Impact Evaluation of Burundi's 2018 Child Benefit Scheme

Kathleen Hopewell

ID: i6312124

United Nations University-MERIT/
School of Business and Economics, Maastricht University

MPP4504 Public Policy Analysis

Eleonora Nillesen

Iman Rajabzadeh

22 December 2022

Table of Contents

Impact Evaluation of Burundi's 2018 Child Benefit Scheme

As of 2016, Burundi had a national Gross Domestic Product (GDP) per capita of 263 U.S. dollars, which placed it as the world's second poorest country (Haidara, 2012). The multidimensionality of economic depravity in Burundi is further underscored by the reality that it was also ranked 184 out of the world's total 188 countries in the 2016 human development index (Sibanda & Villararribas, 2018), making it also one of the world's most developmentally-hindered nations. It is a land-locked, sub-Saharan nation that has long been subjected to the harms of post-Cold War conflicts (Mercier, Ngenzebuke & Verwimp, 2020). The presence of conflict not only initially injures national development, economic security, human capital and the preservation of institutions, but injures them in such a way that can lead to continual degradation of these factors and instigate persistent conflict (Mercier, Ngenzebuke & Verwimp, 2020).

The damages affiliated with poverty extend beyond monetary depravity, but lend themselves, in the longer term, to impede the ability of individuals to meet their basic needs in a sustainable manner, among these needs being food, access to education and access to basic services like sanitary resources and healthcare (United Nations DESA, n.d.). The nature of poverty in Burundi, where the larger part of the nation—the World Food Programme estimated as high as 65 percent—lives below the poverty line, becomes even more dire when acknowledging that upwards of 45 percent of the nation's population is under only 15 years of age (Sibanda & Villararribas, 2018, pp. 1-2). Applying the multidimensional nature of poverty, which necessarily translates into minimized access to basic goods for individuals to meet their most basic needs, to the context of Burundi's extremely young population highlights the need for any poverty reduction strategies to prioritize the needs of Burundi's young, poor demographics.

In line with these notions, UNICEF has implemented a poverty reduction program with the alleviation of the afflictions of poverty on Burundi's children as its generalized core objective. This strategy employed a targeted, rather than universal, rollout of benefits through two main steps to address those most vulnerable to poverty in Burundi: first, a proxy means test to identify the households that are the most poor based on the assets they own or to which they have access/easier access, and second, a categorical targeting strategy that selects households that have children who are under 19 years of age and in which the head of the household is either unemployed or is employed by the nation's informal economy. For the households that did qualify for the program's benefits, based on these two targeting methods, UNICEF distributed cash transfers at a flat rate, and the value of transfers were not dependent upon household size. On average, the transfers measured out to be 15% of the baseline household consumption within the nation.

The issue that remains at hand is whether or not this poverty reduction program actually served to alleviate some of the pressures associated with poverty, among those households that received the treatment, and if so, the extent to which it achieved this alleviation. Further, there is a designated interest in the specific outcomes that this programme had on: food expenditure, food security, education performances of participating households' youngest children and the mental health wellbeing of households' youngest children. These status of these, according to Burundi's government, can serve to indicate whether this program helped to benefit the grander wellbeing of children in Burundi, and further, whether it is a worthwhile program to continue and expand in order to ultimately diminish the harms of poverty among Burundi's children. Each insight into the quality of poverty reduction programs is beneficial to the overall effort to counter poverty and all of its afflictions at large by demonstrating how and where future programs can improve.

Data & Methodology

In order to perform this assessment, Burundi's government provided a dataset collected by their own Comité National de L'Information Statistique (CNIS), or the National Committee on Statistical Information, to be utilized in the evaluation of their targeted child welfare intervention (UN DESA Statistics Division, 2022).It is a panel dataset which offers insight, at the household level, into the sample population via indicators and characteristics of social and economic well-being at both pre-intervention and end-line points in time.

