Causation between Income and Debt in The Cost of Solar panel Installation in Switzerland

Amirreza Sokhankhosh
amirreza.sokhankhosh@gmail.com

Hasti Khajeh
hastikhajeh25@gmail.com

Negin Mohammadi
negiiinn.nm@gmail.com

Sepehr Shahsavar
sepehrshahsavar14@gmail.com

Hossein Khasteh
khasteh@kntu.ac.ir

Abstract

Not only is the equal and fair distribution of opportunities desirable, but it is also demanding. However, studies show that human decisions and actions can be biased toward various groups, which are diversified based on social status, gender, race, etc. In this paper, we assess the fair allocation of funding in a case that, given the same situation, income distribution might affect the outcome, which is financial aid. To do so, we use proposed approaches of causal inference such as matching and IPTW. Moreover, we gather a diverse dataset on Switzerland's social information to help use the said approaches.

1 Introduction

Nowadays, there is no doubt that industrialization can have significant adverse environmental impacts, which lead to climate change, loss of natural resources, air and water pollution, and extinction of species. Energy consumption, especially the energy produced from the burning of fossil fuels, is one of the essential factors that play a role in this matter. Fortunately, numerous solutions have been proposed to undermine the destructive effects of energy consumption of humankind, yet alternative energy sources are the most practical ones. Solar energy might be the most famous alternative energy source since the suggestion of finding a replacement for fossil fuels. Nonetheless, the price of installing this technology should be considered when proposed. Governments around the globe are encouraging the installation of solar panels throughout their countries by offering a proportion of overall cost as financial aid to each household so that they would be able to install the panels on their roofs. By doing so, not only can they provide their own electricity, but also, in some cases, they would be able to sell the remaining produced electricity to the government itself. Hence, in the long run, it may not be only in favor of the environment but also the house owners as well.

Countless factors are considered while predicting the amount of financial aid that the government is supposed to grant each household; Many of which are such as the size and the orientation of the roof with respect to the sun, duration of the day and the season in the location, etc. In this paper, we gathered information about the country of Switzerland, which funds each municipality separately. In the process of research, we aggregated various attributes of this country such as its roofs, income distribution, funding regarded solar-panel installation, and more so that we could assess whether this funding is fair toward different individuals with the same amount of income or not.

We consider the problem of fairness as the fair allocation of opportunities, in which various groups of society should be given the same opportunities no matter what their race, financial status, age, or gender are. In the following sections, with the help of Causal Inference, we seek to evaluate the fairness of given financial aid in Switzerland for the installation of solar panels. In our definition, the allocation of this funding would be fair, if, given the same circumstances on other attributes of

Submitted to 36th Conference on Neural Information Processing Systems (NeurIPS 2022). Do not distribute.

two households, the amount of funding for both of them is the same, even though one of which is wealthier than the other.

Moreover, we introduce similar previous works on this field and then go through a summary of two classic approaches of Causal Inference, Matching, and IPTW, which are the concepts we use to assess the fairness of the problem.

2 Related Work

We got inspired by two research papers, [2] and [5], which are presented by Yuzi He et al and Stephen Lee et all, respectively. The former, [2], proposed a causal framework that uses data subject to reach optimal intervention policies in order to reduce fairness constraints. The latter, [5], introduced DeepRoof, a data-driven approach that uses widely available satellite images to assess the solar potential of a roof. Such estimates can be used to identify ideal location and angle on a roof for installing solar panels.

Hardt et al. [6] suggested a criterion in order to optimally adjust any learned predictor to remove discrimination against a specified sensitive attribute in supervised learning. Also in our paper, we attempt to avoid discrimination against the income attribute. As a result, we both try to achieve equality of opportunity.

Moreover, there has been extensive research on propensity score methods (PSM) for causal inference. Mingxiang Li [7] reviewed PSM literature and provided a procedure for implementing the PSM in different management fields. The potential application of it was further discussed.

Linden [8] used marginal mean weighting through stratification (MMWS), a technique which contains three typical experimental conditions: a binary treatment, an ordinal level treatment, and nominal treatments to remove imbalances. While utilizing such technique, he studied the health care evaluations by examining the pre-post difference in hospitalizations for patients with congestive heart failure following the implementation of a disease management program.

