAM MARGALIT	Explaining Social Policy Preferences: Evidence from the Great Recession	0	0	0	1	0
APSR	2013	107	1	JENS HAINMUELLER, DOMINIK HANGARTNER	Who Gets a Swiss Passport? A Natural Experiment in Immigrant Discrimination	0	0	0	0	1
APSR	2014	108	2	TARIQ THACHIL	Elite Parties and Poor Voters: Theory and Evidence from India	0	0	0	1	0
APSR	2014	108	2	RORY TRUEX	The Returns to Office in a "Rubber Stamp" Parliament	0	0	1	0	0
APSR	2014	108	3	JEREMY FERWERDA, NICHOLAS L. MILLER	Political Devolution and Resistance to Foreign Rule: A Natural Experiment	0	0	1	0	0
APSR	2014	108	3	CHRIS TAUSANOVITCH, CHRISTOPHER WARSHAW	Representation in Municipal Government	0	0	1	0	0

Table A.1: *(continued)*

Journal	Year	Vol.	Issue	Author	Title	IPW	CBPS	EB	PSM	GM
APSR	2014	108	4	LISA HULTMAN, JACOB KATHMAN, MEGAN SHANNON	Beyond Keeping Peace: United Nations Effectiveness in the Midst of Fighting	0	0	0	1	0
APSR	2015	109	1	ARIEL R. WHITE, NOAH L. NATHAN, JULIE K. FALLER	What Do I Need to Vote? Bureaucratic Discretion and Discrimination by Local Election Officials	0	0	0	1	0
APSR	2015	109	2	MONICA DUFFY TOFT, YURI M. ZHUKOV	Islamists and Nationalists: Rebel Motivation and Counterinsurgency in Russia's North Caucasus	0	0	0	1	0
APSR	2015	109	3	ARUN AGRAWAL, ASHWINI CHHATRE, ELISABETH R. GERBER	Motivational Crowding in Sustainable Development Interventions	0	0	0	1	0
APSR	2017	111	1	MICHAEL TOUCHTON, NATASHA BORGES SUGIYAMA, BRIAN WAMPLER	Democracy at Work: Moving Beyond Elections to Improve Well-Being	0	0	0	1	0
APSR	2017	111	3	JAMES BISBEE, JENNIFER M. LARSON	Testing Social Science Network Theories with Online Network Data: An Evaluation of External Validity	1	0	0	0	0
APSR	2018	112	4	PETER SELB, SIMON MUNZERT	Examining a Most Likely Case for Strong Campaign Effects: Hitler's Speeches and the Rise of the Nazi Party, 1927–1933	0	0	0	1	0
APSR	2018	112	4	MATTHEW BLACKWELL, ADAM N. GLYNN	How to Make Causal Inferences with Time-Series Cross-Sectional Data under Selection on Observables	1	0	0	0	0

Table A.1: *(continued)*

Journal	Year	Vol.	Issue	Author	Title	IPW	CBPS	EB	PSM	GM
APSR	2018	112	4	DANIEL M. BUTLER, HANS J.G. HASSELL	On the Limits of Officials' Ability to Change Citizens' Priorities: A Field Experiment in Local Politics	1	0	0	0	0
APSR	2019	113	2	VOLHA CHARNYSH	Diversity, Institutions, and Economic Outcomes: Post-WWII Displacement in Poland	1	1	0	0	0
APSR	2019	113	2	ADAM ZELIZER	Is Position-Taking Contagious? Evidence of Cue-Taking from Two Field Experiments in a State Legislature	1	0	0	0	0
APSR	2020	114	3	OMAR WASOW	Agenda Seeding: How 1960s Black Protests Moved Elites, Public Opinion and Voting	1	1	0	1	0
APSR	2021	115	3	NANCY ARRINGTON, LEEANN BASS, ADAM GLYNN, JEFFREY K. STATON, BRIAN DELGADO, STAFFAN I. LINDBERG	Constitutional Reform and the Gender Diversification of Peak Courts	0	0	0	1	0
APSR	2021	115	3	KEVIN MORRIS	Turnout and Amendment Four: Mobilizing Eligible Voters Close to Formerly Incarcerated Floridians	0	0	0	0	1
APSR	2021	115	4	WILLIAM MINOZZI, GREGORY A. CALDEIRA	Congress and Community: Coresidence and Social Influence in the U.S. House of Representatives, 1801-1861	1	0	0	0	0
APSR	2022	116	1	JAMES BISBEE, DAN HONIG	Flight to Safety: COVID-Induced Changes in the Intensity of Status Quo Preference and Voting Behavior	1	1	0	0	0

Table A.1: *(continued)*

Journal	Year	Vol.	Issue	Author	Title	IPW	CBPS	EB	PSM	GM
JOP	2008	70	3	Mark S. Copelovitch, David Andrew Singer	Financial Regulation, Monetary Policy, and Inflation in the Industrialized World	0	0	0	1	0
JOP	2008	70	3	Cindy D. Kam, Carl L. Palmer	Reconsidering the Effects of Education on Political Participation	0	0	0	1	0
JOP	2009	71	1	Daniel J. Hopkins	The Diversity Discount: When Increasing Ethnic and Racial Diversity Prevents Tax Increases	0	0	0	0	1
JOP	2010	72	3	Eric Schickler, Kathryn Pearson, Brian D. Feinstein	Congressional Parties and Civil Rights Politics from 1933 to 1972	0	0	0	0	1
JOP	2010	72	4	Michaela Mattes, Greg Vonnahme	Contracting for Peace: Do Nonaggression Pacts Reduce Conflict?	0	0	0	0	1
JOP	2011	73	3	Alexander K. Mayer	Does Education Increase Political Participation?	0	0	0	1	1
JOP	2011	73	3	Will Bullock, Joshua D. Clinton	More a Molehill than a Mountain: The Effects of the Blanket Primary on Elected Officials' Behavior from California	0	0	0	0	1
JOP	2011	73	3	John Henderson, Sara Chatfield	Who Matches? Propensity Scores and Bias in the Causal Effects of Education on Participation	0	0	0	1	1
JOP	2012	74	1	Steven E. Finkel, Jeremy Horowitz, Reynaldo T. Rojo-Mendoza	Civic Education and Democratic Backsliding in the Wake of Kenya's Post-2007 Election Violence	0	0	0	1	0

