Who benefits from South African Child Support Grant: The role of gender and birthweight

Adeola Oyenubi

ERSA working paper 781

May 2019

Economic Research Southern Africa (ERSA) is a research programme funded by the National Treasury of South Africa.

Who benefits from the South African Child Support Grant: the role of gender and birthweight

Adeola Oyenubi *

April 12, 2019

Abstract

Stunting (low height-for-age) is known to be a good proxy for a child's wellbeing. Several studies have suggested that the South African Child Support Grant (CSG) reduces stunting in benefiting children. However, all of these studies have estimated the impact of the CSG on the mean of the height-for-age distribution. This paper investigates how this benefit varies across the quantiles of the height-for-age distribution.

The result suggests that the positive effect at the mean is driven by children in the high quantiles and this group of children are more likely to be girls that did not experience low birth weight at birth. I argue that the CSG has not been able to address the malnutrition inequality that disadvantage male children and children born with low birthweight. The finding in this paper suggests that addressing low birthweight may potentially increase the impact of CSG across the distribution of height-for-age score. In the concluding remarks, I discuss how the pregnancy grant proposal can help mitigate this problem.

Keywords: unconditional cash transfer, child health, causal inference and nutrition

JEL: I38 H53 C21 D13

1 Introduction

The literature has shown that adequate nutrition is important for children. This is because nutritional deprivation and malnutrition early in life have long-term negative consequences on the physical and cognitive development of children (Delany et al, 2008; Walker et al, 2015). Stunting (low height-for-age) which is one of the manifestations of malnutrition is associated with poverty and may be irreversible in older children (Delany et al, 2008)[1]. Furthermore, early childhood

*University of the Witwatersrand, South Africa

[1] A stand of the literature suggests the possibility of catch up growth (e.g. Zhang. et. al, (2016))

stunting is likely to contribute to the intergenerational transmission of poverty (Grantham-McGregor et al, 2008). Walker et al, (2015) found that children born to a stunted parent are likely to have a lower score on the cognitive scale, lower development quotient and are likely to be stunted themselves. These studies suggest that the impact of stunting continues in the next generation of children.

One way to address poverty, which is a root cause of malnutrition, is through social transfers such as child support grants (CSG). The South African Child Support Grant (CSG) has been found to boost child nutrition as measured by children's height-for-age (Aguero et al, 2006; Coetzee, 2013 and Grinspun, 2016). However, these studies focus on the impact of the CSG on the mean of the height-for-age distribution. From the policy perspective, this may be misleading if the effect of the CSG is heterogeneous across the height-for-age distribution. In other words, the effect of CSG may vary at different quantiles of the height-for-age distribution. For example low-birthweight increases the risk of stunting in children (Rahman et al, 2016; Aryastami et al, 2017), and early childhood stunting may not be reversible (Duflo, 2003), these taken together implies that poor children that do not start benefiting from the CSG early may not record significant improvement in their height-for-age score. The effect of CSG on stunting for such children may be very different from the effect of CSG for children that do not experience low birthweight or that start receiving CSG early even though they experienced low birthweight.

Existing studies on the impact of CSG on height-for-age of children ignore the role of low birthweight and its differential impact on child health by gender (see Aguero et al, (2006) and Coetzee, (2013) for example). Whereas, results from the medical literature suggests that boys show more postnatal complications as a result of low birthweight. This phenomenon is referred to as the male disadvantage hypothesis. According to this hypothesis, boys are the weaker sex and are more sensitive to adverse environmental factors during gestation, infancy and childhood (Kirchengast & Hartmann, 2009). It is therefore important to investigate the effect of these variables (gender and low birthweight) on the height-for-age of vulnerable children benefiting from the CSG. This information may assist policymakers to fine-tune the delivery of this policy for enhanced impact.

Therefore, in this study, I estimate the quantile effects of CSG on its beneficiaries with particular emphasis on the role of gender and low birthweight. To do this the entropy balancing method (Hainmueller, 2012 & 2013) is used to balance the first three moments of all the relevant covariates. The weights that balance the covariates are then used to estimate the unconditional quantile effect of the CSG.

The results show that children at the top of the height-for-age distribution are driving the positive result observed at the mean. This result holds even after controlling for low birthweight. I further investigate if there are differences between the characteristics of beneficiaries at either end of the outcome distribution. The result shows that in the distribution of benefiting children, the bottom quintile contains significantly more boys and children with low birth weight compared to the top quintile. To investigate gender differences in the effect of the CSG the analysis is disaggregated by gender. The result suggests that CSG has no significant effect on height-for-age for male children. In fact, all of the significant quantile effects holds only for female children. The implication of this is that inequality in malnutrition that favour the girl child is not being adequately addressed by the CSG. I discuss a policy recommendation that could address this trend in the concluding remarks.

The rest of the paper is organized as follows; section 2 discusses the methods and the data used. Sections 2.1 and 2.2 briefly discuss entropy balancing and why the unconditional quantile regression is preferred to the conditional quantile regression in estimating the quantile effects. Section 2.3 discusses the estimation of the caregiver[2] motivation variable, which captures caregivers' eagerness to take up CSG while section 2.4 explores the summary of the data used. Section 3 present the results and section 4 discusses the robustness of the results. Finally, section 5 concludes.

2 Methods and Data

Many studies that have investigated the impact of CSG (in the South African context) have noted that apart from observed covariates that may confound the impact estimate of the CSG, caregiver motivation is a key factor (Oyenubi, 2018; Coetzee, 2011 & 2013). Motivation captures the eagerness of caregivers to take up the grant. Higher caregiver motivation is therefore associated with a lower delay in taking up of the grant. This variable is important because it influences the treatment effect through the length of time (dosage) a child benefits from the CSG. Those with high motivation and consequently high dosage are likely to experience more benefit relative to those with low motivation (Oyenubi, 2018).

One problem with controlling for caregiver motivation is that motivation is unobserved and therefore has to be estimated. To estimate this variable existing research use delay before applying for CSG to recover caregiver motivation. This is achieved by modelling the delay as a function of the child's age and location (rural versus urban) using censored regression. However, Coetzee (2011 & 2013) found that when the estimated motivation is included in the set of covariates under Propensity Score Matching (PSM), it becomes difficult to balance the distribution of covariates in the binary treatment case. This is because by construction motivation in the treatment group will be higher than motivation in the control group since this variable is a function of delay in applying for the CSG. This observation and the fact that dosage of treatment matters is the reason why Coetzee (2013) considered the Generalized Propensity Scores (GPS) in estimating the impact of the CSG. Under the GPS framework, CSG is seen as a continuous treatment, while this is a valid approach[3], Oyenubi (2018) argues that instead of relying on PSM balance can be optimized using Genetic Matching (GenMatch) while still controlling for caregiver motivation in

[2] The primary caregiver is the person that takes care of the needs of the child without payment.

[3] Note that the GPS rules out the quantile treatment approach.

the binary treatment case. PSM seeks to balance the propensity score density, however, balance in the propensity score density may not translate into balance in the other covariates. GenMatch, on the other hand, is a machine learning approach to optimizing balance. GenMatch directly balances the distribution of all covariates (including the propensity scores). This approach has been shown to outperform PSM in terms of mean square error (Diamond and Sekhon, 2013).

