DOES MARRIAGE AFFECT MENTAL HEALTH FOR AMERICANS?

TEAM CAUSAL CRUSADERS: DEEPANSHU MALHOTRA, VIBHA NANDA, YUANXI FU

Source: Javier Sánchez Mingorance www.javiindy.com

Background

In recent years, the relationship between marriage and mental health has become a critical area of study, especially as societal norms and structures evolve. Research shows varying impacts of marriage on an individual's psychological well-being, with some studies suggesting positive effects, such as emotional support and improved life satisfaction, while others point to potential stressors and negative consequences (Grundström et al., 2021;

Jace & Makridis, 2021; Uecker 2012). However, disentangling the causal effect of marriage on mental health is notoriously difficult. For one thing, people who chose to get married are believed to be different from people who stay out of marriage (Stutzer & Frey, 2006). While prior quantitative studies often used linear regression, exploring the causal relationship between marriage and mental health through the lens of causal inference with machine learning could provide clearer insights into the complexities of this relationship.

We collaborated on all stages for the project, from selecting variables to deploying models. We used WhatsApp to communicate, Google Colab and GitHub to share code. We use Zoom meetings to make big decisions, such as confounding variable selections.

Data

We used the General Social Survey (GSS) data which is a nationally representative survey of adults in the United States conducted since 1972 (Davern et al., 2024). The GSS collects data on contemporary American society in order to monitor and explain trends in opinions, attitudes and behaviors. We used the panel data version [GSS 2016-2020 Panel (Release 1a, April 2022)]() in STATA format and read it into python with pandas. The dataset contains two cohorts of longitudinal data: the first cohort from 2016 with re-interviews in 2020 and the second cohort in 2018 with re-interviews in 2020. The data is provided in wide format, with variables from 2016 suffixed with "1a", from 2018 suffixed with "1b" and from 2020 suffixed with "_2". There are **2348** instances from the 2018 cohort and **2867** instances from the 2016 cohort.

	samptype	yearid	fileversion	panstat	wtssall_1a	wtssall_1b	wtssall_2	wtssnr_1a	wtssnr_1b	wtssnr_2	...	sprtlrgr_2	sprtpurp_2	poltrtblk_2	poltrths
0	2016	20160001	GSS 2020 Panel Release 1 (May 2021)	1	0.956994	NaN	1.085009	1.260478	NaN	1.443929	...	7.0	7.0	2.0	
1	2016	20160002	GSS 2020 Panel Release 1 (May 2021)	1	0.478497	NaN	0.542504	0.630239	NaN	0.721964	...	7.0	7.0	4.0	
2	2016	20160003	GSS 2020 Panel Release 1 (May 2021)	0	0.956994	NaN	NaN	1.260478	NaN	NaN	...	NaN	NaN	NaN	N
3	2016	20160004	GSS 2020 Panel Release 1 (May 2021)	1	1.913987	NaN	2.170018	2.520956	NaN	2.887858	...	5.0	4.0	4.0	
4	2016	20160005	GSS 2020 Panel Release 1 (May 2021)	0	1.435490	NaN	NaN	1.890717	NaN	NaN	...	NaN	NaN	NaN	N

Figure 1. A glimpse of the GSS 2016-2020 Panel dataset

Methods

Treatment variable selection

The "marital status" consists of five different statuses: married, widowed, divorced, separated, and never married. We made two choices. First, we only kept married and never married for simplicity. Second, we forced widowed, divorced, separated, and never married into T = 0.

Outcome variable selection

We identified two outcome variables. One variable, “mntlhlth”, is days of poor mental health in the past 30 days. Notably, this variable is biased towards negative mental health conditions. The other, “hlthmntl” is how the respondent rates their mental health, including mood and ability to think, with 1 as “excellent” and 5 as “poor.” This one is balanced but only measured in the 2018 cohort. We will experiment with both outcome variables.

Changes to the plan

We planned to use the values from 2020 (the reinterview) to establish the cause (marriage) preceded the effect (mental health). However, neither “mntlhlth” nor “hlthmntl” were measured in 2020, and we had to use values from 2016 or 2018. Because the two variables mark evaluations of recent mental health conditions (last 30 days or at the time of the interview), we can still claim that marriage preceded mental health.

Confounding variable selection

The GSS dataset contains over 4000 variables (codebook). We started by filtering out the variables with more than 50% missing values, which directly got us down to 120 variables. At the end, we settled for five confounding variables covering five important aspects: finance, religious attitude, level of education, socioeconomic status (SES), and race. For each aspect, we select one variable.