Table A.1: *(continued)*

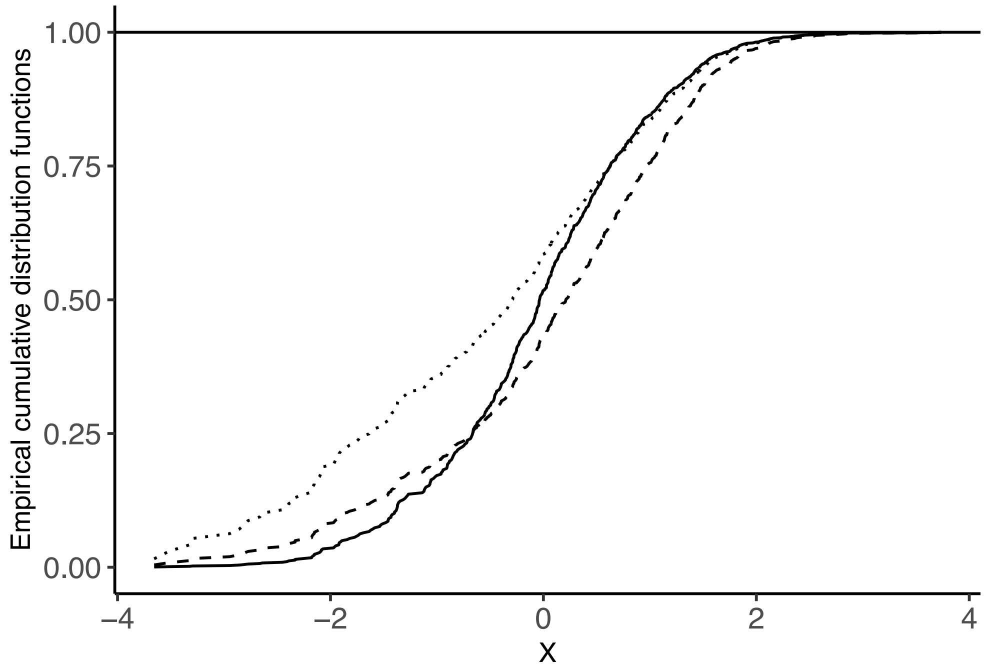
Journal	Year	Vol.	Issue	Author	Title	IPW	CBPS	EB	PSM	GM
JOP	2014	76	3	Joshua D. Kertzer, Kathleen E. Powers, Brian C. Rathbun, Ravi Iyer	Moral Support: How Moral Values Shape Foreign Policy Attitudes	0	0	1	0	0
JOP	2015	77	1	Alberto Chong, Ana L. De La O, Dean Karlan, Leonard Wantchekon	Does Corruption Information Inspire the Fight or Quash the Hope? A Field Experiment in Mexico on Voter Turnout, Choice, and Party Identification	1	0	0	0	0
JOP	2015	77	1	Ryan E. Carlin, Mason Moseley	Good Democrats, Bad Targets: Democratic Values and Clientelistic Vote Buying	0	0	1	0	0
JOP	2015	77	2	Scott A. MacKenzie	Life before Congress: Using Precongressional Experience to Assess Competing Explanations for Political Professionalism	0	0	0	0	1
JOP	2015	77	3	Jens Hainmueller, Andrew B. Hall, James M. Snyder Jr.	Assessing the External Validity of Election RD Estimates: An Investigation of the Incumbency Advantage	0	0	1	0	1
JOP	2015	77	3	Reed M. Wood, Christopher Sullivan	Doing Harm by Doing Good? The Negative Externalities of Humanitarian Aid Provision during Civil Conflict	0	0	0	0	1
JOP	2015	77	3	Rachael K. Hinkle	Legal Constraint in the US Courts of Appeals	0	0	0	1	0
JOP	2016	78	1	Arturas Rozenas	Office Insecurity and Electoral Manipulation	0	0	0	0	1
JOP	2017	79	1	Alexander Hertel-Fernandez	American Employers as Political Machines	0	0	0	1	0

Table A.1: *(continued)*

Journal	Year	Vol.	Issue	Author	Title	IPW	CBPS	EB	PSM	GM
JOP	2017	79	1	Patrick Emmenegger, Paul Marx, Dominik Schraff	Off to a Bad Start: Unemployment and Political Interest during Early Adulthood	0	0	0	1	0
JOP	2017	79	3	Luke J. Keele, Paru R. Shah, Ismail White, Kristine Kay	Black Candidates and Black Turnout: A Study of Viability in Louisiana Mayoral Elections	0	0	0	0	1
JOP	2018	80	2	Lucy Barnes, Avi Feller, Jake Haselswerdt, Ethan Porter	Information, Knowledge, and Attitudes: An Evaluation of the Taxpayer Receipt	1	0	0	0	0
JOP	2018	80	4	Larry M. Bartels	Partisanship in the Trump Era	1	0	0	0	0
JOP	2019	81	1	Albert H. Fang, Andrew M. Guess, Macartan Humphreys	Can the Government Deter Discrimination? Evidence from a Randomized Intervention in New York City	1	0	0	0	0
JOP	2019	81	1	Pablo Beramendi, Mark Dincecco, Melissa Rogers	Intra-Elite Competition and Long-Run Fiscal Development	0	0	0	1	0
JOP	2019	81	3	Rory Truex, Daniel L. Tavana	Implicit Attitudes toward an Authoritarian Regime	0	0	1	0	0
JOP	2020	82	1	Christina J. Schneider	Public Commitments as Signals of Responsiveness in the European Union	0	0	1	0	0
JOP	2020	82	4	Omer Yair, Raanan Sulitzeanu-Kenan, Yoav Dotan	Can Institutions Make Voters Care about Corruption?	0	0	1	0	0
JOP	2020	82	4	Jean Lachapelle	No Easy Way Out: The Effect of Military Coups on State Repression	1	1	0	0	0
JOP	2021	83	1	Jennifer Chudy	Racial Sympathy and Its Political Consequences	1	0	0	0	0
JOP	2021	83	3	Florian Foos, Peter John, Christian Müller, Kevin Cunningham	Social Mobilization in Partisan Spaces	1	0	0	0	0

Table A.1: *(continued)*

Journal	Year	Vol.	Issue	Author	Title	IPW	CBPS	EB	PSM	GM
JOP	2021	83	4	Jonathan S. Krasno, Donald P. Green, Costas Panagopoulos, Dane Thorley, Michael Schwam-Baird	Campaign Donations, Judicial Recusal, and Disclosure: A Field Experiment	1	0	0	0	0
JOP	2021	83	4	Melina R. Platas, Pia J. Raffler	Closing the Gap: Information and Mass Support in a Dominant Party Regime	1	0	0	0	0
JOP	2021	83	4	Livia Isabella Schubiger	State Violence and Wartime Civilian Agency: Evidence from Peru	0	0	0	0	1
JOP	2021	83	4	Jason Douglas Todd, Edmund J. Malesky, Anh Tran, Quoc Anh Le	Testing Legislator Responsiveness to Citizens and Firms in Single-Party Regimes: A Field Experiment in the Vietnamese National Assembly	0	0	0	0	1
JOP	2022	84	1	Quintin H. Beazer, Ora John Reuter	Do Authoritarian Elections Help the Poor? Evidence from Russian Cities	0	0	1	0	0
JOP	2022	84	1	Abby Córdova, Helen Kras	State Action to Prevent Violence against Women: The Effect of Women's Police Stations on Men's Attitudes toward Gender-Based Violence	0	0	0	1	0
JOP	2022	84	2	Ryan Hübert, Ryan Copus	Political Appointments and Outcomes in Federal District Courts	1	0	0	0	0



Notes: This figure shows the empirical cumulative distribution functions (ECDF) of the covariate X. The solid line represents the empirical CDF for the target group ($\mathcal{S} = \{i \mid R_i \in \{0,1\}\}$), and the remaining two lines represent the weighted ECDFs of the weighted group ($\mathcal{S}_1 = \{i \mid R_i = 1\}$) where weights are estimated by the proposed method (the dashed line) and the MLE (the dotted line).

Figure B.1: The empirical cumulative distribution functions

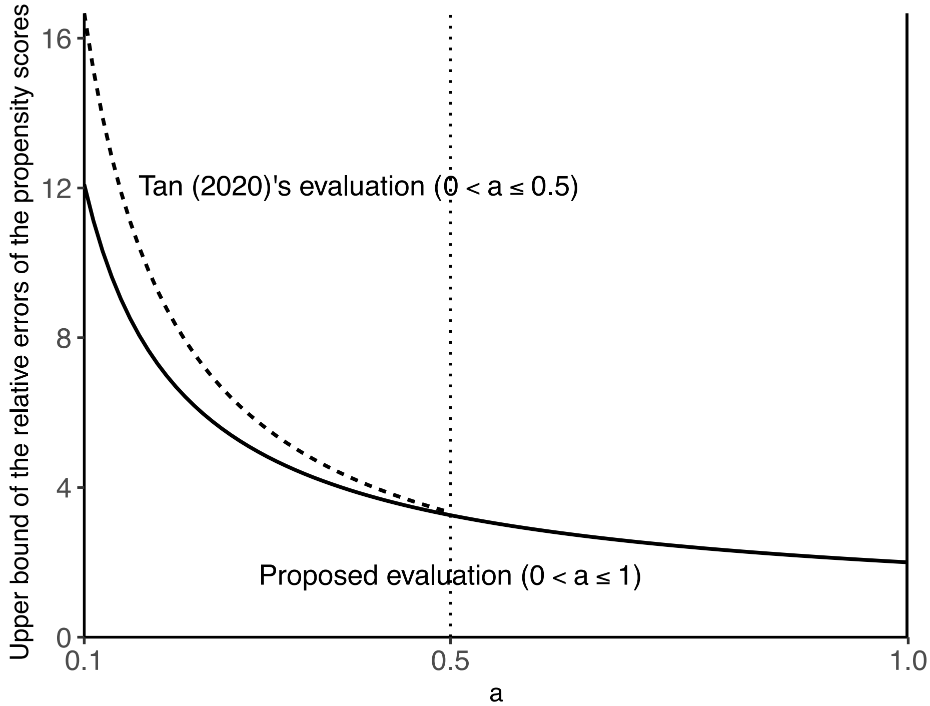
B Distributional Balance in the Numerical Example

Figure B.1 shows the empirical distribution functions of covariate X. The proposed method (the dashed line) approximates the target distribution (the solid line) better than the MLE (the dotted line).

