Journal of Statistical and Econometric Methods, vol.7, no.2, 2018, 43-69
ISSN: 2241-0384 (print), 2241-0376 (online)
Scienpress Ltd, 2018

Distributive and Quantile Treatment Effects: Imputation Based Estimators Approach

Paul. B. Kenfac Dongmezo[1], P. N. Mwita[2] and I. R. Kamga Tchwaket[3]

Abstract

This paper develops two new classes of estimators measuring the distributive effects of a treatment on a population. Using imputation methods, empirical quantile and bootstrap simulations, we managed to define and study the properties of the two classes. The first class is Imputation Based Treatment Effect on distribution based on rank preservation assumption, basically the effect of treatment on the distribution of potential outcome. The second class is Imputation Based Quantile Treatment Effect which, according to this work is supposed to be the true Quantile Treatment Effect since no rank preservation assumption is made. The second class is based on the fact that each quantile before the treatment is tracked after the treatment and the estimator compares the same group before and after. The first class of estimators (for example the one generated by k-Nearest Neighbors imputation method) performs well as classic Quantile Treatment Effect

[1] Pan African University Institute for Basic Sciences, Technology and Innovation (PAUISTI), Kenya.
[2] Machakos University, .PO Box 136-90100; Machakos, Kenya.
[3] Institut Sous régional de Statistique et d'Economie Appliquée (ISSEA)
P.O Box 294; Yaoundé, Cameroon.

Article Info: *Received*: January 23, 2018. *Revised* : March 15, 2018
Published online : April 5, 2018

given the simulation result. When applied to Lalonde real data set, it performs better than classic Quantile Treatment Effect and Firpo's semi parametric estimator especially for middle quantiles. Also, we found that there is a significant difference between the two classes of estimators meaning that the bias caused by rank preservation assumption is quite significant.

Mathematics Subject Classification: 62E15

Keywords: Bias, Distribution, Estimator, Imputation, Treatment, Quantile.

1 Introduction

Quantitative Impact Evaluation techniques have become the core of impact evaluation process. Most of the projects or programs implemented are in need of a strong and complete evaluation process from design, monitoring to final assessment. Nowadays, saying that a program has had an impact is no more enough, policy makers are interested in the quantitative value of the impact. In the literature, most of the quantitative impact evaluation methods focus on the average impact of the program (ATE) of a given population. If we take for example the difference in difference method, it gives you the average impact assuming that the potential outcome is known before and after the implementation of the program. Presenting the average impact as the effect of the program is a way of assuming that the effect of the program is homogenous therefore, the impact is made constant across all units. It is a very strong assumption that cannot be possible in real life because units are always different. Consequently, there is heterogeneity in the effects of the program. Meanwhile, the average effect of the program can be low or higher because of some extreme values of effects. In fact, as units are different or even if they are almost the same accordingly to covariates collected, they could not react to the project the same way. Some may have a greater impact and others a lower impact. As consequence, the average impact can be affected by some units considered atypical. Since ATE is subjected to strong hypothesis and sensitive to extreme value of effect, but up to now cannot be put aside because classical literature is not able to isolate them, various researchers have been working on new methods to overcome those weaknesses of ATE.

To fill the drawback of ATE, to respond to the needs of policy makers and to give more precisions about the effect of a given program, researchers have started to work on methods that can inform on how program has affected a whole population or a specific sub group of population in the targeted population. These methods are called by Imbens et al. ([1]) quantile and distributional effects methods. These methods aim to come up with the effect of the program in sub group of the targeted population, for a specific section of the distribution of the potential outcome, or more to give the effect on the whole distribution of potential outcome. First works done in this area of research in statistics were in the seventies. Doksum ([2]) and Lehman ([3]) were the first to define the quantile treatment effect. Later on, other researchers like Bitler et al. ([4]) estimate the quantile treatment effects in a randomized evaluation of a job training program. Firpo ([5]) developed methods for estimating QTE in observational studies given unconfoundedness. Abadie et al. ([6]) and Chernozhukov et al. ([7]) studied quantile treatment effects in instrumental variables settings and many others. All these works were done in the classical context of impact evaluation that is known in the literature. Despite the precision given by these new methods, it is still difficult, in the literature, to obtain individual effect consequently the true effect on the distribution. So many assumptions are often made before coming up with acceptable impact.

In fact, comparing the quantiles of the distribution before and after the treatment does not give the effect on the distribution if rank preservation is not taken as an assumption. As a reminder, rank preservation assumption for QTE globally states that if a unit is in the quantile before the treatment assignment, after the treatment it will stay in the same quantile. That is a very strong assumption that is difficult to achieve no matter the population and the program. Other researchers have made a weaker assumption which is the monotonicity of the program meaning that for two units, the treatment will keep their position in the final distribution. All these assumptions are difficult to verify and are not really achievable in real life. A way of bypassing these assumptions is to obtain individual effect then compute quantile or distributional effects. A method like propensity score matching can help in this but with a small sample and all the weaknesses leaning on the quality of the control group, the results are not stable and depend on the control group. Following the work recently done by Kenfac et al ([8]) where they managed to estimate counterfactual using imputation methods and obtain good ATE or even better ATE in specific contexts, they finally defined what they called Imputation Based Treatment Effect (IB-TE) from which individual effects can be drawn.

This paper uses the results of [8] to propose a new approach of obtaining quantile treatment effects and impact on the distribution. Using IB-TE estimators, with the good imputation method associated with bootstrap, we can achieve that goal. In fact, having individual effect by completing sample with an estimate of counterfactual means that each unit of the sample can be ranked and follow before and after the treatment. Using this method can help to obtain the true quantile treatment effects without any assumption and better effects on distribution. The advantage of this approach is that no major assumption is made regarding the assignment process or regarding the rank preservation assumption which are sources of bias in classical literature. Meanwhile using imputation methods in context where more than half the observation is missing; a bias might be generated though according to the results obtain by Kenfac et al ([8]), is less than the bias due to assumption taken in classical methods. The results obtained are compared to the one from classic methods not completing the sample.