Oyenubi (2018)'s result shows that it is possible to balance the distribution of covariates that includes the estimated motivation variable under a method that seeks to optimize balance in all covariates rather than the propensity scores alone. Lack of balance in relevant covariates after matching on estimated propensity scores may signal misspecification of the propensity score equation (Caliendo and Kopeinig, 2008).

More recent methods side-line this problem by focusing on balancing the relevant covariates directly (e.g. Diamond and Sekhon (2013)) or estimating propensity scores that incorporates the balancing condition (Imai and Ratkovic, 2014). Another strand of the literature views the balancing problem as a calibration problem (Hainmueller, 2012 & 2013, Chan et al, 2016). Under the calibration approach, weights are calculated for each unit so that the weighted treatment and control groups satisfy prespecified balancing conditions. This approach has also been shown to deliver better performance than the PSM in terms of mean square error (Hainmueller, 2012).

In this paper, I use entropy balancing method (Hainmueller, 2012 & 2013) to balance the distribution of covariates including the estimated motivation variable. This method balances the mean, variance and skewness of the distribution of covariates. The balancing weights from entropy scheme can be combined with any estimator that one may wish to use to estimate treatment effect (Hainmueller, 2012). Therefore, we use the balancing weights to recover the unconditional quantile effect of the CSG on its beneficiaries.

2.1 Entropy weights

Entropy balancing is a pre-processing procedure that allows researchers to create balanced samples for the subsequent estimation of treatment effects (Hainmueller, 2012). This is achieved by reweighting the covariate distributions such that the reweighted data satisfy a set of specified moment conditions. Preprocessing reduces model dependency for the subsequent analysis of the treatment effects in the pre-processed data using standard methods such as regression (Abadie and Imbens, 2011). Similar to methods in the survey adjustment literature (Deming and Stephan 1940; Ireland and Kullback 1968), entropy balancing is based on a maximum entropy-reweighting scheme. It allows exact balance on the first, second and possibly higher moments. In this analysis, we use entropy reweighting to balance the mean, variance and skewness of the covariate distributions as implemented in Hainmueller (2013). The balancing weights are such that they satisfy the specified balancing conditions and remains as close as possible to uniform base weights to prevent loss of information and retain efficiency.

At the mean, we are interested in estimating the average treatment effect on the treated (ATT) given by

$$ATT = E(Y(1)|D = 1) - E(Y(0)|D = 1) \tag{1}$$

where $D = 0\,or\,1$ represent the treatment status (1 for treated and 0 for control). $Y(1)|D$ is the outcome given the treatment status. While $Y(1)|D = 1$ can be estimated from the treated sample, $Y(0)|D = 1$ is unobserved since it represents the counterfactual outcome i.e. the outcome of the control observations had they been treated. Rosenbaum and Rubin (1983) show that assuming selection on observables and overlap condition is satisfied for all covariates x in the support of the treatment distribution $f_{x|D=1}$ we can write

$$ATT = E(Y(1)|D = 1) - \int E(Y|X = x, D = 0)f_{x|D=1}(x)dx \text{ ... ...} \tag{2}$$

The last term in equation (2) is the covariate-adjusted mean. This requires adjusting the covariate distribution in the control group so that it is similar to the covariate distribution in the treatment group. This will make the treatment indicator orthogonal to the covariates. The second term in equation (2) can be estimated as

$$E(Y\widehat{(0)|D} = 1) = \frac{\sum_{\{i|D=0\}} Y_i w_i}{\sum_{\{i|D=0\}} w_i} \text{... ... ...} \tag{3}$$

where w_i, $i = 1, \text{... ... ..}n_o$ are the weights that balance the distribution of covariates across treatment arms and n_o is the number of control observations. These weights can be propensity score weights. To estimate ATT using propensity score weights control units receive weights given by

$$w_i = \frac{\hat{p}(x_i)}{1 - \hat{p}(x_i)}$$

where $\hat{p}(x_i)$ is the propensity score. As mention earlier, the main problem with this approach is that the equation that estimate $\hat{p}(x_i)$ is unknown and can be mis-specified. When this happens, the weights will not balance the distribution of covariates and this will bias the treatment effect estimate.

Under entropy balancing, the weights are estimated by minimizing the entropy distance metric

$$\min_{w_i} H(w) = \sum_{\{i|D=0\}} w_i \log\left[\frac{w_i}{q_i}\right] \text{... ... ...} \tag{4}$$

subject to balancing and normalizing constraints

$$\sum_{\{i|D=0\}} w_i c_{ri}(X) = m_r \text{... ... ...} \tag{5}$$

$$\sum_{\{i|D=0\}} w_i = 1 \text{... ... ...} \tag{6}$$

$$\text{with } r \in 1, ...R \text{ and } w_i \geq 1 \text{ for all } i$$

$q_i = 1/n_0$ is the base weights, $c_{ri}(X) = m_r$ describes the set of R balance constraints imposed on the moments of the reweighted control group. The entropy balancing scheme then searches for a set of weights $W = (w_1, ..., w_{n_0})$ that minimizes equation (4) subject to the balancing and normalizing constraints of equations (5 and 6). Details of the numerical implementation can be found in Hainmueller (2013).

2.2 Conditional versus unconditional quantile effects

The weights recovered from the optimization in the previous section can be used to estimate the ATT but it can also be used to recover the unconditional quantile effects (since these estimators are based on regression equations). Another alternative is to use them to recover the conditional quantile effect estimates, which are defined conditionally on the value of the regressors. Unconditional quantile estimates, on the other hand, are independent of the values of the covariates. For example, if we are interested in a low quantile, conditional quantile effect will summarize the effect for individuals with a relatively low height-for-age score in each included covariate class even if their absolute height-for-age score is high. The unconditional effect, on the other hand, will estimate the effect for observations with a relatively low absolute height-for-age score. This is important for the interpretation of the quantile results. Unconditional quantile effect summarizes the effect of treatment for the entire population; they represent the difference between unconditional quantile of the treated outcome and the non-treated outcome for the population of interest. Therefore, they are often of interest to policymakers because of the ease of interpretation (Frolich and Melly, 2008). This is in contrast to conditional quantile effects that change with the set of conditioning variables.

The distinction between conditional and unconditional quantile effects is also important if we are interested in who seats where in the outcome distribution. Conditional effects may change the placement of individuals in the outcome distribution based on the included covariates while individuals retain their placement in the unconditional quantile. Therefore, with the unconditional quantile effect, one can find out who the gainers and the losers are based on their quantiles.

For the quantile effect, I use the unconditional quantile estimator of Firpo et al, (2009). The method involves estimating the RIF (Regression Influence Function) of any distributional functional of interest. Since we are interested in quantiles the RIF is given by

$$RIF(Y_i; q_\tau) = q_\tau + \frac{\tau - 1[Y_i < q_\tau]}{f_Y(q_\tau)} \tag{7}$$

where Y_i is the outcome, q_τ is the τ^{th}quantile of Y_{it}, $1[Y_i < q_\tau]$ is an indicator variable that denotes when Y_i is below q_τ and $f_Y(q_\tau)$ is the estimated density at q_τ. The RIF is then regressed on the covariates. In essence, our estimation involves including entropy weights in each (quantile) regression to recover the unconditional quantile effects.