C Proofs

C.1 Proof of Proposition 3

The bound in Proposition 3 of this study is slightly tighter ($\approx 1.63/a < 5/(3a)$) than that of Tan (2020) when $a = 1/2$, but it gets much tighter as a gets small: $\approx 4/(3a)$ when $a = 1/5$, for example.



Notes: This figure compares the evaluation of an upper bound of the relative errors of the propensity scores. The dashed line represents the evaluation of Tan (2020), which evaluates the upper bound for special cases where $0 < a \leq 1/2$. The solid line represents the proposed evaluation, which evaluates it more exactly for more general cases where $0 < a \leq 1$.

Figure C.1: Comparison of the evaluation of an upper bound of the relative errors of the propensity scores

Figure C.1 shows the difference between these two evaluations of the upper bound of the relative errors of the propensity scores for $0.1 \leq a \leq 1$.

I provide the proof of Proposition 3, which is a generalization of Proposition 2 of Tan (2020).

Proof. Let $r = \frac{\pi(\mathbf{X})}{\pi(\mathbf{X},\hat{\beta})}$ and $\kappa = \frac{a(1-a+a\log(a))}{a^2-2a+1}$. Since $\kappa > 0$ over $0 < a < 1$, it is sufficient to show that for $a \in (0,1]$, if $0 < r \leq a^{-1}$, then:

$$(r-1)^2/2 \leq r - 1 - \log(r) \quad \text{for } a = 1 \tag{31}$$

$$\kappa(r-1)^2 \leq r - 1 - \log(r) \quad \text{for } 0 < a < 1. \tag{32}$$

First, if $0 < r \leq 1$, then $r - 1 - \log(r) \geq (r-1)^2/2 \geq \kappa(r-1)^2$. The first inequality holds because $r-1-\log(r)-(r-1)^2/2$ decreases to 0 over $0 < r \leq 1$. This is sufficient to prove (31), thus the following proof concentrates on the proof of (32) under the condition of $0 < a < 1$. To prove the second inequality, it suffices to show $k_1(a) = (a-1)^2 - 2a(1-a+a\log(a)) \geq 0$, where the derivative of $k_1(a)$ is $-4(1-a+a\log(a))$. Thus, $k_1(a)$ is decreasing over $0 < a < 1$, which implies that $k_1(a) > 0$.

Second, the function $k_2(r) = r-1-\log(r)-\kappa(r-1)^2$ has a derivative of $(1-r)(2r\kappa-1)/r$. Thus, $k_2(r)$ is increasing over $1 \leq r \leq (2\kappa)^{-1}$ and then decreasing when $r > (2\kappa)^{-1}$. Since $k_2(1) = 0$, it is sufficient to show that $k_2(a^{-1}) = a^{-1} - 1 + \log(a) - a^{-2}\kappa(1-a)^2 \geq 0$ for $0 < a < 1$. Substituting $\kappa = \frac{a(1-a+a\log(a))}{a^2-2a+1}$ gives $k_2(a^{-1}) = 0$ for any $0 < a < 1$. □

C.2 Proof of Proposition 4

For completeness, I provide the proof of (17), which is also found in Tan (2020) though I corrected a typo.

Proof. By $\mathbb{E}[\delta^2] = \mathbb{E}[\delta]^2 + \mathbb{V}(\delta)$ for a random variable δ, one has

$$\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}\frac{R_iY_i}{\pi(X_i,\hat{\beta})} - \mu\right)^2\right] = \mathbb{E}\left[\left(\frac{R}{\pi(\mathbf{X},\hat{\beta})} - 1\right)Y\right]^2 + \frac{1}{n}\mathbb{V}\left(\frac{R}{\pi(\mathbf{X},\hat{\beta})}Y\right). \tag{33}$$

By the Cauchy–Schwartz inequality, the first term is no greater than $c\,\xi(\hat{\beta})$:

$$\mathbb{E}\left[\left(\frac{R}{\pi(\mathbf{X},\hat{\beta})} - 1\right)Y\right]^2 \leq \xi(\hat{\beta})\mathbb{E}[Y^2]. \tag{34}$$

By the law of total variance, one has

$$\begin{aligned}
\frac{1}{n}\mathbb{V}\left(\frac{R}{\pi(\mathbf{X},\hat{\beta})}Y\right) &= \frac{1}{n}\mathbb{V}\left(\mathbb{E}\left[\frac{R}{\pi(\mathbf{X},\hat{\beta})}Y \,\middle|\, \mathbf{X}\right]\right) + \frac{1}{n}\mathbb{E}\left[\mathbb{V}\left(\frac{R}{\pi(\mathbf{X},\hat{\beta})}Y \,\middle|\, \mathbf{X}\right)\right] && (35)\\
&= \frac{1}{n}\mathbb{V}\left(\frac{\pi(\mathbf{X})}{\pi(\mathbf{X},\hat{\beta})}\mathbb{E}[Y \mid \mathbf{X}]\right) + \frac{1}{n}\mathbb{E}\left[\frac{\pi(\mathbf{X})(1-\pi(\mathbf{X}))}{\pi(\mathbf{X},\hat{\beta})^2}Y^2\right] && (36)\\
&\leq \frac{1}{n}c\,\mathbb{E}\left[\frac{\pi(\mathbf{X})^2}{\pi(\mathbf{X},\hat{\beta})^2}\right] + \frac{1}{n}\left(\frac{1}{\rho} - 1\right)c\,\mathbb{E}\left[\frac{\pi(\mathbf{X})^2}{\pi(\mathbf{X},\hat{\beta})^2}\right] && (37)\\
&\leq \frac{c}{n\rho}(1 + \xi(\hat{\beta})), && (38)
\end{aligned}$$

where the second and third (in)equalities hold by the missing-at-random assumption. From

(33), (34) and (38), one gets (17). $\square$

D Equivalence of the Loss Functions of the Proposed Method and Entropy Balancing

In Section 3.6, I pointed out that the DBW and entropy balancing use the same loss function (4) with different link functions, where the DBW uses the logistic function and the entropy balancing uses the exponential function as shown in Wang and Zubizarreta (2020). For completeness, I provide a proof for this notion.

The entropy balancing is a method to estimate weights that minimizes the entropy divergence between the estimated and base weights ($w_{\text{base},i} = 1 \ \forall i$):

$$\operatorname*{arg\,min}_{w_{\text{EB},i}} \sum_{i}^{n} R_i w_{\text{EB},i} \log(w_{\text{EB},i}), \tag{39}$$

subject to the following covariate balancing conditions:

$$\sum_{i=1}^{n} \left(R_i w_{\text{EB},i} - 1\right) X_i = 0. \tag{40}$$

Adding a constant ($-\sum_i^n R_i w_{\text{EB},i} = -n$) to (39) and using the Lagrangian function, the solution to this constrained optimization problem is obtained as the solution to the following unconstrained optimization problem:

$$\operatorname*{arg\,min}_{w_{\text{EB},i},\ \beta_{\text{EB}}} \sum_{i}^{n} \left\{ R_i w_{\text{EB},i} \log(w_{\text{EB},i}) + R_i w_{\text{EB},i} X_i^\top \beta_{\text{EB}} - X_i^\top \beta_{\text{EB}} - R_i w_{\text{EB},i} \right\}, \tag{41}$$

where β_{EB} are the Lagrangian multipliers. Taking the derivative with respect to $w_{\text{EB},i}$ provides:

$$R_i\{\log(w^*_{\text{EB},i}) + X_i^\top \beta^*_{\text{EB}}\} = 0 \quad \text{for all } i. \tag{42}$$

Thus, the solution $w^*_{\text{EB},i}, \beta^*_{\text{EB}}$ satisfies (40) and the following equations:

$$\log(w^*_{\text{EB},i}) = -X_i^\top \beta^*_{\text{EB}} \quad \text{for all } i. \tag{43}$$

Using (43), the solution to (41) is obtained as the minimizer of the following function:

$$\underset{\beta_{\mathrm{EB}}}{\arg\min} \sum_{i}^{n} \{-X_i^{\mathsf{T}} \beta_{\mathrm{EB}} + R_i \exp(-X_i^{\mathsf{T}} \beta_{\mathrm{EB}})\}, \tag{44}$$

where I added a constant ($2 \sum_i^n R_i w_{\mathrm{EB},i} = 2n$). This function is equivalent to the loss function of the DBW (4) by letting $\pi_{\mathrm{EB}}(X_i, \beta_{\mathrm{EB}}) = 1/w_{\mathrm{EB},i} = \exp(X_i^{\mathsf{T}} \beta_{\mathrm{EB}})$ denote propensity scores with the exponential function for a link function.