These variables, however, do not offer clear-cut measurements of the evaluation's outcomes of interest; that is, again, food expenditure, food security, education performances of participating households' youngest children and the mental health wellbeing of households' youngest children. It is a convention to utilize proxy measures in assessments of poverty, and especially of food security, due to the abstract nature of these issues, which renders them inherently immeasurable on their own (World Food Programme, 2009, p. 27). The dataset does offer variables which can be used as proxy measurements of the indicators, and these are: total consumption of the household to measure food expenditure and security, the test scores of the household's youngest child to measure educational performance and a hope index of households' youngest children to measure the mental health wellbeing of the youngest child.

This dataset provided by the government, although sparse, does contain sufficient information to perform an evaluation of the program's impact through the modality of a quasi-experimental method. Among the three possible modalities of such an evaluation, this evaluation employs the propensity score matching methodology of evaluation due to having variables at both points in time and due to having more baseline characteristics than endline, which facilitates the matching of treated and non-treated households which are included in the sample survey. In short, propensity score matching is an "algorithm" (Essama-Nssah, 2006, p 5) which matches individuals in the treatment group of the intervention with non-treated individuals, and calculates the probability of the individuals being selected into program participation dependent upon observable characteristics which are measured at the baseline of the program. Ideally, the matching should yield an unbiased estimation of the average treatment effect on the treated individuals because the effect is, again, ideally, dependent upon those baseline characteristics. Additionally, given the nature of the program's targeting strategy, which is categorical and, therefore, not at-random, there is a need to employ a methodology that can still derive a certain level of causality between the program and its impacts without randomization (Essama-Nssah, 2006). Further, because the treatment selection is not completely random, there is a need to recognize the characteristics that are similar among those who are selected for treatment to understand more fully the true effects of the program (Essama-Nssah, 2006). Propensity score matching is a helpful technique of non-experimental impact evaluation which fits these constraints at hand.

In order to conduct this evaluation, there are three core steps. First, a binary model of program participation must be established; this means to select the characteristics, among the total amount of those provided, that offer information on baseline characteristics of the sample which are determinants of program participation. These characteristics, too, will be the guides for matching treated and non-treated households. The characteristics used in this instance were: the age of the household's youngest child, a binary indicator of whether or not the household's parents are alive, a binary indicator of whether the household is located in a rural area at the baseline, an asset index measurement from the baseline, the household size at the baseline, the score of the youngest child's math test at the baseline and the household's total consumption at the baseline. There is a need to not include too many variables in this step, otherwise the heightened dimensionality of potential selection makes difficult the ability to match treated with non-treated individuals due to overspecification. Second, the region of common support must be defined, and the binary model of program participation must be assessed for balance, as balance is a necessary requirement for a substantive and accurate evaluation. The region of common support refers to the overlap of distributions of propensity scores, which are assigned via the binary model, among groups of treated and non-treated households.

The region of common support, too, speaks to the similarity of these groups, and for this purpose, the more similarities between groups, the better, as it provides a better insight into the program's impact. Thirdly is matching treated and non-treated groups, utilizing one of four methods: nearest-neighbor matching, radius matching, stratification matching and kernel matching. One of these methods is selected, and the remaining three are also performed later on for robustness checks to ensure that the results of the evaluation are reliable.

Results & Analysis

Applying the methodology of the propensity score matching technique to this intervention offers more specific insights into how the treatment of these cash transfers on the treatment group, based on the characteristics that inclined them to be selected for treatment.

Descriptive Statistics

The dataset has a total of 3,632 observations, where each observation is household level data. Of these, 490 households received treatment and 3,142 did not. As Table 1 indicates, those in the treated group have a higher amount of consumption on average than those in the non-treated group, and the overall consumption average at the baseline. This, however, does not give perfect insight into the poverty of these households because they may have higher consumption due to other factors, like having more members in the household. These same findings hold up at the endline, as well. This suggests that Burundi's poor may not necessarily consume less.