Fan Li [9] proposed a unified propensity score weighting framework for estimating causal effects with multiple treatments. Riegg [10] calls attention to the problem of omitted variable bias (OVB) in research on the causal effect of financial aid on college-going. Shiba et al. [11] prepared an overview of causal inference from observational data, matching and inverse probability weighting which are two major propensity score (PS) based methods. Moreover, they introduced common pitfalls and tips for applying the PS methods. Sekhon [12] provided a review which focused on matching methods for estimating causal effects using observational data. Robins [13] compared the strengths and weaknesses of marginal structural models (MSMs) versus structural nested models (SNMs) as tools for causal inference.

Recently, there has been a growing literature of fairness in decision making. There are case studies on social work and health care policy such as, Corbett-Davies et al. [14] formulate the fair decision making as an optimization problem whereas some limitations of fairness are considered. Chouldechova et al. [15] worked on developing, confirming, fairness inspecting, and deploying a risk prediction model and discussed the result of their analysis.

3 Methodology

We briefly review matching and inverse probability of treatment weighting (IPTW), which are the algorithms we use in this paper, and then move on to the data and results.

3.1 Matching

Matched sampling is a method of data collection and organization designed to reduce bias and increase precision in observational studies, i.e. in those studies in which the random assignment of treatments to units (subjects) is absent [1]. In our data-set, in which treatment is not randomly assigned within features between subjects, we cast matching sampling to reduce bias and eliminate the effect of confounders.

To expand it in detail, suppose we are given N observations indexed with $i = 1, 2, ..., N$, consisting of tuples of data of the form of (X_i, y_i^{obs}, t_i). Here X denotes features of the observation, y^{obs} is the observed outcome, and binary variable t indicates whether the observation came from the treated group ($t = 1$), or the control ($t = 0$). We assume that each observation i has two potential outcomes: the controlled outcome $y_i^{(0)}$ and the treated outcome $y_i^{(1)}$, but we only observe one outcome $y^{obs} = y_i^{(t_i)}$. In addition, we assume that given features X, both of the potential outcomes $y^{(0)}$, $y^{(1)}$ are independent of treatment assignment t [2]. This condition is called the ignorability assumption.

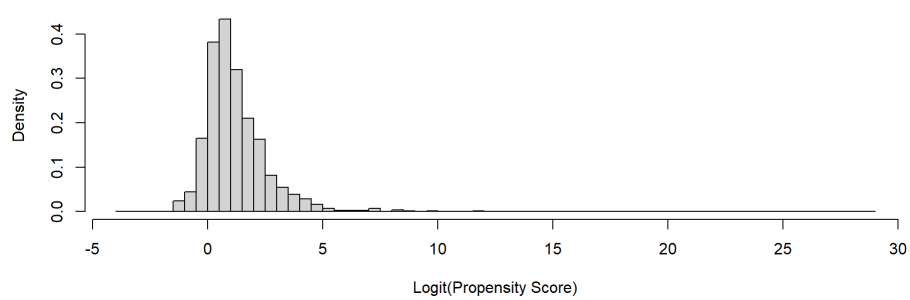
$$(y^{(0)}, y^{(1)}) \perp t|X$$

The goal of matching is to match treated subjects and control subjects on the set of covariates that we identified as sufficient to control for confounding, which we discuss in the Data section. Instead of matching in a greedy approach, we use a value called the propensity score.The propensity score was first defined by Rosenbaum and Rubin in 1983 as 'the conditional probability of assignment to a particular treatment given a vector of observed covariates' [3; 4]. The propensity score gives the probability (ranging from 0 to 1) of an individual being exposed (i.e. assigned to the intervention or risk factor) given their baseline characteristics. The aim of the propensity score in observational research is to control for measured confounders by achieving balance in characteristics between exposed and unexposed groups [4]. To put it simply, the propensity score is simply the probability of receiving treatment, given covariates. If two subjects have the same propensity score, they are just as likely to be found in the treatment group, so if we restrict to a subpopulation of subjects who have the same propensity score, there should be a balance in the two treatment groups, considering we assumed ignorability (that treatment is randomized given X).

