Matching & Synthetic Control: Case Studies

Minerva University

CS130: Statistical Modeling

Prof. Diamond

April 7th, 2024

Darfur Violence & Matching

Rubin Causal Model

1. **Units**: The unit within the context of this Rubin Causal Model is a citizen.
2. **Treatment**: The treatment is direct harm (defined as injured or maimed). Citizens who were directly harmed are referred to have received the treatment and vice versa.
3. **Potential outcomes**: The potential outcomes for the citizens within the Darfur violence context refers to their willingness to either make peace with the perpetrators and call an end to the violence or to seek vengeance. The index measuring pro-peace attitude encapsulates the observed potential outcome, while the potential outcome in case of not getting directly harmed by the violence (treatment) is unobserved.
4. **Predictors:** The main predictors (covariates) of relevance are the villages that the citizens belong to and their gender (females were believed to be subjected to greater severity of violence). Other covariates include age, occupation as herder or farmer (two binary variables), and whether or not the citizens voted in the past.
5. **No-interference between units**: The pro-peace index of a citizen (outcome) depends only on whether or not they were directly harmed by violence (treatment) and not on the treatment or outcome of other units.

Research Question

Research question: Does being subject to direct harm (injury or maiming) during an attack on the village impact the citizens' outlook on peace?

The research question passes Rubin's standard for causal questions as there is more than one potential outcome with a clear definition of causal effect - the difference in pro-peace outlook between being directly harmed and not directly harmed. The treatment under investigation can be conceptually manipulated i.e. it is possible to conceptually create a world where a treated unit did not receive the treatment. Additionally, there is no interference between units with regard to treatment and outcomes.

Predictors of Choice

- **Female**: Females were targeted to more severe forms of violence such as sexual assault, thereby impacting treatment assignment.
- **Past Voted**: It is likely that past participation in the political right of voting led to the identification of being in opposition to the government, therefore impacting treatment.

Propensity Score Matching and Regression

- **Balance:**

  The minimum t-test p-value for both the predictors, female and pastvoted, is 1 which is a significant increase from minimum p-values before matching of 0.16084 and 0.36437 for female and pastvoted respectively.

  The standard mean difference between the control and treatment group reduced from -7.9978 before matching to 0 after matching for female and 5.1898 to 0 for pastvoted, indicating a perfect balance between treated and control groups.

  While it is important to note that the balance only accounts for two observed covariates, the balance achieved is satisfactory enough for farthing analysis.

- **Model Summary:**
  The linear regression model (lm) examines the relationship between the peace factor (*peacefactor*) and predictors (*directlyharmed*, *female*, *pastvoted*) after propensity score matching.

- **Estimates:**

  - <u>Intercept</u>: The baseline peace factor score for individuals not directly harmed (directly harmed = 0) is estimated at *0.45089*. The standard error (Std. Error = *0.02952*) associated with this estimate reflects the variability of the intercept coefficient. This means that if we were to estimate the baseline peace factor score repeatedly from different samples, we would expect the estimated value to vary around *0.45089* by approximately *0.02952* units on average. The very small p-value (p < *<2e-16*) indicates strong evidence against the null hypothesis of no effect by the absence of treatment.

  - <u>Directly Harmed:</u> For each additional individual who was directly harmed (compared to those who were not directly harmed), we expect the peace factor score to increase by approximately *0.02824* units on average. The coefficient estimate (*0.02824*) for being directly harmed has a standard error of *0.02653*, which indicates the variability in the coefficient estimate. The p-value (p = *0.287*) associated with this coefficient suggests that this predictor is not statistically significant at the conventional significance level (set at *0.05*), meaning we fail to reject the null hypothesis that being directly harmed has no effect on *peacefactor*.

- Female: The coefficient estimate *(-0.29346*) for being female indicates that, on average, females have a peace factor score that is lower by approximately *0.29346* units compared to males, holding other variables constant. The standard error (SE = *0.02463*) associated with this estimate reflects the precision of this effect. The very small p-value ($p < <2e-16$) provides strong evidence against the null hypothesis that gender has no effect on the peace factor score. Therefore, we can conclude that being female is significantly associated with a lower peace factor score.

- Past Voted: The coefficient estimate (*0.01310*) for past voting behavior suggests that, on average, individuals who have a history of voting have a peace factor score that is higher by approximately *0.01310* units, controlling for other variables. The standard error (SE = *0.02592*) associated with this estimate indicates the precision of this effect estimate. However, the p-value ($p = 0.613$) is not statistically significant at the conventional significance level (set at 0.05). This suggests that there is insufficient evidence to conclude a linear relationship between past voting behavior and the peace factor score in the studied population.

- **Model Fit:**

  The adjusted R-squared value (*0.1659*) indicates that the model explains about 16.59% of the variance in the peace factor score.

Sensitivity Analysis of PSM Regression

In the sensitivity analysis performed on the regression model *peacefactor ~ directlyharmed + female + pastvoted*, the aim is to assess potential biases caused by unobserved confounding factors. Here's a breakdown of the key sensitivity statistics and their interpretations:

- **Partial R2 of Treatment with Outcome:**

  The partial R-squared of the treatment variable (*directlyharmed*) with the outcome (*peacefactor*) is *6e-04* (0.06%). This statistic indicates that an extreme confounder, orthogonal to the covariates, which explains all the residual variance of the outcome (100%), would need to explain at least *0.06%* of the residual variance of the treatment to completely nullify the observed estimated effect.

- **Robustness Values (RVq=1 and RVq=1, alpha=0.05):**

  - RVq=1: Unobserved confounders that explain more than 2.34% of the residual variance in both the treatment and the outcome are strong enough to bias the point estimate of directlyharmed to zero (100% bias of the original estimate). Conversely, confounders that explain less than 2.34% of the residual variance are not strong enough to nullify the estimated effect.

- RVq=1, alpha=0.05: Unobserved confounders that explain more than 0% of the residual variance in both the treatment and the outcome can bias the estimate of *directlyharmed* to a range where it is no longer statistically different from zero (100% bias of the original estimate) at the significance level of alpha = 0.05. Confounders explaining less than 0% of the residual variance are insufficient to render the estimate statistically insignificant. The extremely low baseline for a confounder to bias the estimate of *directlyharmed* shows that our estimate of the treatment is very weak and subject to biased effects of unobserved confounders.

Genetic Matching I and Regression

- **Balance**

  The mean standard difference for variable *female* between the control and treatment group went down from *-7.9978* to *-7.1101* before and after matching though the T-test p-value decreased from *0.16084* to *0.090045*.

  For variable *pastvoted*, the mean standard deviation between the control and treatment groups decreased significantly from *5.1898* to *0.59705* and the minimum T-test p-value increased from *0.36437* to *0.87336* after matching.

  The lowest minimum p-value observed is for variable *female* at 0.090045.

- **Model Summary**
  - The linear regression model (lm) examines the relationship between the peace factor (*peacefactor*) and predictors (*directlyharmed, female, pastvoted*) after genetic matching.

- **Estimates**
  - Intercept: The estimated intercept is *0.527961*, indicating the expected value of *peacefactor* when all predictors (*directlyharmed, female, pastvoted*) are zero. The small standard error of *0.002245* and low p-value of *<2e-16* show that the estimate is statistically significant.

  - Directly harmed: The coefficient estimate for *directlyharmed* is *0.139680*. This suggests by being directly harmed, the *peacefactor* index increases by *0.139680* units, when other predictors are held constant. The small standard error of

0.002132 suggests high precision in this estimate, and the extremely low p-value (<2e-16) signifies strong evidence against the null hypothesis, indicating a significant positive effect of being directly harmed on *peacefactor*.

  - Female: The coefficient estimate for females is *-0.403079*. This indicates that being female is associated with a decrease of *0.403079* units in *peacefactor*, compared to males, controlling for other predictors. The small standard error *0.002119* underscores the precision of this estimate, and the very low p-value (< *2e-16*) indicates strong evidence against the null hypothesis, highlighting the significant impact of gender on *peacefactor*.

  - Past voted: The coefficient estimate for *pastvoted* is *-0.269868*. This suggests that past voting behavior is associated with a decrease of *0.269868* units in *peacefactor*, holding other predictors constant. The small standard error of *0.002200* reflects the precision of the estimate and the low p-value of *<2e-16* provides strong evidence against the null hypothesis, highlighting that having voted in the past has a negative effect on the peace factor index.

- **Model Fit:**

The adjusted R-squared value (*0.3614*) indicates that the model explains about 36.14% of the variance in the peace factor score.