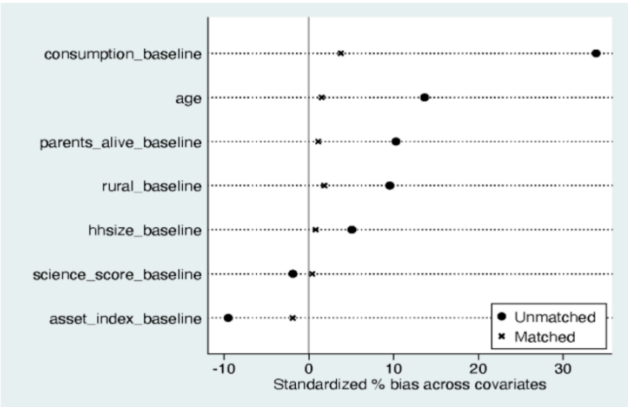
Table 1: Mean and Median Consumption at Baseline, by Treatment Status

	Observations	Mean	Median
Not Treated	3,142	409,538	409,611
Treated	490	410,242.90	410,257
Total	3,632	409,633.39	409,703

Table 2: Mean and Median Consumption at Endline, by Treatment Status

	Observations	Mean	Median
Not Treated	3,142	410,098	410,120
Treated	490	411,101	411,159
Total	3,632	410,232.95	410,285

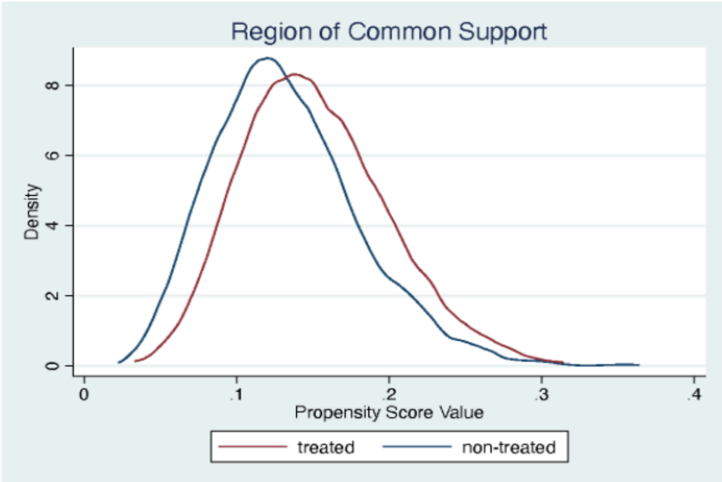
Figure 1: Standardized % of Bias Across Covariates Before and After Matching



The seemingly counterintuitive nature of the above descriptive statistics of the treated and non-treated groups' consumption is also captured in Figure 1, which demonstrates how biased the covariates are, before and after matching treated with non-treated households. The consumption baseline variable maintains a substantially, comparatively large amount of bias prior to matching, meaning that using consumption on its own as a condition of selection for treatment would not yield an unbiased estimate of the probability that a household is selected for treatment. In other words, this means that consumption at the baseline is not, on its own, truly indicative of the nature of poverty as it exists in this context. Thus, the other six covariates were included so as to reduce the bias of selection which would have occurred if consumption at the baseline was used on its own.

After matching treated and non-treated households, the binary model of treatment yielded a region of common support from propensity score .03286087 to propensity score .31447134. This means that the region of common support, using these variables for the model, captured 3,627 of the households in the sample and left only 5 households out of it. This overlap is visualized in This indicates that, based on these variables for matching, the treated and non-treated groups are heavily similar. This ultimately benefits the analysis due to the goal of it being that the effects of the treatment are measured as an outcome of the observable characteristics rather than on receiving the treatment alone.