In a randomized trial we have $P(A = 1|X) = P(A = 1) = 0.5$, but in most cases, our trials are not or cannot be randomized, thus we estimate $P(A = 1|X)$ using logistic regression. We first fit a model using outcome and covariates, then from that model, we get the predicted probability of each subject.

The propensity score π is scalar - each subject has only one propensity score, hence the matching problem is simplified, in that we are only matching on one variable. In this paper, the logit (log-odds) of the propensity score, $logit(\pi)$, is used because it is unbounded, and it differentiates values more vividly than the propensity score itself. This contrast can be seen in figure 1.

Figure 1: Logit of Fitted Propensity Score



To ensure that we are not accepting any bad matches, a $caliper$ can be used. A caliper is the maximum distance we are willing to tolerate between 2 subjects in a match. The distance between observations i, j within the covariate X given that S is the covariant of X, can be computed as:

$$D(x_i, x_j) = \sqrt{(x_i - x_j)^T.S^{-1}.(x_i - x_j)}$$

3.2 Inverse Probability of Treatment Weighting

Despite its various uses in practical studies, the Propensity Score Matching (PSM) has some flaws. PSM creates a subdataset of matches, based on the propensity score of each subject in the original dataset, and then discard the rest; hence we end up with a small proportion of the original data representing all of it. Doing so eliminates some features that might have changed the difference of potential outcomes, which is the value we are interested in.

Inverse Probability of Treatment Weighting (IPTW) is an alternative method for estimating causal inference. In short, IPTW involves two main steps. First, the probability—or propensity—of being exposed, given an individual's characteristics, is calculated. This is also called the propensity score. Second, weights for each individual are calculated as the inverse of the probability of receiving its actual exposure level. The application of these weights to the study population creates a pseudopopulation in which measured confounders are equally distributed across groups [4]. In contrast to matching treated and control subjects on a select group of confounders, the IPTW approach uses the entire dataset and can address a very large number of confounding variables. Each subject in the dataset is assigned a weight based on the likelihood that they would be exposed to the treatment effect under investigation. Applying this weight when conducting statistical tests or regression models reduces or removes the impact of confounders. In this paper, subject i is assigned to a weight w_i corresponding to 1 over the probability of treatment, which is the propensity score π_i.

$$w_i = 1/\pi_i$$

Applying these weights creates a pseudo-population where everyone is equally likely to be treated regardless of their X value. Figure 2 shows the difference of Standard Mean Differences value in the both weighted and raw datasets. It is visible that the weighted dataset is extremely balanced, since the SMD value of its features are less than 0.05, where as there is a visual imbalance in the original dataset, especially in features like Area, and Generated Electricity.

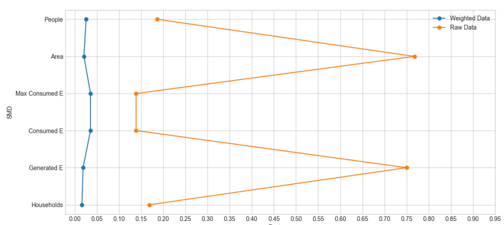


Figure 2: Standard Mean Differences of Features Plot

	Raw Dataset			Weighted Dataset		
	Treatment			Treatment		
	0	1		0	1	
n	583	1651	SMD	2264.95	2227.21	SMD
Households (mean (SD))	0.29 (0.45)	0.21 (0.41)	0.168	0.23 (0.42)	0.23 (0.42)	0.016
Generated E (mean (SD))	0.16 (0.37)	0.49 (0.50)	0.75	0.42 (0.49)	0.41 (0.49)	0.019
Consumed E (mean (SD))	0.53 (0.50)	0.46 (0.50)	0.139	0.49 (0.50)	0.48 (0.50)	0.036
Max Consumed E (mean (SD))	0.53 (0.50)	0.46 (0.50)	0.139	0.49 (0.50)	0.48 (0.50)	0.036
Area (mean (SD))	0.15 (0.36)	0.49 (0.50)	0.767	0.41 (0.49)	0.40 (0.49)	0.021
People (mean (SD))	0.08 (0.26)	0.03 (0.18)	0.187	0.03 (0.18)	0.04 (0.19)	0.026