E Algorithm for the Distribution Balancing Weighting with the Normalization and Regularization

This section extends the difference of the convex functions algorithm for estimating the DBW explained in Section 4.2 to incorporate the normalization and regularization. First of all, I decompose the non-convex loss function of (22) into the difference of the two convex functions as follows:

$$h(\beta) \equiv \sum_{i=1}^{n} \log(\pi(X_i, \beta)) + \frac{\sum_{i=1}^{n} (1 - \pi(X_i, \beta))}{n_0} \sum_{i=1}^{n} \left(\frac{R_i}{\pi(X_i, \beta)} - 1 \right) + \lambda ||\beta_{-1}||^2 \tag{45}$$

$$= h_1(\beta) - h_2(\beta) \tag{46}$$

$$h_1(\beta) = g_1(\beta) - \frac{\sum_{i=1}^{n} \log(\pi(X_i, \beta))}{n_0} \sum_{i=1}^{n} \left(\frac{R_i}{\pi(X_i, \beta)} - 1 \right) + \lambda ||\beta_{-1}||^2 \tag{47}$$

$$h_2(\beta) = g_2(\beta) - \frac{\sum_{i=1}^{n} (1 - \pi(X_i, \beta) + \log(\pi(X_i, \beta))}{n_0} \sum_{i=1}^{n} \left(\frac{R_i}{\pi(X_i, \beta)} - 1 \right) + \sum_{i=1}^{n} \left(\frac{R_i}{\pi(X_i, \beta)} \right). \tag{48}$$

Using this decomposition, the algorithm repeats the following two steps until convergence. First, as the majorization step, it constructs a surrogate function $v(\beta, \beta_t)$ for iteration t as follows:

$$v(\beta, \beta_t) = h_1(\beta) - \{h_2(\beta_t) + h_2'(\beta_t)^{\mathsf{T}} (\beta - \beta_t)\}, \tag{49}$$

where β_t is the initial value or values obtained in the previous iteration and $h_2'(\beta) = \frac{\partial h_2(\beta)}{\partial \beta}$. Second, as the minimization step, it estimates β_{t+1} such that minimizes (49) with respect to β while keeping β_t fixed. This minimization is easily conducted because the objective surrogate function is convex.

The R package dbw implements this algorithm for estimating the DBW.

F Related Literature

To improve the IPW estimation, several studies proposed methods that directly balance the prespecified moments of covariates between the target and weighted groups. Several methods balance finite prespecified moments with or without the propensity score model (Graham, De Xavier Pinto, and Egel 2012) (Hainmueller 2012; Imai and Ratkovic 2014; Tan 2020; Vermeulen and Vansteelandt 2015), another method balances infinite moments with nonparametrically estimated weights (Chan, Yam, and Zhang 2016), and other methods balance the kernel functions (Hazlett 2020; Kallus 2020; Wong and Chan 2017; Zhao 2019). Other studies propose to balance moments approximately up to prespecified levels and minimize the dispersion of the estimated weights (Wang and Zubizarreta 2020; Zubizarreta 2015). However, when the propensity score model, if it exists, is misspecified, the unknown true outcome model needs to lie in a linear space spanned by the balanced functions for consistency (Fan et al. 2021; Zhao and Percival 2017; Zhao 2019; Zubizarreta 2015), which makes it difficult to determine moment conditions in a finite sample. In contrast, the proposed method minimizes the imbalance in the multivariate covariate distribution (Section 3.3), whose importance is acknowledged in several studies (Hainmueller 2012; Li, Morgan, and Zaslavsky 2018; Zhao 2019; Zubizarreta 2015). This allows it to minimize an upper bound of the mean squared error of the parameter of interest (Section 3.5). In the matching framework, the coarsened exact matching shares this spirit of balancing the multivariate covariate distribution (Iacus, King, and Porro 2011, 2012).

The nonparametric methods, such as the nonparametric propensity score estimation with the sieve logistic regression (Hirano, Imbens, and Ridder 2003), the nonparametric moment balancing weight estimation (Chan, Yam, and Zhang 2016), and the kernel balancing weight estimation (Hazlett 2020; Kallus 2020; Wong and Chan 2017; Zhao 2019), also ensure that the covariate distribution will be balanced if the number of balancing conditions grows adequately as the number of the observations increases. However, their performance is not guaranteed with a finite sample and heavily depends on the selected hyper-parameters, and tuning hyper-parameters is still a difficult and unsolved open problem (Wong and Chan 2017; Zhao 2019; Zubizarreta 2015). In contrast, the proposed method is a parametric method that minimizes the distribution imbalance, which leads to other attractive characteristics such as minimizing the upper bound of the MSE of the parameter of interest and the double robustness property (Section 3.5, 4.1, and 4.3).

Recently, Huling and Mak (2020) proposes a nonparametric weight estimation that minimizes the energy distance in the covariate distribution between target and weighted groups. This approach shares the same motivation with the proposed method in that the multivariate covariate distribution, instead of the prespecified moments of covariates, should be balanced. However, it cannot be doubly robust because it does not have a propensity score model, nor does it have the attractive properties that the proposed method has, especially in a finite sample.

Lastly, the entropy balancing method proposed by Hainmueller (2012) is closely related to the proposed method. The entropy balancing method estimates weights to balance the selected moments of covariates between the target and weighted groups without modeling propensity scores explicitly while minimizing the reverse KL divergence between uniform weights and estimated weights. However, it is later shown that it has an implicit propensity score model (Zhao and Percival 2017), and it is modeled using the exponential function, i.e., $\pi(X_i) = \exp(X_i^\intercal \beta)$, in the estimation of the average outcome (Wang and Zubizarreta 2020). Interestingly, the solution to the dual problem is the minimizer of the loss function that is equivalent to the proposed method (Wang and Zubizarreta 2020). Thus, the entropy balancing method for the average outcome estimation minimizes the KL divergence between the true and estimated weights. However, the important difference is that the propensity score model of the entropy balancing method is always misspecified by construction, and it is no longer doubly robust for the estimands other than the average treatment effect on treated (Zhao 2019). This is because it implicitly models the propensity scores as the exponential function of the linear predictors, and thus they are unbounded above and cannot be interpreted as the conditional probability of the response.

G Details of the Empirical Analysis

This section presents details of the study of foreign occupation and insurgency in France analyzed briefly in Section 6. First, like recently proposed estimators such as the calibrated weighting and bias-reduced doubly robust estimators (Chan, Yam, and Zhang 2016; Tan 2020; Vermeulen and Vansteelandt 2015), the nDBW estimator estimates two propensity scores for the ATE estimation. Each of them is used for estimating the inverse probability weights for the treatment group and control group. To optimize the hyper-parameters in the nDBW estimator, I use a grid-search approach with the range $[0, 0.2]$ for the regularization parameter λ and adopt the values based on the upper bound of bias.