Sensitivity Analysis of GenMatch I Regression

In the sensitivity analysis performed on the regression model *peacefactor ~ directlyharmed + female + pastvoted*, the aim is to assess potential biases caused by unobserved confounding factors. Here's a breakdown of the key sensitivity statistics and their interpretations:

- Partial R2 of Treatment with Outcome:

  - This statistic *(0.0476*) indicates that for an extreme confounder explaining 100% of the residual variance of the outcome, it would need to explain at least *4.76%* of the residual variance of the treatment to fully negate the observed estimated effect.

- Robustness Values (RVq=1 and RVq=1, alpha=0.05):

- These values (0.2 and 0.1947, respectively) signify the strength of unobserved confounders needed to reduce the estimated effect to zero or to a level where it is no longer statistically different from zero. If unobserved confounders explain more than 20% of the residual variance of both the treatment and outcome, they can nullify the estimated effect (q = 1), or render it statistically insignificant (q = 1, alpha = 0.05).

Genetic Matching II and Regression

- **Balance**

  The mean standard difference for variable female between the control and treatment group went down from -7.9978 to -0.38116 before and after matching while the T-test p-value also increased from 0.16084 to 0.31731.

  For variable pastvoted, the mean standard deviation between the control and treatment groups decreased significantly from 5.1898 to 0and the minimum T-test p-value increased from 0.36437 to 1 after matching.

  The lowest minimum p-value observed is for variable female at 0.31731.

- **Model Summary**
  - The linear regression model (lm) examines the relationship between the peace factor (*peacefactor*) and predictors (*directlyharmed, female, pastvoted*) along with the interaction between the predictor *wouldvote* and *pastvoted* after genetic matching.

- **Estimates**
  - Intercept: The coefficient estimate for the intercept is 0.440665 with a standard error of 0.003814. The intercept estimate indicates the baseline peacefactor value given all predictors (including directlyharmed) is 0. The small standard error and highly significant p-value of <2e-16 show that the estimate is statistically significant.

  - Directly harmed: The coefficient estimate for directlyharmed is 0.063553. This suggests that being directly harmed is associated with an increase of 0.063553 units in the peacefactor index, when other predictors are held constant. The standard error of 0.002725 reflects the precision of this estimate, and the highly significant p-value (<2e-16) indicates a strong positive effect of being directly harmed on peacefactor.

    - Female: The coefficient estimate for females is -0.386211. This indicates that being female is associated with a decrease of 0.386211 units in peacefactor, compared to males, controlling for other predictors. The standard error of 0.002978 underscores the precision of this estimate, and the highly significant p-value (<2e-16) indicates a substantial impact of gender on peacefactor.

    - Past voted: The coefficient estimate for pastvoted is -0.096035. This suggests that past voting behavior is associated with a decrease of 0.096035 units in peacefactor, holding other predictors constant. The standard error of 0.003580 reflects the precision of the estimate, and the highly significant p-value of (<2e-16) provides strong evidence against the null hypothesis, indicating a negative effect of past voting on the peacefactor index.

    - Would vote: The coefficient estimate for wouldvote is 0.065233, suggesting that planning to vote in the future is associated with an increase of 0.065233 units in peacefactor. The standard error of 0.004603 indicates the precision of this estimate. Despite the slightly lower t-value compared to other predictors, the p-value remains highly significant (<2e-16), indicating a positive effect of intending to vote on peacefactor.

    - Pastvoted:Wouldvote: This interaction term's estimate (0.030328) suggests that the effect of intending to vote (wouldvote) on peacefactor varies depending on past voting behavior (pastvoted). The highly significant p-value (<2e-16) indicates that this interaction is unlikely to be due to random chance. Therefore, for individuals with a history of past voting, the intention to vote (wouldvote) is associated with a moderate increase in peacefactor.

- **Model Fit:**

The adjusted R-squared value (*0.1277*) indicates that the model explains about 12.77% of the variance in the peace factor score.

Sensitivity Analysis of GenMatch II Regression

In the sensitivity analysis performed on the regression model *peacefactor ~ directlyharmed + female + pastvoted + wouldvote:pastvoted*, the aim is to assess potential biases caused by unobserved confounding factors. Here's a breakdown of the key sensitivity statistics and their interpretations:

- Partial R2 of treatment with outcome (0.0448):
  - This represents the proportion of variance in the outcome (peacefactor) explained by the treatment variable (directlyharmed), after accounting for other covariates. An extreme confounder would need to explain at least 4.48% of the residual variance of the treatment to fully negate the observed effect.

- Robustness Values (q = 1):
  - Unobserved confounders that explain more than 19.44% of the residual variance of both the treatment and the outcome are strong enough to entirely nullify the estimated effect of directlyharmed. This indicates the threshold at which potential biases could significantly impact the estimate.

Result Summary

The Genetic Matching II analysis reveals significant findings regarding predictors' impact on the peace factor score. After matching, there is an improved balance between groups, with key covariates like female and past-voted showing minimal mean differences and non-significant p-values.

In the regression model, directly harmed is positively associated with peace factor (coefficient = 0.063553, $p < 2e-16$), while being female (coefficient = -0.386211, $p < 2e-16$) and past voted (coefficient = -0.096035, $p < 2e-16$) are negatively associated. Intending to vote (would vote) shows a positive effect (coefficient = 0.065233, $p < 2e-16$), especially when interacting with past voting behavior (coefficient = 0.030328, $p < 2e-16$).

- Direct harm and peace factor

  For individuals directly harmed by violence in Darfur, their peace factor score increases by approximately 0.063553 units on average compared to those who have not been directly harmed. This positive coefficient suggests that individuals who have been directly harmed by violence in Darfur tend to have higher peace factor scores. This could indicate resilience or a desire for peace among those directly affected by the conflict.

- Female and peace factor

  Female individuals in Darfur have peace factor scores that are lower by approximately 0.386211 units on average compared to male individuals, controlling for other factors. The negative coefficient for being female indicates that women in Darfur generally report lower peace factor scores compared to men. This finding may reflect the disproportionate impact of violence and insecurity on women, as well as potential gender-specific barriers to peace.

- Voting history and peace factor

  Individuals with a history of voting in Darfur have peace factor scores that are lower by approximately 0.096035 units on average compared to those who have not voted in the past, after adjusting for other variables. The negative association with past voting behavior suggests that individuals with a history of voting in Darfur may have lower peace factor scores. This finding could be interpreted in various ways, such as disillusionment with the political process or skepticism about the efficacy of voting in promoting peace.

- Intention to vote and peace factor

  Expressing an intention to vote in the future (would vote) is associated with an increase in peace factor score by approximately 0.065233 units on average. Individuals who express an intention to vote in the future (would vote) exhibit higher peace factor scores. This positive effect underscores the potential for civic engagement and democratic participation to contribute positively to attitudes toward peace in Darfur.

- History of voting, intention to vote, and peace factor

  Among individuals with a history of past voting, expressing an intention to vote in the future is associated with an additional increase in peace factor score by approximately 0.030328 units on average. The positive interaction effect suggests that the intention to vote (would vote) has an even stronger positive impact on peace factor scores among individuals with a history of past voting. This finding highlights the potential cumulative effect of continued civic engagement on fostering attitudes conducive to peace.

The model explains 12.77% of the variance in peace factor score, with sensitivity analysis indicating the potential impact of unobserved confounders, where a confounder would need to explain more than 19.44% of residual variance to nullify the effect of directlyharmed. These results underscore the nuanced relationships between predictors and peacefactor, emphasizing the importance of matching techniques and cautious interpretation in observational studies.

Basque Country & Synthetic Control

Summary Findings

The synthetic control analysis conducted for Andalucía reveals compelling insights into the impact of an intervention on economic outcomes. The estimated treatment effect, with a Mean Square Prediction Error (MSPE) of 0.0006735709, indicates a statistically significant positive effect of the intervention on GDP per capita.

The synthetic Andalucía, constructed from a weighted combination of control regions, closely mirrors the pre-intervention characteristics of Andalucía, validated by the weights assigned to contributing regions (w.weights), such as Navarra (60.7%), Baleares (Islas) (15.4%), and Extremadura (23.2%). Key predictor variables such as education levels, investment, and sector-specific indicators (notably industry and services) significantly contribute to explaining the treatment effect, as indicated by the weights assigned to these variables (v.weights). Overall, the low prediction error (Loss V = 0.0006735709) and robust model performance underscore the reliability of the synthetic control method in capturing the causal impact of policy interventions on regional economic dynamics, offering valuable insights for policy evaluation and decision-making.