2.3 Estimation of caregiver motivation

The estimation of the unobserved caregiver motivation follows the work of Oyenubi (2018). Caregiver motivation for the treated observation is estimated as the number of years between the birthday of the child and the day CSG was first received for the child. For control observations, the delay is approximated by the number of years between the birthday of the child and the date the caregiver applied for the CSG on behalf of the child (where the data is available). If the date of application is not available or CSG has never been applied for on behalf of the child, delay is equal to the age of the child. To accommodate the control observations without data on application date, censored regression is employed to estimate the expected delay equation. These observations are regarded as being right censored, with a variable censoring point that is equal to the age of the child (since an application has not been made for these children but may be made sometime in the future).

The expected delay is calculated as the ordinary least squares prediction of the delay as a function of the child's age, relationship to primary caregiver and location (rural or urban). The difference between actual and expected delay is then standardized[4] to arrive at a variable that represents the unobserved variation in caregiver's motivation to apply for CSG.

2.4 Data

We use the wave 1 data of the National Income Dynamic Study (NIDS). NIDS is a nationally representative panel data, with the first wave undertaken in 2008. Eligibility for CSG is determined by age and a means test[5] as it was applied in 2008. The age and means test conditions are used to identify children who should benefit from CSG. The treatment group is defined in two ways. In the first sample (Sample 1), the treatment group consists of children for whom CSG is being received currently. Note that this excludes children who are living in a household where another child is benefiting from CSG (these children will belong to the control group along with eligible children in non-benefiting households). In the second sample (sample 2), a child is assigned to the treatment group if it is indicated that the CSG is currently being received for the child or another child who lives in the household (similar to Coetzee (2011) & Oyenubi (2018)). Control group for Sample 2 consists only of children in eligible households that are not receiving the CSG. We use the second sample to check the robustness of the result in the first sample since the definition of a household in NIDS survey

[4] Calculated as delay minus mean of delay over standard deviation of delay.

[5] At the time of the NIDS 2008 survey a caregiver (who does not have to be a family member of the child) must have a monthly income below R800 in urban areas or R1,100 in rural areas (see Coetzee (2011)).

required that members of a household share food from a common source[6].

We restrict the sample to only black children since they are the majority in the population and genetic factors may affect the height-for-age score of children differently across all race groups. Table 1 presents the summary statistics for sample 1. Caregivers of benefiting children delay for about 2 years on average before assessing the grant while caregivers of eligible non-benefiting children delay for around 7 years on average. The implication of this is that motivation is significantly higher in the treatment group compared to the control group. However even in the treated group, the average delay is relatively high because stunting does not change rapidly, and it may be irreversible in children older than two years (Cogill, 2003).

On average, there are significant differences at the mean for most of the covariates. Included variables include caregiver characteristics; marital status, employment status, caregiver years of education, age, health-seeking behaviour in form of the number of times the child has visited the hospital in the last 12 months[7] and the relationship of the primary caregiver to the child. Household characteristics; household location (rural or urban), type of dwelling (informal or otherwise), the gender of head of household, availability of electricity, water, telephone, toilet and household food expenditure per adult equivalent[8]. Child characteristics; gender and age of the child. Table A1 in the appendix presents the summary statistics for sample 2. The differences across groups are similar.

3 Results

We start from the mean estimate of CSG on height-for-age. Table 2 presents the summary statistics after applying entropy weights.

Note that birth weight is excluded from this analysis[9]. The result shows that the entropy method balances the mean, variance and skewness of all variables as required. Columns 1 and 2 of table 3 show the treatment effect for samples 1 and 2.

At the mean, the treatment effects are 24% and 19% of standard deviation for samples 1 and 2 respectively. These effects are significant at 1% and 10% for samples 1 and 2 respectively. Note that we include other covariates (i.e. covariates listed in table 2) in this regression for efficiency but the results are not shown in the table since our focus is on the effect of the CSG. These results indicate that the CSG has a positive effect on height-for-age at the mean. It also shows that the effect in sample 1 is stronger than the effect in sample 2, suggesting that caregivers might be favouring the children for whom the CSG is received even though households tend to pool resources. However, this may also

[6] See "http://www.nids.uct.ac.za/documentation/faqs/data-about-nids"

[7] We use a dummy variable for this. The variable is equal to 1 ids the child never visited the hospital in the last 12 months and 0 otherwise.

[8] Food expenditure per adult equivalent scale (AES)=

[9] We add indicator for low birthweight in subsequent analysis to isolate its relationship with height-for-age after accounting for the effect of the treatment.

be an artefact of the fewer observations in sample 2. Either way at the mean, the effect of CSG is significant across samples.

We now turn to the question of heterogeneity of the effect of the CSG across the height-for-age distribution. Figure 1 shows the unconditional quantile treatment effects using RIF regressions (note that the quantile regressions were weighted by the entropy weights). The results[10] show that children at the high quantiles of the height-for-age distribution drive the positive effect at the mean.

A significant positive effect of about 50% of standard deviation exists at the highest quantiles while there is no significant effect for children below the 40th quantile of the height-for-age distribution.

The implication of this result is that the mean effect does not apply to all children that are receiving the CSG. Children with high height-for-age score benefit more from CSG than those with low height-for-age score.

One can argue that policymakers will be more interested in making a difference for children at the lower quantiles since they are more likely to be stunted. We note that children at the 20^{th} centile of the unconditional outcome distribution have a height-for-age score of about -2 standard deviations[11], this means children below the 20^{th} centile are actually stunted and receipt of the CSG does not appear to improve this condition.

One could look at this dynamics across the height-for-age distribution in a number of ways. This could be the effect differences in the motivation of caregivers at different parts of the height-for-age distribution. If this is the case then there should be a significant difference in the motivation of caregivers of children at low (unconditional) quantiles relative to those at higher (unconditional) quantiles. Another plausible explanation is that children at the low quantiles may have experienced low birthweight and have not recovered from it even though they are getting better nutrients (relative to non-beneficiaries) because of the CSG. If this is the case then low birthweight should have a negative sign and be significant when included in the regression shown in columns 1 and 2. It could also be some combination of these two effects.

To examine the second proposition (low birthweight at low quantiles) we include a dummy variable that indicates low birthweight[12] in the regression shown in columns 1 and 2 of table 3. The result is shown in columns 3 and 4 of table 3. The CSG treatment effect increased in both samples to 25% and 22% of the standard deviation in samples 1 and 2 respectively, this effect is significant at the 1% level. More importantly, low birth weight has a negative sign and is significant in explaining variation in height-for-age scores. In sample 1, low birth weight is associated with a reduction of 45% of standard deviation in the outcome variable, while in sample 2 it is associated with a reduction of 75% of standard deviation. This means that after accounting for the effect of the

[10] The treatment effect is also estimated using conditional quantile estimator. The result is presented in figure 1A of the appendix.

[11] Stunting is defined as height-for-age score less than -2 standard deviations of World Health Organization's reference standard.

[12] A child is born with low birthweight if the weight of the child at birth is below 2.5kg

CSG, low birthweight explains significant variation in the height-for-age score. The difference in effect in the two samples also suggests that the effect of low birthweight is stronger in eligible households that are not benefiting from CSG (sample 2). Note that the inclusion of low birthweight indicator did not change the result in figure 1. The quantile results that include low birthweight indicator is shown in figure A2 of the appendix and it shows a similar pattern to figure 1 (i.e. quantile effects without low birthweight).

