UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG

A quantitative impact analysis of South African Child Support Grant on Child welfare

A Research Report submitted in partial fulfilment of the Degree of

Master of Commerce (Economic Science) in the School of Economic and Business Sciences,

University of the Witwatersrand

by

Lerato Eunice Mthembu
Student No: 561362

Supervised by Dr. Adeola Oyenubi

Word Count: 13 317

28 February 2019

ACKNOWLEDGEMENTS

I owe deep gratitude to many people who have supported me from my undergraduate years to now, as a master's student. Those who have supported by equipping me with the right resources to get me through varsity, and those who have supported me with the intangible things. Firstly, getting into varsity wouldn't have been possible if it wasn't for the foundation I received from ProMaths in my Matric year. A programme which dedicates its time in improving the performance of high school learners from underprivileged communities in South Africa. I qualified to do my tertiary studies because of the programme and the connections it provided me with. Therefore I extend my gratitude to Mr Setloagane Manchidi and Mr Tumelo Mabitsela. I also owe my gratitude to NSFAS, Studietrust and Rothschild for making sure that my tuitions were paid in those years.

This thesis would really not have been possible without my supervisor, Dr. Adeola Oyenubi who diligently and patiently walked me through this work. I have a deep sense of respect and appreciation for the time and effort he put into my work, for going over and above what was expected of him. The exposure I got from this experience has equipped me for my future endeavors.

My heartfelt gratitude goes to my mother, who raised me and my six siblings as a single parent. Her strength, teachings and prayers have carried me through all these years. The sacrifices that she has done for us are innumerable, the ones we got to see and those that we didn't.

Dedication:

I owe my deepest gratitude to my husband and our little girl. For allowing me to take time that was meant to be spent with them to work on the research. To my husband, I am humbled by the love you have shown me throughout. You stayed up all night with me in libraries, at work and at home while I was in pursuit of my dreams. Even more, you allowed me to leave work just so I could complete my studies. Thank you for everything that you have done.

Truly it all wouldn't have been possible if it wasn't for the God almighty for holding my hand through it all. ***To God be the Glory***

GLOSSARY

ABSTRACT

This paper investigates the impact of the South African Child Support Grant (CSG) on the beneficiary's height-for-age z-score (HAZ). We make use of data from the National Income Dynamics Study (NIDS), the wave 3 dataset. Using Propensity Score Matching approach the effect of the CSG is positive but statistically insignificant and relatively small. This popular technique rely on assumptions that often do no hold for observational studies. Furthermore, it is susceptible to misspecification of the propensity score equation which could bias the results. This paper therefore look into a technique that address these limitations and can assess the treatment effect robustly and with more precision.

We apply genetic matching algorithm, namely GenMatch. GenMatch is an iterative search algorithm that uses distance metrics to optimize covariate balance in the process of estimating the treatment effect. It automates the search process without the need of manual intervention to achieve the best balance. This algorithm is applied using two balance measures namely, the entropic distance metric and the standardized difference in means. The former compares distributions while the latter compares the first two moments (means and variances) of distributions. The results showcase the significance of utilizing a method that automates the process of optimizing balance and the influence of balance measures on the resulting treatment effect estimate. Specifically, we found that the estimate of the effect of the CSG is larger and more precise than the one reported in the literature.

1. INTRODUCTION

According to data released by Stats SA in 2015, more than half of South Africans still lived below the poverty line[1], with the most vulnerable group being children between the ages of 0-17 years. This suggests that children are the most affected group. As a result, the need for the South African government to address poverty, inequality and unemployment continues to be an urgent and compelling requirement. The government's effort in combating these socio-economic conditions are reflected in the amount of fiscal expenditure that goes towards the social grants. This study focuses on the Child Support Grant (CSG). Over 12 million CSG grants are paid monthly and this figure continues to grow as the number of beneficiaries increase each year. This explains why an ongoing evaluation of the CSG is of particular importance as the government expands on the monetary value that goes towards these grants annually. Therefore, it is critical to investigate whether the implementation of social grants has been effective in improving the wellbeing of the poor in the country.

27% of South African children under the age of 5 are nutritionally stunted (Mqadi, 2017). World Health Organization (WHO) defines children as stunted if their height-for-age is more than two standard deviations below the WHO Child Growth Standards median. This is as a result of insufficient food, poor nutrition for a long period of time, repeated infection, and inadequate psychosocial stimulation (WHO, 2015). Over two decades into democracy the prevalence of stunting has remained stubbornly high. Stunting signals that a child's growth and development is not progressing as it should. This has long run effects if the situation of the child does not change, and could lead to irreversible damages to the child. The child will be deprived from reaching their full cognitive potential, which impacts on schooling performance, and later on educational and employment opportunities (Merwe, 2017). Furthermore, studies show that these children are more likely to live in poverty as adults and are 3 times more likely to give birth to stunted children (Mqadi, 2017).

The most critical time for a child's development is the first 1000 days of their life (i.e. from conception to 2 years). This is the period which stunting could be prevented. This study

[1] At that time (2015) the food poverty line was R441 per person per month (pppm), lower-bound poverty line was R647 (pppm) and the upper-bound poverty line was R992 (pppm). Adjustment to these figures was performed in 2017 to R531, R758 and R1 138 respectively (Lehohla, 2017).

investigates the impact of the South African CSG on the child beneficiary's height-for-age for children under the age of 14. The reason for this cut off age was not only to produce results that are comparable to the studies we follow (Coetzee (2011 & 2013) and Oyenubi (2018)), but to also consider possibilities of catch up growth that may occur after the age of 2. Desmond and Casale (2017) provide evidence in their study, that there are chances for children who are stunted at 2 years to recover from stunting by the age of 5.

The CSG is designed in such a way that treatment is received by children who meet the eligibility criterion, therefore the assignment mechanism is not random. A straightforward estimation of the treatment effect between children who are grant recipients and those who do not receive the grant is prone to selection bias. Bias could arise because the apparent difference in the outcome between the two groups may depend on the characteristics that influenced whether or not the child received the grant. Hence the need to minimize bias by using an econometric method that can create conditions that are similar to the ones we would expect from a randomized experiment. Therefore, to assess the treatment effect of the programme we have to find the counterfactual outcome with minimum bias. This outcome represents what would have been the outcome of the children that received the grant had they not received it. Thus, the comparison group is used to find matches that will be comparable on all observed covariates to the treatment group. In our analysis, this is constructed by using children that did not receive the CSG but satisfy the eligibility conditions.

This paper follows the work of Coetzee (2011 & 2013) and Oyenubi (2018) who both made use of wave 1 dataset from the National Income Dynamics Study (NIDS) data. They estimate the impact of the CSG on height-for-age z-scores (HAZ) as the outcome variable. Our study assesses whether the findings in both studies remains valid in wave 3 of NIDS data. Both studies emphasize that the Propensity Score Matching (PSM) is not effective in evaluating the effect of the CSG. The results under this approach (for wave 1 data) with binary treatment variable, shows that there is a positive effect albeit small and not statistically significant. Ceasing further investigation from this point would give inaccurate conclusions whether the cash transfers have been able to achieve the objectives set out for the CSG. The authors henceforth extended their studies to explore other estimators that they believe capture the treatment effect more efficiently. Coetzee (2011 & 2013) opt for modeling CSG as a continuous treatment while Oyenubi (2018) maintain the assumption of binary treatment but use a better

matching strategy. This study looks into the techniques proposed by Oyenubi (2018) which we will discuss later in this section.

In order to re-estimate the treatment effect in Coetzee (2011) and Oyenubi (2018), we recreate a binary treatment variable that is comparable to these studies. We only consider children under the age of 14 who were either from households where all children received the grant or all children did not receive grant but met the eligibility conditions. By so doing, we eliminated the control group members who partially benefited, by virtue of living in a household with children who are recipients. We also drop treatment group members who had their benefit diluted, by having to share the benefit with other children who are non-recipients of the grant. Furthermore, we refined the treatment group, as in Oyenubi (2018), to only include children who have at minimum received treatment for 34% of their lives. This was done to eliminate members that would dilute the treatment effect estimates because of low dosage[2]. Our data consist of 450 control and 2960 treatment members before matching, this gives us a ratio of 6.5 treatment to control member. We discuss the implication of this ratio in this section and in our results.

The PSM method is used as the baseline technique which will be later compared with other estimates. This method was proposed by Rubin and Rosenbaum (1984) as a solution to reduce the bias when estimating the treatment effect. The PSM has underlying assumptions which provide a way to achieve balance so that causal inference can be made. The assumptions make sure that, after conditioning on the propensity scores (the probability of an individual receiving treatment) of individuals none of the selection bias remains (Coetzee, 2011). We apply this method under the binary treatment case to re-estimate the findings from Coetzee (2011) and Oyenubi (2018) for wave 3 of the NIDS dataset. We found no convincing results which confirms results in the aforementioned studies for wave 1 data. The treatment effect (in the binary treatment case) is 5% of the standard deviation and statistically insignificant at 5% level. Moreover, after matching under the PSM the number of control group members in the matched sample reduced to 281, consequently the ratio of treated to control members increased to 10.5. As mentioned in (Oyenubi, 2018), this in turn give rise to significantly inflated standard error. Failure of the PSM to provide statistically significant results compels us to seek other statistical

[2] By dosage we mean duration of receipt of the CSG

means to assess the effect of the CSG. As mentioned, if further methods are not explored these results would suggest that the programme is not as effective as we would expect.