Figure 3: Standard Mean Differences of Features Table

3.2.1 Marginal Structural Models

In comparison to matching, the IPTW approach can be used to estimate parameters from more general and complicated causal models. These models can be an option when the study is not just interested in the average cause or effect. Marginal structural models (MSMs) are a new class of causal models for the estimation, from observational data, of the causal effect of a time-dependent exposure in the presence of time-dependent covariates that may be simultaneously confounders and intermediate variables.

Suppose X_i is a single variable of the observation i that modifies the effect of treatment $A = a$. A linear MSM model with effect modification can be formulated as below.

$$E(Y^a|X_i) = \psi_0 + \psi_1 a + \psi_2 X_i + \psi_3 X_i a\,; \qquad a = 0, 1$$

According to the formula, if we compute parameters of the MSM model, which are $\psi_i \; for \; i = 0 \; to \; 3$, we could evaluate the mean of potential outcomes for a given value of X_i.

$$E(Y^1|X_i) - E(Y^0|X_i) = \psi_1 + \psi_3 X_i$$

A more generalized formulated of the MSM can be written as the following in which $g()$ represents a link function and $h()$ is a function specifying parametric linear form of a and X_i.

$$g\{E(Y^a|V)\} = h(h, X_i; \psi)$$

In conclusion of the IPTW approach, in this paper, we first estimate propensity scores, then we apply weights to the dataset. After creating balance within features of the data in hand, we specify the MSM of interest and use the programming language R to fit a weighted generalized linear model.

Before we move on to the results we get from these two approaches, we explaining the dataset we gather and the features we consider in this paper.

4 Data

To build a model that can provide us with an answer to the problem of causality, we have to prepare a sufficient amount of data to be able to train it. The majority of this data is gathered from different sources and brought together and maintained by us. In this paper, the data we possess contains solar-panel installation features and economic attributes of each household within various block groups, such as *Cantons*, *Districts*, and *Municipalities* of Switzerland. This data holds 26 Cantons, 148 Districts, and 2322 Municipalities that each has 15 features. The features of each block group include *Name*, *Geometry*, *Base Funding*, *Collector Promotion*, along with *Average Income*, *Number*, *Average Generated Electricity*, *Average* and *Maximum Electricity Consumption*, *Area*, and *Population* of **Households**. In addition *Total Funding*, *Total Price*, *Debt*, and *Debt Rate* are calculated.

In detail, *Geometry* is the shape of each block group; *Base Funding* is the base amount of funding that the government will provide each household to install solar-panels; *Collector Promotion* is the amount of funding which will be given to each household per each panel.

Based on the data presented *here*, for manufacturer **3S Solar Plus AG**, and Model **Mega Slate**[1], we gathered values of *Base Funding* and *Collector Promotion*. Based on this website funding for some cantons are not available at the moment, hence these values were put 0 in the data-set. According to these two features, and considering the area of each panel to be $1.64m^2$, we have:

$$Total\ Funding = Base\ Funding + (Area\ /\ 1.64)\ .\ Collector\ Promotion$$

Just as *Total Funding*, *Total Price* also can be calculated considering the price of each panel, assumed to be 154 CHF, and the number of needed panels.

$$Total\ Price = (Area\ /\ 1.64)\ .\ 154$$

Simply put, *Debt* is the difference of *Total Price* and *Total Funding*. Same goes for *Debt Rate* which is *Debt* over *Total Price*.