The rest of the paper is organized as follows: section two discusses the main and recent methods used in literature to measure distributional effects and quantile treatment effects. Section three presents the new class of distributional and quantile treatment effect based on work done in [8]. Section 4 is dedicated to simulations studies. Section 5 present a small application using a classic data set.

2 Measuring Distributional Treatment Effects in the Literature

There are two branches of the literature dealing with distributional treatment effects and both of them are related to distributions. The first branch uses the joint distribution of potential outcome before and after the treatment $F(Y_0, Y_1)$ to address the problem of distributional effects of treatment. The second branch uses the marginal distribution of potential outcome before $F(Y_0)$ and after $F(Y_1)$. This literature review is presented according to these two branches knowing that the second branch is the more developed because of its usefulness compared to the first one.

2.1 Joint Distribution of Potential Outcome Approach

This strand of the literature was not much developed. Few researchers have work on estimating the distributional treatment effect using a joint distribution of the two potential outcomes (Y_0, Y_1) that is going to be used to compute the difference in quantiles between the two marginal distributions of potential outcome. Hoeffding ([9]) and Fréchet ([10]) pioneer this part of the literature with their work on probability distribution. Then, using their results in the aim of assessing distributional impact of a program, Heckman et al. ([11]) and Heckman and Smith ([12], [13]) found the joint distribution of (Y_0, Y_1) using the marginal distribution of Y_0 and Y_1 in a randomized control experiment, a practical case. Later on, researcher such as Aakvik et al. ([14]) used joint distribution to identify treatment effects of discrete outcome when the treatment effects are heterogeneous. An improvement of that work can be read in Aakvik et al. ([15]). Carneiro et al. ([16],

[17]) in their research proposed an approach to estimate directly or to bound the distribution using the method common in factor analysis but applied to model counterfactual distributions. If potential outcomes are generated by a low dimensional set of factors, then it is possible to estimate the distributions of factors and generate distributions of the counterfactuals. Here, low dimensional refers to the number of factors relative to the number of measured outcomes. For a summary and other development of this branch of the literature, see Section 2 of Abbring et al. ([18]).

2.2 Marginal Distribution of Potential Outcome Approach

Works presented here lean on estimation of the marginal distribution of potential outcome before the treatment and after the treatment. Lehmann ([3]) and Doksum ([2]) were the first to define the quantile treatment effect as the difference between the same quantile in the distribution after and before the treatment under rank preservation assumption. The definition that they used, coming from a shift function between two distributions under nonlinear assumption, is at the core of the development around quantile treatment effects.

Abadie et al. ([19]) used a new instrumental variable approach that measure program impact on quantiles of the distribution of potential outcome. At the same time, assuming heterogeneity in the effects of the program, the monotonicity assumption developed by Imbens and Angrist ([20]) or the uniformity assumption presented in Heckman and Vytlacil ([21]), they estimated the Local Quantile Treatment Effect (LQTE) which is a kind of proxy of LATE in the classical IV literature. They used the identification results in Abadie ([22]); see also Imbens and Rubin ([23]) for identification of the marginal potential distributions of compliers when no covariates are present, and Abadie ([24]) for bootstrap tests of distributional treatment effects in a same framework. Chernozhukov and Hansen ([7]) also proposed an IV model for quantile treatment effects in the presence of endogeneity and under rank invariance assumption. Later on, more development in this area were done by [25]. In line with IV quantile regression, see also [26] for their work on estimating the distributional effects of an endogenous treatment that varies at the group level when there are group-level unobservable.

Heckman and Vytlacil ([21]) also proposed a non-parametric estimators of treatment effects using Marginal Treatment Effect assuming at the same time heterogeneity in choice and response. They developed what is called local instrumental variable. Later on, Carneiro and Lee ([27]) extended that method to the estimation of not only means, but also distributions of potential outcomes. The newly developed method is illustrated by applying it to changes in college enrolment and wage inequality using data from the National Longitudinal Survey of Youth of 1979.

Athey and Imbens ([28]) proposed an estimation of quantile treatment effect under the assumption of difference in difference methods meaning data are available before and after the treatment for all units (kind of panel data analysis) and under rank preservation assumption. Their approach provides an estimate of the entire counterfactual distribution of outcomes that would have been experienced by the treatment group in the absence of the treatment and likewise for the untreated group in the presence of the treatment. Firpo ([5]) proposed a semi parametric estimator of QTE assuming that selection to treatment is based on observable characteristics. Root-N consistency, asymptotic normality, and achievement of the semiparametric efficiency bound was shown for that estimator. Ping Yu ([29]) proposed an estimator of marginal quantile treatment effects meaning conditional quantile on the covariates and rank in the distribution. The base of his work is the following definition of marginal quantile treatment effect $\Delta_\tau^{MQTE}(x,u_D) = Q_{Y_1|X,U_D}(\tau|x,u_D) - Q_{Y_0|X,U_D}(\tau|x,u_D)$. After taking rank preservation assumption, he developed sharp bounds for the quantile treatment effect with and without the monotonicity assumption.

Most of the literature presented here focuses on the IV method and its derivatives. Also, many assumptions are considered for implementation of such methods. The main and common one is unconfoundedness which is difficult to verify. Another one at the core of assessment of distributional effects of treatment is heterogeneity of the response which is more logical given heterogeneity of population. With the IV method, there is a need of existence for good instruments satisfying some conditions that is not always easy to verify. Rank preservation assumption is as well assumed. Violation of these assumptions can significantly affect those methods especially when results are subjected to their verification.