Coetzee (2011 & 2013) and Heinrich et al. (2012) argue that the inconclusive results can be attributed to the fact that the binary PSM ignores the variation in dosage that the child received of the CSG. Coetzee (2013) advanced her analysis to consider a continuous treatment estimator that takes into account the variation in dosage. The author use the Generalized Propensity Score (GPS) to estimate the dose-response function of receiving the CSG. The author recorded positive and statistically significant gains in HAZ for children who received the grant for 30-60% of their lives[3]. The maximum gain was for children who have been receiving the grant for 40% of their lives in comparison to those who received it for 30% of their lives. The recorded increase in HAZ is 4% of the standard deviation. Although Coetzee (2013) reported statistically significant results under the GPS approach, the size of the effect remains relatively small. Furthermore, the effect was only positive for doses between 30 and 60% of the child's life. Therefore, one would argue that the findings in Coetzee (2011 & 2013) do not sufficiently answer weather the issue of insignificant results lie with the method used or the binary treatment assumption?

This paper maintains the condition of binary treatment. However, we took steps to mitigate the effect of treatment dosage in the definition of treatment and control group members. Our data excludes control members who partially benefit, treated members who received less than 34% dosage of CSG and additionally those whose benefits is diluted. We apply genetic matching algorithm (namely GenMatch) to estimate the average treatment effect on the treated (ATT). This algorithm is a multivariate matching method that uses an iterative search algorithm to optimize covariate balance. The focus on optimizing balance is important for bias reduction in treatment effect estimates. Therefore, this automated process is expected to improve the covariate balance in the matched data and the precision of the estimate (Oyenubi, 2018). Our results are similar to the results of Oyenubi (2018) in wave 1, the estimate of the effect of CSG under GenMatch is larger than the effect estimate under PSM method in wave 3. We find an estimated effect of 14% of the standard deviation when the default balance measure of GenMatch is used. Moreover the estimates (under all balance measures used) are statistically significant at 1% level when matching is not accounted for. GenMatch conveniently automates

[3] Note that this informs the choice of our 34% cut-off

the search process without the need of manual intervention to iteratively check covariate balance in the matched sample. The implication is that GenMatch is robust to the misspecification of the propensity score equation (Diamond & Sekhon, 2013). This is in contrast to the PSM method that require the researcher to manually adjust the specification of the propensity score equation until acceptable level of balance is achieved. The results we find under this approach highlight the significance of using a better matching technique.

The GenMatch algorithm enables the researcher freedom to define the balance measure to be used. To exploit this, we applied the Entropy distance metric proposed as the balance measure by Oyenubi (2018). This measure is a distributional balancing measure and it is expected to better improve covariate balance relative to when the default GenMatch balancing measure is used (Oyenubi, 2018). The measure takes into consideration the shape of the covariate distribution in both treatment arms and can be applied for both discrete and continuous variables. Using a measure that considers covariate distribution rather than a few moments leads to stronger effect estimates (Oyenubi, 2018). We record an effect of 31.5% of the standard deviation when the entropy metric replaced the GenMatch default balance measure. These results were statistically significant at 1% level when matching is not accounted for. The apparent increase in size of the effect can be attributed to the change in balance measure used. Our analysis, demonstrates the significance of balancing and matching measures as important in estimating the treatment effect of a policy or programme.

This paper proceeds as follows: Section 2 will go through the literature review. Which will cover the history of the South African CSG and the various studies that investigated the impact of the programme. Then Section 3 will explain in detail the methodologies that we applied in our analysis. This is followed by section 4 which details the construction of our data. Thereafter, we cover our results in section 5. Then lastly, we have the concluding section, which discuss our findings and some recommendations.

2. LITERATURE REVIEW

2.1. South African Child Support Grant Background and the Caregiver Motivation

In 1994 the South African Government inherited a fragmented social security system that was racially bias and constructed solely in the interests of the apartheid beneficiaries (Overseas Development Institute, 2006). The system needed extensive policy and administrative reform to rebuild it as an anti-poverty strategy that is inclusive of the vulnerable and majority living in marginalised conditions. The country was faced with several challenges in administering the programme and making sure that social assistance grants benefit the intended parties, hence much has changed over the years in trial and error to improve the system to its current level (Kelly, 2017). Current figures suggest that over 17 million social grants are paid monthly and almost 70% of these are Child Support Grants. Social Grants are instrumental in improving the standards of living by redistributing wealth to create a more equitable society.

The Child Support Grant (CSG) was introduced in 1998 and has since undergone administrative and several policy changes to redefine eligibility over the years. As a result, coverage has expanded as more and more people qualify for it. Age eligibility was initially restricted to children under 7 years old when the programme was reintroduced in 1998. It was subsequently adjusted in 2003 to accommodate children up to their 9th birthday, in 2004 it was extended up to 11 years, and in 2005 it went further up to their 14th year. On the 1st of January 2010, eligibility was again adjusted up, to cover children born after 1 January 1994 and remaining eligible until their 18th birthday, whereas those born prior to the date fell out of the programme and lost eligibility at 14 (d'Agostino, et al., 2017).

Caregivers who are eligible to receive the cash transfer are required to meet the means test[4] that allows them to receive the grant on behalf of the intended beneficiary. Currently the value of the Child Support Grant is R410 per month, when it was introduced it was merely R100 per month (d'Agostino & Scarlato, 2016)[5]. The budget spent towards social grants rises annually.

[4] Household income was capped at the nominal level of R800 in urban areas and R1,100 in rural areas for 10 years. However, in 1999, the government changed this rule to one which considered only the income of the primary caregiver plus her/his partner (d'Agostino, et al., 2017). On the 1st April 2017, an unmarried caregiver whose annual income is below R45,600 qualified for the social grant. The means for married caregivers their joint annual income is required to be below R91,200 to qualify for the social grant.

[5] In the current budget (2019) the CSG is set to increase to R420 in April and R430 in October.

This continues to grow over the years as child population rises in conjunction with the efforts by the government to get eligible children registered.

The disruption caused by apartheid economy not only brought intergenerational poverty but it still shapes many family structures in South Africa today. This has perpetuated the psychology of single parenthood that emanated from the displacement of families in that era, hence paternal roles continued to be absent in poverty stricken communities. Today the majority of children live apart from their biological fathers and often the adult women of the household provide care for the children and other home duties (d'Agostino & Scarlato, 2016). A child's primary caregiver could therefore be any elder in the family that looks after and lives with the child. This implies that the role of a caregiver is not necessarily restricted to the parents of the child. As a result, many unobserved factors do contribute to motivation for a caregiver to register the child to receive CSG. Their earnestness to apply for the grant on behalf of the child would determine the duration which the child receives the CSG and by extension the effect of the CSG.

To receive the grant the caregiver must supply several documentations to proof that they are the primary caregiver, furthermore the caregiver need to provide proof of immunization and proof of efforts to secure employment or to join a development programme (Kelly, 2017). This becomes administratively burdensome and could be discouraging for the caregiver to get the child registered. Especially if they live in settlements whereby access to public services and necessary facilities is difficult and prolongs the process of getting the child registered. The effect of these factors requires one to exercise caution in handling the bias that may arise from caregiver motivation in estimating the impact of CSG. It is important that the estimators used in this kind of setting are free from bias as much as possible (Oyenubi, 2018). We discuss how we control for this unobservable variable, caregiver's motivation, later under the data section.

2.2. Review of the studies that investigated the CSG

Heinrich, et al. (2012) conducted a study using PSM, Generalized propensity score (GPS) and Double Difference (diff-in-diff) method to evaluate the impact of the CSG on several outcomes to measure effectiveness of the programme. The study found that Children who were enrolled in the CSG at birth completed significantly more grades of schooling than children who were

enrolled at age six. The authors argue that the binary PSM ignores the effect that the dosage of grant the child received has on the child's welfare. They provide further evidence that the receipt of the CSG reduces adolescent absences from school, particularly for male adolescents and that CSG significantly reduces six main risky behaviors – sexual activity, pregnancy, alcohol use, drug use, criminal activity and gang membership (Heinrich, et al., 2012).

Similarly, Coetzee (2013) uses the PSM method to analyse data from the first wave of the NIDS data. The author investigates the impact of CSG on a range of welfare variables, for children under the age of 14. Under the binary treatment construction, the author found no convincing results (i.e. results were statistically insignificant). For this reason Coetzee (2013) then extended the analysis by using the GPS method. This method assumes weak unconfoundedness and considers the effect of different levels of treatment. It is an extension of propensity score matching methods which was developed by Hirano and Imbens (2004). The method allows for impact assessment of the duration of programme participation on the outcomes of interest. Hirano and Imbens (2004) show how it is possible to remove the bias relating to treatment dosage by introducing an unconfoundedness assumption similar to the one upon which the binary PSM method is based. The results improved when the binary assumption was relaxed by considering the duration the child was enrolled in the programme (dosage). Coetzee (2013) recorded positive and statistically significant gains in HAZ for children who received the grant for 30-60% of their lives. However, her study found maximum gains in HAZ for children who received CSG for 40% of their live relative to those who receive the grant for 30% of their live. This effect is relatively small, at 4% of the standard deviation.

