Proceedings of the 3rd International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/6/20230827

Study on whether marriage affects depression based on causal inference

Haoran Zhou[1,5], Junliang Lu[2,6], Ziyu Li[3,7], Xinyi Zhang[4,8]

[1]Software College, Northeastern University, Shenyang, 110167, China
[2]School of Mathematical Sciences, University of Nottingham, Nottingham, NG7 2RD, United Kingdom
[3]School of Mathematical Sciences, Dalian University of Technology, Dalian, 116081, China
[4]Chengdu Shude High School, Chengdu, 610066, China

[5]20195430@stu.neu.edu.cn
[6]dannyzl234@outlook.com
[7]3533682676@qq.com
[8]2962288045@qq.com

Abstract. This paper applies causal-based machine learning algorithms to evaluate the causal effect of marriage on depression. The paper verifies the reinforcement of adopting causal inference through the relationship between causality and correlation and confounding bias and selection bias. In this paper, we firstly implement meta learner to estimate and analyse the causal effects. Considering the influence of confounding factors, we utilize two stages of least squares estimation and deep IV estimation based on instrumental variables to fully evaluate the causal effects. The evaluation of linear and nonlinear models shows different results, which is worthy of discussion in future studies. In conclusion, people in the rural region who get married are slightly less likely to get depressed in the future.

Keywords: depression, confounding bias, selection bias, meta-learner, instrumental variable estimation.

1. Introduction

1.1. Aim

The global prevalence of depression is high, and depression has gradually become an important cause of endangering people's health across the country. Depression is closely related to higher suicide risk rates, with nearly 800,000 people dying each year from depression. Therefore, in order to predict or detect patients at high risk of depression as early as possible, taking a comprehensive assessment and early prevention and treatment are needed, and further referral examinations are also necessary; it can alleviate the suffering of depressed patients and reduce medical expenses, and improve the quality of national medical services. In the past, most prediction studies of depression used logistic regression algorithms to establish models, and these methods have certain defects and deficiencies. Some scholars systematically integrated algorithms and deep learning to evaluate and predict depression-related factors in big data sets, but the results are still modest.

This study uses the data of patients with depression to analyze the factors influencing the occurrence of depression, uses meta learner and combined with some of the previous algorithms and causality analysis to predict the occurrence of depression, evaluates the utility of machine learning models in identifying depression-related risk factors, and provides a theoretical basis for machine learning algorithms to better apply to the clinical diagnosis and prediction of depression.

1.2. Depression

In 1948, the World Health Organization (WHO) Declaration of Establishment proposed a definition of physical health: health is not only the absence of disease and weakness but also includes physical, mental, and social perfection [1]. However, with the development of sociomedical models, mental health problems such as anxiety, depression, psychological stress, and fear have gradually gained attention. Around 450 million people worldwide have mental health problems in developing countries [2]. For instance, in 2015, the WHO pointed out that more than 3 million people in China die prematurely every year, more than 70 percent of white-collar workers in mainland cities are at sub-health levels, and more than 60 percent are overworked. As a result, their health levels are significantly low [3].

The term depression has existed in human history since the mid-20th century, and it is still a popular topic until now that has been extensively explored and studied by scholars from various countries; the study of it has gone through a long and bumpy historical road. Depression is a mental state of low mood and aversion to activity, which affects more than 280 million people of all ages (about 3.5% of the global population). The previously mentioned study of depression will also naturally include the study of its causes as an essential part of human cognition of depression. It can cause severe symptoms that affect how you feel, think, and handle daily activities, such as sleeping, eating, or working. It is an illness that can affect anyone—regardless of age, race, income, culture, or education. Research suggests that genetic, biological, environmental, and psychological factors play a role in depression.

1.3. Cause of the depression

Although many researchers across the country have analyzed the causes of depression, the cause of depression is still not very clear; most researchers believe that the cause of depression and other related disciplines are closely related, such as biology, psychology, and social and environmental sciences.

Even depression can not be generalized; different causes or personal physiological conditions can lead to different symptoms, but several factors are still very prominent and worth studying. Gender is a non-negligible factor in a depressive mood. Studies have shown that women have specific genes, such as the 5-serotonin gene, the tryptophan hydroxylase gene, and the brain-derived neurotrophic factor gene. In addition, women are more likely to have psychological inferiority, sensitivity, low self-evaluation, and other negative emotions in adolescent development, and the unique physiological structure makes rapid changes in the level of sex hormones in women's postpartum and premenstrual periods, especially the change of estrogen levels is more likely to promote depression [4].