Tables G.1–G.2 present the details of the results, where the third column shows the upper bound of bias and the last four columns show the hyper-parameters used for the nDBW DR estimator. First, when the distance from the demarcation line is small, all the estimators produce similar point estimates, standard errors, and upper bound of bias. However, as the distance increases, the MLE DR, calibrated weighting DR, and CBPS DR estimates diverge from those of the nDBW DR and entropy balancing DR estimators, and their standard errors increase, which reach their peaks when the distance is 27.5. When the distance becomes larger, the calibrated weighting DR estimator does not produce estimates due to the convergence issue, which is indicated by the blanks for corresponding spaces in the table.

Table G.1: Results for the study of political devolution and resistance activities

	ATE	s.e.	upper bound of bias	λ_t	λ_c
0 < (distance from the demarcation line) < 15					
nDBW DR	0.49	0.21	0.12	0.13	0.05
MLE DR	0.49	0.21	0.13		
Calibrated weighting DR	0.50	0.21	0.12		
CBPS DR	0.50	0.21	0.13		
Entropy balancing DR	0.47	0.21	0.12		
Unweighted	0.46	0.23	0.18		
2.5 < (distance from the demarcation line) < 17.5					
nDBW DR	0.51	0.19	0.12	0.11	0.06
MLE DR	0.51	0.20	0.13		
Calibrated weighting DR	0.52	0.19	0.12		
CBPS DR	0.53	0.20	0.13		
Entropy balancing DR	0.50	0.20	0.12		
Unweighted	0.48	0.21	0.19		
5 < (distance from the demarcation line) < 20					
nDBW DR	0.30	0.18	0.13	0.10	0.03
MLE DR	0.30	0.20	0.14		
Calibrated weighting DR	0.31	0.19	0.13		
CBPS DR	0.29	0.20	0.14		
Entropy balancing DR	0.31	0.19	0.13		
Unweighted	0.28	0.19	0.21		
7.5 < (distance from the demarcation line) < 22.5					
nDBW DR	0.22	0.25	0.14	0.11	0.03
MLE DR	0.19	0.24	0.15		
Calibrated weighting DR	0.19	0.29	0.15		
CBPS DR	0.16	0.24	0.16		
Entropy balancing DR	0.20	0.21	0.15		
Unweighted	0.14	0.16	0.25		
10 < (distance from the demarcation line) < 25					
nDBW DR	0.15	0.23	0.15	0.09	0.02
MLE DR	0.16	0.26	0.16		
Calibrated weighting DR	0.17	0.29	0.16		
CBPS DR	0.16	0.28	0.17		
Entropy balancing DR	0.13	0.20	0.15		
Unweighted	0.09	0.16	0.26		

Notes: This table presents the results for the municipalities within various distances from the demarcation line estimated with the various methods, where the third column shows the upper bound of bias and the last two columns show the hyper-parameters used for the nDBW DR estimator.

Table G.2: Results for the study of political devolution and resistance activities

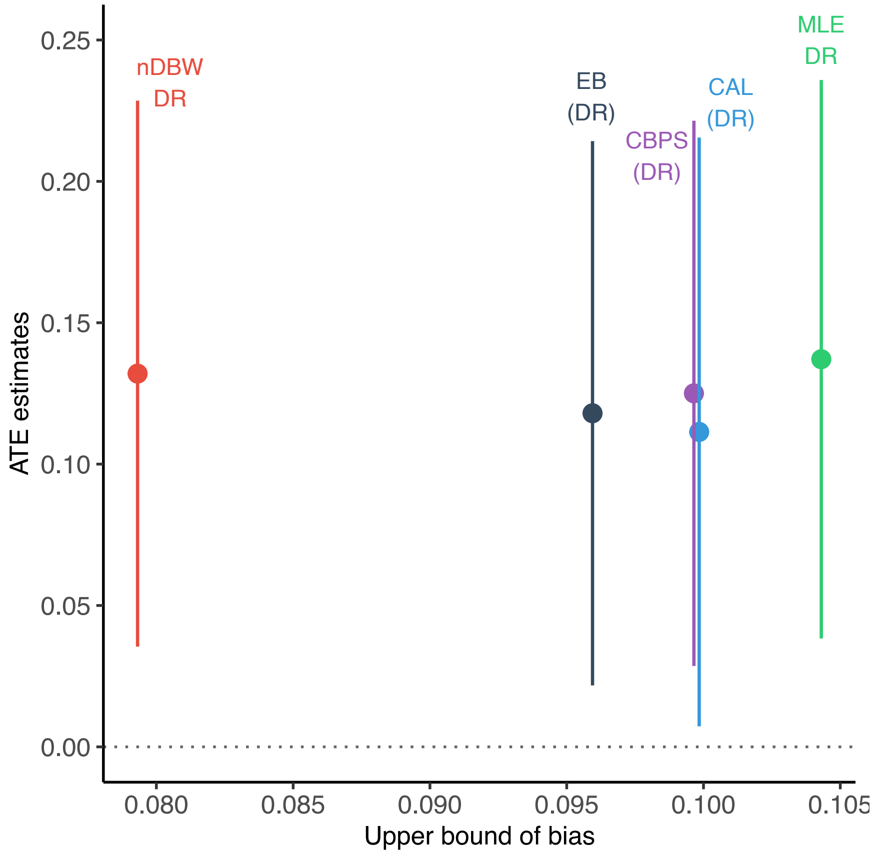
	ATE	s.e.	upper bound of bias	λ_t	λ_c
12.5 < (distance from the demarcation line) < 27.5					
nDBW DR	0.25	0.34	0.16	0.09	0.03
MLE DR	0.40	0.52	0.19		
Calibrated weighting DR	0.33		0.20		
CBPS DR	0.44	0.58	0.19		
Entropy balancing DR	0.22	0.27	0.17		
Unweighted	0.09	0.16	0.28		
15 < (distance from the demarcation line) < 30					
nDBW DR	0.06	0.14	0.17	0.10	0.04
MLE DR	0.03	0.13	0.20		
Calibrated weighting DR					
CBPS DR	−0.01	0.13	0.21		
Entropy balancing DR	0.05	0.17	0.19		
Unweighted	0.08	0.13	0.30		
17.5 < (distance from the demarcation line) < 32.5					
nDBW DR	0.10	0.12	0.19	0.09	0.03
MLE DR	0.09	0.14	0.24		
Calibrated weighting DR					
CBPS DR	0.07	0.13	0.25		
Entropy balancing DR	0.12	0.14	0.22		
Unweighted	−0.02	0.13	0.32		
20 < (distance from the demarcation line) < 35					
nDBW DR	0.00	0.13	0.20	0.09	0.04
MLE DR	0.02	0.15	0.23		
Calibrated weighting DR					
CBPS DR	−0.05	0.18	0.28		
Entropy balancing DR	0.04	0.17	0.26		
Unweighted	−0.07	0.11	0.36		

Notes: This table presents the results for the municipalities within various distances from the demarcation line estimated with the various methods, where the third column shows the upper bound of bias and the last two columns show the hyper-parameters used for the nDBW DR estimator. The blanks for the calibrated weighting DR estimator indicate that it does not estimate weights because of the convergence issue.

H Another Empirical Study: The Persuasive Effect of the News Media on Voting Choice

This section examines the performance of the nDBW estimator with a study of the persuasive effect of the news media on voting choice. Ladd and Lenz (2009) studies the effect of newspaper endorsements on voting choice by exploiting a rare opportunity during the campaign for the 1997 UK general election, where a prominent British newspaper, the *Sun*, switched its support from Conservative to Labour, and three smaller newspapers, the *Daily Star*, *Independent*, and *Financial Times*, switched from no endorsement to a Labour endorsement, while other newspapers did not change their support. They use data from several waves of the British Election Panel Study 1992–1997 to compare the treated respondents who read one of the above newspapers and untreated respondents who did not read them. The outcome variable is voting choice in the 1997 election reported in the post-election survey.