Oyenubi (2018) who followed Coetzee (2013) closely proposes a different matching approach, the GenMatch, along with the entropy metric (a balance measure). The author compares the performace of this approach to the PSM method. He explores the binary case further, however refines the treatment group to only include beneficiaries that received CSG for at least 34% of their lives. This was to remove members in the treatment group that would decrease the average effect of the treatment, and as a result, dilute the effect of the CSG (Oyenubi, 2018). This also eliminates the assumption that all eligible children are enrolled in the programme once the policy change is implemented. His analysis proved that CSG has a positive impact on the HAZ of children who received the grant for at least a third of their lives. The author found an effect of 16% of the standard deviation, when using the GenMatch algorithm with the Entropy metric. His results also illustrated that the GenMatch measures the treatment effect with more precision

than the PSM method. Moreover, he demostrated that the entropy distance metric captures balance better than other alternatives applied in literature.

Both Coetzee (2013) and Oyenubi (2018) highlighted how motivation of the caregiver to register the child for the programme influences the child's treatment dosage and consequently the effectiveness of the CSG. As demonstrated by Heinrich et al. (2012), children who enjoy high dosage of CSG exhibited better results than children who were almost similar to the control group in terms of the dosage received. Heinrich, et al., (2012) mention that some children who should have been enrolled in the CSG following the extension of age eligibility were disconnected from the programme. This was because of lack of knowledge from the social welfare offices about the policy changes. Moreover, the high burdens placed on caregivers in the re-application process discouraged participation of poorer households. The administrative burdens may vary geographically or politically, and thus have an influence on the take-up rates so that take-up rates may be different across these factors (Heinrich, et al., 2012). We can say that motivation is influenced by myriad factors. As a result, this unobserved factor (caregiver motivation) is an important source of bias that influences the effectiveness of the CSG. As mentioned, we elaborate on how we will control for caregiver motivation in the later section.

d'Agostino, et al., (2017) took a different approach from the aforementioned authors, by examining the causal effects of the programme across birth cohorts using a Regression-Discontinuity Design (RDD). The authors considered the change in age eligibility for the CSG that was implemented on 1st January 2010, for children born on or after 1 January 1994 to qualify until their 18th birthday whereas those born before 1994 lost their eligibility at age 14. This resulted in discontinuity with different birth cohorts, those born after 1 January 1994 and those born before the date. Although their results showed that the CSG did not lead to significant changes in the dietary habits of the beneficiary households, the study did prove that the cash transfers provided by the CSG significantly increased the total food expenditure per adult equivalent (d'Agostino, et al., 2017). Overall, these studies show that the CSG has some positive effects on a range of outcome variables especially those that have to do with nutrition.

It is imperative to continue the efforts of finding statistical techniques that can assess treatment effect with the most precision and minimum bias. The South African CSG is a complex case to assess. The cash transfer is not conditional to any outcome it seeks to promote and it is often shared with other household members. This dilutes the potential effect it could have on child's

wellbeing. Moreover, the length of time the child receives the CSG (dosage) varies amongst the recipients. This is influenced by unobserved factors[6] one which we highlight is caregiver motivation. The inclusion of this unobserved variable to the model causes balancing property not to hold. This is as a result of the levels of caregiver's motivation being vastly different between the treated and untreated children (Coetzee, 2013). Therefore the complexity of the CSG requires the researcher to place more caution when estimating the effect of the grant on beneficiary's wellbeing. This justifies how members of our treatment arms were defined (as discussed in the data description section) to support the consideration of a binary treatment case in assessing treatment effect. However, we use a better matching method that automates the process of optimizing covariate balance in estimating the treatment effect. It also motivates the use of a balance measure that not only considers discrete and continuous variables, but also different shapes of the covariate distribution. We discuss all these techniques in detail in the later section.

3. THE DATA DESCRIPTION

3.1.Controlling for unobserved variable - The Caregiver Motivation

Caregiver's motivation is considered an unobserved variable. This variable plays an important role in determining the dosage of treatment the child receives. Caregiver motivation is considered to be correlated with the duration of time the caregiver took to apply for the CSG. Less motivated caregivers delaying longer to apply in comparison to highly motivated caregivers who would apply early. Oyenubi (2018) argues that a delay in applying for the programme affects the impact via the extent of time the child benefits from CSG. If this unobserved variable is not accounted for in our model, this can bias our results.

To best control for the motivation variable, we used censored regression to estimate the expected delay as proposed by Oyenubi (2018). The observed delay for the control members is based on the time it took to apply for the CSG, which is the time between the child's birth and date of application. Our data had some members in the control group with missing date of application. As a solution for the missing information, we make use of the child's age as a proxy for observed delay. However, this may not be an accurate indication of delay. Exlcuding this pool would further reduce the number of observations in our control group. We therefore

[6] Refer to the Appendix Table 10 for the reasons provided by caregiver for no application of child support grant

use censored regression where these observations are regarded as being right censored with a variable censoring point that is equal to the age of the child.

The observed delay for the treated group was taken as time between child's birth and the date they started receiving grant. The way in which this variable is created, already shows how different it is between the treated and untreated group. The mean value of delay is more likely to be high for the control group than it would for the treated members, and this would filter into the estimation of the caregiver motivation. Note in Table 1.2 that the mean value of caregiver's motivation is negative for control group members. This suggests that some of the comparison group children may not be comparable with the treated group children in terms of caregiver motivation. This will have implication for the ratio of treated to control observations.

We estimate expected delay using censored regression as a function of observed delay, caregiver's gender, child age and the geographical area which they reside. Thereafter, the difference between observed delay and expected delay is calculated and subsequently standardized. This then produced what would represent the unobserved variation in caregiver's motivation to apply for CSG.

Table1.2. Summary of Caregiver's Motivation for the different Treatment status

Treatment Status	**Observations**	**Mean**	**SD**	**Min.**	**Max**
Treated	2 960	0.078	0.430	-1.799	1.827
Control	450	-1.520	0.886	-3.236	1.246

3.2.Data Summary Statistics

Our analysis makes use of wave 3 dataset, which was collected in the year 2012, by the National Income Dynamics Study (NIDS), this data is a national representative survey. For the purpose of this study, we only retained observations for children under the age of 14 who were eligible to receive the CSG at the time when the survey was completed. This is in order for the assessment to be comparable to that of Coetzee (2013) and Oyenubi (2018). In addition to that, we filtered it down to those who lived in households whereby either all children benefited or did not benefit. This is the base which was used to construct our treatment and control group. Furthermore, we refined the treatment group to only include children who have at least received treatment for 34% of their lives. This was to eliminate members who would distort our

treatment effect. The numbers then dwindled down to 3,410 children of which 2,960 was treated and 450 was not treated. This ratio of control to treatment members (i.e. 6.5) means that each control observation will be matched about 7 times on average. This in turn means that the standard error of the estimate will be inflated (Oyenubi, 2018) and explains in part why PSM's estimate is not significant. Moreover, this could bias our result, if the sample size is too small for the propensity score to eliminate the conditional bias (Diamond & Sekhon, 2013)[7]. We elaborate further on this implication when the PSM method is applied.

Height-for-age z score (HAZ) is used as an outcome variable in our analysis as it is widely considered a good measure to capture long term effect of malnutrition in children. Coetzee (2013) mentions that this variable is a suitable measure of nutritional wellbeing and long-run health of children, measuring since birth the accumulated investment in children's nutrition and health. We summarized the data in Table 1.1 to show the means and standard deviations for all the variables in use. The table only contains the eligible subjects. Column 2 contains the summary of all the treated children and columns 3 contains the summary statistics of the control group. Whilst, column 4 shows the difference in means of characteristics across treatment status for the treated and control. Based on the t test (*ttest* command on stata) the means difference for the two groups appear to be statistically significantly different from 0 in most of the covariates in use. The means were only similar for caregiver's years of education, child gender and household head gender.

HAZ is slightly different between the untreated and treated children. We note some significant difference between the two eligible groups. Children in the control group have caregivers who are more likely to be married, of other ethnic background than African, older and employed. Similarly, these children live in households that are more likely to have access to water, telephone, with lesser number of household residents and living on higher household expenditure than children in the treated group.

[7] This authors show that GenMatch reduces bias and Mean Square Error relative to PSM which makes it a better approach under these conditions.