Age and educational attainment are also essential influencers of depression, with older adults and people with low education at higher risk of depression [5]. The physical functions of middle-aged and older adults are also gradually declining, and various diseases, especially chronic diseases, will increase. When the elderly suffer from pain or mobility problems caused by physical diseases, they are often accompanied by negative emotions such as distress and negativity, which can easily cause depression, and depression will make chronic diseases such as high blood pressure, diabetes, heart disease, and other conditions worsen, forming a vicious circle.

As mentioned above, there are still many other reasons for depression. In this paper, specific relevant factors, such as age 、marriage, are selected for in-depth analysis.

1.4. EDA

This is an Exploratory Data Analysis to analyze and investigate data sets and summarize their main characteristics. This is a heat map to show the relationship between various variables.

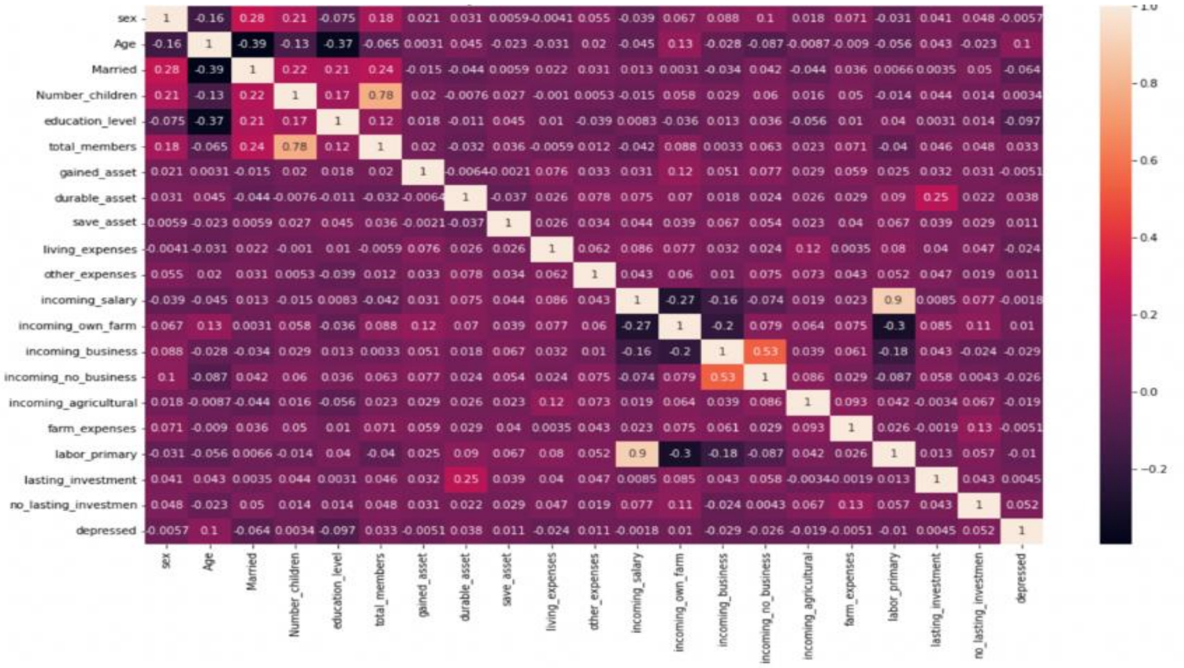

Figure 1. Heatmap of variable correlations.

Through this heat map, we found that sex, marital status, and education level have a more significant relationship with depression. Then we continue to make further analysis.

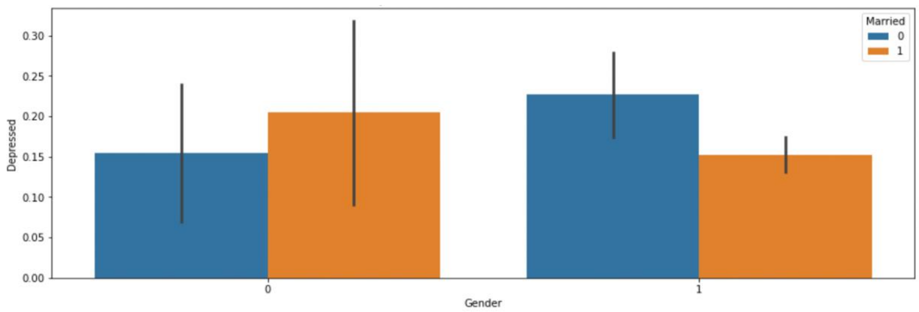

Figure 2. Depressed Vs. Marital Status VS Gender.