$$Debt = Total\ Price - Total\ Funding$$
$$Debt\ Rate = Debt / Total\ Price$$

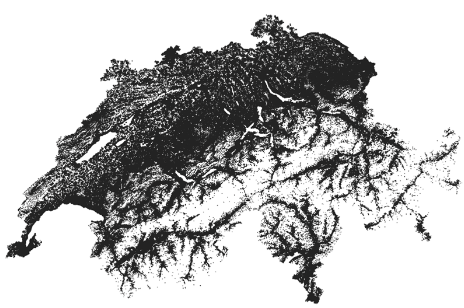


Figure 4: Roofs

Figure 5: Households

Features such as *Geometry*, *Average Generated Electricity*, and *Area* per household are deducted from *Maps of Switzerland - Swiss Confederation*. This interface is an online website full of various maps about Switzerland. The data we select from this interface is *Solar Energy - Suitability of Roofs* , which contains roofs of Switzerland with their suitability for solar-panel installation categorized from 1 to 6. This data also provides area, generated electricity, and geometry of each roof. With respect to the geometry of each block group, we assign each roof to a block group, whether canton, district, or municipality, and calculate the average of the listed features per household. Hence, we end up with Average Generated Electricity and Area per household for each block group.

[1]TKN for Collector and TKN of the plant are 0.123 kW. For more information see *here*

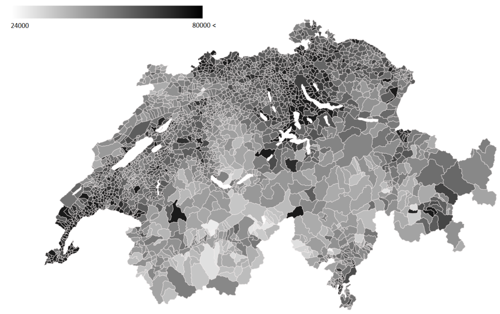


Figure 6: Income

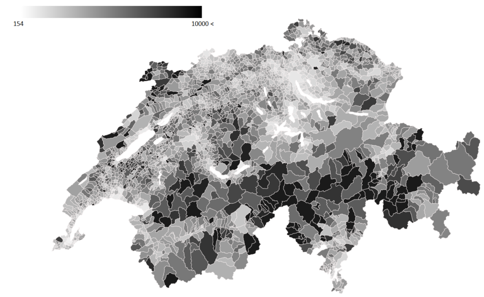


Figure 7: Generated Electricity

The income data for each house, meaning the income of the people living in a house, is derived from *here*. This site specifies statistics on direct federal tax (DBST), broken down by natural and legal persons. In this paper, the average income section in each block group has been used. The number of houses in each block group is extracted from *Permanent resident population*. To sum up, using the number of houses and the population of each of them, the average family population in each block group has been calculated.

As it is apparent from Figures 6 and 7, as *Income* in one region (block group) increases, *Area* and *Generated Electricity*, which are highly correlated with each other due to the formula that they are based on, decrease. Although, the subjects are not randomly assigned within features of the data-set, and we cannot deduct **causation** from **correlation**. Thus, as we proceed in the result section, we calculate the causal inference value to ensure that there does not exist any bias in our data-set.

5 Results

In this paper, we assume that the treatment is the income of each municipality in our data, and to give it a binary value we define a threshold ,th, which differs in a range of 48,000 CHF to 98,000 CHF. This numbers are not random, but they are 20,000 CHF greater and less than the average income of each municipality in our data. Based on th who ever has an income value greater than th is considered "Rich" and vice versa. The outcome of our study can be either the $DebtRate$ or $Debt$.

According to the ignorability assumption, given features X, both of the potential outcomes $y^{(0)}$, $y^{(1)}$ are independent of treatment assignment t. To demonstrate this assumption, in causal inference studies, we evaluate some features considering the fact that they might be confounders. In this paper, our features are *Households, Generated Electricity, Population*, etc which we discussed in the Data section in explicit details. Concerning these features, we continue computing the causal inference value using the approaches that we went over in the Methodology section.

5.1 Matching

As it was mentioned before, the goal of matching is to match the treated subjects and the control on the set of covariates which might be considered confounders. After fitting a regression model to predict the propensity score of each subject, we match with a caliper 0.1 to ensure we are not considering bad matches. Figure 4 shows the distribution of features between the control and the treated in the matched subjects. As it can be seen, the closeness of plots between the two groups in each of the features is an evidence on behalf of good matching which is what we are after.

After matching on the confounders, we can compute the causal inference value just by calculating the mean differences between the treated and control groups. We loop through the threshold th to manipulate rich and poor groups, and then compute the causal inference value.