Table I: Covariate Balance

	Treated	Synthetic	Sample Mean
school.illit	8.398	9.032	11.155
school.prim	83.678	82.856	80.727
school.med	5.078	5.291	5.450
school.high	2.846	2.822	2.668
invest	21.577	20.722	21.413
special.gdpcap.1960.1969	3.700	3.698	3.573
special.sec.agriculture.1961.1969	21.726	25.974	21.328
special.sec.energy.1961.1969	6.278	2.680	5.246
special.sec.industry.1961.1969	22.780	22.722	22.401
special.sec.construction.1961.1969	7.832	7.606	7.238
special.sec.services.venta.1961.1969	34.290	34.037	36.677
special.sec.services.nonventa.1961.1969	7.096	6.986	7.112
special.popdens.1969	24.040	50.029	104.439

Table 1: Table outlining the covariate balance between treated and synthetic Andalucia .

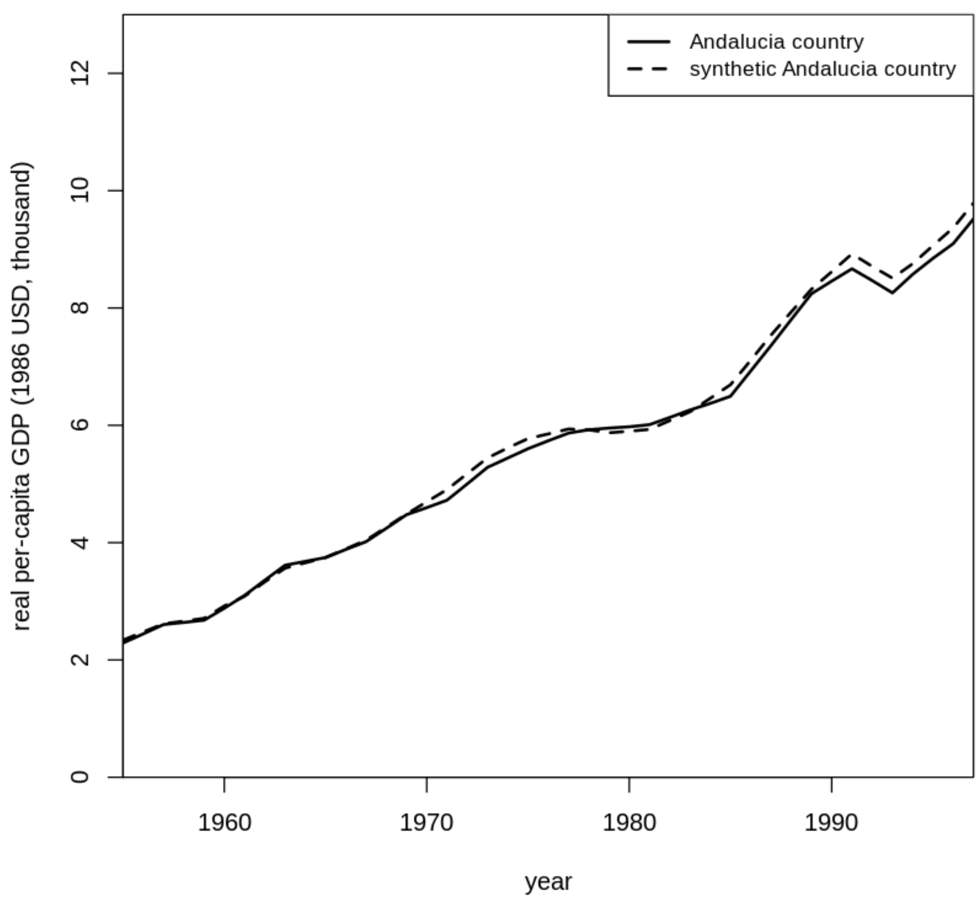
Table II: Unit-Weights

	w.weights	unit.names	unit.numbers
2	0.000	Andalucia	2
4	0.003	Principado De Asturias	4

5	0.154	Baleares (Islas)	5
6	0.000	Canarias	6
7	0.000	Cantabria	7
8	0.003	Castilla Y Leon	8
9	0.000	Castilla-La Mancha	9
10	0.000	Cataluna	10
11	0.000	Comunidad Valenciana	11
12	0.232	Extremadura	12
13	0.000	Galicia	13
14	0.000	Madrid (Comunidad De)	14
15	0.000	Murcia (Region de)	15
16	0.607	Navarra (Comunidad Foral De)	16
18	0.000	Rioja (La)	18

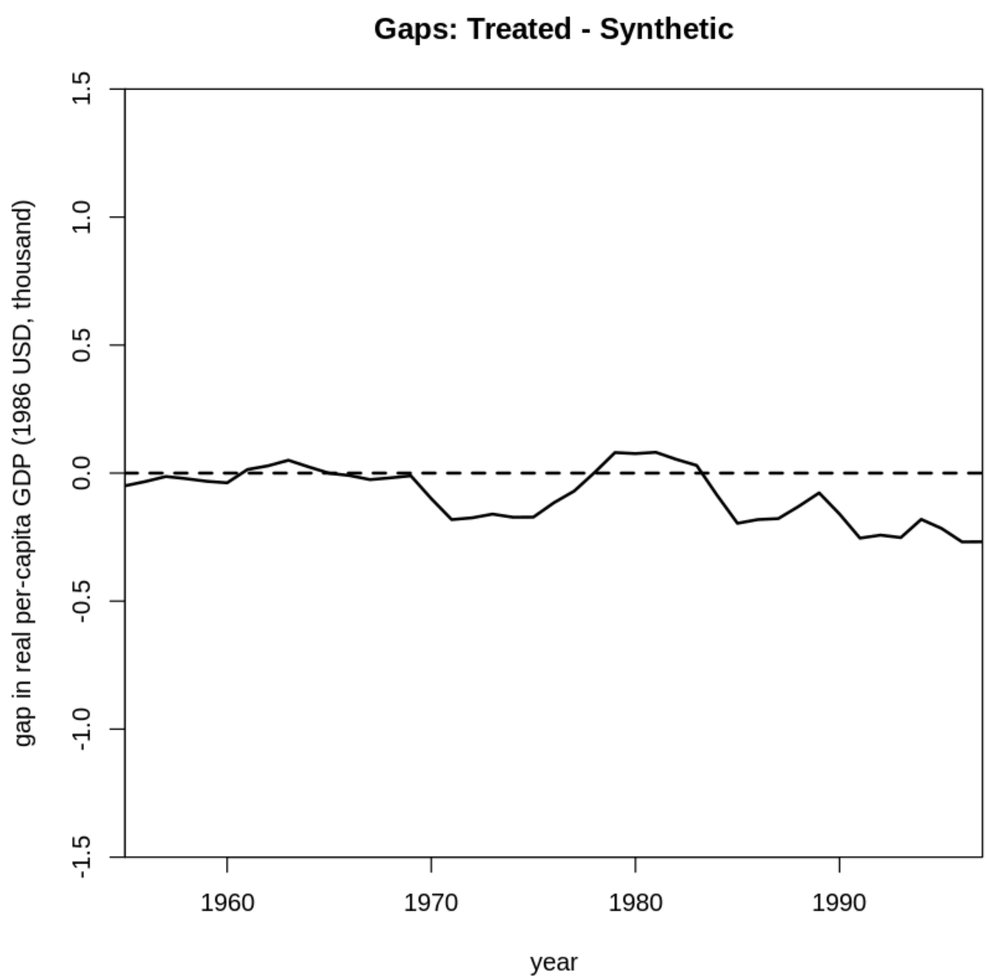
Table 2: Table outlining the donor pool units used to construct synthetic Andalucia with corresponding weights.

Plot I: Path Plot



Plot 1: Line plot comparing Andalucia versus synthetic Andalucia. It is evident that the synthetic Andalucia closely matches actual Andalucia in terms of economic performance over the years.

Plot II: Gaps Plot



Plot 2: Line plot illustrating the gap in real per capita GDP between actual and synthetic Andalucia over the years.

Link to Code

https://sle-collaboration.minervaproject.com/?minervaNotebookId=clb55k1st007t0j4f0k3x5d0o&userId=11935&name=Alina+Salman&avatar=https%3A//s3.amazonaws.com/picasso.fixtures/Alina_Salman_11935_2021-09-05T19%3A27%3A47.604Z&readOnly=0&isInstructor=0&signature=1fcd796af922e68e338b61dbc9bbb1784f9b782005403e5d37e6abd6065e372b

AI Support

I used ChatGPT with syntax help in R and ensure my summarized findings aligned with the results in the code.