It is found that male depression before marriage is more potent than that after marriage, while female depression before marriage is weaker than that after marriage.

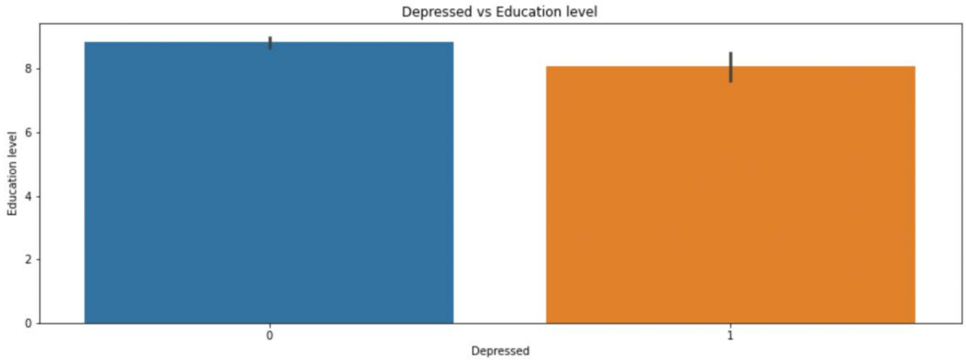

Figure 3. Depressed Vs Education Level.

It is found that higher-educated individuals are less depressed.
Based on the Exploratory Data Analysis, we conduct follow-up research.

1.5. ATE and ATT
ATE is the expectations of the effects of the treated and untreated objects, and ATT is the treated objects use some ways to find objects like them as substitutes and see the difference in their effects.

2. Motivation

2.1. Machine learning prediction
This subsection briefly displays how a machine learning algorithm works to predict depression, which serves as a comparison to causal machine learning. The first step is data preprocessing, which drops all objects with missing values and columns of ID since they make no contributions to prediction. The model selected is logistic regression, the classification method specified is one-vs-rest(OvR), the penalty method specified is L2 regularization and the algorithm specified is L-BFGS. At the same time, other parameters are default values in the scikit-learn function. From the receiver-operating-characteristic curve, it can be told that this model does not fit very well but has excellent robustness, and the accuracy score of this model is 0.83924

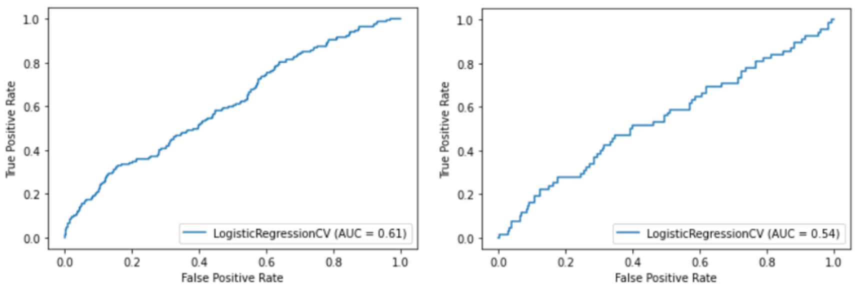

Figure 4. Receiver-operating-characteristic curve.

In addition, a random forest classifier is applied to view how these features contribute to the model. The influence of age and education level is relatively solid and consistent with our domain knowledge. Nevertheless, the impact of the number of children is pretty high, which is quite abnormal and might be caused by some unidentified confounders. Since machine learning is not well enough to build a fit model, there is a need for causal inference.