Table 1.1.: *Data Summary Statistics*									
	All				***Eligible Treated***		***Eligible Control***		***Treated vs. Control***
Variable	*Mean*	*SD*	*Min*	*Max*	*Mean*	*SD*	*Mean*	*SD*	*Mean diff*
Caregiver Characteristics									
marital	0.26	0.44	0	1	0.23	0.42	0.45	0.50	0.23***
years of education	8.49	3.65	0	15	8.46	3.63	8.69	3.76	0.23
african	0.88	0.32	0	1	0.90	0.30	0.77	0.42	-0.13***
coloured	0.11	0.31	0	1	0.09	0.29	0.18	0.39	0.08***
asian	0.00	0.06	0	1	0.00	0.04	0.01	0.11	0.01***
employed	0.29	0.45	0	1	0.27	0.45	0.41	0.49	0.14***
gender	0.95	0.21	0	1	0.96	0.18	0.86	0.35	-0.11***
Age	36.54	12.72	15.3	91.7	35.92	12.29	40.62	14.58	4.69***
motivation	-0.13	0.75	-3.24	1.83	0.08	0.43	-1.52	0.89	-1.60***
Child Characteristics									
gender	0.51	0.50	0	1	0.50	0.50	0.52	0.50	0.02
relationship with caregiver	0.80	0.40	0	1	0.80	0.39	0.71	0.45	-0.09***
HAZ	-1.03	1.60	-5.93	5.99	-1.06	1.59	-0.82	1.67	0.24***
Age	5.40	2.66	0.4	10.2	5.42	2.65	5.27	2.76	-0.16
african	0.89	0.32	0	1	0.90	0.29	0.76	0.43	-0.14***
coloured	0.11	0.31	0	1	0.09	0.29	0.19	0.39	0.10***
asian	0.00	0.06	0	1	0.00	0.04	0.01	0.11	0.01***
HH Characteristics									
electricity	0.79	0.40	0	1	0.78	0.41	0.86	0.35	0.08***
water	0.58	0.49	0	1	0.56	0.50	0.72	0.45	0.16***
telephone	0.03	0.17	0	1	0.02	0.15	0.10	0.30	0.08***
expenditure	2829.63	3755.52	221	69747.95	2510.65	2685.86	4928.52	7380.50	2417.87**
household size	6.43	3.23	2	29	6.66	3.33	4.91	1.88	-1.75***
household head Gender	0.51	0.50	0	1	0.50	0.50	0.52	0.50	0.02
geography	0.40	0.49	0	1	0.38	0.49	0.54	0.50	0.16***

**** Statistically significant at α=0.01 ** statistically significant at α=0.05*

4. METHODOLOGY

4.1. Propensity Score Matching (The PSM)

The propensity score matching method was proposed by Rubin and Rosenbaum (1983) as a method to reduce the bias in the estimation of treatment effects with observational data. The estimation of the treatment effect may be biased due to confounding factors. The PSM offers a way to estimate treatment effects while controlling for these confounding factors. The success of this method depends on how good a match the control group is to the group that received treatment. If the distributions of observed confounders are not similar after matching on an estimated propensity score, the propensity score will fail to remove the conditional bias (Diamond & Sekhon, 2013).

4.1.1. Theoretical Framework and Underlying Conditions of The Propensity Score Matching

The propensity score is the probability of each child within the sample receiving the grant. In this study, the propensity score is estimated using the logistic regression (logit). The sample in use only considers children who are eligible to receive the grant. This makes the group of children that are treated and untreated similar in the observed characteristics. The similarity in the observable covariates between the two groups would potentially assist to find overlap in the propensity scores of treated and control members, as opposed to members from the ineligible population (Coetzee, 2011).

The main idea here is to solve the missing information problem in order to find the average treatment effect on the treated (ATT). We have to find the counterfactual outcome that would represent the outcome had the treated not received treatment. In order to achieve this, we identify members in the control group that are similar in observed covariates (i.e. with the least difference in propensity scores) to the treated members. These members are matched following a particular matching estimator. In this study we follow 1:1 nearest-neighbour (NN) matching estimator with replacement. For us to be able to capture the ATT with reduced bias using the PSM method, the following assumptions have to hold.

Firstly, there has to be a region of overlap between the propensity scores of the treated and control units in order to have a meaningful estimation of the treatment effect (Coetzee, 2011). This has been referred to in the literature as the "region of common support".

- *Overlap or Common Support Condition:* All subjects have a positive probability of being treated, that is;

$$0 < P(D = 1|X) < 1 \qquad (2)$$

The estimate can be seriously biased if a sufficient overlap does not exist between the propensity scores of the treated and untreated units (Coetzee, 2011). The second condition required to hold is;

- *Unconfoundedness Assumption:* This is also referred to as the conditional independence assumption. Given a set of observable covariates X which are not affected by treatment, potential outcomes are independent of treatment assignment, that is;

$$(Y_0, Y_1) \perp D|x \quad \rightarrow \quad (Y_0, Y_1) \perp D|p(x) \qquad (1)$$

Rubin and Rosenbaum (1983) show that the conditional independence condition also holds with the use of propensity scores. The propensity score must be such that the conditional distribution of the covariates in *x,* given the propensity score *p(x),* must be the same for treated *(D=1)* and control *(D=0)* units (i.e. $D \perp x|p(x)$). This also implies that covariate distribution will be *balanced.* Therefore, for units with the same propensity score, selection into treatment is random, and these units should be identical in terms of the observable characteristics in X (Coetzee, 2011). This assumption suggests that treatment assignment and the outcome are conditionally independent given the true propensity score (Diamond & Sekhon, 2013). It also takes into account the fact that in observational studies, the observable variables that affect the outcome may be distributed differently across treatment arms thus confounding the treatment effect (Diamond & Sekhon, 2013). That is to say, the same factors influencing selection into treatment could in turn potentially influence the outcome variable. Hence conditioning on the propensity score will remove all biases due to observable covariates.

If the above conditions are met, it allows for the untreated members to be used to estimate an unbiased counterfactual outcome for the treatment group. Since in a true sense the

counterfactual mean for the treated is not observed, i.e. $E[Y_0|D = 1]$. Evoking the above identifying assumptions that solve selection bias allows us to use the mean for the untreated, i.e. $E[Y_0|D = 0]$, as the counterfactual mean. Therefore, we can estimate ATT as follows;

$$ATT = E[Y_1|D = 1] - E[Y_0|D = 1] \quad (3)$$

4.1.2. The Problem with the PSM method

When we control for caregiver motivation under the PSM approach it become difficult to find balance between the treatment arms. This is due to how this variable is created, the levels of caregiver's motivation would be vastly different between treated and untreated children. Therefore the inclusion of this variable causes the balancing property not to hold. As noted before, caregiver motivation is a very strong predictor of treatment status which causes the overlap assumption to be violated (Coetzee, 2011)[8]. This could in turn could increase bias. If we exclude this unobservable variable, the model may not comply with the conditional unconfoundedness assumption discussed above (Coetzee, 2011). The inclusion or exclusion of this variable give rise to bias when the PSM method is applied. The balancing condition under the PSM method is commonly assessed by comparing the means and variances of covariates (Oyenubi, 2018). This is a limitation in the use of PSM when working on complex cases like the CSG, especially when the covariates have nonellipsoidal distributions.

Generally the correct propensity score model is unknown, and if the model is misspecified, it can increase bias even if the selection on observables assumption holds (Diamond & Sekhon, 2013). Considering that there is some human element involved when selecting observable covariates to estimate the propensity score, this opens up possibilities of misspecification. Diamond and Sekhon (2013) argue that if after matching on an estimated propensity score, the distributions of observed confounders are still not comparable across treatment arms, the propensity score must be misspecified or the sample size is too small for the propensity score to eliminate the conditional bias. In other words, a misspecified propensity score model may increase the imbalance of some observed variables post matching, even worse if the covariates

[8] Refer to Figure 3.1 and 3.2 in the Appendix (C). The graphs show the common support (overlap) assumption violated when the caregiver's motivation is included, and the difference when it is not included.

have nonellipsoidal distributions. Considering that the propensity score is a balancing score, covariate imbalance after propensity score matching is of concern. It is therefore the reason why it is important to evaluate the covariate balance post matching and iteratively modify the propensity model accordingly until the best balance is achieved (Rubin & Rosenbaum, 1984). The PSM requires manual intervention for this to be achieved. Without a computational processing that will automate the repetitive search for maximum balance, it would be difficult to manually find the ultimate best balance. Diamond and Sekhon (2013) argue that GenMatch can solve this problem.

4.2.Genetic Matching Algorithm

4.2.1. The Algorithm

To explain how the *Matching* occurs in GenMatch we need to first discuss the process which the *genetic algorithm* follows in achieving the optimal balance. The GenMatch runs a *genetic search* which is an automated *iterative algorithm*. The algorithm performs optimization in a situation where the objective function does not satisfy usual properties like convexity and continuity (Ponsich , et al., 2013). An evolutionary algorithm (EA) is a class of metaheuristic that relies on emulating Darwinian's theory (i.e. "survival of the fittest" approach) to optimize an objective function. The process starts with a population of random solutions to the problem at hand. Each of these solutions are made up of elements called genes (i.e. a solution is a combination of genes). These solutions are evaluated with the loss or fitness function and ranked according to their performance. Then this population of solutions must evolve into a new and better solution to the problem at hand, i.e. they form another generation of solutions. This is achieved by a number of genetic operations namely elitism, crossover and mutation (among others). Elitism allows the best performing solutions (i.e. a proportion of the previous solutions) to move into the next generation without any change. Crossover combine solutions from the old population to form new solutions (offsprings) while mutation randomly changes genes in existing solutions to form new solutions. These genetic operations improve the average fitness in each successive generation relative to the previous one. This process continues until it is no longer possible to improve the fitness values of the current population.

In the context of matching Diamond and Sekhon (2013) apply this algorithim to the problem of finding optimal balance. This problem can be viewed as that of optimizing an objective

function that does not have the usual properties. We are interested in optimizing balance in a multivariate case. The loss function is therefore the objective function to be optimized is a function of a balance measure. The default loss function which is used in this study is called lexical optimization[9].

4.2.2. GenMatch Formulae

We first have to discuss what *Mahalanobis distance (MD)* does because the GenMatch is a generalized form of this distance metric. The distance metric measures the multivariate distance between individuals in different groups. Specifically, the distance between the covariates X_i *and* X_j for individuals from the two treatment groups, is given by.

$$MD(X_i, X_i) = \sqrt{\left(X_i - X_j\right)^T S^{-1}(X_i - X_j)} \qquad (4)$$

Where S is the sample covariate matrix of X whereas X^T is the transpose of the matrix X (Diamond & Sekhon, 2013).