In [2]

```
install.packages("sensemakr")
install.packages("Matching")
install.packages("rgenoud")
```

Run Code

Out [2]

```
Updating HTML index of packages in '.Library'

Making 'packages.html' ...
 done

Updating HTML index of packages in '.Library'

Making 'packages.html' ...
 done

Updating HTML index of packages in '.Library'

Making 'packages.html' ...
 done
```

+ New Code   + New Text

In [3]

```
library(sensemakr)
library(Matching)
library(rgenoud)
data("darfur")
```

Run Code

Out [3]

```
See details in:

Carlos Cinelli and Chad Hazlett (2020). Making Sense of Sensitivity: Extending Omitted Variable Bias. Journal of
the Royal Statistical Society, Series B (Statistical Methodology).

Loading required package: MASS

##
##  Matching (Version 4.10-14, Build Date: 2023-09-13)
##  See https://www.jsekhon.com for additional documentation.
##  Please cite software as:
##   Jasjeet S. Sekhon. 2011. ``Multivariate and Propensity Score Matching
##   Software with Automated Balance Optimization: The Matching package for R.''
##   Journal of Statistical Software, 42(7): 1-52.
##


##  rgenoud (Version 5.9-0.10, Build Date: 2023-12-13)
##  See http://sekhon.berkeley.edu/rgenoud for additional documentation.
##  Please cite software as:
```

```
##   ``Genetic Optimization Using Derivatives: The rgenoud package for R.''
##   Journal of Statistical Software, 42(11): 1-26.
##
```

+ New Code   + New Text

In [5]

```r
# Extract covariates (predictors) and treatment indicator
Tr <- darfur$directlyharmed

# This extracts the treatment indicator variable (Tr) from the 'darfur' dataset.
# 'directlyharmed' is a binary variable indicating whether individuals were directly harmed.

# Specify the logistic regression model formula
model_formula <- Tr ~ female + pastvoted

# Define the logistic regression model formula.
# Here, 'Tr' is the response variable (treatment indicator),
# and 'female', 'village', and 'pastvoted' are predictors used in the model.

# Fit the logistic regression model
darfur_model <- glm(formula = model_formula, data = darfur, family = binomial(link = "logit"))

# Fit a logistic regression model using the specified formula and data.
# The 'family = binomial(link = "logit")' part specifies a logistic regression with a logit link function,
# suitable for binary response variables like 'Tr'.

# Extract the fitted values (predicted probabilities) from the logistic regression model
X <- darfur_model$fitted.values

# Extract the predicted probabilities (fitted values) of the logistic regression model.
# These probabilities estimate the likelihood of being in the treatment group (Tr = 1) for each observation.

# Perform nearest neighbor matching using Match function
matched_data <- Match(Y = NULL, Tr = Tr, X = X,
                      estimand="ATT", M = 1, replace = FALSE)

# Perform nearest neighbor matching based on propensity scores.
# 'Y = NULL' indicates that there's no outcome variable specified (unsupervised matching).
# 'Tr' is the treatment indicator variable.
# 'X' contains the estimated propensity scores (predicted probabilities).
# 'estimand = "ATT"' specifies that we are estimating the Average Treatment Effect on the Treated (ATT).
# 'M = 1' means each treated unit is matched with exactly one control unit.
# 'replace = FALSE' ensures that each control unit is used only once in the matching process.

# Check covariate balance using MatchBalance
match_balance <- MatchBalance(
  Tr ~ female + pastvoted,
  data = darfur,
  match.out = matched_data
```

```
  nboots = 500
)

# Assess covariate balance after matching.
# 'Tr ~ female + village + pastvoted' specifies the covariates included in the balance check.
# 'data = darfur' specifies the dataset.
# 'match.out = matched_data' is the output from the matching process to be evaluated.
# 'nboots = 500' specifies the number of bootstrap samples used for balance assessment.
# The function evaluates whether the distribution of covariates is similar between treatment and control groups
# after performing the matching based on propensity scores.

# The resulting 'match_balance' object can be used to examine the balance of covariates
# and assess the effectiveness of the matching procedure.


summary(matched_data)

```

Run Code

Out [5]

```
***** (V1) female *****
                       Before Matching          After Matching
mean treatment........    0.43289                 0.43289
mean control..........    0.47256                 0.43289
std mean diff.........    -7.9978                       0

mean raw eQQ diff.....   0.039698                       0
med  raw eQQ diff.....          0                       0
max  raw eQQ diff.....          1                       0

mean eCDF diff........   0.019832                       0
med  eCDF diff........   0.019832                       0
max  eCDF diff........   0.039665                       0

var ratio (Tr/Co).....     0.9855                       1
T-test p-value........    0.16084                       1


***** (V2) pastvoted *****
                       Before Matching          After Matching
mean treatment........    0.65784                 0.65784
mean control..........     0.6332                 0.65784
std mean diff.........     5.1898                       0

mean raw eQQ diff.....   0.024575                       0
med  raw eQQ diff.....          0                       0
max  raw eQQ diff.....          1                       0

mean eCDF diff........   0.012323                       0
med  eCDF diff........   0.012323                       0
max  eCDF diff........   0.024646                       0
```

+ New Code   + New Text

R (384MB RAM) | Edit   Run All Cells   Kernel Ready |

In [6]

```r
# Extract indices of control units that were matched to the treated units
match_data_indices <- matched_data$index.control[matched_data$index.treated]

# Subset the original dataset to include rows corresponding to matched treated and control units
match_data <- darfur[c(matched_data$index.treated, match_data_indices), ]

# Display the first few rows of the matched data subset
head(match_data)

# Fit a linear regression model using matched data
matched_regression <- lm(peacefactor ~ directlyharmed + female + pastvoted, data = match_data)

# Display summary of the regression model
summary(matched_regression)

# Perform sensitivity analysis using sensemakr package
darfur.sensitivity <- sensemakr(
  model = matched_regression,             # Specifying the regression model to analyze
  treatment = "directlyharmed",           # Variable representing the treatment
  benchmark_covariates = "female",        # Covariate(s) to benchmark against
  kd = 1:3,                               # Range of values for number of top matches for treatment
  ky = 1:3,                               # Range of values for number of top matches for outcome
  q = 1,                                  # Number of treated units to evaluate
  alpha = 0.05,                           # Significance level for sensitivity analysis
  reduce = TRUE                           # Whether to reduce sensitivity analysis results
)

```

Run Code

Out [6]

	wouldvote	peacefactor	peace_formerenemies	peace_jjindiv	peace_jjtribes
	<dbl>	<dbl[,1]>	<dbl>	<dbl>	<dbl>
106	1	0.4951777	1	0	0
107	0	0.0000000	0	0	0
110	0	0.0000000	0	0	0
116	0	0.0000000	0	0	0
124	0	0.2020084	0	0	0
132	1	0.7949184	1	0	1

R (384MB RAM) | Edit	Run All Cells	Kernel Ready | |

```
lm(formula = peacefactor ~ directlyharmed + female + pastvoted,
    data = match_data)

Residuals:
    Min      1Q  Median      3Q     Max
-0.4922 -0.1988 -0.1574  0.3027  0.8295

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)     0.45089    0.02952  15.273   <2e-16 ***
directlyharmed  0.02824    0.02653   1.064    0.287
female         -0.29346    0.02463 -11.916   <2e-16 ***
pastvoted       0.01310    0.02592   0.506    0.613
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3278 on 740 degrees of freedom
  (314 observations deleted due to missingness)
Multiple R-squared:  0.1659,	Adjusted R-squared:  0.1625
F-statistic: 49.05 on 3 and 740 DF,  p-value: < 2.2e-16
```

+ New Code + New Text

In [7]

```
# Display the sensitivity analysis results
darfur.sensitivity
summary(darfur.sensitivity)
```