The common problem of these two branches of the literature is the counterfactual. In fact, most of these methods suffer from the fact that to estimate marginal distribution or joint distribution of potential outcome, the full set of observations is needed. Given that it is not possible to observe Y_0 and Y_1 at the same time, the previous methods suffer from that incompletion even if the method of estimation of the distribution is good. This paper, in order to simplify the problem of impact evaluation and from a statistical point of view solve, as Kenfac et al ([8]), the problem of distributional effects from the source as Rubin highlighted, assuming only that counterfactual is a missing value that can be estimated by specific methods according to the assignment process.

3 Imputation Based Distributional Treatment Effects

This section develops the framework of Imputation Based Distributional Effect of a treatment, then shows how estimators are derived from imputation methods, their properties and what their specificities are compared to others in the same literature. Some tools useful for development of estimators such as quantiles are also briefly presented.

3.1 IB-TE Framework Description

The approach developed in this paper is inspired by the work of [8]. In their work, their aim was to find out if imputation methods can be used to solve the problem of impact evaluation in a better manner. Considering counterfactual as a missing value from a statistical point of view, different methods are used to estimate it. As a consequence, selection bias, unconfoundedness and rank preservation assumption which are some of important assumptions that should be considered and verified when assessing distributional effects of a program, are taken into account by the imputation method. Otherwise, bootstrap associated to the imputation method is supposed to soften the bias due to not taking into account those assumptions.

The framework of IB-TE estimators is the following. We consider a treatment assignment defined by T_i, equal to 1 if unit i is treated and 0 otherwise. The potential outcomes observed are noted $Y_{i,0}$ if unit i is not treated and $Y_{i,1}$ if unit is treated. Assuming that we have a global sample of n units n_1 are treated, $n-n_1$ assumed as control group, after imputation the two distributions are given by:

Treated group: $TrCom = \left(Y_{1,1}, Y_{2,1}, ..., Y_{n_1,1}, \tilde{Y}_{n_1+1,1}, ..., \tilde{Y}_{n,1}\right)$

Control group: $CoCom = \left(\tilde{Y}_{1,0}, \tilde{Y}_{2,0}, ..., \tilde{Y}_{n_1,0}, Y_{n_1+1,0}, ..., Y_{n,0}\right)$

where potential outcomes with tilde are estimated counterfactual by a given imputation method. Let’s also consider a set of covariates X corresponding to characteristics of different units.

For the imputation methods, three imputation methods will be mainly considered from the works of [8] which are Multiple Imputation, Quantile Regression in case we have data before and after the treatment and finally linear regression (deterministic or random). In additions, other imputation methods like k-NN will be tested as well.

The definition of empirical quantile given below will be used in this paper.

Definition: Quantile

Let Y be a real valued random variable; and Let $0 < p < 1$ be a probability. The

p^{th} quantile is the smallest number $y \in \mathbb{R}$ such that $P(Y > y) \le 1 - p$.

Assuming that Y is a continuous distribution which is often our case in IE framework, the p^{th} quantile can be defined in many ways as:

- ✓ The p^{th} quantile is the number $y \in \mathbb{R}$ such that $P(Y > y) = 1 - p$;
- ✓ The p^{th} quantile is the number $y \in \mathbb{R}$ such that $P(Y \le y) = p$;
- ✓ Considering the distribution function of Y being $F(y) = P(Y \le y) \quad \forall\, y \in \mathbb{R}$; the p^{th} quantile is the number $y \in \mathbb{R}$ such that $F(y) = p$
- ✓ The p^{th} quantile is $Q(p) = F^{-1}(p), \quad p \in (0,1)$.

In this work, the quantiles used for simulation will be deciles ($p = t \times \frac{1}{10}, \quad t = 1, 2, ..., 9$) and for application can be centiles ($p = t \times \frac{1}{100}, \quad t = 1, 2, ..., 99$) depending on the size of the sample.

Besides the theoretical definition of quantile, the one that is going to be used here is the definition of the empirical quantile.

Given a series of data $y_1, y_2, ..., y_n$ with a common distribution, we want to estimate the p^{th} quantile $Q(p)$ with p as a probability. The p^{th} quantile is such y that $P(Y \le y) = p$. For any y we can estimate $P(Y \le y)$ using frequencies, the estimate is therefore:

$$\hat{p} = \frac{\#(y_i : y_i \le y)}{n}$$

The empirical quantile is such that $\hat{p}$ is close to p. Practically; to compute the empirical quantile, two steps have to be taken: (i) consider an integer m such that $p \approx m/n$ (we may round pn to the closest integer); (ii) the estimate of $Q(p)$ is the m^{th} largest observation in the series.

3.2 Effect across distribution and subgroups: IB-TED

Given that the counterfactual is estimated as a missing value, the sample of the treatment group and control group are completed. For treatment group, the sample is completed by an estimated value of what would have been the potential outcome for those who are not treated but considered as control. The same thing is done when considering those who are not treated in the same population. Using those estimations, individual effects is given by $\hat{\Delta}_i = (Y_{i1} - \tilde{Y}_{i0})T_i + (\tilde{Y}_{i1} - Y_{i0})(1 - T_i)$.

Imputation Based Treatment Effect on Distribution (IB-TED)

To obtain the imputation based effect of the treatment on the distribution, two samples have to be compared: Treatment completed by imputation $TrCom = \left(Y_{1,1}, Y_{2,1}, ..., Y_{n_1,1}, \tilde{Y}_{n_1+1,1}, ..., \tilde{Y}_{n,1}\right)$ and control group completed $CoCom = \left(\tilde{Y}_{1,0}, \tilde{Y}_{2,0}, ..., \tilde{Y}_{n_1,0}, Y_{n_1+1,0}, ..., Y_{n,0}\right)$. The comparison is made using quantile difference, or mean and median effects of distribution of effects given by $\left(\hat{\Delta}_1, \hat{\Delta}_2, ..., \hat{\Delta}_n\right)$. Given a probability τ we define:

$$IB-TED(\tau) = \tau\left(Y_{1,1}, Y_{2,1}, ..., Y_{n_1,1}, \tilde{Y}_{n_1+1,1}, ..., \tilde{Y}_{n,1}\right) - \tau\left(\tilde{Y}_{1,0}, \tilde{Y}_{2,0}, ..., \tilde{Y}_{n_1,0}, Y_{n_1+1,0}, ..., Y_{n,0}\right)$$

With $\tau()$ as the quantile of order τ of the distribution in brackets.