To address bias due to differences in other attributes than the treatment, Ladd and Lenz (2009) balances the treated and untreated groups on the following pre-treatment variables using a genetic matching, and (Hainmueller 2012) addresses this problem by using the entropy balancing in the reanalysis. The control variables include various measures for a respondent's prior evaluation of the Labour Party and Conservative Party (such as prior party support, prior voting choice, prior party identification, etc.), prior ideology, socioeconomic status (such as education, income, working-class identification, race, profession, region, etc.), age, gender, political knowledge, etc. Using these control variables, I estimate the average treatment effect (ATE) with the nDBW DR, MLE DR, calibrated weighting DR, CBPS DR, and entropy balancing DR estimators and their jack-knife standard errors with the 192 treated and 1,219 untreated respondents. To optimize the hyper-parameters in the nDBW estimator, I use a grid-search approach with the range $[0, 0.3]$ for the regularization parameter λ and adopt the values based on the upper bound of bias, which yields $\lambda_t = 0.28$ and $\lambda_c = 0.01$.



Notes: This figure shows the upper bound of bias (Hazlett 2020) for various estimators on the x-axis and their ATE estimates with the 95% confidence intervals on the y-axis. The red, navy, purple, blue, and green lines indicate the nDBW DR, entropy balancing DR, CBPS DR, calibrated weighting DR, and the MLE DR estimators.

Figure H.1: Results for the persuasive effect of the news media on voting choice

Figure H.1 presents the results, where the x-axis shows the upper bound of bias for the estimators (Hazlett 2020), and the y-axis shows their ATE estimates with the 95% confidence intervals. The red, navy, purple, blue, and green lines indicate the nDBW DR, entropy balancing DR, CBPS DR, calibrated weighting DR, and the MLE DR estimators. In this case, all the estimators provide similar ATE estimates (0.11–0.14). Regarding the quality of the distributional balance of these estimators, the nDBW DR estimator effectively constrains the upper bound of bias, which confirms its superiority in distribution balancing.

I The Full Results of the Simulation Studies

The full results of the simulation studies are presented in Table I.1–I.6.

Table I.1: Simulation results: Linear outcome model 1

	Type A PS coefficients				Type B PS coefficients			
	$n = 200$		$n = 1000$		$n = 200$		$n = 1000$	
	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Correct PS model								
nDBW	−0.60	2.86	−0.09	1.33	0.76	2.95	0.03	1.34
MLE	−0.27	11.61	−0.04	4.90	−0.03	7.67	−0.02	2.82
CBPS	1.94	4.54	0.37	1.76	2.04	4.73	0.30	1.78
Calibrated weighting	0.06	2.57	0.00	1.15	0.04	2.60	−0.06	1.16
Entropy balancing	0.06	2.57	−0.01	1.15	0.04	2.60	−0.06	1.16
True propensity score	0.05	23.28	0.39	10.74	−0.45	17.95	−0.26	8.12
Unweighted	−9.78	10.39	−9.97	10.09	10.05	10.68	9.95	10.07
nDBW DR	0.54	2.89	0.17	1.31	0.74	2.80	0.09	1.20
MLE DR	0.36	3.65	0.01	1.74	0.44	3.74	0.00	1.76
CBPS DR	0.49	3.25	0.07	1.59	0.38	3.08	0.02	1.46
Calibrated weighting DR	0.54	2.97	0.12	1.41	0.42	2.74	0.04	1.22
Entropy balancing DR	1.15	3.16	0.91	1.64	1.40	3.10	1.11	1.66
True propensity score DR	0.37	3.57	0.04	1.81	0.62	4.08	0.05	1.96
Imputation	−0.55	3.32	−0.82	1.74	4.93	5.79	4.90	5.08
Misspecified PS model								
nDBW	−3.21	4.77	−3.70	4.08	3.48	4.64	1.10	1.75
MLE	20.93	143.01	44.47	292.98	0.54	6.75	0.53	2.51
CBPS	1.03	4.84	−2.01	2.94	6.18	7.55	3.86	4.19
Calibrated weighting	−2.14	3.83	−2.76	3.14	2.31	3.70	2.23	2.58
Entropy balancing	−1.52	3.56	−1.95	2.46	3.79	4.80	3.77	4.00
True propensity score	0.21	23.15	−0.28	10.51	0.18	18.23	−0.08	8.37
Unweighted	−10.02	10.64	−9.97	10.10	9.92	10.51	10.01	10.14
nDBW DR	−1.98	3.73	−2.61	2.99	2.49	3.81	1.79	2.18
MLE DR	−5.97	22.52	−16.30	129.23	3.11	4.49	3.07	3.39
CBPS DR/BRDR	−2.73	4.36	−3.57	3.97	3.07	4.39	3.33	3.63
Calibrated weighting DR	−2.14	3.83	−2.76	3.14	2.31	3.70	2.23	2.58
Entropy balancing DR	−1.52	3.56	−1.95	2.46	3.79	4.80	3.77	4.00
True propensity score DR	0.31	3.66	0.09	1.77	0.47	4.09	0.12	1.97
Imputation	−0.62	3.39	−0.82	1.70	4.84	5.67	4.96	5.14

Notes: This simulation compares the performance of various methods for estimating propensity scores and (inverse probability) weights by investigating combinations of six versions of the true outcome model (linear 1, linear 2, quadratic 1, quadratic 2, exponential 1, and exponential 2) and two versions of coefficients for the true propensity score model (type A and B) with the two different numbers of observations ($n = 200$ and $n = 1000$). For each estimation method, I use two propensity score model specifications (correct and misspecified) and report the bias and RMSE for each in the table.

Table I.2: Simulation results: Linear outcome model 2

	Type A PS coefficients				Type B PS coefficients			
	$n = 200$		$n = 1000$		$n = 200$		$n = 1000$	
	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Correct PS model								
nDBW	−0.16	3.77	−0.05	1.76	0.28	3.77	−0.01	1.77
MLE	−0.12	10.57	−0.07	4.51	−0.10	9.87	0.00	3.71
CBPS	2.03	5.45	0.35	2.15	2.01	5.50	0.33	2.16
Calibrated weighting	0.11	3.47	−0.02	1.58	−0.02	3.47	−0.04	1.56
Entropy balancing	0.11	3.47	−0.02	1.58	−0.03	3.47	−0.04	1.56
True propensity score	0.12	22.19	0.35	10.17	−0.52	19.90	−0.27	8.98
Unweighted	−3.59	6.18	−3.76	4.37	3.76	6.29	3.73	4.34
nDBW DR	0.07	3.64	0.03	1.65	0.62	3.66	0.08	1.63
MLE DR	−0.06	4.89	−0.13	2.19	0.43	5.14	0.02	2.35
CBPS DR	0.06	4.40	−0.08	2.01	0.34	4.37	0.04	2.01
Calibrated weighting DR	0.10	3.77	−0.05	1.72	0.37	3.73	0.06	1.67
Entropy balancing DR	0.02	3.85	−0.08	1.74	1.18	3.91	0.95	1.91
True propensity score DR	−0.12	5.19	−0.09	2.48	0.63	5.56	0.07	2.62
Imputation	−3.21	5.78	−3.36	4.02	5.25	6.86	5.24	5.58
Misspecified PS model								
nDBW	−4.55	6.45	−4.62	5.03	4.04	6.08	1.97	2.79
MLE	17.20	127.21	36.83	234.68	1.47	9.01	1.39	3.63
CBPS	−0.42	5.86	−3.24	4.13	7.43	9.35	4.56	5.16
Calibrated weighting	−3.55	5.59	−3.73	4.18	3.27	5.33	3.15	3.66
Entropy balancing	−3.49	5.70	−3.51	4.03	4.54	6.33	4.45	4.86
True propensity score	0.19	22.11	−0.22	10.02	0.20	20.02	−0.16	9.28
Unweighted	−3.88	6.33	−3.74	4.34	3.69	6.13	3.73	4.36
nDBW DR	−3.36	5.48	−3.56	4.02	3.27	5.32	2.58	3.17
MLE DR	−6.05	17.36	−9.29	62.15	4.34	6.43	4.25	4.73
CBPS DR/BRDR	−4.42	6.42	−4.53	4.99	4.44	6.44	4.49	4.95
Calibrated weighting DR	−3.55	5.59	−3.73	4.18	3.27	5.33	3.15	3.66
Entropy balancing DR	−3.49	5.70	−3.51	4.03	4.54	6.33	4.45	4.86
True propensity score DR	−0.32	5.28	−0.02	2.42	0.47	5.75	0.07	2.63
Imputation	−3.43	5.91	−3.36	3.95	5.23	6.83	5.24	5.60

Notes: This simulation compares the performance of various methods for estimating propensity scores and (inverse probability) weights by investigating combinations of six versions of the true outcome model (linear 1, linear 2, quadratic 1, quadratic 2, exponential 1, and exponential 2) and two versions of coefficients for the true propensity score model (type A and B) with the two different numbers of observations ($n = 200$ and $n = 1000$). For each estimation method, I use two propensity score model specifications (correct and misspecified) and report the bias and RMSE for each in the table.