Run Code

Out [7]

```
Sensitivity Analysis to Unobserved Confounding

Model Formula: peacefactor ~ directlyharmed + female + pastvoted

Null hypothesis: q = 1 and reduce = TRUE

Unadjusted Estimates of ' directlyharmed ':
  Coef. estimate: 0.02824
  Standard Error: 0.02653
  t-value: 1.06444

Sensitivity Statistics:
  Partial R2 of treatment with outcome: 0.00153
  Robustness Value, q = 1 : 0.03837
  Robustness Value, q = 1 alpha = 0.05 : 0

For more information, check summary.

Sensitivity Analysis to Unobserved Confounding

Model Formula: peacefactor ~ directlyharmed + female + pastvoted

Null hypothesis: q = 1 and reduce = TRUE
-- This means we are considering biases that reduce the absolute value of the current estimate.
-- The null hypothesis deemed problematic is H0:tau = 0

Unadjusted Estimates of 'directlyharmed':
  Coef. estimate: 0.0282
  Standard Error: 0.0265
  t-value (H0:tau = 0): 1.0644
```

```
  Partial R2 of treatment with outcome: 0.0015
  Robustness Value, q = 1: 0.0384
  Robustness Value, q = 1, alpha = 0.05: 0

Verbal interpretation of sensitivity statistics:

-- Partial R2 of the treatment with the outcome: an extreme confounder (orthogonal to the covariates) that
explains 100% of the residual variance of the outcome, would need to explain at least 0.15% of the residual
variance of the treatment to fully account for the observed estimated effect.

-- Robustness Value, q = 1: unobserved confounders (orthogonal to the covariates) that explain more than
3.84% of the residual variance of both the treatment and the outcome are strong enough to bring the point
estimate to 0 (a bias of 100% of the original estimate). Conversely, unobserved confounders that do not
explain more than 3.84% of the residual variance of both the treatment and the outcome are not strong enough
to bring the point estimate to 0.

-- Robustness Value, q = 1, alpha = 0.05: unobserved confounders (orthogonal to the covariates) that explain
more than 0% of the residual variance of both the treatment and the outcome are strong enough to bring the
estimate to a range where it is no longer 'statistically different' from 0 (a bias of 100% of the original
```

+ New Code   + New Text

In [9]

```r
X <- cbind(darfur$peacefactor + darfur$female + darfur$pastvoted)

genout <- GenMatch(Tr=Tr, estimand="ATT", X = X, M=1, pop.size=1000)
summary(genout)
match_out <- Match(Tr = Tr, estimand="ATT", X=X, Weight.matrix=genout)
summary(match_out)
gen_matched_balance <- MatchBalance(Tr ~ darfur$female + darfur$pastvoted, data = darfur,
match.out=match_out, nboots=100)
```

Run Code

Out [9]

```
Sun Apr  7 18:02:27 2024
Domains:
 0.000000e+00   <=  X1   <=    1.000000e+03

Data Type: Floating Point
Operators (code number, name, population)
        (1) Cloning...........................  122
        (2) Uniform Mutation..................  125
        (3) Boundary Mutation.................  125
        (4) Non-Uniform Mutation..............  125
        (5) Polytope Crossover................  125
        (6) Simple Crossover..................  126
        (7) Whole Non-Uniform Mutation........  125
        (8) Heuristic Crossover...............  126
        (9) Local-Minimum Crossover...........  0

SOFT Maximum Number of Generations: 100
Maximum Nonchanging Generations: 4
Population size       : 1000
Convergence Tolerance: 1.000000e-03

Not Using the BFGS Derivative Based Optimizer on the Best Individual Each Generation.
Not Checking Gradients before Stopping.
Using Out of Bounds Individuals.
```

```
GENERATION: 0 (initializing the population)
Lexical Fit..... 5.638261e-01  1.000000e+00
#unique......... 1000, #Total UniqueCount: 1000
var 1:
best............ 1.000000e+00

              Length Class  Mode
value              2 -none- numeric
par                1 -none- numeric
Weight.matrix      1 -none- numeric
matches       128817 -none- numeric
ecaliper           0 -none- NULL
```

```
mean eCDF diff........    0.019832                 0.016232
med  eCDF diff........    0.019832                 0.016232
max  eCDF diff........    0.039665                 0.032465

var ratio (Tr/Co).....      0.9855                  0.98599
T-test p-value........     0.16084                 0.090045


***** (V2) darfur$pastvoted *****
                       Before Matching          After Matching
mean treatment........     0.65784                  0.65784
mean control..........      0.6332                  0.65501
std mean diff.........      5.1898                  0.59705

mean raw eQQ diff.....    0.024575                 0.030741
med  raw eQQ diff.....           0                        0
max  raw eQQ diff.....           1                        1

mean eCDF diff........    0.012323                 0.015371
med  eCDF diff........    0.012323                 0.015371
max  eCDF diff........    0.024646                 0.030741

var ratio (Tr/Co).....     0.96965                  0.99607
T-test p-value........     0.36437                  0.87336


Before Matching Minimum p.value: 0.16084
Variable Name(s): darfur$female  Number(s): 1

After Matching Minimum p.value: 0.090045
Variable Name(s): darfur$female  Number(s): 1
```

+ New Code + New Text

In [10]
```
summary(gen_matched_balance)
```

Run Code

Out [10]
```
                    Length Class  Mode
BeforeMatching      2      -none- list
AfterMatching       2      -none- list
BMsmallest.p.value  1      -none- numeric
BMsmallestVarName   1      -none- character
BMsmallestVarNumber 1      -none- numeric
AMsmallest.p.value  1      -none- numeric
AMsmallestVarName   1      -none- character
```

+ New Code   + New Text

In [11]

```r
# Extract indices of control units that were matched to the treated units
gen_indices <- match_out$index.control[match_out$index.treated]

# Subset the original dataset to include rows corresponding to matched treated and control units
gen_match_data <- darfur[c(match_out$index.treated, gen_indices), ]

# Display the first few rows of the matched data subset
head(gen_match_data)

# Fit a linear regression model using matched data
gen_regression <- lm(peacefactor ~ directlyharmed + female + pastvoted, data = gen_match_data)

# Display summary of the regression model
summary(gen_regression)

# Perform sensitivity analysis using sensemakr package
darfur.sensitivity.gen <- sensemakr(
  model = gen_regression,            # Specifying the regression model to analyze
  treatment = "directlyharmed",          # Variable representing the treatment
  benchmark_covariates = "female",       # Covariate(s) to benchmark against
  kd = 1:3,                              # Range of values for number of top matches for treatment
  ky = 1:3,                              # Range of values for number of top matches for outcome
  q = 1,                                 # Number of treated units to evaluate
  alpha = 0.05,                          # Significance level for sensitivity analysis
  reduce = TRUE                          # Whether to reduce sensitivity analysis results
)

# Display the sensitivity analysis results
darfur.sensitivity.gen
```

Run Code

Out [11]

	wouldvote	peacefactor	peace_formerenemies	peace_jjindiv	peace_jjtribe
	<dbl>	<dbl[,1]>	<dbl>	<dbl>	<dbl>
106	1	0.4951777	1	0	0
106.1	1	0.4951777	1	0	0
106.2	1	0.4951777	1	0	0
106.3	1	0.4951777	1	0	0

	wouldvote	peacefactor	peace_formerenemies	peace_jjindiv	peace_jjtribe
	<dbl>	<dbl[,1]>	<dbl>	<dbl>	<dbl>
106.4	1	0.4951777	1	0	0
106.5	1	0.4951777	1	0	0

```
Call:
lm(formula = peacefactor ~ directlyharmed + female + pastvoted,
    data = gen_match_data)

Residuals:
     Min       1Q   Median       3Q      Max
-0.66764 -0.25809  0.00531  0.14499  1.00531

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)     0.527961   0.002245  235.22   <2e-16 ***
directlyharmed  0.139680   0.002132   65.52   <2e-16 ***
female         -0.403079   0.002119 -190.21   <2e-16 ***
pastvoted      -0.269868   0.002200 -122.64   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3076 on 85874 degrees of freedom
Multiple R-squared:  0.3615,    Adjusted R-squared:  0.3614
F-statistic: 1.62e+04 on 3 and 85874 DF,  p-value: < 2.2e-16

Warning message in ovb_partial_r2_bound.numeric(r2dxj.x = r2dxj.x[i], r2yxj.dx = r2yxj.dx[i], :
“Implied bound on r2yz.dx greater than 1, try lower kd and/or ky. Setting r2yz.dx to 1.”

Sensitivity Analysis to Unobserved Confounding

Model Formula: peacefactor ~ directlyharmed + female + pastvoted

Null hypothesis: q = 1 and reduce = TRUE

Unadjusted Estimates of ' directlyharmed ':
  Coef. estimate: 0.13968
  Standard Error: 0.00213
  t-value: 65.52284

Sensitivity Statistics:
  Partial R2 of treatment with outcome: 0.04761
  Robustness Value, q = 1 : 0.19999
  Robustness Value, q = 1 alpha = 0.05 : 0.19465

For more information, check summary.
```