Imputation Based Treatment Effect on Distribution in sub population (IB-TED ($X = x$))

Assuming that covariates are available, the effect can be defined for a subgroup of the population on interest. For example, if it is suspected that a treatment can have different effects depending on the sex, the model is developed generally but the distributional effect according to the sex can be computed.

Given a probability τ and subgroup defined by $X = x$, we define:

$$IB-TED(\tau \mid X = x) = \tau\left(Y_{1,1}, Y_{2,1}, ..., Y_{n_1,1}, \tilde{Y}_{n_1+1,1}, ..., \tilde{Y}_{n,1} \mid X = x\right) - \tau\left(\tilde{Y}_{1,0}, \tilde{Y}_{2,0}, ..., \tilde{Y}_{n_1,0}, Y_{n_1+1,0}, ..., Y_{n,0} \mid X = x\right)$$

If the subsample $X = x$ is all taken, the $IB-ATE(X = x)$ can be computed and compared to the general $IB-ATE$.

Properties of this class of estimators will be derived empirically using simulations and bootstrap. By increasingly sampling the population, the bias, consistency and convergence will be studied.

3.3 Effects on Quantiles : IB-QTE

To obtain the effect of the treatment on the distribution, the quantiles are considered to be the ones affected by the distribution. Taking just the difference between quantile of two groups does not give the true QTE but just the effect on the distribution. By assuming rank preservation or rank invariance for most of the methods in the literature, helps to ensure that when the difference of two quantiles is done, we obtain the treatment effect since the same units are compared in treatment and control. In the method developed here, such assumption is not made. Given that we have individual effects, we can track all the units of a specific quantile of the distribution before the treatment and find its associate treated potential outcome after the treatment to form the after-treatment quantile group. Two steps have to be taken to identify the IB-QTE:

(i) Identify the quantile of belonging for each unit in the distribution of the control group given by the following sample $\left(\tilde{Y}_{1,0}, \tilde{Y}_{2,0}, ..., \tilde{Y}_{n_1,0}, Y_{n_1+1,0}, ..., Y_{n,0}\right)$ and identify the value of the quantile as well which probably correspond to the value of potential outcome of given unit;

(ii) For each quantile group identified previously, create the image group in the treatment sample given by $\left(Y_{1,1}, Y_{2,1}, ..., Y_{n_1,1}, \tilde{Y}_{n_1+1,1}, ..., \tilde{Y}_{n,1}\right)$;

(iii) The average or median change observed in the quantile control group compared to the image group in the treatment sample is the IB-QTE.

Let's τ be a probability such that $0 < \tau < 1$, let $Q_0(\tau)$ the τ^{th} quantile group and $q_0(\tau)$ the value of the τ^{th} quantile in the completed distribution of control group (CoCom) and $Q_1(\tau)$ the τ^{th} quantile group and $q_1(\tau)$ the value of the

τ^{th} quantile in the completed distribution of treated group (TrCom). Let also assume that $q_0(\tau)$ is attained for the i_0 unit in the CoCom distribution.

1) The Imputation Based Quantile Treatment is defined by:

$$IB-QTE\left(Q_0(\tau)\right)=Y^*_{i_0 1}-q_0(\tau)=Y^*_{i_0 1}-Y^*_{i_0 0}$$

with $$Y^*_{i1}=\begin{cases} Y_{i1} & if\ \ the\ unit\ i\ is\ treated \\ \tilde{Y}_{i1} & if\ \ the\ unit\ i\ not\ treated \end{cases}$$ and

$$Y^*_{i0}=\begin{cases} Y_{i0} & if\ \ the\ unit\ i\ is\ not\ treated \\ \tilde{Y}_{i0} & if\ \ the\ unit\ i\ is\ treated \end{cases}$$

2) The Imputation Based Average Quantile Treatment Effect is defined by:

$$IB-AQTE\left(Q_0(\tau)\right)=E\left(Y^*_{i1}\,\middle|\, i\in Q_0(\tau)\right)-E\left(Y^*_{i0}\,\middle|\, i\in Q_0(\tau)\right)$$

3) The Imputation Based Median Quantile Treatment Effect can be as well defined by:

$$IB-MedQTE\left(Q_0(\tau)\right)=Med\left(Y^*_{i1}\,\middle|\, i\in Q_0(\tau)\right)-Med\left(Y^*_{i0}\,\middle|\, i\in Q_0(\tau)\right)$$

With *Med* as the median of the group in bracket in other term the 0.5^{th} quantile.

3.4 Comparisons and Differences

The two classes of estimators defined previously both access the distributional effect of the program or treatment but both of them are not QTE. The IB-TED is the effect of the treatment on the whole distribution of potential outcome. Its value is the answer to the question of how much the distribution of the potential outcome shifted positively or negatively depending on the sign of the effect. A unit in the first decile can find itself in the 8th decile it doesn’t matter as long as it contributes to shift upward the distribution. The IB-QTE is really the Quantile Treatment Effect since it measured on how a given quantile changed after the treatment. All the units in a specific quantile are tracked and they are put in the same group after the treatment. That group will then be compared to the initial quantile where they were before the treatment. The difference obtained is what is called in this study the true QTE. To obtain that result, no assumption is made and it is possible since imputation can help us to compute the individual effect as stated in [8].

Therefore, the main difference in the first one is the effect on the distribution while in the second one is the real effect on the quantile defined before the treatment assignment.