- degree: respondent's highest degree
- hhrace: race of household
- satfin: satisfaction with financial situation
- neisafe: how safe interviewer thinks neighborhood is safe (we think this variable can be used to approximate SES)
- relpersn: whether the respondent considers oneself a religious person

Changes to the plan

We dropped a few variables initially selected.

- adults: household members 18 yrs and older
- earners: how many in family earned money

We dropped them because they are likely the consequence of marriage and therefore, strong predictors of marriage.

- wrkstats: labor force status

We dropped wrkstats because it has 8 levels, and each requires a one-hot-encoding. Since we were encountering problems in model building, we decided to drop it. Also, the finance confounding variable "satfin" can cover some of the effects from labor force status, for example, employed versus unemployed.

It occurred to us that work may be a source of stress that can affect people's mental health and their ability to get or stay married. For this reason, we identified one more variable called "stress" (how often does r find work stressful). This variable was measured in 2018 but missing one balot in 2016. However, we did not include this variable in the current study due to the constraint of time.

Preprocessing

We created four datasets from the GSS 2016-2020 Panel dataset to be used for training the PS models.

1. 2016_all_marital (2795 instances)
2. 2016_marital_trim (1974 instances)
3. 2018_all_marital (2310 instances)
4. 2018_marital_trim (1645 instances)

Suffixed “all_marital” indicates that we kept instances with “marital” as 1 to 5 (married, widowed, divorced, separated, and never married). Suffixed “marital_trim” indicates that we only kept instances with “marital” as 1 (married) or 5 (never married). Instances with any NAs for any of the treatment and confounding variables were dropped.

We created six datasets from the GSS 2016-2020 Panel dataset to be used for training the DR models

1. 2016_mntlhlth (1067 instances)
2. 2018_mntlhlth (1393 instances)
3. 2018_hlthmntl (2296 instances)

In these three datasets, the treatment variable takes “married” as 1 and the rest (“widowed”, “divorced”, “separated”, and “never married”) as 0. The suffix of the datasets’ names indicate different outcome measures. For the 2016 cohort, hlthmntl was not measured, and therefore, we do not have 2016_hlthmntl. For the five confounding variables, only “hhrace” is one-hot-encoded, and the rest were used was they were. Rows with any NA for any of the treatment, outcome, or confounding variables were dropped.

4. 2016_mntlhlth_marital_trim (760 instances)
5. 2018_mntlhlth_marital_trim (1057 instances)
6. 2018_hlthmntl_marital_trim (1633 instances)

In these three datasets, the treatment variable takes “married” as 1 and “never married” as 0. For the five confounding variables, only “hhrace” is one-hot-encoded, and the rest were used was they were. Rows with any NA for any of the treatment, outcome, or confounding variables were dropped.

Doubly Robust (DR) Estimator

We used the doubly-robust estimator (Bang & Robins, 2005; Funk et al., 2011). The estimator includes two parts: an propensity score (PS) model that predicts the likelihood of an instance belonging to the treatment group based on all confounding variables; and an outcome model that predicts the outcome based on the treatment variable, the confounding variables, and a weighting derived from the propensity scores. The weighting scheme is inverse probability of treatment weighting (IPTW), where the treated instance is weighted by 1/PS and the untreated weighted by 1/(1-PS).

Model training and evaluation

We trained four PS models and six DR models (Table 1). The difference between the number of PS models and DR models is due to the fact that “2018_mntlhlth” and “2018_hlthmntl” share the same PS model, and “2018_mntlhlth_marital_trim” and “2018_mntlhlth_marital_trim” share the same PS model. We used Random Forest for the PS model to model the complex, nonlinear relationship between the confounding variables and the treatment variable. We used weighted linear regression as the outcome model for it can provide a p-value and confidence intervals (CI) for the average treatment effect (i.e., the regression coefficient of the treatment variable).

Table 1. Datasets and Models		
PS models		
Dataset	PS model	
2016_all_marital	2016_all_marital_PS	
2016_marital_trim	2016_marital_trim_PS	
2018_all_marital	2018_all_marital_PS	
2018_marital_trim	2018_marital_trim_PS	
DR models		
Dataset	PS model	DR model
2016_mntlhlth	2016_all_marital_PS	2016_mntlhlth_DR

2018_mntlhlth	2018_all_marital_PS	2018_mntlhlth_DR
2018_hlthmntl		2018_hlthmntl_DR
2016_mntlhlth_marital_trim	2016_marital_trim_PS	2016_mntlhlth_marital_trim_DR
2018_mntlhlth_marital_trim	2018_marital_trim_PS	2018_mntlhlth_marital_trim_DR
2018_hlthmntl_marital_trim		2018_hlthmntl_marital_trim_DR

For training the PS model, we split the dataset 80/20. We used the 80% to choose model parameters using a 5-fold CV and compare the accuracy between the train and test datasets to make sure we did not overfit. The DR model was fitted with weighted linear regression using all instances and evaluated using R^2 and adjusted R^2.