The GenMatch is a generalized form of Mahalanobis distance. Unlike equation 4 GenMatch make use of

$$GMD(X_i, X_i, W) = \sqrt{\left(X_i - X_j\right)^T \left(S^{-\frac{1}{2}}\right)^T W S^{-\frac{1}{2}}(X_i - X_j)} \qquad (5)$$

The difference here from the normal MD is that there is a weight matrix *W* added. $S^{-1/2}$ is the Cholesky decomposition of S, that is $S=(S^{-1/2})(S^{-1/2})^T$, and $(S^{-1/2})^T$ is the transposition of it. The diagonal elements in the weight matrix *W* are the only non-zero elements, the others are restricted to zero (Diamond & Sekhon, 2013). The diagonal elements is the weight attached to each covariate to produce the distance $GMD(X_i, X_i, W)$ this distance is then used for matching (similar to propensity scores). These weight matrices represent the random solutions mentioned earlier. These solutions are then used to estimate equation 5, and for each solution, balance is checked in the matched sample produced by using the distance defined in 5.

[9] To be explained latter

Lexical optimization sorts all balance statistics from the most discrepant to the least, a solution that comes from the weight matrix W_a is preferred to another solution W_b, if W_a produces more balance in the matched sample. If multiple sets of weights (W) result in the same maximum discrepancy, the second largest discrepancy is examined to choose the best weight. This process continues iteratively until all ties are broken (Sekhon, 2011). All covariates are assigned weight based on their contribution in achieving the best overall balance. This is the main aim of the GenMatch, to find covariate weights which optimize post-matching covariate balance.

The researcher can decide to match not only on covariates but propensity score as well. Therefore, *X* in equation (5) may be replaced with *Z*, where *Z* is a matrix consisting of both the propensity score, *p(X)*, and the underlying covariates *X* (Diamond & Sekhon, 2013). The propensity score will be treated similar to the other covariate, it will be allocated with the appropriate weight based on its contribution towards achieving the best balance. If the optimal balance is achieved by matching only on propensity score, other variables will be assigned with zero weight. Therefore, the GenMatch will be equivalent to propensity score matching (Diamond & Sekhon, 2013). It is also possible for the propensity score to be assigned zero weight, in this case GenMatch is equivalent to Mahalabolis distance matching. Our analysis operated the GenMatch exclusive of the propensity score. We deliberate on this in our results.

4.3.Entropy as the Balancing Measure

In recent times, there has been a growing attention in developing nonparametric measures to capture the difference between conditional distributions in a more complex, and usually nonlinear system. Information theory offers useful concepts that can be used to measure the divergence or "distance" between probability distributions. One concept that is widely used in information theory, is the entropy, which possesses various desirable properties. It is considered an appropriate measure of distance and it performs well in identifying dependence even in nonlinear time series (Granger, et al., 2003). Furthermore, this measure can be applicable for both discrete and continuous variables, and it is considered a flexible nonparametric measure for conditional dependence (Granger, et al., 2003). Moreover, the entropy is dimension-less as it can be easily applied to univariate and multivariate contexts. Interested readers can refer to the work of (Oyenubi, 2018) on the concept of relative entropy,

which this paper makes use of. Here we will only discuss the two-step procedure required when applying the entropy measure (S_ρ) to compare two distributions.

First, we estimate the densities to be compared, f_1 and f_0, then measure the distance between them. The resulting distance measure are sensitive to any errors emanating from the first step (Oyenubi, 2018). The equation below is the formulae for Entropy

$$S_\rho = \frac{1}{2}\int_{-\infty}^{\infty}(f_1^{\frac{1}{2}} - f_0^{\frac{1}{2}})^2\, dx \qquad (7)$$

Following Maasoumi & Wang (2013) and Granger, et al. (2003), we consider a kernel based implementation of the above equation (7), which gives us kernel density estimates of f_1 and f_0, that is, $\hat{f}_1^{\frac{1}{2}}$ and $\hat{f}_0^{\frac{1}{2}}$ respectively. Therefore transforms the equation to

$$\hat{S}_\rho = \frac{1}{2}\int_{-\infty}^{\infty}\left(\hat{f}_1^{\frac{1}{2}} - \hat{f}_0^{\frac{1}{2}}\right)^2 dx \qquad (8)$$

In order to achieve this, the choice of bandwidth and kernel becomes essential in assuring that the distance measure in the second step is a reliable one (Oyenubi, 2018). Similar to what these authors, Maasoumi & Wang (2013) and Granger, et al. (2003), we use the Gaussian kernel and a robust version of the "normal reference rule-of-thumb" bandwidth ($= 1.06\min(\sigma, \frac{IQR}{1.349})n^{-\frac{1}{5}}$) to get the kernel density estimates. σ represents the standard deviation of the variable whose density is being estimated and *IQR* is the interquartile range. The choice of kernel happens not to be as essential as the choice of bandwidth in the kernel density estimation (Oyenubi, 2018). For discrete distributions we then use

$$\hat{S}_\rho = \frac{1}{2}\sum(\hat{p}_1^{\frac{1}{2}} - \hat{p}_0^{\frac{1}{2}})^2 \qquad (9)$$

Oyenubi (2018) argue that since this measure is sensitive to differences in distribution it is a better measure to use to quantify imbalance than alternatives that are being used in the literature.

5. MAIN RESULTS

Our analysis is performed using two different estimators, the PSM and GenMatch, to capture the treatment effect of the programme using wave 3 of the NIDS data. Under GenMatch we use three balance measures the default measure, Standardized Difference in Mean (SDM) and entropy distance metric. Here we go through our findings from the different analysis comparing the effectiveness of each method in capturing the effect. We first assess the results from the different matching method to highlight the importance of an effective matching technique. Then later expand this to also consider how a good balance measure can improve the precision of our results.

Worth noting before we proceed into the results, our data composition is slightly different to the studies which we follow, Coetzee (2013) and Oyenubi (2018). That is, our treatment group consist of member who follow a strict treatment protocol. In other words, those with high dosage of treatment and little risk of diluted effect.

5.1.ESTIMATING UNDER DIFFERENT MATCHING METHODS

5.1.1. Analysis for the Propensity Score Matching

We use the logistic regression (a probability model) to estimate the propensity score. The below equation shows the variables used to estimate the propensity scores for each member in the sample (the conditional probability of receiving treatment). The choice of variables used to control for observed characteristics of our sample are based on the studies we follow, Coetzee (2011 & 2013) and Oyenubi (2018). We experienced a challenge in finding a specification that would satisfy the balancing condition (as implemented in the "*pscore*" stata command). This was a manual trial and error process which we ended up settling on the below specification without interacting variables. As mentioned before, adding caregiver motivation into the model makes it difficult to find balance with the use of PSM method.

$$P(X) = Pr\,(CSG=1) = \alpha_0 + \beta_1\, electricity + \beta_2\, marital + \beta_3\, water + \beta_4\, telephone + \beta_5\, caregiver\ age + \beta_6\, caregiver\ years\ of\ education + \beta_7\, african + \beta_8\, coloured + \beta_9\, asian + \beta_{10}\, employed + \beta_{11}\, motivation + \beta_{13}\, expenditure + \beta_{14}\, household\ size + \beta_{15}\, household\ head\ gender + \beta_{16}\, child\ relationship\ with\ caregiver$$

Table 2 presents the results from the logit estimation. As already noted, caregiver's motivation has a crucial role on determining the duration the child receives grant, and in turn on the effect of the CSG. The variable that captures caregiver's motivation appears to mirror this, in that, the coefficient is positive and statistically significant at 1% level. Which indicate the expected positive relationship, that is, with a highly motivated caregiver the child is more likely to be in the treatment group. We also note that children whose caregivers are married and/or employed are less likely to receive the CSG. Furthermore, children whose caregiver's are African or Coloured are more likely to receive treatment. Looking at the variables controlling for household characteristics, only household expenditure, water and telephone indicate a negative and statistically significant relationship with being treated. This would imply that children living in household with high expenditure and access to water and telephone are less likely to receive treatment. The results also show that the household size is statistically significant at 1% level, with a positive relationship. This suggest that children from households that have more people living in it are more likely to receive the benefit of treatment.

***Table 2**: Logit estimates of percentage treatment*

Variables	***Coefficient***	***Standard Error***
electricity	*-0.3933*	*0.2648*
marital	*-0.7785****	*0.2031*
water	*-0.5710****	*0.2195*
telephone	*-0.9205***	*0.4125*
caregiver age	*0.0152*	*0.0108*
caregiver years of education	*-0.0018*	*0.0332*
african	*3.6998***	*1.4509*
coloured	*3.5334***	*1.4674*
asian	*3.2861*	*1.8236*
employed	*-0.5951****	*0.2010*
motivation	*3.8404****	*0.1886*
household expenditure	*-0.0001****	*0.0000*
household size	*0.3474****	*0.0469*
household head gender	*0.0890*	*0.1860*
relationship with caregiver	*0.3975*	*0.3214*
constant	*-1.1888*	*1.5826*

** Significant at 5% level, *** significant at 1% level.*

In this paper we make use of the 1:1 NN matching estimator[10] with replacement for all the analysis. The average treatment effect of the treated (ATT) is calculated using the Stata command "*psmatch2*" and "*teffects psmatch*[11]". The results are presented in *Table 3*, in which the treatment effect is positive, it is 5.6% of standard deviation. The results from this analysis are not very different to Coetzee (2013) and Oyenubi (2018) results using wave 1 dataset. The conclusion remains similar, in that, the treatment effect is not statistically significant at 5% level. The size of the effect also remains relatively small. Note that the number of control group member in the matched dataset is now 281 (this increases the ratio of treated to control members to 10.5 which significantly inflate the standard error under PSM). This partly explains why the estimate under this method is not statistically significant.