4 Simulations and Summary of Results

The main objective of this section is to use simulations to test our hypothesis that when using imputation before applying empirical quantiles can lead to better estimators than using empirical quantiles straight forward or any other method. In this section, a description of simulation procedure and parameters is done, then simulations are performed under Random Assignment (MCAR missingness) hypothesis and under Deterministic Assignment (MAR or NMAR missingness) hypothesis.

4.1 Description of Simulation Procedure

The aim of this simulation is to show that using the appropriate imputation method to estimate counterfactual; we can come up with better distributional effects including quantile treatment effects.

The simulation recreates a hypothetical situation where a treatment (project or program) has to be assigned in a population with all the parameters being known. For example, if the assignment process is well known, the potential outcome is known, decision to treat everyone or not to treat everyone can be taken so that computation of the true impact of the project on the distribution of outcome can be done easily. In brief, all parameters are mastered and they can be modified to obtain different results according to the objectives fixed. Therefore, for a given assignment process, simulation results will tell which imputation method gives closer results to the true distributional impact.

In the simulation exercise, a data base of 10,000 cases is generated. The potential outcome (Y_b) and the covariates before the treatment assignment or before the program are generated. Since we are in the simulation, a situation where by all units are not treated (Y_{2NT}) and a situation where all units are treated (Y_{2T}) is simulated at the same time. From that, the true QTE is computed in the overall population as follows $QTE_{true}(\tau)=\tau(Y_{2T})-\tau(Y_{2NT})$ where τ is a given probability. The quantile treatment effect computed from the population is basically the true QTE given the way the data are simulated.

The next step of simulation is to create the treatment variable (T), by deciding which case is treated and which case is not treated according to the assignment process chosen. In this study, two situations are considered: Random assignment leading to MCAR missingness mechanism and Controlled assignment leading to MAR or NMAR missingness mechanism. If $T=1$ the case is treated and if $T=0$ the case is not treated. From this stage of simulation, the classic and most used distributional impact of treatment in the literature under rank preservation assumption can be computed by: $QTE_{Class\_true}(\tau)=\tau(Y_{2T}\,|T=1)-\tau(Y_{2NT}\,|T=0)$, it is just the difference between quantile in the treated group and the same quantile for the control group.

The potential outcome in the reality is now generated in the variable Y_a as follows:

- For non-treated ($T=0$), $Y_a=Y_{2NT}$, the value of Y_{2NT} is just reported when $T=0$;
- For treated ($T=1$), $Y_a=Y_{2T}$, the value of Y_{2T} is just reported when $T=1$.

Using Y_a the potential outcome with missing values are generated (Y_T and Y_{NT}), if you are treated, what would have happened if they were not treated is a missing value and also if they are not treated, what would have happened if you were treated is consider as a missing value.

- ✓ For treatment case, create Y_T as follows: report all observations of potential outcome for treated and for non-treated considered as missing values: $Y_T = Y_a$ if the unit is treated and $Y_T = .\ (miss)$ if the unit is not treated and has to be imputed;
- ✓ For non-treatment case, create Y_{NT} as follows: report all observations of potential outcome for non-treated and for treated consider as missing values: $Y_{NT} = Y_a$ if the unit is not treated and $Y_{NT} = .\ (miss)$ if the unit is treated and has to be imputed.

The classical quantile treatment effect can also be computed using Y_a the potential outcome in the reality as follow: $QTE_{Class\_true}(\tau) = \tau(Y_a | T = 1) - \tau(Y_a | T = 0)$. One of the major weakness of this estimator is that it assumes the rank preservation assumption.

Knowing that in our simulation we hypothetically have all potential outcomes, the real estimators of the distributional impact can be computed as well as an estimation of values considered as missing. The aim being to estimate missing values using imputation methods then compute imputation based distributional impact. Those estimators will be compared to the distributional estimators computed using the population and to the ones obtained by classical methods (classical empirical quantiles).

Under the large class of existing imputation methods, the chosen ones in this study are: Mean imputation, Random imputation, Linear regression imputation (deterministic and random), Nearest Neighbour imputation, Multiple Imputation, Maximum Likelihood imputation, Propensity score matching imputation and finally Quantile regression imputation which is not commonly used.

To test the performance of our computed Imputation Based Distributional

Treatment Effect estimators (IB-QTE and IB-TED), the average bias is computed as follows: ($AvrgBias = E(\hat{\theta} - \theta)$) and compared to the one for classical method. Of course, this is done under a bootstrap procedure of 1000 replications which will give us variance and standard deviation.

Globally, the simulation shows that most of our estimators are biased no matter the assumptions and the assignment process and the bias decreases slowly. Checking out the variances of each estimator, fortunately it tends to zero as the sample size increases and for almost all IB-QTE estimators they are smaller than classic QTE estimators. The conclusion is straight, IB-QTE and IB-TED estimators can be used to estimate distributional effect of a treatment.

4.2 Random Assignment Hypothesis Results (MCAR)

Results of this section are twofold; the first result is the quantile treatment effect in both cases reduced sample and sample completed using imputation methods under rank preservation assumption. The second result will be classical quantile treatment effect under rank preservation assumption compared to IB-QTE without assuming rank preservation.