Figure 2: Kernel Density Lines Region of Common Support



Robustness Check

As mentioned in an earlier section, there is a need to attempt multiple modes of matching to ensure that the evaluation technique yields robust results regarding the average treatment effect on the treated group. Addressing the first outcome of concern, food expenditure, the consumption endline variable was used as a proxy in assessment. As demonstrated in Table 3, there was a statistically significant effect of the program on the treated households' food expenditure, indicated through the statistically significant t-values across all four of the matching techniques, at the 1% significance level. Because the same indicator was utilized for the food security outcome, it yielded the same results, as demonstrated in Table 4.

Table 3 : Propensity Score Matching, Testing Impact on Food Expenditure

	N. Treated	N. Control	ATT	Std. Error	T-Value
Nearest Neighbor Method	490	439	533.472	133.192	4.005
Stratification Method	487	3137	466.239	66.763	6.983
Radius Method	490	3137	933.918	94.783	9.853
Kernel Method	490	3137	798.413	89.416	8.929

Table 4 : Propensity Score Matching, Testing Impact on Food Security

	N. Treated	N. Control	ATT	Std. Error	T-Value
Nearest Neighbor Method	490	439	533.472	133.192	4.005
Stratification Method	487	3137	466.239	66.763	6.983
Radius Method	490	3137	933.918	94.783	9.853
Kernel Method	490	3137	798.413	89.416	8.929

These outputs signify that, among the treated group, the program had an average treatment effect within the range of 533.472 to 933.918; this means that, based on the nearest neighbor mode of matching, for those who were treated, their consumption increased, on average, by 533.472 units.

In terms of the educational outcome, the treatment did not have as great of an effect on the educational performance as it did on the food outcomes. The variable used in this instance was the endline science test score of the youngest child, as a proxy of education performance, and among the treated group, had a mean of -.01734475 , which speaks to why the average treatment effects in this output are so small.

Table 5 **:** Propensity Score Matching, Testing Impact on School Performance of Youngest Child

	N. Treated	N. Control	ATT	Std. Error	T-Value
Nearest Neighbor Method	490	439	0.051	0.029	1.772*
Stratification Method	487	3137	0.066	0.021	3.181***
Radius Method	2	3135	-0.303	0.071	-4.282***
Kernel Method	490	3137	0.065	0.023	2.876***

All of these average treatment effects are statistically significant, as indicated by the t-values which are all statistically significant at the 1% level. This output can be interpreted to mean that, for those that received treatment, using the nearest neighbor matching technique, their youngest child's test score increased, on average, by 0.051 units. Based on the average test score of the treated group, this seems to be a significant benefit of the treatment.

Finally, the outcome regarding the mental health well-being of the youngest child demonstrates that there was a statistically significant effect on the hope index of the youngest child at the endline by the treatment, indicated by the t-values which are all statistically significant at the 1% significance level.

Table 6 **:** Propensity Score Matching, Testing Impact on Mental Health Wellbeing of Youngest Child

	N. Treated	N. Control	ATT	Std. Error	T-Value
Nearest Neighbor Method	490	439	0.307	0.067	4.601***
Stratification Method	490	3137	0.321	0.046	7.029***
Radius Method	490	3137	0.283	0.046	6.168***
Kernel Method	490	3137	0.294	0.047	6.273***

Utilizing the output from the nearest neighbor matching technique, for those who received treatment, their youngest child's hope index value increased, on average, by 0.307 units. This is regarding a variable that ranged from 2.099 to 7.672 with a mean of 4.2703, so it might not be as great of an impact as on above outcomes, but it is still significant within the context of that variable's measurement.

Heterogeneous Impact

Overall, this evaluation also demonstrated that the average effect treatment had on the treated group was different among different demographics of the treated population, most namely being those households that were treated and were located in rural and non-rural areas at the baseline and those that were treated and had heads that were unemployed or not unemployed.