+ New Code + New Text

In [12]

```
summary(darfur.sensitivity.gen)
```

Run Code

Out [12]

```
Sensitivity Analysis to Unobserved Confounding
```

```
Null hypothesis: q = 1 and reduce = TRUE
-- This means we are considering biases that reduce the absolute value of the current estimate.
-- The null hypothesis deemed problematic is H0:tau = 0

Unadjusted Estimates of 'directlyharmed':
  Coef. estimate: 0.1397
  Standard Error: 0.0021
  t-value (H0:tau = 0): 65.5228

Sensitivity Statistics:
  Partial R2 of treatment with outcome: 0.0476
  Robustness Value, q = 1: 0.2
  Robustness Value, q = 1, alpha = 0.05: 0.1947

Verbal interpretation of sensitivity statistics:

-- Partial R2 of the treatment with the outcome: an extreme confounder (orthogonal to the covariates) that
explains 100% of the residual variance of the outcome, would need to explain at least 4.76% of the residual
variance of the treatment to fully account for the observed estimated effect.

-- Robustness Value, q = 1: unobserved confounders (orthogonal to the covariates) that explain more than 20%
of the residual variance of both the treatment and the outcome are strong enough to bring the point estimate
to 0 (a bias of 100% of the original estimate). Conversely, unobserved confounders that do not explain more
than 20% of the residual variance of both the treatment and the outcome are not strong enough to bring the
point estimate to 0.

-- Robustness Value, q = 1, alpha = 0.05: unobserved confounders (orthogonal to the covariates) that explain
more than 19.47% of the residual variance of both the treatment and the outcome are strong enough to bring
the estimate to a range where it is no longer 'statistically different' from 0 (a bias of 100% of the
```

+ New Code + New Text

In [13]

```r
X <- cbind(darfur$peacefactor, darfur$female, darfur$pastvoted, darfur$wouldvote * darfur$pastvoted)
genout2 <- GenMatch(Tr=Tr, estimand="ATT", X = X, M=1, pop.size=1000)
summary(genout2)
match_out2 <- Match(Tr = Tr, X=X, Weight.matrix=genout2)
summary(match_out2)
gen_matched_balance2 <- MatchBalance(Tr ~ darfur$female + darfur$pastvoted, data = darfur,
match.out=match_out2, nboots=100)

```

Run Code

Out [13]

```
Sun Apr  7 18:42:47 2024
Domains:
 0.000000e+00   <=  X1   <=    1.000000e+03
 0.000000e+00   <=  X2   <=    1.000000e+03
 0.000000e+00   <=  X3   <=    1.000000e+03
 0.000000e+00   <=  X4   <=    1.000000e+03

Data Type: Floating Point
Operators (code number, name, population)
        (1) Cloning..........................  122
        (2) Uniform Mutation.................  125
        (3) Boundary Mutation................  125
        (4) Non-Uniform Mutation.............  125
        (5) Polytope Crossover...............  125
```

```
        (7) Whole Non-Uniform Mutation........  125
        (8) Heuristic Crossover...............  126
        (9) Local-Minimum Crossover...........  0

SOFT Maximum Number of Generations: 100
Maximum Nonchanging Generations: 4
Population size       : 1000
Convergence Tolerance: 1.000000e-03

Not Using the BFGS Derivative Based Optimizer on the Best Individual Each Generation.
Not Checking Gradients before Stopping.
Using Out of Bounds Individuals.

Maximization Problem.
GENERATION: 0 (initializing the population)
Lexical Fit..... 5.216167e-02  3.173112e-01  3.173112e-01  1.000000e+00  1.000000e+00  1.000000e+00

              Length Class  Mode
value             8  -none- numeric
par               4  -none- numeric
Weight.matrix    16  -none- numeric
matches       57138  -none- numeric
ecaliper          0  -none- NULL


Estimate...  0
SE.........  0
T-stat.....  NaN
p.val......  NA

Original number of observations..............  1276
Original number of treated obs...............  529
Matched number of observations...............  529
Matched number of observations  (unweighted).  19046


***** (V1) darfur$female *****
                       Before Matching          After Matching
mean treatment........    0.43289                  0.43289
mean control..........    0.47256                  0.43478
std mean diff.........    -7.9978                 -0.38116

mean raw eQQ diff.....   0.039698               5.2504e-05
med  raw eQQ diff.....          0                        0
max  raw eQQ diff.....          1                        1

mean eCDF diff........   0.019832               2.6252e-05
med  eCDF diff........   0.019832               2.6252e-05
max  eCDF diff........   0.039665               5.2504e-05

var ratio (Tr/Co).....     0.9855                  0.99898
T-test p-value........    0.16084                  0.31731


***** (V2) darfur$pastvoted *****
                       Before Matching          After Matching
mean treatment........    0.65784                  0.65784
```

+ New Code   + New Text

In [18]

```
# Extract indices of control units that were matched to the treated units
gen_indices2 <- match_out2$index.control[match_out2$index.treated]

# Subset the original dataset to include rows corresponding to matched treated and control units
```

R (384MB RAM) | Edit   Run All Cells   Kernel Ready |

```r

# Display the first few rows of the matched data subset
head(gen_match_data2)

# Fit a linear regression model using matched data
gen_regression2 <- lm(peacefactor ~ directlyharmed + female + pastvoted + wouldvote*pastvoted, data = gen_match_data2)

# Display summary of the regression model
summary(gen_regression2)

# Perform sensitivity analysis using sensemakr package
darfur.sensitivity.gen2 <- sensemakr(
  model = gen_regression2,            # Specifying the regression model to analyze
  treatment = "directlyharmed",          # Variable representing the treatment
  benchmark_covariates = "female",       # Covariate(s) to benchmark against
  kd = 1:3,                              # Range of values for number of top matches for treatment
  ky = 1:3,                              # Range of values for number of top matches for outcome
  q = 1,                                 # Number of treated units to evaluate
  alpha = 0.05,                          # Significance level for sensitivity analysis
  reduce = TRUE                          # Whether to reduce sensitivity analysis results
)

# Display the sensitivity analysis results
darfur.sensitivity.gen2
```

Run Code

Out [18]

	wouldvote	peacefactor	peace_formerenemies	peace_jjindiv	peace_jjtribe
	<dbl>	<dbl[,1]>	<dbl>	<dbl>	<dbl>
106	1	0.4951777	1	0	0
106.1	1	0.4951777	1	0	0
106.2	1	0.4951777	1	0	0
106.3	1	0.4951777	1	0	0
106.4	1	0.4951777	1	0	0
106.5	1	0.4951777	1	0	0

```
Call:
lm(formula = peacefactor ~ directlyharmed + female + pastvoted + 
    wouldvote * pastvoted, data = gen_match_data2)
```

```
Residuals:
     Min       1Q   Median       3Q      Max
-0.56945 -0.11801 -0.02197  0.08917  0.97803

Coefficients:
                     Estimate Std. Error  t value Pr(>|t|)
(Intercept)          0.440665   0.003814  115.551  < 2e-16 ***
directlyharmed       0.063553   0.002725   23.326  < 2e-16 ***
female              -0.386211   0.002978 -129.693  < 2e-16 ***
pastvoted           -0.096035   0.003580  -26.829  < 2e-16 ***
wouldvote            0.065233   0.004603   14.171  < 2e-16 ***
pastvoted:wouldvote  0.030328   0.005612    5.404 6.55e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3068 on 61979 degrees of freedom
Multiple R-squared:  0.3597,    Adjusted R-squared:  0.3597
F-statistic:  6964 on 5 and 61979 DF,  p-value: < 2.2e-16


Sensitivity Analysis to Unobserved Confounding

Model Formula: peacefactor ~ directlyharmed + female + pastvoted + wouldvote *
    pastvoted

Null hypothesis: q = 1 and reduce = TRUE

Unadjusted Estimates of ' directlyharmed ':
  Coef. estimate: 0.06355
  Standard Error: 0.00272
  t-value: 23.32625

Sensitivity Statistics:
  Partial R2 of treatment with outcome: 0.0087
  Robustness Value, q = 1 : 0.08941
  Robustness Value, q = 1 alpha = 0.05 : 0.08222

For more information, check summary.
```

+ New Code + New Text

In [16]

```
1 summary(darfur.sensitivity.gen2)
```

Run Code

Out [16]

```
  Partial R2 of treatment with outcome: 0.0448
  Robustness Value, q = 1: 0.1944
  Robustness Value, q = 1, alpha = 0.05: 0.189
```