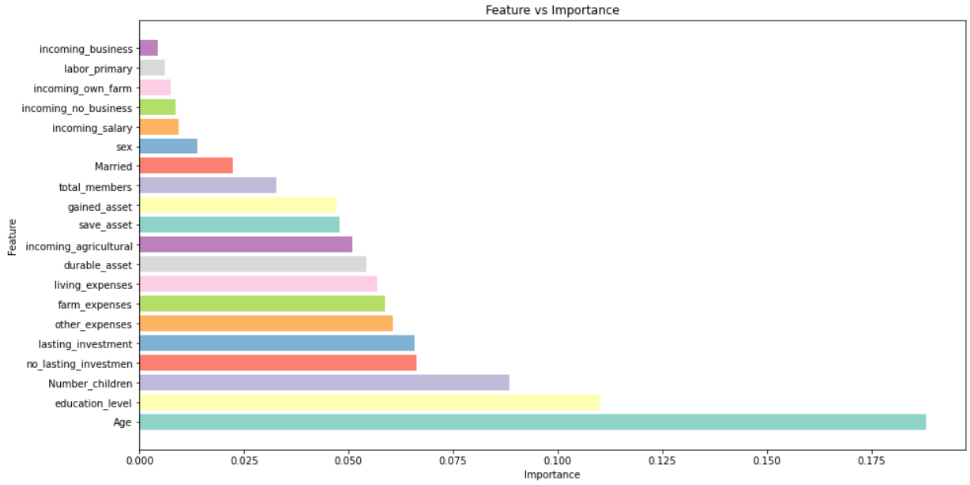

Figure 5. Feature importance.

2.2. Correlation and causality

Causality is different from correlation. Firstly, correlation does not necessarily imply causality. Two events are possibly observed to be correlated but do not have causal relationships. For example, assume it is observed during the same period that population increases in one city which is the leading cause, is verified to be economic development in this city, and the frequency of rain increases in another city which is verily due to climate change. Therefore, these two events are observed to be positively correlated, but they do not have causal relationships. Though it is pretty evident in this case that correlation might not imply causality, in other circumstances, it can be misleading so that that correlation might be treated as causality. For example, some researchers bring up knowledge more broadly.

In contrast, others bring up knowledge more specifically among those who think that increasing knowledge may be why memory improves with age. However, the generic knowledge hypothesis is not supported by the findings of the current studies [6]. Although knowledge and memory correlate positively with age, it is proved that increasing knowledge does not cause memory improvements with age in the general knowledge term.

When referring to the differences between correlation and causality, correlation does not necessarily imply causality is usually mentioned, which is cited in many textbooks. However, it is worth mentioning that causality does not necessarily imply correlation. For example, assume that smoking causes cancer and select a group of people which comprises smoking and non-smoking person and observe whether they have cancer. Consequently, it might be observed that there is only a weak correlation between smoking and cancer, and the conclusion could be drawn that smoking does not strongly correlate to cancer. The possible explanation could be that these people who smoke could have healthier bodies compared to those who do not smoke, so they would have a lower possibility of having cancer if they did not smoke. This is the counterfactual situation. Smoking increases the possibility of having cancer, therefore, in real situations, the possibility of having cancer is approximately the same for these two groups of people, and they are observed to have a weak correlation. Generally speaking, the health condition of people offsets the causal effects of smoking.

2.3. Confounding bias and selection bias

Causal machine learning could have better effects than classic machine learning because it can identify confounding and selection bias and take more specific debiasing measurements. Confounding bias means a confounding variable acts as the common cause for treatment and outcome, so they form a false causal effect. Roughly speaking, two kinds of confounding bias exist, as shown in Figure 3. In the left causal graph, confounding bias could fail the causal effect, while in the right causal graph, the causal

DOI: 10.54254/2755-2721/6/20230827

effect is only biased. Confounding bias has been identified before, but might be in other names, such as "causal fork", "classical confounder," and causal confounder [7]. In epidemiology, there is a causal model that illustrates confounding bias where cultural factors are the common cause of when humans began making and imbibing alcoholic drinks and caffeine assumption [8]. This depression dataset would be adopted to depict confounding bias; the treatment variable would be "Married", and the outcome variable would be "depressed". Assuming that income would influence people's decision of marriage and depression. In this circumstance, income is a common cause for "Married" and "Depressed", and leads to confounding bias.

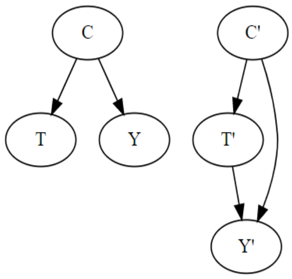

Figure 6. Confounding bias.

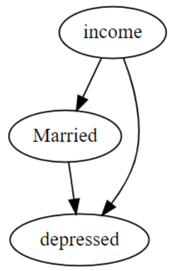

Figure 7. Confounding bias in depression.