Table 3: *Treatment Effect Estimate using the PSM (1:1)*

	normal S.E (matching not accounted for)	***with AI robust S.E (matching accounted for)***
Estimate	*0.0563*	*0.0563*
Standard Error	*0.3851*	*0.2899*
p.value	*0.8829*	*0.8460*

5.1.2. Analysis for the Genetic Matching (the default method)

Here we compare the findings from the different matching methods, the GenMatch with the default loss function and the results from the PSM method presented in the previous analysis. Note that the results under the GenMatch are estimated differently to Oyenubi (2018). In this study we use GenMatch to estimate the treatment effect without the inclusion of the propensity scores. The propensity scores become a nuisance when added to GenMatch as part of the covariates. The scores were not contributing towards finding balance in the algorithm, this shows in the weight which they were allocated, i.e. 0.10 weight (0%). Adding the scores to the system also translated into results that are insignificant. As a results, in this paper the GenMatch is ran without including the propensity score to the system.

[10] Refer to the Appendix for estimated effect under 1:2 NN matching estimator

[11] The difference between the two commands is that; the former produces standard errors that do not take into account that the propensity score is estimated, whereas the latter takes into account the fact that propensity scores are estimated rather than known when calculating standard errors.

Table 4: *Treatment effect Estimate using the GenMatch (with the default balance measure)*

	normal S.E (matching not accounted for)	***with AI robust S.E (matching accounted for)***
Estimate	*0.1461*	*0.1461*
Standard Error	*0.0438*	*0.1532*
T-stat	*3.3373*	*0.9536*
p.value	*0.0008*	*0.3403*

Our results presented in *Table 4* show the estimated treatments effect when the default balance measure of the GenMatch is used. Under this method, the treatment effect is 14% of the standard deviation and statistically significant at 1% level when the standard error does not account for the matching process. The effect has increased in size and in addition it is statistically significant relative to the results in table 3. Furthermore, the standard errors (S.E) from this method are smaller than those recorded under the PSM estimates (note that under GenMatch the number of matched control observation is 450). Compared to the continuous treatment estimate from Coetzee (2013), we record a treatment effect that is larger in size and statistically significant. These results show that we can still find convincing results even under the binary case. The change in the results under the GenMatch, can be attributed to the fact that this technique seeks to find the optimal balance in the process of estimating the treatment effect.

5.2. ESTIMATING UNDER THE GENMATCH WITH DIFFERENT BALANCE MEASURES

In this section we assess the results under the different balance measures used to define the loss function in the GenMatch algorithm. This is an extension from the default balance measure from the previous sub section. Here we consider Standardized difference in mean (SDM) and the Entropy distance metric. These balance measures differ by the moments which they use to quantify balance and we will go through them in the results.

5.2.1. The Standardized Difference in Mean (SDM) as balance measure

Table 5 presents the treatment effect estimate when using the standardized difference in means as a balance measure under GenMatch. The size of the effect has significantly improved from the previous estimates. With this balance measure we find the estimated effect of 29% of

standard deviation. Moreover it is statistically significant at 1% level when matching is not accounted for. It is evident from these results alone, that the balance measure in use have an effect on the resulting estimate. The standardized difference selects optimal weights that attempt to balance mean and variance of the covariates while the default measures optimizes the mean of p-values in the t-test and KS test (Oyenubi, 2018).

Table 5: *Treatment effect Estimate using GenMatch (with the Standardized difference in means)*

	normal S.E (matching not accounted for)	***with AI robust S.E (matching accounted for)***
Estimate	*0.2920*	*0.2920*
Standard Error	*0.0408*	*0.2949*
T-stat	*7.1499*	*0.9903*
p.value	*0.0000*	*0.3220*

5.2.2. The Entropy distance metric as balance measure

We present the treatment effect estimate of the GenMatch approach applied with the Entropy as the balance measure in Table 6. The magnitude of the effect has slightly improved with this balance metric. The treatment effect is 31.5% of the standard deviation, this estimates also remains significant at 1% level.

Table 6: *Treatment effect Estimate using the GenMatch (with the Entropy)*

	normal S.E (matching not accounted for)	***with AI robust S.E (matching accounted for)***
Estimate	*0.3155*	*0.3155*
Standard Error	*0.0409*	*0.2782*
T-stat	*7.7203*	*1.1340*
p.value	*0.0000*	*0.2568*

While the Standardized difference in means balances two moments, the Entropy measure will generate weights that attempt to balance the distribution of covariates. Oyenubi (2018) argues that this is an important property which the distance metric offers, for continuous variables such as the motivation variable. With this measure, the weight will take into consideration the shape of the covariate distribution in both treatment arms. As mentioned before, this measure

can be applicable for both discrete and continuous variables, and it is considered a flexible nonparametric measure for conditional dependence. With all things fixed, it is apparent from the difference in size of the estimated effect in the two analysis that a better balance measure can increase precision in our estimate.

It is a common practice in many applied studies to use robust standard errors to obtain unbiased standard errors under cases of heteroscedasticity. Even in case of homoscedasticity, robust standard errors are still considered unbiased though inefficient. As a conservative measure, this study further considered this technique in the different matching methods applied. Albeit the effect being larger and significant when using the normal standard error test on GenMatch, t-stats using robust standard errors infer that the effect is statistically not significant. Based on these results, even though the effect is there it may not have much power. It is however worth noting that the sample size used in this study may have not been sufficiently large. According to Woolridge (2012) the t-stats obtained using robust regression might have distributions that are not close to the t distribution if the sample size is small. Therefore this could throw off our inference. Using robust standard errors, the t-statistics obtained only have distributions which are similar to the exact t-distributions if the sample size is large. Therefore we can argue in favor of always reporting and making inference using only the robust standard errors in cross-sectional studies when the sample size is large.

5.3.EVALUATING COVARIATE BALANCE

The following tables indicates whether each of the variables is balanced for each of the matching techniques applied. We look at balance pre and post matching.

5.3.1. When Matching under the PSM

To test for balance after matching on the propensity scores, we used the stata command "*pstest*". The null hypothesis for this test is that the two covariate means are equal in the treated and control groups, this would also indicate that the groups are balanced. From the table, it can be seen that the balancing property is not satisfied for the majority of the variables. We note balance to only hold for caregiver's years of education and gender of the household head variables before matching, this is however not the case for most of the variables after

matching[12]. Furthermore, the percentage in bias increased post matching in majority of the variables. The most tolerated percentage bias[13] is under 10% post matching, we find this to hold for variables such as telephone, electricity, caregiver's age, caregiver's ethnicity, household expenditure and child relationship with caregiver.

Table 8: *Covariate Balance under the Propensity Score Matching*

Variables	Unmatched (U)	Mean		%bias	%reduct bias	t-test	
	Matched (M)	Treated	Control			T	p>t
motivation	U	0.0783	-1.5202	229.7		61.54	0.000
	M	0.0783	0.29903	-31.7	86.2	-16.57	0.000
employed	U	0.2757	0.41111	-28.8		-5.91	0.000
	M	0.2757	0.47905	-43.2	-50.2	-16.51	0.000
marital	U	0.2280	0.45333	-48.9		-10.33	0.000
	M	0.2280	0.41115	-39.7	18.7	-15.4	0.000
caregiver's years of education	U	8.4628	8.6933	-6.2		-1.25	0.212
	M	8.4628	9.4601	-27	-332.7	-11.68	0.000
caregiver's age	U	35.9230	40.616	-34.8		-7.35	0.000
	M	35.9230	36.363	-3.3	90.6	-1.39	0.165
household head gender	U	0.5044	0.52222	-3.6		-0.7	0.481
	M	0.5044	0.55507	-10.1	-184.2	-3.91	0.000
child relationship with caregiver	U	0.8078	0.71333	22.3		4.64	0.000
	M	0.8078	0.77162	8.5	61.7	3.42	0.001
african	U	0.9030	0.76667	37.3		8.54	0.000
	M	0.9030	0.8723	8.4	77.5	3.75	0.000
coloured	U	0.0946	0.18444	-26.1		-5.78	0.000
	M	0.0946	0.12568	-9	65.4	-3.82	0.000
asian	U	0.0020	0.01333	-13		-3.78	0.000
	M	0.0020	0.00169	0.4	97	0.3	0.763
electricity	U	0.7845	0.86222	-20.5		-3.81	0.000
	M	0.7845	0.75946	6.6	67.9	2.29	0.000
water	U	0.5598	0.72	-33.8		-6.45	0.000
	M	0.5598	0.61926	-12.6	62.9	-4.66	0.000
telephone	U	0.0206	0.1	-33.8		-9.15	0.000
	M	0.0206	0.00777	5.5	83.8	4.18	0.000

[12] Also see Figure 4 in the Appendix under section C

[13] The standardized percentage bias is shown prior and post matching, along with the achieved percentage reduction in abs (bias). The standardized % bias is the % difference of the sample means in the treated and untreated as a percentage of the square root of the average of the sample variances in the treated and untreated groups (Rubin & Rosenbaum, 1983).

expenditure	U	2510.6000	4928.5	-43.5		-13.03	0.000
	M	2510.6000	2815.8	-5.5	87.4	-4.63	0.000
household size	U	6.6547	4.9089	64.6		10.88	0.000
	M	6.6547	5.1635	55.2	14.6	21.99	0.000

5.3.2. When Matching with the GenMatch Algorithm

Table 9 presents the results for covariate balance between the two treatment groups when we use the Entropy and the SDM as balance measures. Under each balancing measure, balance for unmatched and matched along with the difference between the two samples is shown. We note that balance improves on most covariates under the entropy distance metric, and only worsens for covariates such as caregiver's years of education and age. Similarly when we use the SDM balance improves for the majority of the covariate. It worsens for the variables that controls for caregiver's employment status, years of education, household head gender and child gender. These results for balance show that there is generally an improvement in balance for more variables under the GenMatch approach.