Results

PS model results

Table 2 shows the train set and test set accuracy of our four PS models. The closeness between the train set and test set accuracies support that our PS models are not overfitted. Figure 1 shows the PS score distribution from the four models.

Table 2. PS Model Results

PS model	Train set accuracy	Test set accuracy
2016_all_marital_PS	0.6588	0.6225
2018_all_marital_PS	0.6569	0.5909
2016_marital_trim_PS	0.7169	0.6962
2018_marital_trim_PS	0.7280	0.6809

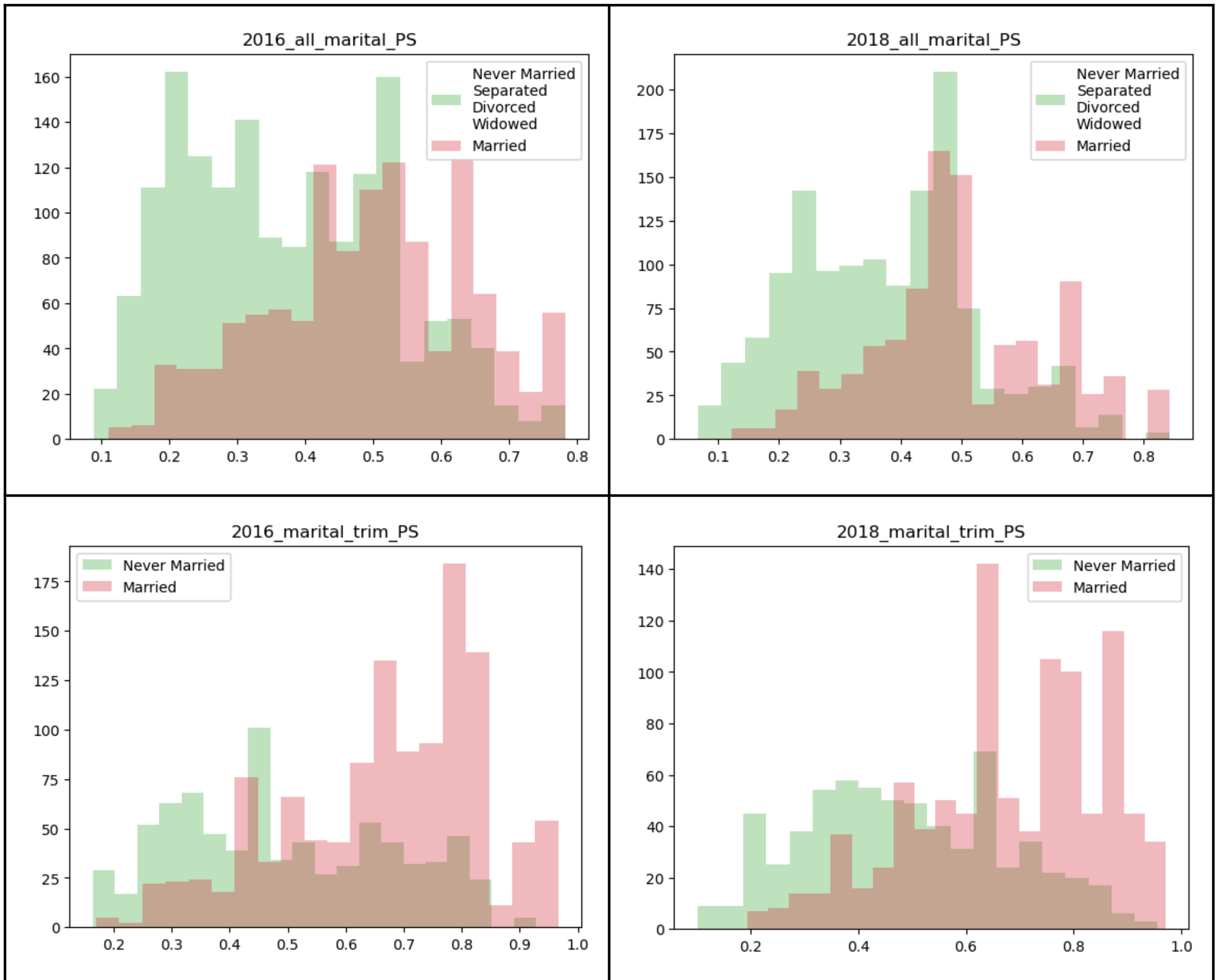

Figure 2. PS Distributions from Four PS Models

DR model results

To our surprise, all six DR models converge on the conclusion that marriage ***positively*** influences mental health. The regression coefficients were all negative and significant. For “mntlhlth,” a negative coefficient indicates reduction in days of poor mental health. For “hlthmntl,” a negative coefficient indicates moving towards better mental health. However, a large proportion of variance was not explained by the variables, as seeing from R^2 and adjusted R^2.