Table I.3: Simulation results: Quadratic outcome model 1

	Type A PS coefficients				Type B PS coefficients			
	$n = 200$		$n = 1000$		$n = 200$		$n = 1000$	
	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Correct PS model								
nDBW	−2.26	6.28	−0.55	2.81	−0.45	5.77	−0.21	2.62
MLE	−0.78	18.19	−0.06	8.02	−0.21	14.70	0.00	5.60
CBPS	1.24	7.26	0.24	3.34	2.05	7.21	0.32	2.92
Calibrated weighting	−1.55	6.33	−0.37	2.97	−0.75	5.78	−0.22	2.57
Entropy balancing	−3.74	6.92	−3.03	4.04	−3.47	6.50	−3.06	3.91
True propensity score	−0.17	30.73	0.54	14.26	−0.60	25.64	−0.31	11.65
Unweighted	−6.85	8.93	−6.91	7.36	7.16	9.77	6.94	7.52
nDBW DR	−2.16	6.08	−0.48	2.66	−0.86	5.67	−0.20	2.53
MLE DR	−1.43	7.45	−0.31	3.53	−0.53	6.95	−0.14	2.96
CBPS DR	−1.58	6.63	−0.39	3.11	−0.55	6.18	−0.14	2.79
Calibrated weighting DR	−1.78	6.15	−0.45	2.81	−0.64	5.74	−0.17	2.56
Entropy balancing DR	−3.68	6.86	−2.72	3.79	−2.13	5.92	−1.67	2.94
True propensity score DR	−1.63	7.58	−0.30	3.71	−0.78	6.88	−0.22	3.15
Imputation	−5.81	8.39	−5.03	5.72	−2.89	6.78	−2.73	3.81
Misspecified PS model								
nDBW	−3.28	6.61	−1.88	3.11	1.69	6.42	0.24	2.68
MLE	34.36	244.59	80.25	611.06	−1.47	13.75	−1.46	5.38
CBPS	2.53	7.42	1.01	3.17	5.40	9.63	2.23	3.91
Calibrated weighting	−1.83	6.22	−0.43	2.73	0.66	6.30	1.07	2.87
Entropy balancing	−3.63	6.87	−2.40	3.59	−0.95	6.30	−0.62	2.70
True propensity score	0.29	30.65	−0.23	14.00	0.27	26.02	−0.15	12.01
Unweighted	−6.95	9.03	−6.85	7.30	6.85	9.42	6.99	7.55
nDBW DR	−1.89	6.17	−0.47	2.66	0.32	6.24	1.00	2.83
MLE DR	1.71	25.10	17.40	174.96	1.02	7.76	1.34	3.51
CBPS DR/BRDR	−2.04	6.71	−0.09	3.13	1.32	7.02	1.21	3.21
Calibrated weighting DR	−1.83	6.22	−0.43	2.73	0.65	6.30	1.07	2.87
Entropy balancing DR	−3.63	6.87	−2.40	3.59	−0.95	6.30	−0.62	2.70
True propensity score DR	−1.57	7.53	−0.27	3.68	−0.78	7.05	−0.23	3.11
Imputation	−5.76	8.26	−5.00	5.67	−2.94	6.89	−2.73	3.78

Notes: This simulation compares the performance of various methods for estimating propensity scores and (inverse probability) weights by investigating combinations of six versions of the true outcome model (linear 1, linear 2, quadratic 1, quadratic 2, exponential 1, and exponential 2) and two versions of coefficients for the true propensity score model (type A and B) with the two different numbers of observations ($n = 200$ and $n = 1000$). For each estimation method, I use two propensity score model specifications (correct and misspecified) and report the bias and RMSE for each in the table.

Table I.4: Simulation results: Quadratic outcome model 2

	Type A PS coefficients				Type B PS coefficients			
	$n = 200$		$n = 1000$		$n = 200$		$n = 1000$	
	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Correct PS model								
nDBW	−0.70	5.75	−0.19	2.63	−2.19	6.25	−0.50	2.78
MLE	−0.76	14.18	−0.09	5.78	−0.45	19.10	0.16	8.00
CBPS	1.57	6.72	0.32	2.93	1.65	7.56	0.34	3.35
Calibrated weighting	−0.98	5.73	−0.21	2.60	−1.40	6.32	−0.32	2.95
Entropy balancing	−3.64	6.57	−3.02	3.89	−3.63	6.87	−3.00	4.01
True propensity score	−0.46	26.59	0.33	12.16	−0.60	30.81	−0.20	13.66
Unweighted	6.75	9.34	6.93	7.52	−6.90	8.88	−6.88	7.33
nDBW DR	−1.60	6.20	−0.39	2.69	−3.03	6.61	−0.71	2.86
MLE DR	−1.11	7.63	−0.12	3.43	−1.53	9.63	−0.17	4.80
CBPS DR	−1.33	6.87	−0.23	3.16	−1.75	7.36	−0.33	3.74
Calibrated weighting DR	−1.50	6.38	−0.32	2.84	−2.10	6.51	−0.49	3.06
Entropy balancing DR	−4.30	7.49	−3.42	4.36	−5.09	7.81	−4.02	4.86
True propensity score DR	−1.26	7.52	−0.18	3.54	−1.87	10.22	−0.31	5.16
Imputation	−1.11	7.12	−0.26	3.16	−11.23	12.50	−10.89	11.16
Misspecified PS model								
nDBW	5.57	9.08	6.27	7.02	−7.14	9.20	−4.12	4.94
MLE	23.21	109.17	42.96	259.51	−10.78	16.41	−10.53	11.55
CBPS	7.26	11.09	3.79	5.21	−4.53	8.87	−8.22	8.92
Calibrated weighting	4.67	8.46	5.46	6.30	−6.40	8.63	−5.40	6.05
Entropy balancing	2.30	7.29	3.15	4.41	−9.16	10.71	−8.50	8.88
True propensity score	0.39	26.63	−0.44	11.70	0.24	31.11	0.06	14.22
Unweighted	7.05	9.65	6.94	7.54	−7.00	9.00	−6.85	7.34
nDBW DR	4.31	8.24	5.18	6.03	−6.63	8.71	−4.70	5.38
MLE DR	10.72	31.50	26.53	186.73	−7.23	10.21	−6.63	7.37
CBPS DR/BRDR	5.69	9.57	6.97	7.83	−7.15	9.54	−7.01	7.61
Calibrated weighting DR	4.67	8.46	5.46	6.30	−6.40	8.63	−5.40	6.05
Entropy balancing DR	2.30	7.29	3.15	4.41	−9.16	10.71	−8.50	8.88
True propensity score DR	−1.04	7.63	−0.25	3.47	−1.90	10.36	−0.31	5.24
Imputation	−0.99	7.21	−0.23	3.15	−11.37	12.64	−10.88	11.17

Notes: This simulation compares the performance of various methods for estimating propensity scores and (inverse probability) weights by investigating combinations of six versions of the true outcome model (linear 1, linear 2, quadratic 1, quadratic 2, exponential 1, and exponential 2) and two versions of coefficients for the true propensity score model (type A and B) with the two different numbers of observations ($n = 200$ and $n = 1000$). For each estimation method, I use two propensity score model specifications (correct and misspecified) and report the bias and RMSE for each in the table.