In terms of the first proposition: motivation differs for children at low and high quantiles. I exploit the fact that the effects are not conditioned on covariates. This means that children's placement in the height-for-age distribution does not change during estimation. In other words, those with high values of height-for-age are the ones experiencing significant positive effect and children with low values of height-for-age are the ones not experiencing significant benefit relative to eligible non-beneficiaries. This means the (unconditional) outcome distribution of treated observations can be divided into quintiles to examine if there are differences in the covariates of children at either end of the height-for-age distribution (i.e. the first and fifth quintile). The result is shown in table 4. Although the result shows that on average children in quintile 5 have caregivers with higher motivation than children in quintile 1, the difference in caregiver motivation is not significant. This suggests that the significant effect at the high quantiles may not be due to the difference between caregivers' motivation at different parts of the outcome distribution.

The result also shows that children in quintile 5 are more likely to have older, married and more educated caregivers who live in a rural area and are more likely to be the child's father or mother. They are more likely to live in a male-headed household that has a toilet, is without electricity[13] and spends more on food per adult equivalent. However, like motivation, all these differences are not significant.

Covariate differences that are significant include the child's gender, availability of water, access to telephone and low birthweight. The first quintile has a higher percentage of children (4% higher) that experienced low birthweight relative to the fifth quintile. Given that low birthweight is associated with stunting (Espo et al, 2002; Rahman et al, 2016; Aryastami et al, 2017), and children that experience stunting early in life may not recover from it (Delany et al, 2008), birthweight at the low quintile may be a factor driving the non-significant effect at the lower quantiles[14]. Since all of the variables that show significant difference are dummy variables, differences at the mean will suffice to describe the difference between the two quantiles. However, the cumulative distribution of the (raw) birthweights is provided in the appendix (Figure A3) and it shows that the distribution of birthweight for quantile 5 dominates the distribution of birthweight in quantile 1.

Table 4 throws up a few more possibilities. The fact that quintile 5 children are more likely to live in households that have access to water suggests that they

[13] perhaps because they are in rural areas

[14] The effect of low birthweight may be compounded by the lower motivation at the lower quantiles.

are living in cleaner environments than children at the lower quantiles. This is because stunting captures multiple dimensions of children's health, development and the environment where they live (Wamani et al, 2007). There could also be a gender dimension to the effect. Wamani et al, (2007) noted that boys are more stunted than girls in Sub-Saharan Africa, similar results have been found by other studies conducted in African (Espo et al, 2002; Wamani et al, 2004; Zere and McIntyre, 2003). This observation agrees with the gender entry in table 4 which shows that the low quintile contains more male children (11% more) than the high quintile. We, therefore, investigate the gender dimension in the next section.

3.1 Gender differences in height-for-age score

If stunting is more prevalent in male children than female children, then the quantile results might be driven by the proportion of male children in the low quintile. To investigate this we re-estimate the quantile effects by gender while still controlling for birthweight. The results are shown in figures 2 and 3. CSG has no significant effect across the height-for-age distribution for male children, on the other hand, female children above the 40^{th} quantile experience significant improvement in height-for-age.

This effect increases for female children as we approach the higher quantiles with the highest effect being about 50% of standard deviation. We note that this is exactly the dynamics observed in figure 1.

The implication of these results is that the effect in figure 1 is purely driven by female children. We note that Zere and McIntyre (2003) who used South African data collected in 1993 noted that there is inequality in height-for-age across gender that favours the girl child. CSG was introduced in 1998 and has since expanded in coverage over the years. While one can argue that it has led to a significant improvement in the wellbeing of benefiting children, our result suggests that it is not having the expected impact on malnutrition inequality that favours the girl child.

Even though table 4 shows that male children are more at the lower quantile, the results in figures 2 and 3 suggest that this is not what is driving the result at the lower quantiles. CSG has no significant effect on the height-for-age of children at the lower quantiles irrespective of their gender. Furthermore, at the high quantiles, only the girls experience a significant benefit.

I then examine the role of low birthweight in the gender-specific regressions (at the mean). Table 5 shows the effect of low birthweight and CSG on height-for-age scores by gender. For the boys, there is a significant negative relationship between low birthweight and the height-for-age scores while the effect of CSG is not significant. On the other hand, for girls, CSG has a positive and significant effect on height-for-age while low birth weight is not a significant predictor of the outcome. Even though the prevalence of low-birthweight is higher among female children in our sample (57% for girls and 43% for boys) low birth weight affect the height-for-age of male children significantly while it does not have a significant relationship with the outcome for female children.

This is not surprising since the literature on the effect of low birthweight suggests that it has a more detrimental effect on male compared to female children (Roy et. al, 2014). Existing literature on the effect of CSG also shows that receiving CSG early in life significantly boost child height especially for girls (Grinspun, 2016). These findings suggest that when it comes to the effect of low birth weight on height-for-age, girls are likely to recover better than boys. The implication is that a reduction in the prevalence of low birthweight may unlock a stronger positive effect for the CSG as far as the height-for-age of the boy child is concerned. This may reduce the inequality in the effect of CSG observed in figure 1.

4 Robustness

In this section, I examine the robustness of the results. The use of entropy weights comes with a caveat. The higher the level of imbalance in the covariate distributions, the further weights have to be adjusted to meet the balancing constraints (Hainmueller, 2013). In situations where there is limited overlap across treatment arms, the solution may involve extreme adjustments to the weights of some control units. Large weights increase the variance of the resulting estimate and may create a situation where the analysis relies too heavily on a small number of highly weighted controls[15]. One way to deal with this problem is to trim the weights. Hainmueller (2013) describes a weight refinement algorithm that can be implemented by iteratively calling the entropy balance search algorithm. In each iteration the set of solution weights w^* is trimmed at the user specified threshold and passed as a vector of starting weights q (instead of the uniformly distributed weights) for the subsequent call. The idea behind this weight refinement algorithm is to lower the variance of the weights.

To implement the algorithm we trim weights by setting weights larger than the 99 percentile of current weight solution to the value of the weight at the 99 percentile. Then the adjusted set of weights are passed as a vector of starting weights to the subsequent call of entropy balance. This iterative process reduced the standard deviation of the weights from 2.1 to 1.7 (after which the variance of weights stabilized). It also help reduce the value of the maximum weight by about 60% while still balancing the first three moments of the distribution of covariates. We then re-estimate the quantile effects with the new set of weights. The result presented in figure 4 shows that the result if figure 1 is robust to extreme weights.

5 Conclusion

This paper explores the impact of the CSG beyond the mean to uncover the dynamics of its effect on the quantiles of the height-for-age distribution. To

[15] We note that similar problem is shared by many pre-processing methods when matching with replacement reuses the good controls many times.

do this we estimate the quantile effect of CSG after balancing the first three moments of the relevant covariate distributions.

Our result shows that in the combined population of boys and girls, children above the 40th percentile of the height-for-age distribution drive the significant positive effect observed at the mean. For children below this threshold, CSG has no significant effect. We also show that while low birth weight is a significant correlate of height-for-age the effect of CSG remains in the combined population even after controlling for low birthweight.