Selection bias means that, in reality, people would not make random choices but have preferences. Sometimes, random experiments are assigned, but people have a preference for making choices, or they have a low possibility of making little choices and assuming an experiment that tests the effects of exposure to advertisements on the shopping behaviors of consumers, where consumers are randomly exposed to merchandises of all qualities. They are given the same chance to buy merchandise of all qualities. However, in reality, consumers are more likely to purchase high-quality products than low-quality ones, which generates bias. This is where selection bias lies.

3. Estimating causal effects

Before we dive into this part, let us first define our notations:

$$Treatment: T \in \{1, 0\}$$

$$The\ potential\ outcome\ with\ treatment: Y_1$$

$$Potential\ outcome\ without\ treatment: Y_0$$

$$Features: X$$

The goal is to estimate the treatment effect:

$$ATE = E[Y_1 - Y_0]$$

However, in reality, when an individual received the treatment, we can only get the Y_1

from that treatment. In order to solve this issue, we can use meta-learners. Meta learners are methods that utilize off-the-shelf machine learning algorithms to predict the average treatment effect, and in this paper, we will use the X learner [9].

3.1. Backgrounds about the X learner

The X learner has two stages and a propensity score model e which is simply the probability of the treatment given features:

$$P(T|X)$$

In the first stage, we split the data into two groups, one that receives the treatment and another that does not. Then, we input each group's features X and labeled outcomes Y_1 and Y_0 to two machine learning models to get the corresponding values:

$$M_0 = E[Y|T = 0, X]$$

DOI: 10.54254/2755-2721/6/20230827

$$M_1 = E[Y|T = 1, X]$$

Now during the second stage, we will input the treatment effects, which are Y_1 and Y_0 in each group; then, we will have:

$$\tau_0 = M_0 - Y_0$$

$$\tau_1 = Y_1 - M_1$$

Then, we need to create two additional machine learning models that take in each group's features and τ to predict:

$$M_{\tau 0} = E[\tau_0 | \mathrm{T} = 0]$$

$$M_{\tau 1} = E[\tau_1 | \mathrm{T} = 1]$$

Eventually, we can get an ultimate model by incorporating propensity scores as weightings:

$$\widehat{M} = M_{\tau 0}\, P(T|X) + M_{\tau 1}\big(1 - P(T|X)\big)$$

3.2. Meta Learner experiments with settings

First and foremost, the propensity score model we used is the logistic regression model, where we input all features and treatments and try to predict $P(T|X)$. On the other hand, we used LightGBM, which is a speed-up version of gradient boosting decision tree algorithms, and we set the maximum tree depth to be 3 and the minimum number of data needed in a leaf as 30 [10].

In addition, we also need to select a treatment. In recent studies, we found it interesting that young adults who got married are less likely to get depressed than single adults [11]. Hence, we chose marriage as our treatment and tried to study the causal effect of marriage on depression. We also divided people into three different age groups. People between 17 to 35 years old are considered young adults, people between 36 to 55 years old are considered middle-aged adults, and people above 56 years old are considered seniors. Note that these three groups are turning into binary values.

3.3. Experiment results

After implementing the x learner, we eventually get the value of the average treatment effect as -0.036. This indicates that if the people received the treatment (get married in this case), they are less likely to get depressed. The result is consistent with the psychology experiments conducted before [11]. However, what is each feature's importance in our model to estimate the ATE? To dive into this, we draw the SHAP value plot, listed each feature's importance in determining the average treatment effect [12]:

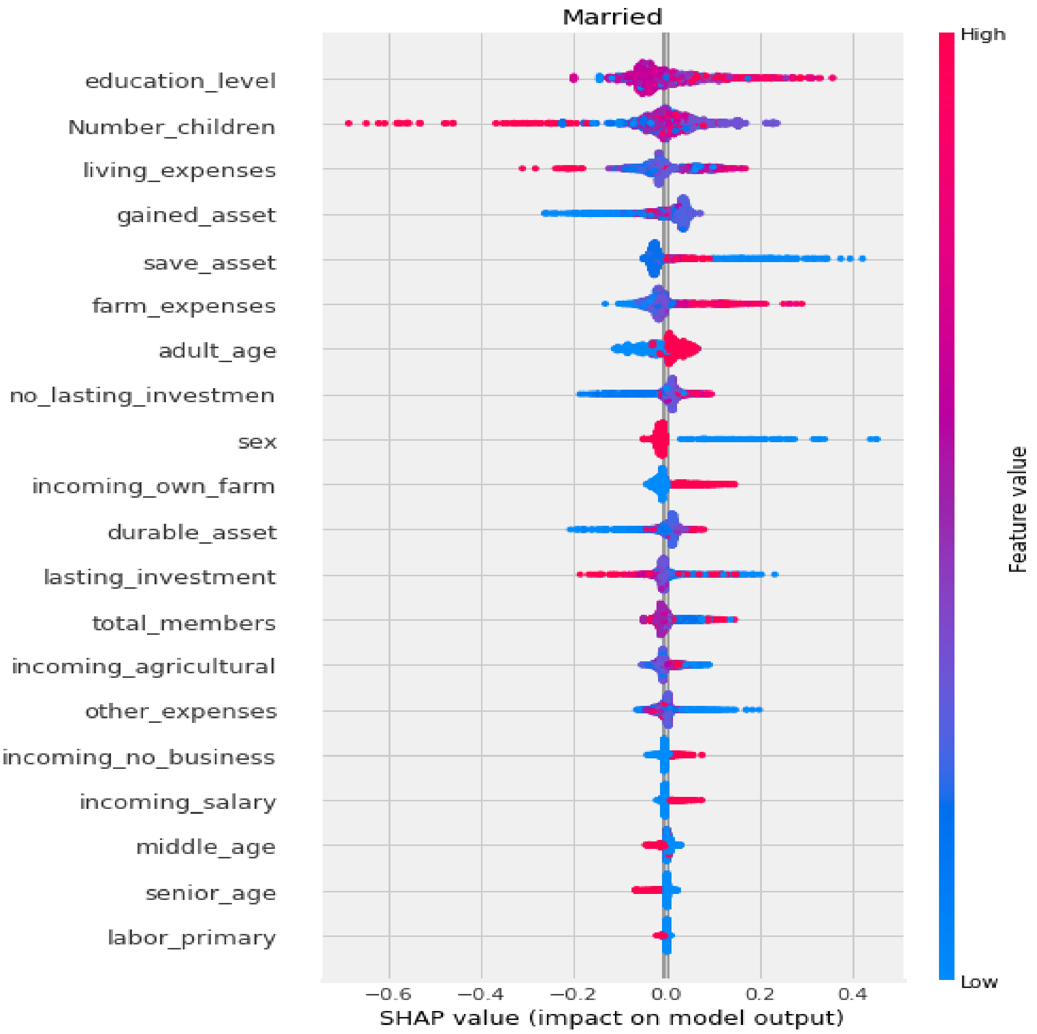

Figure 9. SHAP Value.

As we can see in the above plot, the most exciting part is how education level, number of children, age groups, and sex affect people's mental health problems. Here, as the education level increase, the average treatment effect will tend to increase. In other words, for people with high education levels, when these people get married, it will increase the chance that they are depressed. However, on the other hand, it is also engrossing that more children in the family will let the average treatment effect of marriage be lower, resulting in a low probability of getting depressed.

We also want to see the effect that each age group has in determining the causal effect. Since these three age groups are binary values, a high feature value indicates that these people are within these age groups. From the plot, it's clear that neither middle-aged adults nor seniors positively affect the ATE value. Nonetheless, if young adults get married, it's more likely for them to get depressed. This contradicts the result of the study [11] as they indicate that young adults who get married will be less likely to get depressed. One assumption we might have been that those young adults who got married are stressed about making bills and paying for different loans they have. As a result, it is more likely for them to get depressed.

4. Instrumental method

The above approach calculated the causal effect of marriage on depression based on existing variables; however, since it is impossible to list all the factors affecting depression in this dataset, there must be

DOI: 10.54254/2755-2721/6/20230827

some unmeasured confounders affecting both treatments and outcomes. For example, if we choose marriage as treatment, other factors, such as socio-economic factors, may influence marriage and depression. Therefore, we need instrumental variables to solve this problem.

4.1. Instrumental variables estimation

Here is an example, figure1 shows the simple graphical causal model, and here is the linear regression model to calculate the causal effect of marriage on depression.

$$\log(depression)_i = \beta_0 + \kappa\,\mathrm{marriage}_i + \beta\mathrm{status}_i + u_i$$

Clearly, Socioeconomic status is the confounder that we need to control to eliminate bias.

However, we may have the problem that we do not have any good measurement of Socioeconomic status, or we do not even know if there are any other unobserved variables that affect both marriage and depression.