Table I.5: Simulation results: Exponential outcome model 1

	Type A PS coefficients				Type B PS coefficients			
	$n = 200$		$n = 1000$		$n = 200$		$n = 1000$	
	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Correct PS model								
nDBW	−3.99	11.49	−1.08	5.46	−0.54	10.77	−0.38	5.01
MLE	−1.28	29.80	−0.10	13.73	−0.25	23.32	−0.04	8.80
CBPS	1.51	13.34	0.23	6.47	2.93	13.00	0.40	5.39
Calibrated weighting	−2.67	11.55	−0.75	5.74	−1.25	10.71	−0.42	4.84
Entropy balancing	−5.90	12.08	−4.80	6.91	−5.46	11.47	−4.86	6.61
True propensity score	−0.11	49.48	0.84	22.47	−0.77	38.44	−0.47	17.41
Unweighted	−14.23	17.95	−14.34	15.16	14.79	19.57	14.41	15.45
nDBW DR	−3.58	10.77	−0.94	5.06	−1.10	10.49	−0.32	4.81
MLE DR	−2.44	13.31	−0.58	6.69	−0.58	12.93	−0.22	5.52
CBPS DR	−2.64	11.91	−0.75	5.85	−0.67	11.60	−0.22	5.21
Calibrated weighting DR	−2.94	10.97	−0.86	5.31	−0.88	10.69	−0.28	4.82
Entropy balancing DR	−5.56	11.72	−4.10	6.35	−2.60	10.61	−2.05	4.96
True propensity score DR	−2.67	13.96	−0.55	6.98	−0.85	13.00	−0.31	5.90
Imputation	−10.23	15.02	−9.05	10.40	−2.38	11.67	−2.55	5.66
Misspecified PS model								
nDBW	−7.91	13.26	−5.49	7.27	4.72	12.80	1.24	5.36
MLE	57.77	470.18	159.76	1503.70	0.43	23.32	0.37	9.51
CBPS	1.78	12.85	0.54	6.41	10.65	18.11	6.21	8.64
Calibrated weighting	−5.15	12.06	−2.62	6.04	2.61	12.17	3.17	6.08
Entropy balancing	−7.48	13.07	−5.40	7.52	0.83	11.72	1.23	5.15
True propensity score	0.31	48.20	−0.27	22.45	0.38	38.69	−0.28	17.85
Unweighted	−14.52	18.28	−14.28	15.09	14.28	18.98	14.47	15.48
nDBW DR	−5.13	11.96	−2.63	5.88	2.05	11.98	2.78	5.88
MLE DR	0.81	61.38	45.24	601.02	3.69	14.97	4.11	7.91
CBPS DR/BRDR	−5.84	13.06	−2.35	6.95	4.23	13.75	3.96	6.94
Calibrated weighting DR	−5.15	12.06	−2.62	6.04	2.61	12.17	3.17	6.08
Entropy balancing DR	−7.48	13.07	−5.40	7.52	0.83	11.72	1.23	5.15
True propensity score DR	−2.69	13.86	−0.47	7.26	−0.99	13.08	−0.34	5.91
Imputation	−10.18	14.92	−9.04	10.32	−2.47	11.95	−2.50	5.58

Notes: This simulation compares the performance of various methods for estimating propensity scores and (inverse probability) weights by investigating combinations of six versions of the true outcome model (linear 1, linear 2, quadratic 1, quadratic 2, exponential 1, and exponential 2) and two versions of coefficients for the true propensity score model (type A and B) with the two different numbers of observations ($n = 200$ and $n = 1000$). For each estimation method, I use two propensity score model specifications (correct and misspecified) and report the bias and RMSE for each in the table.

Table I.6: Simulation results: Exponential outcome model 2

	Type A PS coefficients				Type B PS coefficients			
	$n = 200$		$n = 1000$		$n = 200$		$n = 1000$	
	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Correct PS model								
nDBW	−1.07	10.63	−0.38	4.98	−3.95	11.52	−1.03	5.34
MLE	−1.29	22.02	−0.17	9.15	−0.89	31.26	0.23	14.07
CBPS	2.04	12.17	0.39	5.46	2.08	13.63	0.37	6.59
Calibrated weighting	−1.76	10.55	−0.41	4.90	−2.50	11.56	−0.68	5.71
Entropy balancing	−5.84	11.51	−4.81	6.59	−5.76	12.05	−4.75	6.83
True propensity score	−0.84	39.62	0.41	18.13	−0.93	48.77	−0.28	22.09
Unweighted	14.03	18.73	14.34	15.41	−14.39	18.03	−14.31	15.13
nDBW DR	−2.85	11.43	−0.77	5.09	−5.32	12.05	−1.41	5.48
MLE DR	−1.91	14.22	−0.20	6.55	−2.77	17.88	−0.33	9.64
CBPS DR	−2.35	12.78	−0.42	6.00	−3.17	13.59	−0.69	7.28
Calibrated weighting DR	−2.67	11.79	−0.60	5.40	−3.78	12.00	−1.00	5.95
Entropy balancing DR	−7.04	13.28	−5.60	7.53	−8.70	13.92	−6.93	8.60
True propensity score DR	−2.06	14.14	−0.31	6.79	−3.27	19.39	−0.55	10.54
Imputation	−0.70	13.16	0.68	5.98	−20.27	22.67	−19.85	20.36
Misspecified PS model								
nDBW	10.55	17.32	11.99	13.47	−12.99	16.97	−7.26	8.89
MLE	33.66	145.54	60.81	362.53	−18.59	27.04	−18.18	19.73
CBPS	12.17	19.73	7.11	9.72	−9.60	16.34	−15.25	16.43
Calibrated weighting	8.63	16.01	10.26	11.93	−11.35	15.56	−9.81	11.04
Entropy balancing	4.84	14.05	6.36	8.68	−16.36	19.28	−15.34	16.07
True propensity score	0.72	39.87	−0.68	17.44	0.49	50.87	0.20	22.79
Unweighted	14.67	19.55	14.42	15.51	−14.46	18.11	−14.24	15.12
nDBW DR	7.97	15.57	9.73	11.41	−11.74	15.76	−8.50	9.80
MLE DR	20.04	59.54	49.55	355.21	−13.17	18.47	−12.22	13.54
CBPS DR/BRDR	10.71	18.34	13.09	14.81	−13.05	17.38	−12.94	14.02
Calibrated weighting DR	8.62	16.01	10.26	11.93	−11.35	15.56	−9.81	11.04
Entropy balancing DR	4.84	14.05	6.36	8.68	−16.36	19.28	−15.34	16.07
True propensity score DR	−1.59	14.40	−0.43	6.71	−3.20	19.84	−0.49	10.52
Imputation	−0.41	13.51	0.78	5.98	−20.46	22.83	−19.83	20.37

Notes: This simulation compares the performance of various methods for estimating propensity scores and (inverse probability) weights by investigating combinations of six versions of the true outcome model (linear 1, linear 2, quadratic 1, quadratic 2, exponential 1, and exponential 2) and two versions of coefficients for the true propensity score model (type A and B) with the two different numbers of observations ($n = 200$ and $n = 1000$). For each estimation method, I use two propensity score model specifications (correct and misspecified) and report the bias and RMSE for each in the table.

References only in Supplementary Materials

Graham, Bryan S., Cristine Campos De Xavier Pinto, and Daniel Egel. 2012. "Inverse Probability Tilting for Moment Condition Models with Missing Data." *Review of Economic Studies* 79(3): 1053–79.

Hirano, Keisuke, Guido W. Imbens, and Geert Ridder. 2003. "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score." *Econometrica* 71(4): 1161–89.

Huling, Jared D., and Simon Mak. 2020. "Energy Balancing of Covariate Distributions." *arXiv preprint arXiv:2004.13962* .

Iacus, Stefano M., Gary King, and Giuseppe Porro. 2011. "Multivariate Matching Methods That Are Monotonic Imbalance Bounding." *Journal of the American Statistical Association* 106(493): 345–361.

Iacus, Stefano M., Gary King, and Giuseppe Porro. 2012. "Causal Inference without Balance Checking: Coarsened Exact Matching." *Political Analysis* 20(1): 1–24.

King, Gary, and Richard Nielsen. 2019. "Why Propensity Scores Should Not Be Used for Matching." *Political Analysis* 27(4): 435–454.

Ladd, Jonathan McDonald, and Gabriel S. Lenz. 2009. "Exploiting a Rare Communication Shift to Document the Persuasive Power of the News Media." *American Journal of Political Science* 53(2): 394–410.