More importantly, we found that the mean and the quantile effects hold only for female children. For male children, CSG has no significant effect on height-for-age at the mean or at the quantiles. We argue that this has a lot to do with low birthweight. Even though the low birthweight is more prevalent among female children, its negative effect is stronger in the population of male children. We also argue that caregiver motivation may be a factor since nutritional deficiencies in the early stages of life may not be reversible (Duflo, 2003).

The implication of this result is that while it is known that the effect of low birth weight continues to manifest in the growth of children, in the case of our sample it seems this effect is stronger for male children. Therefore reducing the prevalence of low birthweight in general for the beneficiaries of CSG will unlock more benefit for the CSG in terms of the height-for-age score of benefiting children. Furthermore, reducing the prevalence of low birthweight for boys, in particular, may reduce the inequality in malnutrition and boost the effect of the CSG across the quantiles of the height-for-age distribution.

There is one more possibility that has not been considered in our analysis. Duflo (2003) found that pension received by grandmothers in South Africa has a significant effect the girl child while it has no effect on the boy child. If similar effect exist for the CSG part of the effect can be due to caregiver bias. However, I believe this is unlikely due to a number of reasons. First, the old age pension is income for pensioners so the caregiver will feel he/she has freedom to spend it as he/she feels appropriate. For this grant the possibility that there is preferential treatment for girls is higher than in the case of the CSG which is attached to each child. It will at least be harder for a caregiver to receive a grant for one child and spend more of it on another child because of gender.

One policy proposal that can mitigate the effect of low birthweight is providing support for women during pregnancy (Chersich, et. al, 2016). The current practice of providing support only when the child is born limits the effectiveness of the support. Evidence in the literature and my result suggests that damage done by maternal deprivation during pregnancy may not be reversed by the CSG, especially for boys. This policy proposal will also help address the problem of delay and caregiver motivation to a large extent. If pregnant women are supported, their children will not lack support in the critical first two years of life. This will also increase the effect of the CSG on child nutrition and outcome later in life.

References

[1] Abadie A, Imbens G (2011). Bias Corrected Matching Estimators for Average Treatment effects. *Journal of Business and Economic Statistics*, 29(1), 1-11.

[2] Aislinn Delany, Zenobia Ismail, Lauren Graham, Yuri Ramkissoon. 2008. Review of the Child Support Grant: Uses, Implementation and Obstacles.

[3] Aguero, J., Carter, M., & Woolard, I. (2006). The impact of unconditional cash transfers on nutrition: The South African Child Support Grant. *Southern Africa labour & development research unit Working paper.*

[4] Aryastami, N. K., Shankar, A., Kusumawardani, N., Besral, B., Jahari, A. B., & Achadi, E. (2017). Low birth weight was the most dominant predictor associated with stunting among children aged 12–23 months in Indonesia. *BMC Nutrition*, *3*(1), 16.

[5] Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of economic surveys*, *22*(1), 31-72.

[6] Chan, K. C. G., Yam, S. C. P., & Zhang, Z. (2016). Globally efficient nonparametric inference of average treatment effects by empirical balancing calibration weighting. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *78*(3), 673-700.

[7] Chersich, M.F., Luchters, S., Blaauw, D., Scorgie, F., Kern, E., Van den Heever, A., Rees, H., Peach, E., Kharadi, S. and Fonn, S., 2016. Safeguarding maternal and child health in South Africa by starting the Child Support Grant before birth: Design lessons from pregnancy support programmes in 27 countries. *South African Medical Journal, 106(12), pp.1192-1210.*

[8] Cogill B (2003). Anthropometric Indicators Measurement Guide. *Food and Nutritional Technical Assistance Project*, Washington DC. 92 pp.

[9] Coetzee, M. (2013). Finding the Benefits: Estimating the Impact of The South African Child Support Grant. *South African Journal of Economics*, *81*(3), 427-450.

[10] Deming WE, Stephan FF (1940). On the Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals Are Known. *Annals of Mathematical Statistics*, 1940, 427-444.

[11] Diamond, A. and Sekhon, J.S., 2013. Genetic matching for estimating causal effects: A general multivariate matching method for achieving balance in observational studies. *Review of Economics and Statistics, 95(3), pp.932-945.*

[12] Duflo, E. (2003). Grandmothers and Granddaughters: Old-Age Pensions and Intrahousehold Allocation in South Africa. *The World Bank Economic Review, 17*(1), 1-25

[13] Espo M, Kulmala T, Maleta K, Cullinan T, Salin ML, Ashorn P (2002). Determinants of linear growth and predictors of severe stunting during infancy in rural Malawi. *Acta Paediatr*, 91**:**1364-1370.

[14] Firpo, S., Fortin, N. M., & Lemieux, T. (2009). Unconditional quantile regression. *Econometrica*, *77*(3), 953-973.

[15] Frolich, M. *Melly. B. (2008). Unconditional quantile treatment effect under endogeneity* (No. 3288). IZA Discussion paper.

[16] Grantham-McGregor, S., Cheung, Y. B., Cueto, S., Glewwe, P., Richter, L., Strupp, B., & International Child Development Steering Group. (2007). Developmental potential in the first 5 years for children in developing countries. *The lancet*, *369*(9555), 60-70.

[17] Grinspun, A. (2016). No small change: The multiple impacts of the Child Support Grant on child and adolescent well-being. *ChildGauge*, 44.

[18] Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, *20*(1), 25-46.

[19] Hainmueller, J., & Xu, Y. (2013). Ebalance: A Stata package for entropy balancing. *Stata journal.*

[20] Ireland CT, Kullback S (1968). Contingency Tables with Given Marginals. *Biometrika*, 55,179-188.

[21] Imai, K., & Ratkovic, M. (2014). Covariate balancing propensity score. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *76*(1), 243-263.

[22] Kirchengast, S. and Hartmann, B., 2009. The male disadvantage hypothesis reconsidered: is there really a weaker sex? An analysis of gender differences in newborn somatometrics and vital parameters. *Journal of Life Sciences, 1(1), pp.63-71.*

[23] Oyenubi, A. (2018). *Quantifying balance for casual inference: An information theoretic perspective* (Doctoral dissertation, University of Cape Town).

[24] Rahman, M. S., Howlader, T., Masud, M. S., & Rahman, M. L. (2016). Association of low-birth weight with malnutrition in children under five years in Bangladesh: do mother's education, socio-economic status, and birth interval matter?. *PloS one*, *11*(6), e0157814.

[25] Roy, P., Kumar, A., Kaur, I.R. and Faridi, M.M., 2014. Gender differences in outcomes of low birth weight and preterm neonates: the male disadvantage. *Journal of tropical pediatrics, 60(6), pp.480-481.*

[26] Wamani H, Tylleskär T, Åstrøm AN, Tumwine JK, Peterson S (2004). Mothers' education but not fathers' education, household assets or land ownership is the best predictor of child health inequalities in rural Uganda. *Int J Equity in Health*, 3:9.

[27] Wamani, H., Åstrøm, A. N., Peterson, S., Tumwine, J. K., & Tylleskär, T. (2007). Boys are more stunted than girls in sub-Saharan Africa: a meta-analysis of 16 demographic and health surveys. *BMC pediatrics*, *7*(1), 17.