Table 9*: Covariate Balance under the GenMatch (with Entropy and SDM)*

Variables	***Entropy***			***SDM***		
	Unmatched	*Matched*	*Difference*	*Unmatched*	*Matched*	*Difference*
motivation	0.5074***	0.0455***	-0.4619	3.1136***	0.3019***	-2.8117
employed	0.0624***	0.0324***	-0.0300	0.2988***	0.3053***	0.0065
married	0.0961***	0.0001	-0.0960	0.5229	0.1074***	-0.4155
caregiver's years of education	0.0208***	0.0298***	0.0090	0.0632	0.3059***	0.2427
caregiver's age	0.0209***	0.0221***	0.0012	0.3718***	0.2638***	-0.1079
household head gender	0.0342	0.0068***	-0.0274	0.0357	0.2306***	0.1950
child's gender	0.0342	0.0006	-0.0274	0.0357	0.2306	0.1950
african	0.1324***	0.0000	-0.1324	0.4320***	0.0000	-0.4320
coloured	0.1044***	0.0000	-0.1044	0.2927***	0.0000	-0.2927
asian	0.2937	0.0000	-0.2937	0.1913***	0.0000	-0.1913
electricity	0.0152***	0.0093***	-0.0059	0.1929***	0.1199***	-0.0730
water	0.0326***	0.0072***	-0.0254	0.3266***	0.0731**	-0.2535
telephone	0.2457***	0.0333***	-0.2124	0.4629***	0.0000	-0.4629
expenditure	0.0914***	0.0465***	-0.0449	0.6595***	0.0525.	-0.6070
household size	0.0688***	0.0470***	-0.0219	0.5504***	0.3064***	-0.2440

*Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

6. CONCLUSION

This paper attempts to capture the effect the CSG has on beneficiary's height-for-age z scores (which is proxy for child wellbeing). Whilst taking into account the complexities researchers are faced with when investigating the South African case. We confirm and expand on the results from Coetzee (2013) and Oyenubi (2018) using a different and more recent dataset. When we assess treatment under the binary case of the PSM our results are similar to that of Coetzee (2013). We find inconclusive evidence even though the effect was positive. The effect is relatively small and statistically insignificant. We expand our analysis to adopt the techniques proposed in Oyenubi (2018). The GenMatch algorithm is applied with various balance measures. In which our results provide evidence that there is a positive and significant effect of the CSG. Moreover, the method produced estimates with standard errors[14] that are smaller than when matching is performed under the PSM. Our results echo the findings by Oyenubi (2018), in that, the results highlighted the importance of using a matching technique that seeks to optimize balance and a balance measure that does not ignore the aspects of covariate balance.

The GenMatch has properties that are normally preferred when estimating the treatment effect. By using this automated process to search for the matched sample that achieves the best balance, it is able to obtain better levels of balance without requiring any intervention from the researcher. That is, to iteratively check covariate balance in the matched samples then refine the covariate weights accordingly till balance is achieved. Even though it is computationally intensive and runs the search automatically, it still provides the researcher leeway to have their own input to the process regarding the data information which they have knowledge of. For example, the covariates to match on, the loss function to be used which is a summary of measures of balance, the matching estimator and the propensity score model to include. It is also robust to the misspecification of the propensity score.

We however ought to mention that, despite the improvements provided by the GenMatch in comparison to the propensity score matching as a matching algorithm, this techniques equally relies on observables provided by the data. As a result renders it just as exposed as the PSM to unobserved factors. The results will be valid under conditional independence assumption.

[14] When matching is not accounted for.

The impact of CSG is wide-ranging, as not only does the child benefit but the household. Therefore it bears change in many aspects of the lives of those living in household that receives the grant. This then demonstrate how CSG plays such a critical role in bettering the lives of many poor South African households. However, for impact analysis sake, this reality also makes the South African case complex to assess, because the CSG programme works as an unconditional cash transfer. That is, there are no set conditions for the caregiver to fulfil to promote a desired outcome in line with the objectives of the programme. Which could therefore be used as a direct outcome for the purpose of assessing the effect of the CSG. Others argue that this could partially explain why some studies reported very little to small effect, which makes it difficult to deliver evidence that support that the cash transfers are primarily spent on improving the lives of benefiting children.

In Brazil, the objectives of the *Familia Escola* are directed towards ending child labour and to have children go to school, while *Progresa* in Mexico is targeted at health and education. This is a common approach in other Latin American countries (Lekezwa, 2011). In turn, school attendance rates, clinic visits, and height and weight for age all quantify the effectiveness of such programmes in these countries (Lekezwa, 2011). In the South African case, the effect is much diluted. As a result, this could partly explain why South Africa is experiencing a high prevalence of stunting. Some reports suggest that, grants aimed at improving nutrition are not as effective as expected. Therefore, this illustrates the need for the programme to be relooked at, to have some sort of direct tie with a clearly defined promotional outcome measure it seeks to advance. We need concentrated efforts committed to nutrition, programmes that would advance healthy eating and living.

We cannot however conclude without highlighting the significance of caregiver's motivation on the extent which the child enjoys the benefit and as a result the effectiveness of the CSG. In this regard, one could recommend a way in which the caregiver's motivation becomes less influential in determining the amount of dosage the child receives. Even if it means we have to bring it closer to conception, literally. The government could run a programme that starts when the mother is in their early stages of pregnancy. For every pregnancy booking session, there should be a free exercise session and wellness teaching on foods to eat and how to keep the baby healthy. Attendance should be incentivized by receiving coupons that would subsidize nutritional foods and only made valid to certain foods and stores. This would help monitor their participation while enforcing eating healthy and being active.

Moreover, the facilities for grant application could be made easily accessible upon child birth by having them conveniently at the hospital for application to happen after the child is born. Furthermore, the systems should adopt innovative ways that would lessen the burden on documentations required to receive the grant. This process could be made to run parallel with the child birth certificate application at the hospital.

7. REFERENCES

Agüero, J. M., Carter, M. R. & Woolard, I., 2007. *The Impact of Unconditional Cash Transfers on Nutrition: The South African Child Support Grant ,* New York: International Poverty Centre.

Angrist, J. D. & Pischke, J.-S., 2008. *Mostly Harmless Econometrics: An Empiricist's Companion.* s.l.:s.n.

Austin, P. C., 2010. Statistical Criteria for Selecting the Optimal Number of Untreated Subjects Matched to Each Treated Subject When Using Many-to-One Matching on the Propensity Score. *American Journal of Epidemiology,* 172(9), p. 1092–1097.

Coetzee, M., 2011. Finding the benefits: Estimating the impact of the South African child support grant. *University of Stellenbosch,* Volume Working Paper 230, pp. 1-41.

Coetzee, M., 2013. Finding the Benefit: Estimating the impact of the South African Child Support Grant. *South African Journal of Economics,* pp. 427-450.

d'Agostino, G. & Scarlato, M., 2016. Gender Inequality in the South African Labour Market: the Impact of the Child Support Grant. *MPRA,* pp. 1-17.

d'Agostino, G., Scarlato, M. & Napolitano, S., 2017. Do Cash Transfers Promote Food Security? The Case of the South African Child Support Grant. *Journal of African Economies,* p. 1–27.

Desmond, C. & Casale, D., 2017. Catch-up growth in stunted children: Definitions and predictors. *PLoS ONE,* 12(12), pp. 1-12.

Diamond, A. & Sekhon, J. S., 2013. Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies. *The Review of Economics and Statistics,* 3(95), pp. 932-945.

Diks , C. & Fang, H., 2017. Transfer Entropy for Nonparametric Granger Causality Detection: An Evaluation of Different Resampling Methods. *MDPI,* 19(372).

Granger, C. W., Maasoumi, E. & Racine, J., 2003. A Dependence Metric for Possibly Nonlinear Processes.

Heinrich, C. et al., 2012. *The South African Child Support Grant Impact Assessment: Evidence from a survey of children, adolescents and their households,* s.l.: Department of Social Development, South African Social Security Agency and UNICEF.

Hirano, K. & Imbens, G. W., 2004. *The propensity score with continuous treatments.* s.l.:John Wiley & Sons, Ltd .

Jehoma, S. & Delany , A., 2016. *Implementation of social grants: Improving delivery and increasing access,* Cape Town: Economic Policy Research Institute UCT.