Table 2. DR Model Results					
DR model	Regression coefficient of the marital status	95% CI	R^2	Adjusted R^2	P-value of the F-statistic
2016_mntlhlth_DR	-1.6874	[-2.541, -0.833]	0.043	0.037	6.00e-08
2018_mntlhlth_DR	-1.4083	[-2.094, -0.723]	0.060	0.055	9.67e-16
2018_hlthmntl_DR	-0.1607	[-0.236, -0.085]	0.086	0.083	5.21e-41
2016_mntlhlth_marital_trim_DR	-1.4292	[-2.400, -0.459]	0.060	0.052	5.73e-08
2018_mntlhlth_marital_trim_DR	-1.3312	[-2.071, -0.592]	0.064	0.058	2.03e-13
2018_hlthmntl_marital_trim_DR	-0.1625	[-0.256, -0.075]	0.085	0.081	6.73e-28

Reflections

The biggest surprise is that we obtained a robust conclusion across different datasets. Process-wise, working with survey data turned out to be more challenging than we expected. We had to change our plan several times because of the constraints of data. The huge amount of variables also makes this research challenging. GSS Data Explorer, which is meant to make variable exploration easier, has its design limitations. For example, it did not produce “hlthmntl” when we searched by keyword “mental health.”

Limitation and future work

We saw signs that linear regression may not be a good outcome model: the proportion of variance explained was low. And in the cases where “mntlhlth” as outcome variables, many confounding variables were not significant – not the case when “hlthmntl” was the outcome variable. If we had more time, we would for sure try other outcome models.

Combining widowed, separated, and divorced with never married was a choice of intuition. However, in a retrospective, the

Code availability

Google Collab Notebook: https://drive.google.com/file/d/1hmEqQv64WcoNtIqvx7oJHLdOS3WkHbno/view?usp=sharing

Github:

https://github.com/yuanxiesa/IS557_machine_learning_team_project/tree/main/final_project_causal_ml

References

Bang, H., & Robins, J. M. (2005). Doubly robust estimation in missing data and causal inference models. *Biometrics*, *61*(4), 962–973. https://doi.org/10.1111/j.1541-0420.2005.00377.x

Davern, Michael; Bautista, Rene; Freese, Jeremy; Herd, Pamela; and Morgan, Stephen L.; General Social Survey 1972-2024. [Machine-readable data file]. Principal Investigator, Michael Davern; Co-Principal Investigators, Rene Bautista, Jeremy Freese, Pamela Herd, and Stephen L. Morgan. Sponsored by the National Science Foundation. NORC ed. Chicago: NORC, 2024: NORC at the University of Chicago [producer and distributor]. Data accessed from the GSS Data Explorer website at gssdataexplorer.norc.org.

Funk, M. J., Westreich, D., Wiesen, C., Stürmer, T., Brookhart, M. A., & Davidian, M. (2011). Doubly robust estimation of causal effects. *American Journal of Epidemiology*, *173*(7), 761–767. https://doi.org/10.1093/aje/kwq439

Grundström, J., Konttinen, H., Berg, N., & Kiviruusu, O. (2021). Associations between relationship status and mental well-being in different life phases from young to middle adulthood. *SSM - Population Health*, *14*, 100774. https://doi.org/10.1016/j.ssmph.2021.100774

Jace, C. E., & Makridis, C. A. (2021). Does marriage protect mental health? Evidence from the COVID-19 pandemic. *Social Science Quarterly*, *102*(6), 2499–2515. https://doi.org/10.1111/ssqu.13063

Stutzer, A., & Frey, B. S. (2006). Does marriage make people happy, or do happy people get married? *The Journal of Socio-Economics*, *35*(2), 326–347. https://doi.org/10.1016/j.socec.2005.11.043

Uecker, J. E. (2012). Marriage and Mental Health among Young Adults. *Journal of Health and Social Behavior*, *53*(1), 67–83. https://doi.org/10.1177/0022146511419206