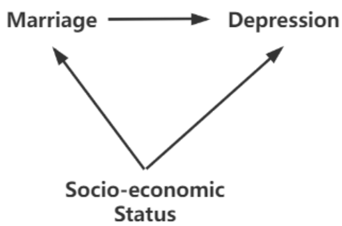

Figure 10. Socio-economic status.

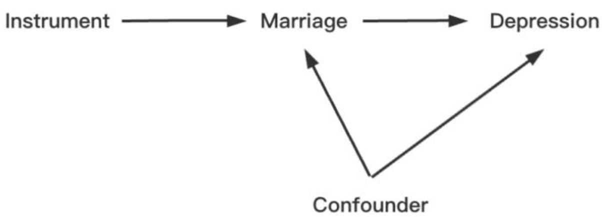

Figure 11. Marriage.

So that is why we need instrumental variable Z, which is uncorrelated with the outcome Y, but is correlated with the treatment T.

We use W for the confounders; the previous formula can be written as:

$$Y_i = \beta_0 + \kappa\,T_i + \beta W_i + u_i$$

However, in this case, we don't have data on confounder W_i, so we can use v_i to replace $\beta W_i + u_i$.

$$Y_i = \beta_0 + \kappa\,T_i + v_i$$

$$v_i = \beta W_i + u_i$$

Since the instrument Z is only correlated with the Y through T, this implies that the covariance of Z and v_i is 0, so here comes this formula:

$$Cov(Z,Y) = Cov(Z, \beta_0 + \kappa\,T_i + v_i) = \kappa Cov(Z,T) + Cov(Z,v) = \kappa Cov(Z,T)$$

Dividing each side by variance Z, and rearranging the terms, we get the causal effect.

$$\kappa = \frac{Cov(Y_i, Z_i)/V(Z_i)}{Cov(T_i, Z_i)/V(Z_i)} = \frac{\text{Reduced Form}}{\text{1st Stage}}$$

We know that covariances divided by variances are regression coefficients, so the numerator is the result from the regression of Y on Z, and the denominator is the result from the regression of T on Z, and they have a special name called Reduced Form coefficient and First Stage coefficient.

Before we use the instrument, we need to ensure it is valid. This implies arguing in favor of the two Instrumental Variables assumptions:

$Cov(Z,T) \neq 0$ This is saying that we should have a solid 1st stage or that the instrument impacts the treatment variable.

$Y \perp Z|T$ This is the exclusion restriction, stating that instrument Z only affects the outcome Y through the treatment T.

In this dataset, we consider gender to be an appropriate instrumental variable since gender affects the marriage rate of individuals, and we explain that the male-to-female ratio in a given region shows that it is unlikely to be 1:1, so that an imbalance in the ratio leads to a greater probability of marriage for one gender. Moreover, gender does not directly affect whether a person is depressed or not and is not correlated with other variables.

We calculate the covariance of treatment and iv, and the result shows that it is not zero, which means they are correlated. As for condition 2, it is not computed but depends on a priori knowledge, and in this problem, gender does not influence other factors such as education level, household income, etc. So, we consider gender as a good iv here.

4.2. Two stages least squares estimation

Generally, 2SLS is referred to as IV estimation for models with more than one instrument and only one endogenous explanatory variable [13,14]. It can be shown that IV estimation equals 2SLS estimation when there is one endogenous and one instrumental variable. Also, 2SLS can be used for models with multiple endogenous explanatory variables if we have the same number of instruments as endogenous variables. In our experiment, we will use 2SLS from the CausalML package to estimate the average treatment effect of marriage on depression.

As for the data pre-processing, we drop the rows which contain null values. And then, for the input X, we will drop the iv, treatment, and outcome. We also drop the "Survey_id" because it is useless there. Furthermore, we define the outcome variable y, treatment variable t, and instrument variable z. Finally, we directly use the API of 2SLS in CausalML to calculate the ATE; the result shows that it is -0.026.

From the 2SLS, we calculate the ATE is about -0.026, which is smaller than the result of the meta-leaner. It means that confounders indeed affect our judgments about the causality of marriage on depression. Moreover, the final result shows that in rural regions, marriage has a negative causal effect on depression, which means if you get married, you will be less likely to get depressed. However, the influence is limited because the causal effect is relatively small.