4.2.1- RA under rank preservation assumption: Effect on the whole distribution (IB-TED)

Under rank preservation assumption, the simulations show that the bias of all estimators (Classic QTE and IB_TED) is decreasing too slowly and will probably never get to zero. Therefore, it is clear that we are working with biased estimators and the best one will be the one with the smallest bias and smallest variance (small and convergent variance). From simulation results, it is clear that IB- TED estimators are far better than Classic QTE in almost all the cases except few cases related to sample size and small quantiles. In fact, for small quantiles (1^{st} and 2^{nd} decile) classic QTE performs as well as the IB- TED, it is even better for the 1^{st} decile than all IB- TED estimators especially for small sample (N<200). Out of those specific cases highlighted previously, IB-QTE estimators are better than classic QTE. For example, the k-NN imputation which is among the best IB-TED gives the best results for large sample (N>200) and no matter the quantile selected. The k-NN IB- TED is always among the three best estimators in our simulations. For small samples, PSM IB-TED and k-NN IB-TED are sharing the first and the second position in term of estimators with the smallest bias. For example, for the sample size of 50 and the 8^{th} decile, the bias for classic QTE is 4.6 while for PSM IB- TED it is 2.5 and for k-NN it is 6.5. For the same sample size and for the 5^{th} decile (the median), the bias is 3.5 for PSM IB- TED and 7 for k-NN.

4.2.2- RA without rank preservation assumption: IB-QTE

Without assuming rank preservation, each unit in a quantile before the treatment will be followed after the treatment and the two groups will be compared to get the true quantile treatment effect based on the distribution of the potential outcome before the treatment. In other words, units in the τ^{th} quantile will be grouped to form the comparison group after treatment: this is how we define the true Quantile Treatment Effect (IB-QTE). From the result of simulation, outcomes are a bit mitigated at first sight, they show globally that IB-QTE are as good as the classic QTE. For small quantiles no matter the sample size, classic QTE are slightly better than most of IB-QTE estimators but for quantile above median, IB-QTE are far better. Taking for example linear regression imputation, for two different sample sizes (50 and 100) the IB-QTE obtained from that is better than classic QTE especially for larges quantiles but less good for small quantiles. For example, for N=50 the bias for IB-QTE is small for the 8^{th} and 9^{th} deciles and smaller than classic QTE bias. For large samples, in general, classic QTE are better than IB-QTE for extreme quantile (1^{st} decile and 9^{th} decile quantile) but for the other ones in between they are not better. The IB-QTE that are good in average are ML IB-QTE, k-NN IB-QTE and MI IB-QTE compared to classic QTE.

4.3Assignment controlled by Variables (NMAR or MAR)

Considering that the assignment process depends on a given variable or combinations of some given variables. When a threshold is established, the population is divided in two parts and the group below the threshold is treated while the group above is not and they are compared with each other.

4.3.1- NMAR and MAR under rank preservation assumption: Effect on the whole distribution (IB-TED)

Assuming NMAR or MAR in simulation and under rank preservation assumption, the simulations results show that for small samples (N<200), IB-TED produced using deterministic imputation and random imputation are good as classical imputation for big quantiles but better for small quantiles. Taking for example the N=50, the bias of the IB-TED for deterministic linear regression is smaller than the bias due to classic QTE for the first eight deciles and only bigger for the 9th decile. For the bigger sample size (N>200), the k-NN IB-TED is the best method for almost all the quantiles. For example, for N=500 the bias for the IB-TED for the first sixth deciles is smaller than the bias for the classic QTE but almost the same and a little bit greater for the last three. In the case of random assignment, on average the best imputation methods are deterministic linear regression IB-TED for small sample and k-NN IB-TED for large samples.

4.3.2- NMAR and MAR without rank preservation assumption: IB-QTE

Assuming that the assignment process is NMAR or MAR, the true QTE called here IB-QTE is estimated using imputation method. The idea here is to follow each quantile after the treatment and obtain the treatment effect on that quantile. Comparison between IB-QTE and classic QTE shows that results are mitigated. In some cases, classic QTE is better while in some others it is IB-QTE which are better. The only constancy is that IB-QTE is far better than classic QTE no matter the sample size for middle deciles (3rd, 4th, 5th, 6th and 7th deciles). Therefore, under NMAR or MAR and without rank preservation assumption, IB-QTE estimators are better in estimating the middle decile of the distributional effect of a treatment. The chosen estimators are k-NN IB-QTE and ML IB-QTE for small sample and k-NN IB-QTE, ML IB-QTE and MI IB-QTE for large sample. As an example of this result, considering N=50, IB-QTE estimators like k-NN, ML and MI are better than classic estimators especially for middle deciles.

5 Applications

After simulations, where the results showed that IB-TED and IB-QTE estimators can perform as well as classic quantile treatment effect estimators otherwise better in some cases, the next step is to apply these results to real set of data since simulation are always questionable.

Firstly, the Lalonde ([30]) data set is considered for application. Lalonde data set (1986) contain the treated and control units from the male sub-sample from the National Supported Work Demonstration. The NSW Demonstration [Manpower Demonstration Research Corporation (MDRC) 1983 was a federally and privately funded program implemented in the mid-1970s to provide work experience for a period of 6-18 months to individuals who had faced economic and social problems prior to enrolment in the program. Those randomly selected to join the program participated in various types of work, such as restaurant and construction work. Pre-intervention variables were collected by the program to allow Lalonde to use control groups, selected using pre-intervention variables to compare and obtain the treatment effect on treated.

Applying our estimators to Lalonde data, the three best IB distributional effects are considered and compared to classic QTE without completing data and to Firpo's ([5]) results on the same data set. Combining bootstrap to imputation methods and applying the empirical quantile shows that estimation of the median impact using IB-TED estimators is closer to the ATE estimator than the Firpo's result and classical QTE.