[28] Walker, S. P., Chang, S. M., Wright, A., Osmond, C., & Grantham-McGregor, S. M. (2015). Early Childhood Stunting Is Associated with Lower Developmental Levels in the Subsequent Generation of Children, 2. *The Journal of nutrition*, *145*(4), 823-828.

[29] Williams, M. J. (2007). The social and economic impacts of South Africa's child support grant. *Economic Policy Research Institute Working Paper*, *40*.

[30] Zere, E., & McIntyre, D. (2003). Inequities in under-five child malnutrition in South Africa. *International journal for equity in health*, *2*(1), 7.

[31] Zhang, R., Undurraga, E.A., Zeng, W., Reyes-García, V., Tanner, S., Leonard, W.R., Behrman, J.R. and Godoy, R.A., 2016. Catch-up growth and growth deficits: Nine-year annual panel child growth for native Amazonians in Bolivia. *Annals of human biology, 43(4), pp.304-315.*

Table 1: Summary statistics (sample 1)

	Treatment group				Control group					
Variable	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean difference	t stat
delay years	2.05	2.07	0.00	9.26	7.67	5.39	0.00	16.05	5.612***	(43.78)
motivation	0.14	1.15	-1.45	4.39	-0.22	0.65	-1.18	3.59	-0.356***	(-10.60)
employed	0.24	0.43	0.00	1.00	0.24	0.43	0.00	1.00	-0.00412	(-0.28)
married	0.35	0.48	0.00	1.00	0.41	0.49	0.00	1.00	0.0543**	(3.29)
Primary caregivers edu	7.36	4.38	0.00	19.00	6.10	4.79	0.00	20.00	-1.266***	(-8.16)
Primary caregivers age	38.16	13.65	14.00	89.00	43.83	17.69	15.00	88.00	5.673***	(10.81)
HH head gender	0.36	0.48	0.00	1.00	0.35	0.48	0.00	1.00	-0.00964	(-0.59)
Child's gender	0.50	0.50	0.00	1.00	0.50	0.50	0.00	1.00	0.000321	(0.02)
electricity	0.64	0.48	0.00	1.00	0.63	0.48	0.00	1.00	-0.00634	(-0.39)
water	0.17	0.37	0.00	1.00	0.19	0.39	0.00	1.00	0.0242	(1.87)
telephone	0.03	0.16	0.00	1.00	0.04	0.19	0.00	1.00	0.00897	(1.51)
toilet	0.20	0.40	0.00	1.00	0.24	0.43	0.00	1.00	0.0478***	(3.42)
Food expenditure (AES)	183.92	120.82	4.51	1704.96	195.55	151.87	4.45	1809.98	11.63*	(2.54)
Child's age	6.64	3.74	0.26	15.25	7.85	5.40	-0.63	16.05	1.204***	(7.89)
rural	0.74	0.44	0.00	1.00	0.72	0.45	0.00	1.00	-0.0274	(-1.81)
Child's relationship to grant recipient	0.74	0.44	0.00	1.00	0.00	0.00	0.00	0.00	-0.743***	(-64.03)
Informal dwelling	0.08	0.27	0.00	1.00	0.08	0.26	0.00	1.00	-0.00345	(-0.38)
Hospital visits (No=1)	0.45	0.50	0.00	1.00	0.49	0.50	0.00	1.00	0.0367*	(2.15)
observations	2170				1418				3588	

t statistics in parentheses * p<0.05, ** p<0.01, *** p<0.001

Table 2: First three moments before and after entropy balancing

Entropy weighting						
	Treat			Control		
	mean	variance	skewness	mean	variance	skewness
motivation	.1414	1.327	.7458	.1413	1.327	.7458
employed	.2429	.184	1.199	.2428	.184	1.2
married	.3521	.2282	.6194	.352	.2283	.6198
Primary caregivers edu	7.365	19.2	-.2611	7.365	19.2	-.2612
Primary caregivers age	38.16	186.4	.8362	38.16	186.5	.8362
HH head gender	.3618	.231	.5754	.3617	.231	.5758
Child's gender	.5046	.2501	-.01843	.5046	.2502	-.01844
electricity	.6373	.2312	-.5713	.6374	.2313	-.5716
water	.1659	.1384	1.796	.1659	.1384	1.797
telephone	.02765	.0269	5.762	.02765	.0269	5.762
toilet	.1959	.1576	1.533	.1958	.1576	1.533
Food expenditure (AES)	183.9	14598	3.71	183.9	14603	3.71
Informal dwelling	.0788	.07263	3.127	.07878	.07263	3.127
Hospital visits	0.4535	0.2479	0.187	0.4535	0.2479	0.187

Note that by definition any difference between the moments are not significant

Table 3: Effect of CSG with and without low birthweight

Dependent variable height-for-age score	(1)	(2)	(3)	(4)
Treatment (Individual)	0.24***		0.25***	
	(0.07)		(0.07)	
Treatment (Household)		0.19*		0.22**
		(0.11)		(0.11)
Low birthweight			-0.45***	-0.74***
			(0.16)	(0.24)
Constant	-1.45***	-1.09***	-1.44***	-1.11***
	(0.18)	(0.31)	(0.18)	(0.31)
Other covariates	included	included	included	included
Observations	2,837	1,976	2,837	1,976
R-squared	0.02	0.02	0.02	0.02

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.
By other covariates I mean the covariates listed in table 2.

Table 4: Covariate difference in the first and last quintile of height-for-age scores for benefiting children

	5th Quintile minus 1st Quintile	
delay years	-0.197	(-1.27)
motivation	-0.0664	(-0.79)
employed	0	(0.00)
married	-0.0440	(-1.21)
Primary caregivers edu	-0.118	(-0.35)
Primary caregivers age	-2.047	(-1.91)
HH head gender	-0.00824	(-0.22)
Child's gender	0.114**	(3.04)
electricity	0.0366	(0.99)
water	-0.0811**	(-2.84)
telephone	-0.0239*	(-2.06)
toilet	-0.00951	(-0.33)
Food expenditure (AES)	-13.09	(-1.53)
Child's age	-0.216	(-0.77)
rural	0.0199	(0.62)
Child's relationship to grant recipient	0.0404	(1.21)
Informal dwelling	-0.0120	(0.63)
Low birthweight	0.0433**	(2.76)
Hospital visits	-0.0142	(-0.38)
Observations	702	

t statistics in parentheses * p<0.05, ** p<0.01, *** p<0.001

Table 5: Effect of CSG with and without low birthweight (by gender)

Dependent variable	(1)	(2)	(3)	(4)
height-for-age score	Girls		Boys	
Treatment(individual)	0.32***	0.33***	0.14	0.13
	(0.09)	(0.09)	(0.10)	(0.10)
Low birthweight		-0.32		-0.57***
		(0.33)		(0.15)
Constant	-1.45***	-1.45***	-1.48***	-1.46***
	(0.23)	(0.23)	(0.24)	(0.24)
Observations	1,401	1,401	1,436	1,436
R-squared	0.03	0.03	0.02	0.03

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Figure 1: Unconditional quantile effect of CSG