It is shown in the figure 5 that there is no correlation between income and debt, since the value of difference between the two groups is less than 0.03 at max, but it is also evident from the figure 6 that the p-value of this study is greater than usual and is about 0.05. The large number of p-values encourages debates on whether we can rely upon the outcome or the correct work of the model. This means that the output cannot be trusted, and the model did not work as planned.

Figure 8: Features of each Matched Subjects

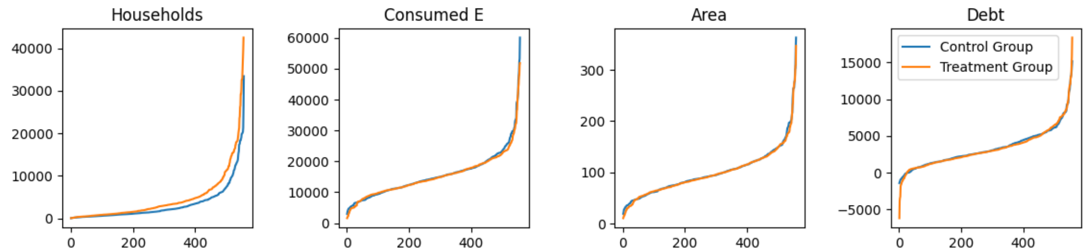


Figure 9: Causality of Income over Debt within Swiss Municipalities (Matching)

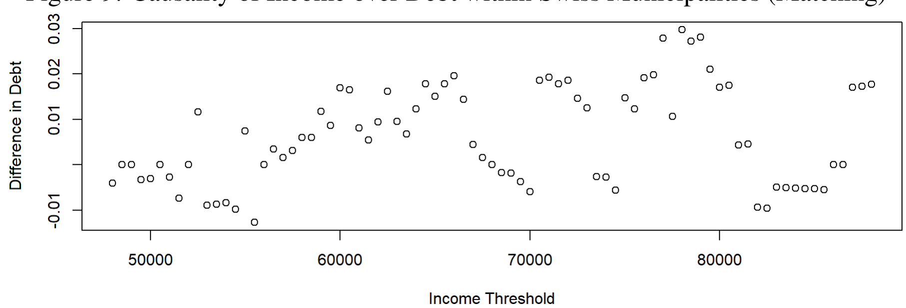
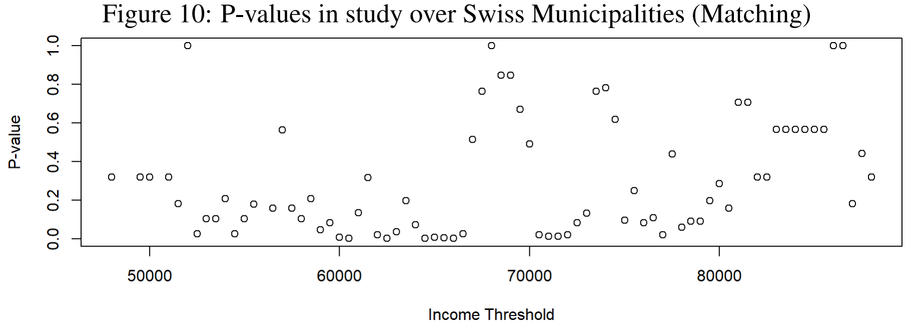


Figure 10: P-values in study over Swiss Municipalities (Matching)



A potential reason for this outcome can be the underlying argument against matching, which is that it discards the data that cannot be matched to any subjects in either the control, or the treated group, hence lots of information might not be considered in the model, which could have changed the answer, and that brings us to the alternative approach, IPTW.

5.2 IPTW and MSM

Based on the algorithm we discussed in the Methodology section, we apply a weight to each subject so the overall balance would be achieved and the effect of each covariate could be overlooked. As Figure 2 and 3 shows this approach works splendidly, because the SMD of each feature is less than 0.05, and that is good sign of balance.