In Table 7 the difference among those who were treated and were located in rural areas is demonstrated, across the four outcomes of interest in this analysis. Though many of the figures were not statistically significant, the outcomes in food expenditure, food security and mental health wellbeing of the youngest child all indicate, with statistical significance at

Table 7 : Comparing Treatment Effects Among Households in Rural and Non-Rural Locations at Baseline

	N. Treated	N. Control	ATT	Std. Error	T-Value
Food Expenditure (Rural)	378	337	550.469	147.542	3.731***
Food Expenditure (Non Rural)	112	97	344.734	290.661	1.186
Food Security (Rural)	378	337	550.469	147.542	3.731***
Food Security (Non Rural)	112	97	344.734	290.661	1.186
School Performance of Youngest Child (Rural)	378	337	0.021	0.033	0.617
School Performance of Youngest Child (Non Rural)	112	97	-0.051	0.064	-0.803
Mental Health of Youngest Child (Rural)	378	337	0.212	0.080	2.649***
Mental Health of Youngest Child (Non Rural)	112	97	0.178	0.125	1.427

the 1% significance level that those households in rural areas reaped more benefits from the treatment than those whose households were not located in rural areas at the baseline.

This difference among groups, too, is present in assessing the average treatment effect on the treated group among those households whose heads were unemployed at the baseline versus those that were not unemployed. Here, again, it is made clear that in the areas of food expenditure, food security and mental health wellbeing of the youngest child, there is a statistically significant difference, at the 1% and 5% significance levels, between those who received treatment with an unemployed household head and those with a non-unemployed household head.

Table 8 : Comparing Treatment Effects Among Households with Unemployed and Non-Unemployed Heads

	N. Treated	N. Control	ATT	Std. Error	T-Value
Food Expenditure (Unemployed)	385	303	698.966	164.237	4.256***
Food Expenditure (Non Unemployed)	105	95	279.426	276.652	1.010
Food Security (Unemployed)	385	303	698.966	164.237	4.256***
Food Security (Non Unemployed)	105	95	279.426	276.652	1.010
School Performance of Youngest Child (Unemployed)	385	303	0.021	0.039	0.538
School Performance of Youngest Child (Non Unemployed)	105	95	0.027	0.059	0.462
Mental Health of Youngest Child (Unemployed)	385	303	0.206	0.089	2.310**
Mental Health of Youngest Child (Non Unemployed)	105	95	0.346	0.145	2.392**

Ultimately, these differences signify the need to angle future policy interventions that aim to reduce poverty at those households in rural areas and with unemployed heads, because they reaped a comparatively larger benefit from the treatment than those that were not rural households or households with an unemployed head.

Policy Recommendations and Conclusion

Recognizing the heightened vulnerability of Burundi's rural and unemployed populations signifies the need for future poverty reduction strategies to acknowledge these demographic vulnerabilities and target them specifically. These two demographics, however, collectively point to a potential opportunity that exists in the overlap of their domains: the agricultural sector of the economy (International Finance Corporation, 2022; Mercier, Ngenzebuke & Verwimp, 2020; Sibanda & Villararribas, 2018). As of November 2022, the World Bank Group noted that despite Burundi's main agricultural exports, which are coffee, tea, cotton and palm oil, make up less than 5% of the nation's GDP, there is a major dependency on these exports in the nation's economy in the areas of employment, national revenues and foreign exchange (International Finance Corporation, 2022, p. 4).

As it exists now, Burundi's agricultural sector does not exploit its own full potential in the areas of employment and diversity of products; it is additionally hindered due to difficulties surrounding the acquisition of rural land to be used in agriculture, lack of knowledge regarding updated sustainable agriculture technologies and lack of local branding, competitiveness and market knowledge (International Finance Corporation, 2022). Ultimately, liberalization of the agricultural sector, to which the government has already expressed their commitment under the current regime, will allow more private enterprises to participate in the public agriculture market (International Finance Corporation, 2022, p. 47). This, in turn, will help to stimulate job creation for the demographics that were identified as the most vulnerable in this evaluation: those in rural areas and those who are unemployed household heads (International Finance Corporation, 2022, p. x). Beyond the fiscal benefits of expanding this sector, overall agricultural sector development could, in the longer term, also serve to alleviate some of the harms associated with food insecurity in Burundi, especially if additional products like animal protein and cereals receive investments from private enterprises as a result of liberalization (International Finance Corporation, 2022, p. 44) as they would allow Burundi to spend less of its finances on importing those products.