R (384MB RAM) | Edit Run All Cells Kernel Ready |

```
estimate to a range where it is no longer 'statistically different' from 0 (a bias of 100% of the original
estimate), at the significance level of alpha = 0.05. Conversely, unobserved confounders that do not explain
more than 18.9% of the residual variance of both the treatment and the outcome are not strong enough to bring
the estimate to a range where it is no longer 'statistically different' from 0, at the significance level of
alpha = 0.05.

Bounds on omitted variable bias:

--The table below shows the maximum strength of unobserved confounders with association with the treatment
and the outcome bounded by a multiple of the observed explanatory power of the chosen benchmark covariate(s).

 Bound Label R2dz.x R2yz.dx      Treatment Adjusted Estimate Adjusted Se
   1x female 0.0151  0.3343 directlyharmed            0.0910      0.0018
   2x female 0.0302  0.6688 directlyharmed            0.0454      0.0013
   3x female 0.0453  1.0000 directlyharmed           -0.0008      0.0000
```

+ New Code + New Text

In [26]

```r
library(Synth)
data(basque)

# Define predictors, dependent variable, and other parameters
predictors <- c("school.illit", "school.prim", "school.med",
                "school.high", "invest")
dependent <- "gdpcap"
treatment_identifier <- 3  # Andalucía's region number in the dataset

# dataprep: prepare data for synth
dataprep.out <-
  dataprep(
  foo = basque
  ,predictors= c("school.illit",
                 "school.prim",
                 "school.med",
                 "school.high",
                 "school.post.high"
                 ,"invest"
                 )
   ,predictors.op = c("mean")
   ,dependent     = c("gdpcap")
   ,unit.variable = c("regionno")
   ,time.variable = c("year")
   ,special.predictors = list(
    list("gdpcap",1960:1969,c("mean")),
    list("sec.agriculture",seq(1961,1969,2),c("mean")),
    list("sec.energy",seq(1961,1969,2),c("mean")),
    list("sec.industry",seq(1961,1969,2),c("mean")),
    list("sec.construction",seq(1961,1969,2),c("mean")),
    list("sec.services.venta",seq(1961,1969,2),c("mean")),
    list("sec.services.nonventa",seq(1961,1969,2),c("mean")),
    list("popdens",1969,c("mean")))
    ,treatment.identifier  = 3
    ,controls.identifier = c(2, 4:16, 18)
    ,time.predictors.prior = c(1964:1969)
```

```r
    ,time.optimize.ssr     = c(1960:1969)
    ,unit.names.variable   = c("regionname")
    ,time.plot           = c(1955:1997)
    )

# 1. combine highest and second highest
# schooling category and eliminate highest category
dataprep.out$X1["school.high",] <-
 dataprep.out$X1["school.high",] +
 dataprep.out$X1["school.post.high",]
dataprep.out$X1                 <-
 as.matrix(dataprep.out$X1[
  -which(rownames(dataprep.out$X1)=="school.post.high"),])
dataprep.out$X0["school.high",] <-
 dataprep.out$X0["school.high",] +
 dataprep.out$X0["school.post.high",]
dataprep.out$X0                 <-
dataprep.out$X0[
 -which(rownames(dataprep.out$X0)=="school.post.high"),]

# 2. make total and compute shares for the schooling catgeories
lowest  <- which(rownames(dataprep.out$X0)=="school.illit")
highest <- which(rownames(dataprep.out$X0)=="school.high")

dataprep.out$X1[lowest:highest,] <-
 (100 * dataprep.out$X1[lowest:highest,]) /
 sum(dataprep.out$X1[lowest:highest,])
dataprep.out$X0[lowest:highest,] <-
 100 * scale(dataprep.out$X0[lowest:highest,],
             center=FALSE,
             scale=colSums(dataprep.out$X0[lowest:highest,])
                                                )

# run synth
synth.out <- synth(data.prep.obj = dataprep.out)

# Get result tables
synth.tables <- synth.tab(
                          dataprep.res = dataprep.out,
                          synth.res = synth.out
                          )

# results tables:
print(synth.tables)

# plot results:
# path
path.plot(synth.res = synth.out,
          dataprep.res = dataprep.out,
          Ylab = c("real per-capita GDP (1986 USD, thousand)"),
```

```r
          Ylim = c(0,13),
          Legend = c("Andalucia country","synthetic Andalucia country"),
          )

## gaps
gaps.plot(synth.res = synth.out,
          dataprep.res = dataprep.out,
          Ylab = c("gap in real per-capita GDP (1986 USD, thousand)"),
          Xlab = c("year"),
          Ylim = c(-1.5,1.5),
          )

```

Run Code

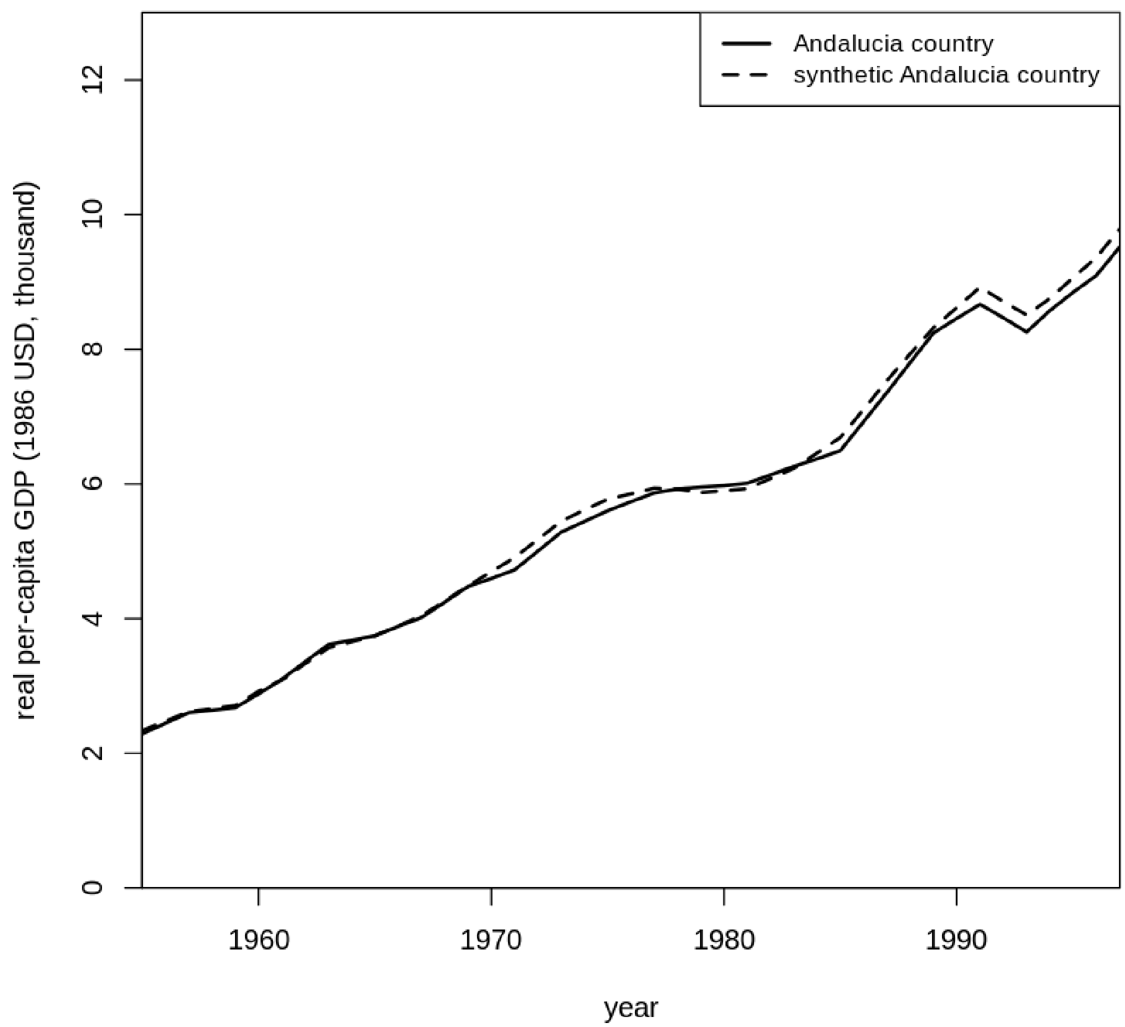
Out [26]