Kelly, G., 2017. *Everything you need to know about social grants: For people who receive a grant or need to receive one.* [Online]

Available at: https://www.groundup.org.za/article/everything-you-need-know-about-social-grants_820/

Lechner, M., Rodriguez-Planas, N. & Fernández Kranz, D., 2015. *Difference-in-Difference Estimation by FE and OLS when there is Panel Non-Response,* s.l.: Institute for the Study of Labor.

Lehohla, P., 2017. *Poverty Trends in South Africa: An examination of absolute poverty between 2006 and 2015,* Pretoria: Statistics South Africa (Stats SA).

Leibbrandt, M., Lilenstein, K. & Shenker, C., 2013. The influence of social transfers on labour supply: A South African and international review. *SALDRU,* pp. 1-36.

Lekezwa, B. I., 2011. *The Impact of Social Grants as Anti-Poverty Policy Instruments in South Africa: an Analysis Using Household Theory to Determine Intra-Household Allocation of Unearned Income,* Stellenbosch: University of Stellenbosch.

Maasoumi, E. & Wang, L., 2013. The Gender Earnings Gap: Measurement and Analysis. *Emory University Department of Economics,* 5(13).

Mawson, N., 2017. *#Budget2017: Inflationary increase for welfare grants.* [Online]
Available at: https://www.iol.co.za/business-report/budget/budget2017-inflationary-increase-for-welfare-grants-7877435

Merwe, M. V. D., 2017. *Daily Marverick.* [Online]
Available at: https://www.dailymaverick.co.za/article/2017-05-02-analysis-nutritional-stunting-and-why-sa-is-coming-up-short/
[Accessed January 2019].

Mqadi, S., 2017. *DGMT.* [Online]
Available at: https://dgmt.co.za/stunting-the-silent-killer-of-south-africas-potential/
[Accessed January 2019].

Overseas Development Institute, 2006. Social Grants South Africa. *Policy Brief 1: Inter-Regional Inequality Facility,* February.

Oyenubi, A., 2018. Quantifying balance for causal inference: An information theoretic perspective. *University of Cape Town.*

Ponsich , A., López Jaimes, A. & Coello Coello, C. A., 2013. A Survey on Multiobjective Evolutionary Algorithms for the Solution of the Portfolio Optimization Problem and Other Finance and Economics Applications. *IEEE,* 17(3), pp. 321 - 344.

Ramos, M. et al., 2014/15. A Quantitative Analysis of the Impacts of Social Grants on. In: *Submission for the 2014/15 Division of Revenue.* s.l.:s.n., pp. 39-74.

Rubin, D. B. & Rosenbaum, P. R., 1983. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Journal of the American Statistical Association,* Volume 70, pp. 41-55.

Rubin, D. B. & Rosenbaum, P. R., 1984. Reducing Bias in Observational Studies Using Sub-Classification on the Propensity Score. *Journal of the American Statistical Association,* 79(387).

Schiel, R., Leibbrandt, M. & Lam, D., 2014. Assessing the impact of social grants on inequality: A South African case study. *World Institute for Development Economics Research,* pp. 1-25.

Sekhon, J. S., 2011. Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching package for R. *Journal of Statistical Software,* 42(7), pp. 1-47.

Villa, J. M., 2016. diff: Simplifying the estimation of difference-in-differences treatment effects. *The Stata Journal,* 16(1), p. 52–71.

WHO, 2015. *World Health Organization (WHO).* [Online]
Available at: https://www.who.int/nutrition/healthygrowthproj_stunted_videos/en/
[Accessed February 2019].

8. APPENDIX

Appendix A

Table 10: Reason for no application of child support grant	***Freq.***	***Percent***
Ineligible because the child is too old	6	0.27
CSG applied for by someone in another household	7	0.31
Cost of application is too high	21	0.93
Child is not eligible as receives a different grant	25	1.11
Caregiver has not heard of CSG	34	1.5
Caregiver does not know how to apply for grant	45	1.99
Don't know	47	2.08
Caregiver cannot apply as not the child's mother	67	2.96
Application process is too complicated	77	3.41
Other reasons	103	4.56
Cannot be bothered	175	7.74
In process of applying	236	10.44
Caregiver doesn't have the right document	297	13.14
Haven't got round to it yet	326	14.42
Child is not eligible as caregiver income is too high	794	35.12

Appendix B – Estimates when matching is performed on 1:2 nearest neighbour matching estimator

Our estimates for all the analysis were performed on 1:1 NN matching estimator with replacement, and we have leeway to increase the number of untreated observations (M) we match to each treated observation , i.e. we use a 1:M matching. Considering that the ratio of the untreated to treated observations in our data was very small. To be specific there are 2960 treated members and 450 control members. Austin (2010) argues in his findings that on average, increasing the number of untreated observations matched to each treated observation raised biasness of the estimated treatment effect; conversely, it tended to result in increased precision. This motivates for the extension of our analysis from this point to explore 1:2 matching.

Table 7.1. *Treatment effect Estimate using the GenMatch with the Entropy (matching 1:2)*

	normal S.E (matching not accounted for)	***with AI robust S.E (matching accounted for)***
Estimate	*0.3406*	*0.3406*
Standard Error	*0.0403*	*0.47055*
T-stat	*8.4583*	*0.72378*
p.value	*0.0000*	*0.4692*

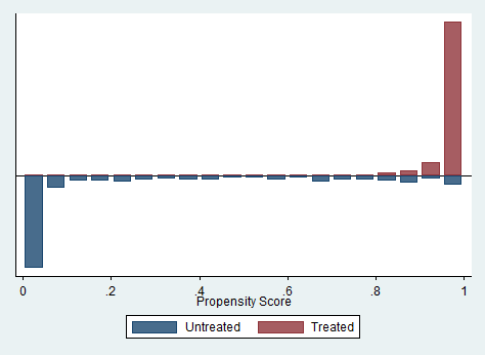
Table 7.2. *Treatment Effect Estimate using the PSM (matching 1:2)*

	normal S.E (matching not accounted for)	***with AI robust S.E (matching accounted for)***
Estimate	*0.0797*	*0.0797*
Standard Error	*0.3691*	*.2850*
p.value	*0.8290*	*0.779*

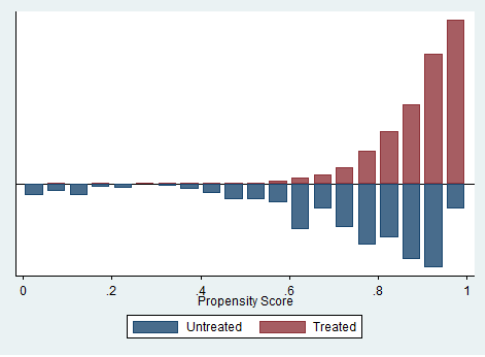
Table 7.1 shows that if we use 1:2 matching in the GenMatch when the entropy is used as a balance measure. The estimated effect increases further in size, the effect is 34% of the standard deviation, even though the standard error remained slightly the same. Similar, we see an increase in the estimated effect of the PSM. The estimated effect increased to 7% of the standard deviation however still remain relatively small and statistically insignificant. Furthermore, the standard errors (S.E) from the PSM method are greater than those recorded under the GenMatch estimates (with different balance measures used) when matching is not accounted for. The results here echo the same findings we have in the paper on the two Matching methods.

Appendix C

Figure 3.1.: *Regions of common support*



**When caregiver motivation is included in the model*



**When caregiver motivation is excluded from the model*

From a visual inspection of these figures, there is a visible difference between the two graphs. One that is on the left shows the distribution of the propensity scores when the model controls for caregiver motivation, whereas the one of the right excludes the variable. The first graph shows that the distribution of the propensity scores is varsity different between the treated and control groups therefore violating the overlap assumption which we highlighted under the PSM. To limit the possibility for bias, the treatment effect is evaluated along the common support for this sample. On the contrary, when this variable is excluded from the model, we notice a change in the common support. In other words, the propensity scores of the untreated almost mirror those for the treated.

Figure 3.2.: *Propensity Score before and after matching (Kernel density)*

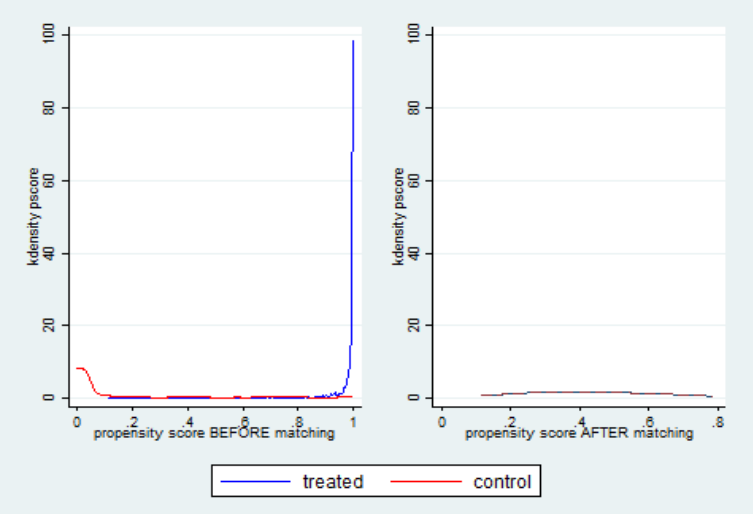


Figure 4: *Covariate Balance under the Propensity Score Matching*