After applying these weights, we fit a marginal structural model, and then compute the value of causal inference between income being the treatment and debt or debt rate being the outcome. Figure 8 and 9 show the causality of income over debt and debt rate. In case of debt rate, the values starts at above 34 percent and as the income threshold gets greater, it lowers with it. Figure 9 shows the p-value of this study, which is calculated for each income threshold. Because this value is small, the outcome we get from the IPTW approach can be more trusted in comparison to matching. These results provide confirmatory evidence that supports the hypothesis of there being a causation relationship between Income and Debt in the gathered data.

Even though we gathered our data after a broad range of search, the causal inference outcome might differ if we consider other covariates in the study.

Figure 11: Causality of Income over Debt within Swiss Municipalities (IPTW)

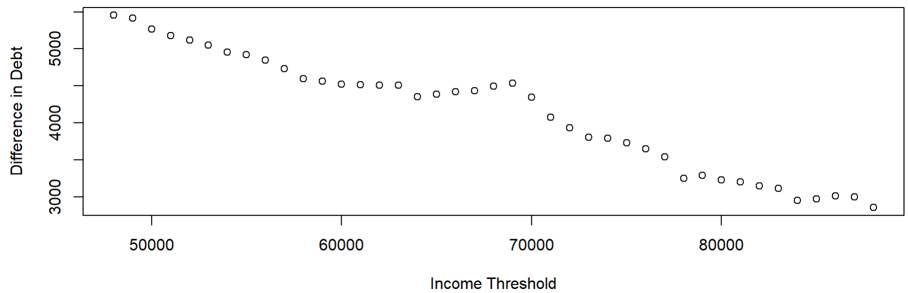


Figure 12: Causality of Income over Debt Rate within Swiss Municipalities (IPTW)

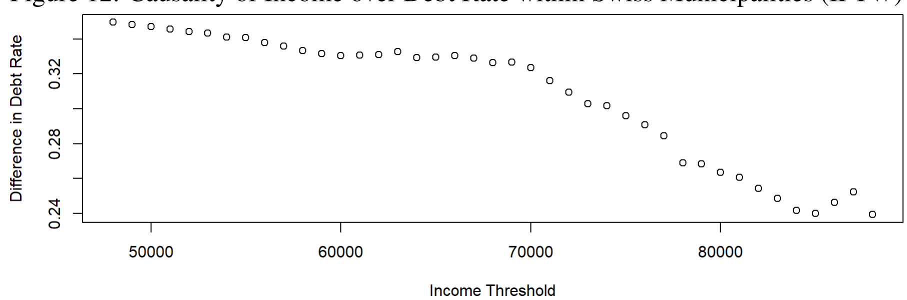
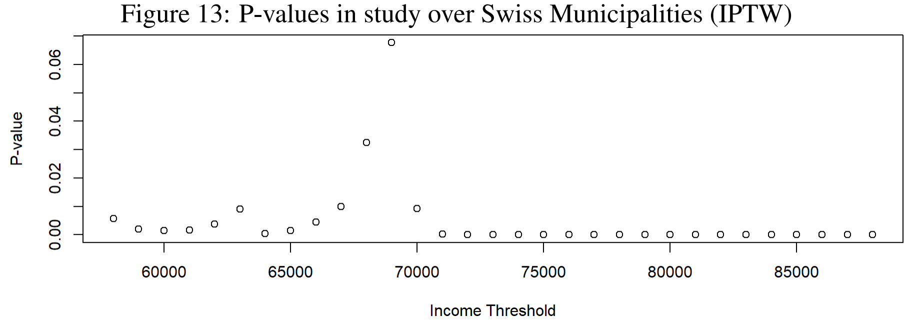


Figure 13: P-values in study over Swiss Municipalities (IPTW)



6 Future Works

In this paper, we answer the question as to whether bias exist in governmental financial aids or not by aggregating attributes of all municipalities of Switzerland in a database and formulating the problem as a Causal Inference problem. Moreover, we compare the outcome of two approaches in Causal Inference and derive a conclusion on them.

While this methodology facilitates policy-makers in their decisions, our work still has imperfections. In the study of Causal Inference, no one can be sure that existing covariates are the only factors that play roles in the outcome. Meaning that there might be other influential attributes to this matter that we do not consider.