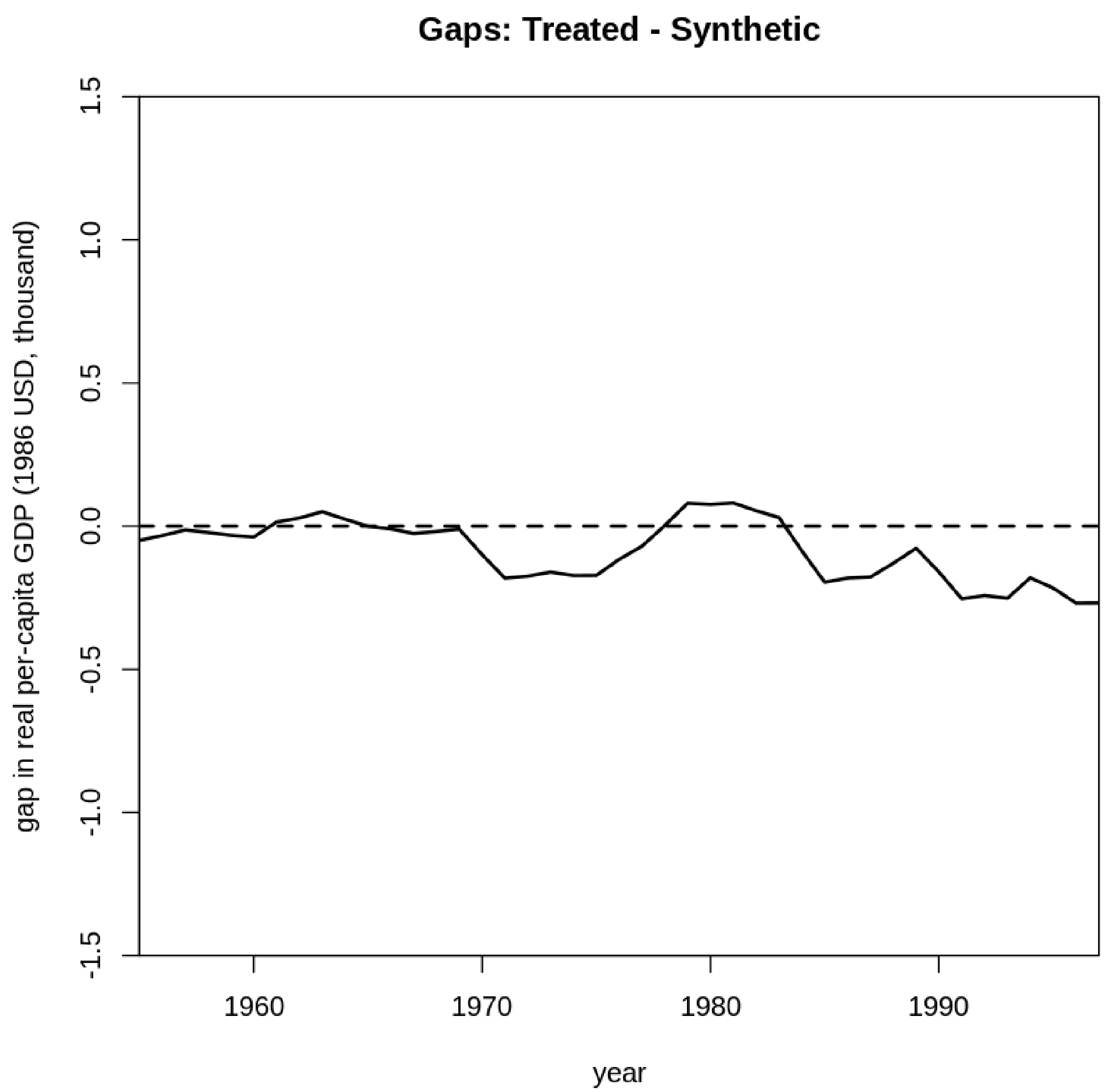
```
school.med                              0.007
school.high                             0.088
invest                                  0.009
special.gdpcap.1960.1969                0.097
special.sec.agriculture.1961.1969       0.001
special.sec.energy.1961.1969            0
special.sec.industry.1961.1969          0.485
special.sec.construction.1961.1969      0.046
special.sec.services.venta.1961.1969    0.132
special.sec.services.nonventa.1961.1969 0.004
special.popdens.1969                    0.105

$tab.w
   w.weights                   unit.names unit.numbers
2      0.000                    Andalucia            2
4      0.003         Principado De Asturias            4
5      0.154              Baleares (Islas)            5
6      0.000                     Canarias            6
7      0.000                    Cantabria            7
8      0.003               Castilla Y Leon            8
9      0.000           Castilla-La Mancha            9
10     0.000                     Cataluna           10
11     0.000         Comunidad Valenciana           11
12     0.232                  Extremadura           12
13     0.000                      Galicia           13
14     0.000          Madrid (Comunidad De)           14
15     0.000             Murcia (Region de)           15
16     0.607 Navarra (Comunidad Foral De)           16
18     0.000                   Rioja (La)           18

$tab.loss
          Loss W       Loss V
[1,] 0.01006354 0.0006735709
```

R (384MB RAM) | Edit   Run All Cells   Kernel Ready | |


Andalucia country
synthetic Andalucia country
real per-capita GDP (1986 USD, thousand)
0
2
4
6
8
10
12
1960
1970
1980
1990
year



+ New Code

+ New Text

In [27]

```
1
```

Run Code

Out [27]

```
Warning message in min(x):
“no non-missing arguments to min; returning Inf”
Warning message in max(x):
“no non-missing arguments to max; returning -Inf”
Warning message in min(x):
“no non-missing arguments to min; returning Inf”
Warning message in max(x):
“no non-missing arguments to max; returning -Inf”
Error in plot.window(...): need finite 'xlim' values
Traceback:

1. plot(treatment_effect, type = "l", col = "blue", xlab = "Year",
 .     ylab = "Treatment Effect", main = "Estimated Treatment Effect for Andalucía")
```

```
.     ylab = "Treatment Effect", main = "Estimated Treatment Effect for Andalucía")
3. localWindow(xlim, ylim, log, asp, ...)
4. plot.window(...)
```

+ New Code + New Text

```
In [44]   1
```

Run Code

Out [44]

darfur {sensemakr}	R Documentation

Data from survey of Darfurian refugees in eastern Chad.

Description

Data on attitudes of Darfurian refugees in eastern Chad. The main "treatment" variable is `directlyharmed`, which indicates that the individual was physically injured during attacks on villages in Darfur, largely between 2003 and 2004. The main

R (384MB RAM) | Edit Run All Cells Kernel Ready |

outcome of interest is `peacefactor`, a measure of pro-peace attitudes.
Key covariates include `herder_dar` (whether they were a herder in Darfur), `farmer_dar` (whether they were a farmer in Darfur), `age`, `female` (indicator for female), and `past_voted` (whether they report having voted in an earlier election, prior to the conflict).

Usage

```
darfur
```

Format

A data frame with 1276 rows and 14 columns.

wouldvote If elections were held in Darfur in the future, would you vote? (0/1)

peacefactor A measure of pro-peace attitudes, from a factor analysis of several questions. Rescaled such that 0 is minimally pro-peace and 1 is maximally pro-peace.

peace_formerenemies Would you be willing to make peace with your former enemies? (0/1)

peace_jjindiv Would you be willing to make peace with Janjweed individuals who carried out violence? (0/1)

peace_jjtribes Would you be willing to make peace with the tribes that were part of the Janjaweed? (0/1)

gos_soldier_execute Should Government of Sudan soldiers who perpetrated attacks on civilians be executed? (0/1)

directlyharmed A binary variable indicating whether the respondent was personally physically injured during attacks on villages in Darfur largely between 2003-2004. 529 respondents report being personally injured, while 747 do not report being injured.

age Age of respondent in whole integer years. Ages in the data range from 18 to 100.

farmer_dar The respondent was a farmer in Darfur (0/1). 1,051 respondents were farmers, 225 were not.

herder_dar The respondent was a herder in Darfur (0/1). 190 respondents were farmers, 1,086 were not.

pastvoted The respondent reported having voted in a previous election before the conflict (0/1). 821 respondents reported having voted in a previous election, 455 reported not having voted in a previous election.

hhsize_darfur Household size while in Darfur.

village Factor variable indicating village of respondent. 486 unique villages are accounted for in the data.

female The respondent identifies as female (0/1). 582 respondents are female-identified, 694 are not.

References

Cinelli, C. and Hazlett, C. (2020), "Making Sense of Sensitivity: Extending Omitted Variable Bias." Journal of the Royal Statistical Society, Series B (Statistical Methodology).
Hazlett, Chad. (2019) "Angry or Weary? How Violence Impacts Attitudes toward Peace among Darfurian Refugees." Journal of Conflict Resolution: 0022002719879217.

[Package *sensemakr* version 0.1.4]

+ New Code + New Text

In [13] 1

Run Code + New Code + New Text

In [14]

R (384MB RAM) | Edit Run All Cells Kernel Ready |