4.3. Deep IV estimation

All the above experiments are based on the assumption that the causality model is linear. To generalize it to a non-linear model, we have used the Deep IV method to explore more possibilities [15].

The Deep IV estimator will fit two models:

- Treatment model: it estimates the distribution of the treatment T given Z and X, using a mixture density network.
- Response model: It estimates the dependence of the response Y on T and X.

We have defined the neural network models and trained them. The final prediction of the ATE is 0.1296.

Although 2SLS and Deep IV give two opposite conclusions, I think we cannot hastily decide that the Deep IV approach is the right one because the problem of overfitting may occur for simple data in neural networks. Also, the instrumental variables may be a week in this dataset. Therefore, I think we still need more comprehensive data for future research.

Acknowledgment
Haoran Zhou and Junliang Lu contributed equally to this work and should be considered co-first authors.

Code availability

- Ziyue Wang: Meta learner analysis and relevant code: https://github.com/ZiyueWang675/causal_inference_depression_dataset.git
- Haoran Zhou: The code for training and evaluating the model and dealing with the data in this paper is available at: https://www.kaggle.com/code/acidd15/causal-inference/edit/run/107248417.
- Junliang Lu: Machine learning algorithm and relevant code is available at: https://github.com/Seaky-Danny/ML-for-depression/commit/110396f27ce67079cfa3c5184a7bc036f97a5545

References

[1] Gao, Wenfu; Lei, chun lei; Wang ligang, " Popularizing psychology and building a healthy China" [J].Proceedings of the Chinese Academy of Sciences, 2016, 31(11): 1187-97.
[2] NADEEM M,ALI A, BUZDAR M A. The association between muslim religiosity and young adult college students' depression, anxiety, and stress [J]. J Relig Health, 2017, 56(4): 1170-9.
[3] Wen Junna; Yang Xudong; Yang Yonghong, "Research Overview on equity of health resources allocation in China" [J]. Medicine and Philosophy (A), 2016, 37(06); 60-4.
[4] KUEHNER C. Why is depression more common among women than among men?[J]. Lancet Psychiatry, 2017, 4(2): 146-58.
[5] Luan Wenjin; Yang fan; Chuan hongli, "Self-assessment of mental health and its influencing factors among the elderly in China"[J]. 2012,42(03):75-83
[6] DeMarie-Dreblow, Darlene. "Relation between knowledge and memory: A reminder that correlation does not imply causality." Child Development 62.3 (1991): 484-498.
[7] Sander, Greenland. "May 2003 - Volume 14 - Issue 3 : Epidemiology." Journals.lww.com, 8 Oct. 2002, journals.lww.com/epidem/Fulltext/2003/05000/Data.
[8] Maclure, Malcolm, and Sebastian Schneeweiss. "Causation of bias: the episcope." Epidemiology (2001): 114-122.
[9] S. R. Künzel, J. S. Sekhon, P. J. Bickel, and B. Yu, "Meta-learners for Estimating Heterogeneous Treatment Effects using Machine Learning," Proc. Natl. Acad. Sci. U.S.A., vol. 116, no. 10, pp. 4156–4165, Mar. 2019, doi: 10.1073/pnas.1804597116.
[10] G. Ke et al., "LightGBM: A Highly Efficient Gradient Boosting Decision Tree," in Advances in Neural Information Processing Systems, 2017, vol. 30. Accessed: Oct. 09, 2022. [Online]. Available: https://proceedings.neurips.cc/paper/2017/hash/6449f44a102fde848669bdd9eb6b76fa-Abstract.html
[11] Uecker JE. Marriage and mental health among young adults. J Health Soc Behav. 2012 Mar;53(1):67-83. doi: 10.1177/0022146511419206. Epub 2012 Feb 9. PMID: 22328171; PMCID: PMC3390929.
[12] S. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions." arXiv, Nov. 24, 2017. doi: 10.48550/arXiv.1705.07874.
[13] Guo, Ruocheng, et al. "A survey of learning causality with data: Problems and methods." ACM Computing Surveys (CSUR) 53.4 (2020): 1-37.
[14] Angrist, Joshua D., and Guido W. Imbens. "Two-stage least squares estimation of average causal effects in models with variable treatment intensity." Journal of the American statistical Association 90.430 (1995): 431-442.

DOI: 10.54254/2755-2721/6/20230827

[15] Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. "Identification of causal effects using instrumental variables." Journal of the American statistical Association 91.434 (1996): 444-455.