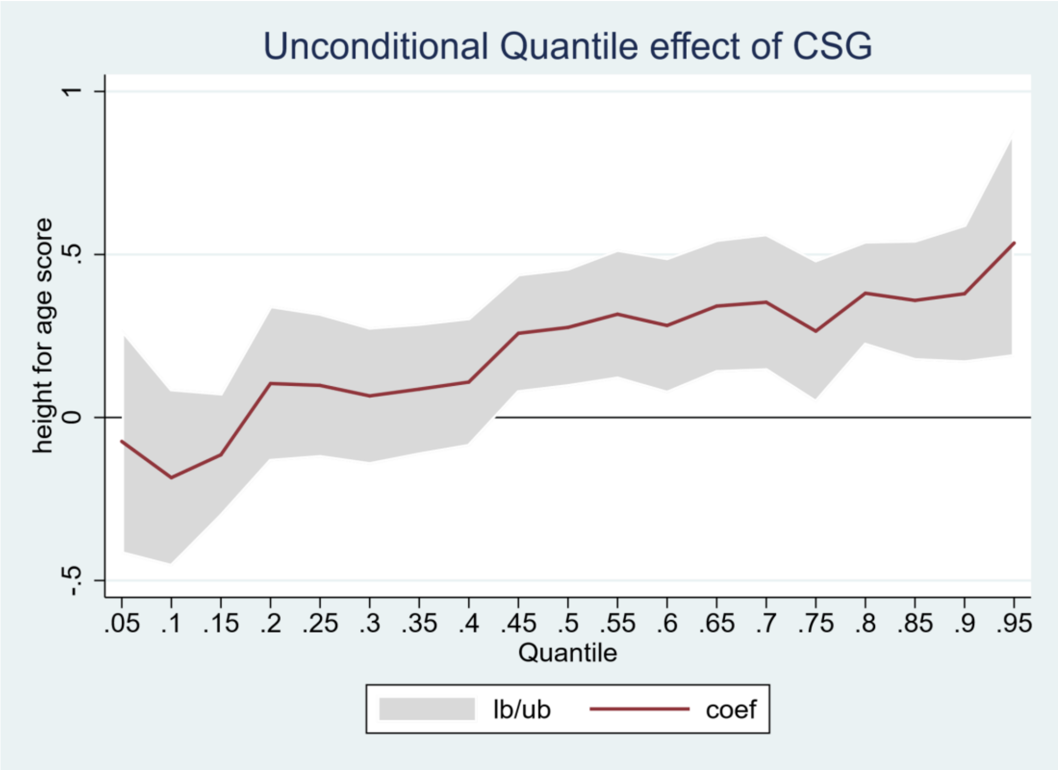


Figure 2: Unconditional quantile effect of CSG (male children)

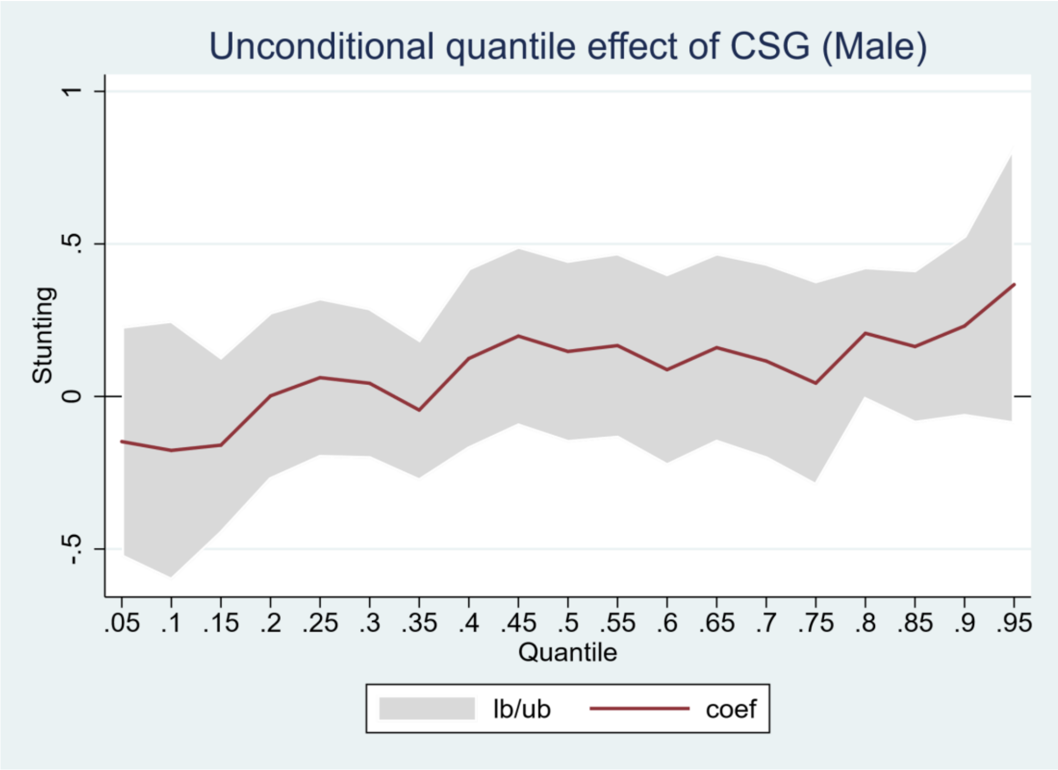


Figure 3: Unconditional quantile effect of CSG (female children)