In fact, looking at classic QTE estimators under rank preservation assumption, the effect of the program is increasing with the deciles. For the first decile, the effect is 0 then for the second, it is still 0 the 3rd decile effect is 943 then the median effect is 1093.5 which is far from the ATE ($1794). Adding bootstrap on it did not change much the result, the largest effect being for the 8th and 9th decile respectively 2273 and 3197 which is basically explosive and too much. Analysis of Firpo's results show also that the effect is increasing with deciles after the 4th decile. The effect is 0 for the 1st and 2nd decile which is not likely to happen, then 711 for the 3rd decile, 21 for the 4th decile meaning that some effects were probably negative. Then comes the median effect which is 1927 quite close to 1794 the ATE but after that the effect becomes explosive, 3879 for the 6th decile, 4517 for the 7th decile to end at 5530 for the 9th which is again not likely to happen. Firpo's method may perform well only for median. Now if we take one of our best IB-TED estimator (k-NN) under rank preservation assumption, the effect is quite uniform across decile with an average difference of 30 units. The tendency is the following: effect of the program is larger for the tail of the distribution of potential outcome (1793.8 for the 2nd decile and 1811.8 for the 9th decile) and quite stable and close to the ATE for the middle deciles (1702.4 for median effect and 1693.2 for 6th decile). This pattern of results is almost the same for all IB-TED estimators computed using imputation methods.

In conclusion, explosiveness of Firpo's results and classic QTE shows that estimators construct using non-parametric approach and using incomplete data set are not convergent practically. They basically show that the effect of training program on earning increases with the decile meaning that the more you earned before the program the more the program will have an effect on you which is counter intuitive given that theoretically effect of the training is most likely to be greater for those who were earning less. From the result using IB-TED, the effect is more stable across the distribution of earnings and bigger for people earning less and people earning more (tail of the distribution) which is more likely to happen than for explosive effect of the training.

6 Conclusion

Following pioneer work of Kenfac et al ([8]) in using imputation methods to derive Average Treatment Effects from individual effects as they have introduced in their paper, this paper has tried to use the same approach to estimate the distributional effect of a program in the population. Given that in the literature most of QTE are done using rank preservation assumption, the work done here managed to overcome that weakness and does not make any such assumptions.

The aim of this research was to come up with a new class of distributional effect estimators performing better than existing ones with less assumptions made. Using the approach developed by Kenfac et al ([8]), two distributional effect estimators can be defined: IB-TED (Imputation Based Treatment Effect on the Distribution) being the effect of the program on the distribution assuming rank preservation assumption and IB-QTE (Imputation Based Quantile Treatment Effect) being the true quantile treatment effect without any rank assumption was found out. The first estimator is compared to all estimator in the literature assuming rank preservation like classic QTE and Firpo's semi parametric QTE and the second one is just to present what is really the quantile treatment effect from our own perspective.

Simulation results shows that IB-QTE and IB-TED are biased just like classic QTE but consistent and can perform as good as existing estimators or even better than existing ones especially classic QTE in some cases and for some specific quantiles. This new class of distributional effects estimators came with the possibility of having the true QTE without any assumption and to estimate more accurately the effect of a treatment on a given distribution.

The estimators constructed have been tested on the famous dataset of Lalonde and the results are more likely to be realistic than the classic QTE and

Firpo's results on the same dataset. According to IB-TED results, the effect of the training is larger for extreme deciles and close to the average for the middle deciles. This result is different from classic QTE results and Firpo's result but theoretically is more likely to happen. This means that IB-TED can be a serious and useful approach in accessing the distributional effects of a program.

Meanwhile, the IB approach is subjected to getting the best imputation method. Better is the imputation method and the best will be the IB estimator computed. As an example, if the imputation method is regression, make sure that the R^2 is good enough for the explanatory variables to explain enough percentage of the missing variable. With the best imputation methods, the IB estimators will be perfect for estimating the treatment effect.

ACKNOWLEDGEMENTS. The first author wishes to thank the African Union for supporting his research at the Pan African University, Institute for Basic Sciences, Technology and Innovation.

References

[1] G.W. Imbens and J.M. Wooldridge, Recent Developments in the Econometrics of Program Evaluation, Journal of Economic Literature, 47, (2008), 5-86.

[2] K. Doksum, Empirical Probability Plots and Statistical Inference for Nonlinear Models in the Two-Sample Case, Annals of Statistics, 2, (1974), 267-277.

[3] E.L. Lehmann, Nonparametrics: Statistical Methods Based on Ranks, San Francisco: Holden-Day (1974).

[4] M.P. Bitler, J.B. Gelbach and H.W. Hoynes, What Mean Impacts Miss: Distributional Effects of Welfare Reform Experiments, American Economic Review, 96, (2006), 988-1012.

[5] S. Firpo, Efficient Semiparametric Estimation of Quantile Treatment Effects, Econometrica, 75(1), (2007), 259-276.

[6] A. Abadie, J.D. Angrist and G.W. Imbens, Instrumental Variables Estimates of the Effect of Subsidized Training on the Quantiles of Trainee Earnings, Econometrica, 70, (2002), 91-117.

[7] V. Chernozhukov and C. Hansen, An IV Model of Quantile Treatment Effects, Econometrica, 73, (2005), 245-261.

[8] D. P. B Kenfac., N. P. Mwita and T. I. R Kamga, Imputation Based Treatment Effect Estimators, Journal of Statistical and Econometric Methods, vol.6, no.3, (2017), 23-64.

[9] W. Hoeffding, Masstabinvariate Korrelationstheorie, Schriften des Mathematischen Instituts und des Instituts Für Angewandte Mathematik der Universität Berlin, 5, (1940), 179-233.

[10] M. Fréchet, Sur Les Tableaux de Corrélation Dont Les Marges Sont Données, Annals University Lyon, Series A, 14, (1951), 53-77.

[11] J.J. Heckman, H. Ichimura and P.E. Todd (HIT), Matching as an Econometric Evaluation Estimator: Evidence From Evaluating a Job Training Programme, The Review of Economic Studies, 64, (1997), 605-654.

[12] Heckman, J.J. and J.A. Smith, Assessing the Case for Randomized Evaluation of Social Programs, in Measuring Labour Market Measures: Evaluating he Effects of Active Labour Market Policies, K. Jensen and P.K. Madsen, eds., Copenhagen: Danish Ministry of Labor, (1993), 35-96.