Furthermore, numerous methods have been proposed in the field of Machine Learning which can be utilized in Causal Inference problems. The propensity score can be predicted using a trained Neural Network which may solve the problem more efficiently.

Lastly, our data is limited to the country of Switzerland; however, interested researches in this field may be able to find the same attributes of separate countries and gather a far larger database than us. In this manner they can come up with a stronger confident over the results of the models.

References

[1] Rubin, D. (2006). Matching to Remove Bias in Observational Studies. In Matched Sampling for Causal Effects (pp. 62-80). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511810725.007

[2] He, Y., Burghardt, K., Guo, S., Lerman, K. (2020). Inherent Trade-offs in the Fair Allocation of Treatments. arXiv. https://doi.org/10.48550/arXiv.2010.16409

[3] PAUL R. ROSENBAUM, DONALD B. RUBIN, The central role of the propensity score in observational studies for causal effects, Biometrika, Volume 70, Issue 1, April 1983, Pages 41–55, https://doi.org/10.1093/biomet/70.1.41

[4] Nicholas C Chesnaye, Vianda S Stel, Giovanni Tripepi, Friedo W Dekker, Edouard L Fu, Carmine Zoccali, Kitty J Jager, An introduction to inverse probability of treatment weighting in observational research, Clinical Kidney Journal, Volume 15, Issue 1, January 2022, Pages 14–20, https://doi.org/10.1093/ckj/sfab158

[5] Stephen Lee, Srinivasan Iyengar, Menghong Feng, Prashant Shenoy, and Subhransu Maji. 2019. DeepRoof: A Data-driven Approach For Solar Potential Estimation Using Rooftop Imagery. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery Data Mining (KDD '19). Association for Computing Machinery, New York, NY, USA, 2105–2113. https://doi.org/10.1145/3292500.3330741

[6] Hardt, M., Price, E., Srebro, N. (2016). Equality of Opportunity in Supervised Learning. arXiv. https://doi.org/10.48550/arXiv.1610.02413

[7] Li, M. (2012). Using the Propensity Score Method to Estimate Causal Effects. Organizational Research Methods. https://doi.org/10.1177/1094428112447816

[8] Linden A. Combining propensity score-based stratification and weighting to improve causal inference in the evaluation of health care interventions. J Eval Clin Pract. 2014 Dec;20(6):1065-71. doi: 10.1111/jep.12254. Epub 2014 Sep 29. PMID: 25266868.

[9] Li, F., Li, F. (2018). Propensity Score Weighting for Causal Inference with Multiple Treatments. arXiv. https://doi.org/10.48550/arXiv.1808.05339

[10] Cellini, Stephanie. (2007). Causal inference and omitted variable bias in financial aid research: Assessing solutions. Review of Higher Education, The. 31.

[11] Shiba, K., Kawahara, T. (2021). Using Propensity Scores for Causal Inference: Pitfalls and Tips. Journal of epidemiology, 31(8), 457–463. https://doi.org/10.2188/jea.JE20210145

[12] Jasjeet S. Sekhon; Opiates for the Matches: Matching Methods for Causal Inference; 2009

[13] Robins, J.M. (2000). Marginal Structural Models versus Structural nested Models as Tools for Causal inference. In: Halloran, M.E., Berry, D. (eds) Statistical Models in Epidemiology, the Environment, and Clinical Trials. The IMA Volumes in Mathematics and its Applications, vol 116. Springer, New York, NY. https://doi.org/10.1007/978-1-4612-1284-3_2

[14] Pierson, E., Feller, A., Goel, S., Huq, A. (2017). Algorithmic decision making and the cost of fairness. arXiv. https://doi.org/10.1145/3097983.309809

[15] Chouldechova, A., Benavides-Prado, D., Fialko, O. amp; Vaithianathan, R.. (2018). A case study of algorithm-assisted decision making in child maltreatment hotline screening decisions. Proceedings of the 1st Conference on Fairness, Accountability and Transparency, Proceedings of Machine Learning Research 81:134-148 Available from https://proceedings.mlr.press/v81/chouldechova18a.html.