Overall, the nature of poverty in Burundi is underlined by its extremely young population, and those who are in rural areas or live in households with an unemployed head experience a heightened risk of food insecurity compared to those who do not have those characteristics. To substantively counter poverty as it exists in Burundi necessitates the prioritization of these demographics in any such efforts. Though it is the world's second poorest nation, Burundi maintains a unique area of opportunity to alleviate its poverty and reverse the stunting of its developmental capacities.

References

Austin, P. C. (2011). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behavioral Research*, *46*(3), 399–424. https://doi.org/10.1080/00273171.2011.568786

Brown, C., Ravallion, M., & van de Walle, D. (2017). Sobering evidence from Africa illustrates how hard it is to target antipoverty efforts well. *Finance & Development*, *54*(4), 4.

Essama-Nssah, B. (2006). *Propensity Score Matching And Policy Impact Analysis—A Demonstration In Eviews*. The World Bank. https://doi.org/10.1596/1813-9450-3877

Haidara, H. (2012). *Enabling poor rural people to overcome poverty in Burundi*. International Fund for Agricultural Development IFAD. https://www.ifad.org/documents/38714170/39150184/Investing+in+rural+people+in+Burundi_e.pdf/5f8915b7-49c9-444e-8fdf-64901d8f615b

International Finance Corporation. (2022). *Creating Markets in Burundi: Leveraging Private Investment for Inclusive Growth.* Country Private Sector Diagnostic, World Bank Group. Retrieved from https://www.ifc.org/wps/wcm/connect/33a8e8c2-b88f-4ed1-b026-3340fcbf6f99/CPSD-Burundi-EN.pdf?MOD=AJPERES&CVID=ogJXFlK

Mercier, M., Ngenzebuke, R. L., & Verwimp, P. (2020). Violence exposure and poverty: Evidence from the Burundi civil war. *Journal of Comparative Economics*, *48*(4), 822–840. https://doi.org/10.1016/j.jce.2020.04.005

OECD (2022), Household spending (indicator). doi: 10.1787/b5f46047-en. Retrieved from https://data.oecd.org/hha/household-spending.htm

Sibanda, R., & Villararribas, V. (2018). *Burundi interim country strategic plan (2018–2020)* (Executive Board Sessions Agenda item 6). World Food Programme. https://executiveboard.wfp.org/document_download/WFP-0000051030

Smith, L. C., & Subandoro, A. (2007). *Measuring Food Security Using Household Expenditure Surveys*. International Food Policy Research Institute. https://doi.org/10.2499/0896297675

United Nations DESA. (n.d.). *Poverty Eradication*. United Nations Department of Economic and Social Affairs: Poverty. Retrieved from https://www.un.org/development/desa/socialperspectiveondevelopment/issues/poverty-eradication.html

United Nations DESA Statistics Division. (2022). *Burundi*. UN Statistics Division. https://unstats.un.org/capacity-development/UNSD-FCDO/burundi/

World Food Programme. (2009) *Comprehensive Food Security and Vulnerability Analysis Guidelines.* World Food Programme Food Security Analysis Service. Retrieved from https://documents.wfp.org/stellent/groups/public/documents/manual_guide_proced/wfp203208.pdf