[13] Heckman, J.J. and J.A. Smith, Evaluating the Welfare State, Econometric and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium, S. Strøm, eds., Cambridge: Cambridge University Press, (1998), 241-318.

[14] A. Aakvik, J.J. Heckman and E. Vytlacil, Semiparametric Program Evaluation Lessons from an Evaluation of a Norwegian Training Program, mimeo, Department of Economics, University of Chicago, (1999).

[15] A. Aakvik, J.J. Heckman and E. Vytlacil, Estimating Treatment Effects for Discrete Outcomes When Responses to Treatment Vary: An Application to Norweigian Vocational Rehabilitation Programs, Journal of Econometrics, 125, (2005), 15-51.

[16] P. Carneiro, K.T. Hansen and J.J. Heckman, Removing the Veil of Ignorance in Assessing the Distributional Impacts of Social Policies, Swedish Economic Policy Review, 8, (2001), 273-301.

[17] P. Carneiro, K.T. Hansen and J.J. Heckman, Estimating Distributions of Treatment Effects with An Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice, International Economic Review, 44, (2003), 361-422.

[18] J.H. Abbring and J.J. Heckman, Econometric Evaluation of Social Programs, Part III: Distributional Treatment Effects, Dynamic Treatment Effects, Dynamic Discrete Choice, and General Equilibrium Policy Evaluation, Handbook of Econometrics, Vol. 6B, J.J. Heckman and E.E. Leamer, eds., Amsterdam: Elsevier Science B.V., Ch. 72, (2007), 5145-5303.

[19] A. Abadie, Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models, Journal of the American Statistical Association, 97, (2002), 284-292.

[20] G.W. Imbens and J.D. Angrist, Identification and Estimation of Local Average Treatment Effects, Econometrica, 62, (1994), 467-475.

[21] J.J. Heckman and E.J. Vytlacil, Structural Equations, Treatment E¤ects, and Econometric Policy Evaluation, Econometrica, 73, (2005), 669-738.

[22] A. Abadie, Semiparametric Instrumental Variable Estimation of Treatment Response Models, Journal of Econometrics, 113, (2003), 231-263.

[23] G.W. Imbens and D.B. Rubin, Estimating Outcome Distributions for Compliers in Instrumental Variables Models, Review of Economic Studies, 64, (1997), 555-574.

[24] A. Abadie, Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models, Journal of the American Statistical Association, 97, (2002), 284-292.

[25] V. Chernozhukov and C. Hansen, Quantile Models with Endogeneity, Annual Review of Economics, 5, (2013), 57-81.

[26] D. Chetverikov, Testing Regression Monotonicity in Econometric Models, mimeo, (2013), UCLA.

[27] P. Carneiro and S. Lee, Estimating Distributions of Potential Outcomes Using Local Instrumental Variables with an Application to Changes in College Enrollment and Wage Inequality, Journal of Econometrics, 149, (2009), 191-209.

[28] S. Athey and G.W. Imbens, Identification and Inference in Nonlinear Difference-In-Differences Models, Econometrica, 74, (2006), 431-497.

[29] P. Yu, Marginal quantile treatment effect, mimeo, Department of Economics, University of Auckland (2014).

[30] R. Lalonde, Evaluating the Econometric Evaluations of Training Programs, American Economic Review, 76, (1986), 604-620.

Appendices

Results of estimations for Lalonde data.

Parameters	N (Sample size)	**350**
	P (Share of treated)	**0,4**

	Deciles	1st dec	2nd dec	3rd dec	4th dec	5th dec	6th dec	7th dec	8th dec	9th dec
Classic QTE reduced sample	Values Pop	0,0	0,0	943,0	1169,5	1093,5	1452,3	1802,5	2273,1	3197,8
	Values Boot	0,0	4,6	929,0	1324,9	1184,3	1493,5	2019,2	2248,1	3121,7

	Deciles	1st dec	2nd dec	3rd dec	4th dec	5th dec	6th dec	7th dec	8th dec	9th dec
Firpo Results	Values	0,0	0,0	711,0	21,0	1927,0	3879,0	4517,0	6027,0	5530,0

	Imputation Methods		Deciles of the completed distribution using different imputation methods								
			1st dec	2nd dec	3rd dec	4th dec	5th dec	6th dec	7th dec	8th dec	9th dec
Imputation Based results (Completed sample)	Mean Imputation (1)	Gen Mean	1765,3	1858,3	1822,2	1805,2	1805,2	1805,2	1805,2	1746,2	1789,0
		Conditional	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Random Imputation (2)	General	1692,0	1742,3	1774,3	1771,6	1810,8	1828,8	1846,2	1864,5	1871,5
		Hot deck	1715,8	1726,9	1787,8	1789,9	1800,0	1814,0	1829,8	1873,1	1906,3
	Deterministic linear regression Imp (3)		1553,2	1526,0	1526,9	1532,9	1554,4	1571,8	1618,6	1714,6	1789,0
	k-NN imputation (4)	V1	1782,2	1793,8	1776,3	1733,2	1702,4	1693,2	1747,6	1788,9	1811,8
		V2	1724,9	1771,5	1744,8	1716,8	1699,1	1704,6	1744,5	1789,8	1811,8
	Random linear regression Imp (5)		1482,1	1527,8	1564,2	1603,9	1630,1	1665,2	1685,7	1732,0	1779,2
	Multiple imputation (6)	Mice	1675,8	1631,7	1619,9	1601,9	1597,8	1598,8	1599,5	1642,4	1707,1
	Maximum Likelihood Imp (7) (MissMech package)	Normal	1455,4	1521,7	1546,8	1584,5	1610,1	1638,7	1669,8	1708,4	1757,6
		Dist-free	1494,8	1526,8	1569,0	1588,2	1623,2	1648,3	1668,6	1706,8	1771,1
	Propensity score matching Imp (8)		1282,6	1259,2	1213,0	1164,6	1014,9	620,4	362,5	287,6	220,7