Appendix

Appendix 1: Propensity Score Comparison

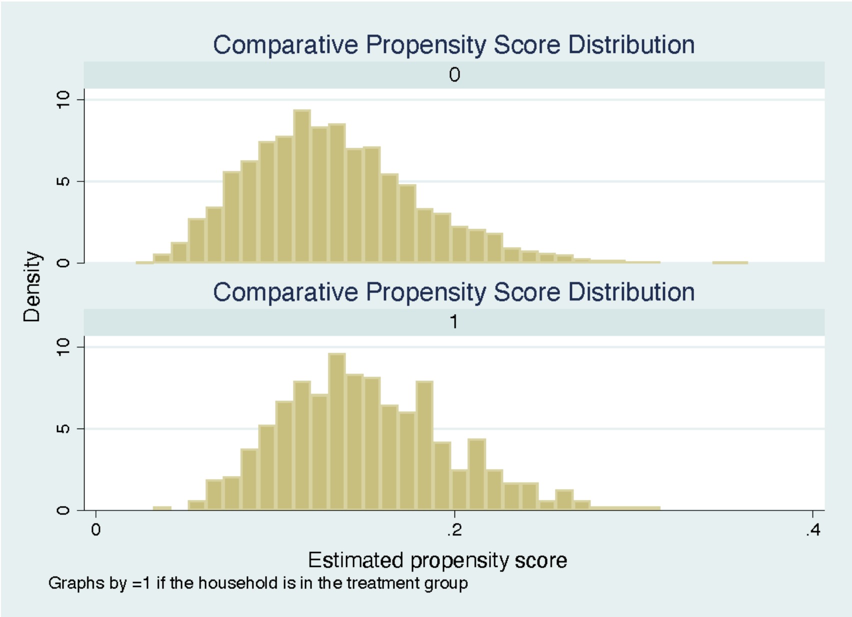


Figure 7: Comparison of Propensity Scores by Treated and Non-Treated Groups

Appendix 2: Propensity Score Comparison 2

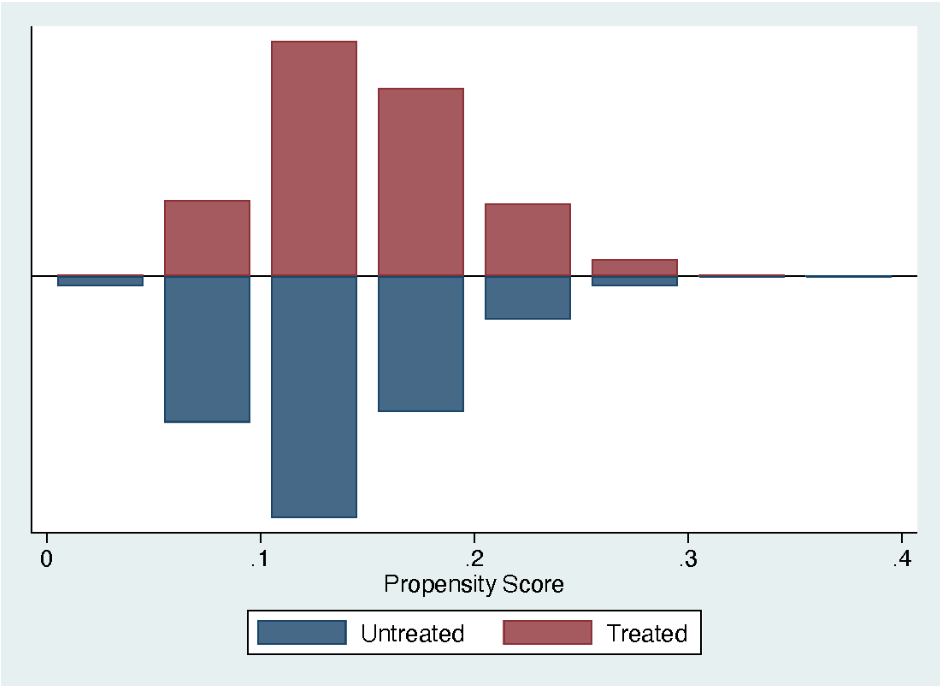


Figure 8: Two-way Vertical Comparison of Propensity Scores by Treated and Non-Treated Groups