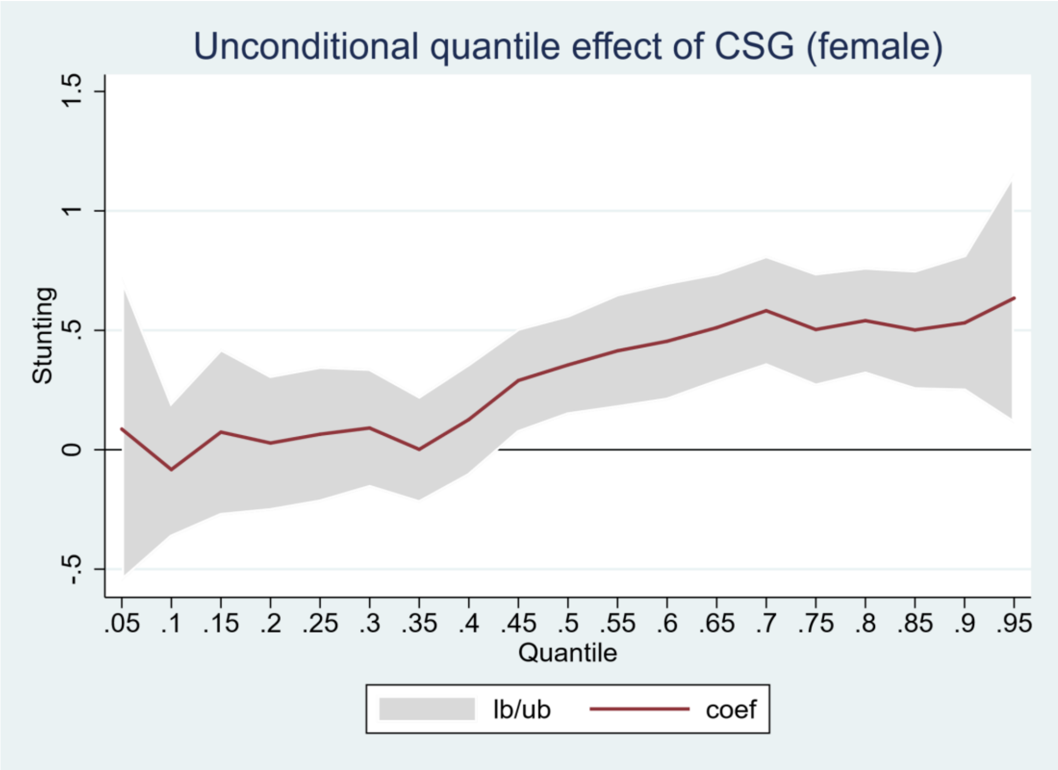
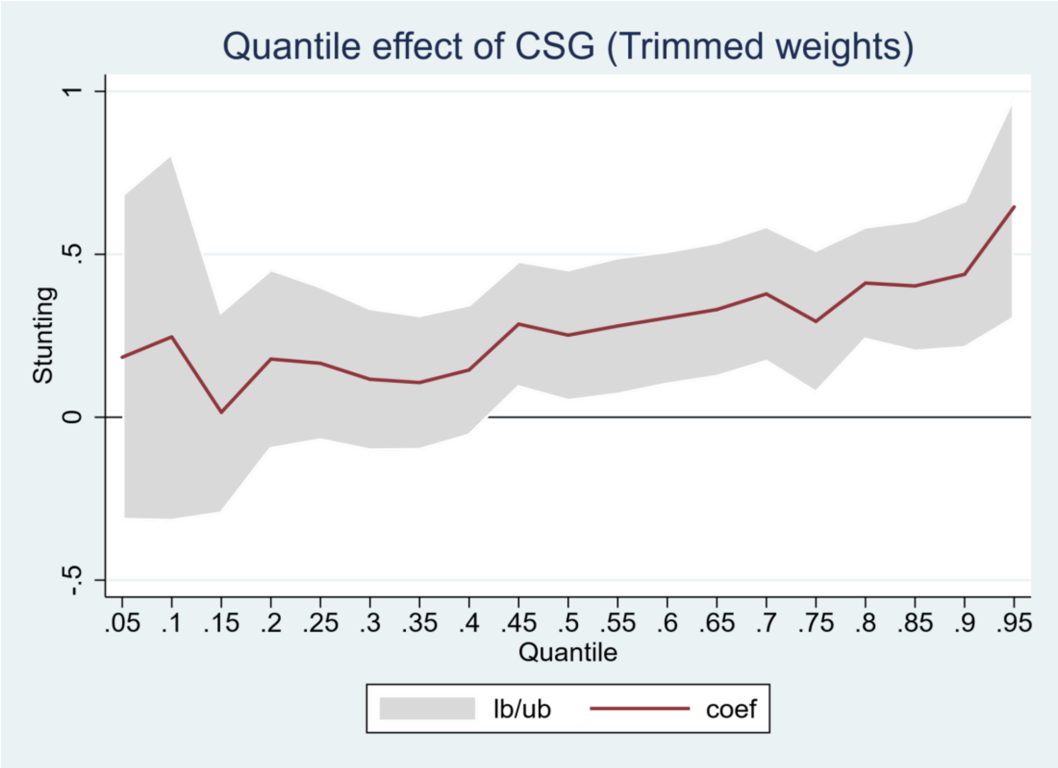


Figure 4: Unconditional quantile effect of CSG (Trimmed weights)



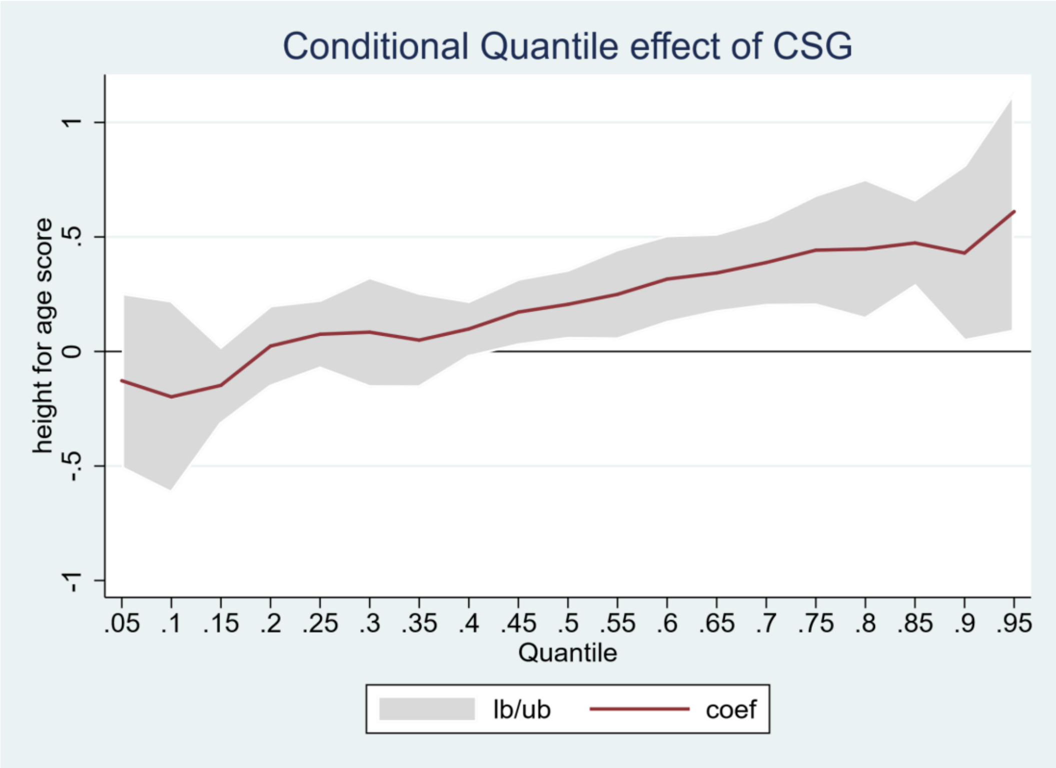
Appendix

Table A1: Summary statistics (sample 2)

	Treatment group				Control group					
Variable	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean difference	t stat
delay years	2.04	2.08	0.00	9.26	7.53	5.19	0.04	15.33	5.479***	(37.24)
motivation	0.13	1.16	-1.45	4.39	-0.20	0.62	-1.11	3.59	-0.330***	(-7.93)
employed	0.25	0.44	0.00	1.00	0.25	0.43	0.00	1.00	-0.00659	(-0.36)
married	0.35	0.48	0.00	1.00	0.41	0.49	0.00	1.00	0.0623**	(3.10)
Primary caregivers edu	7.56	4.34	0.00	19.00	6.30	4.84	0.00	20.00	-1.260***	(-6.69)
Primary caregivers age	38.11	13.61	14.00	89.00	44.60	18.41	15.00	88.00	6.483***	(10.05)
HH head gender	0.38	0.48	0.00	1.00	0.36	0.48	0.00	1.00	-0.0199	(-0.99)
Child's gender	0.50	0.50	0.00	1.00	0.50	0.50	0.00	1.00	-0.00596	(-0.29)
electricity	0.66	0.47	0.00	1.00	0.67	0.47	0.00	1.00	0.00903	(0.46)
water	0.16	0.37	0.00	1.00	0.21	0.41	0.00	1.00	0.0500**	(3.11)
telephone	0.03	0.16	0.00	1.00	0.04	0.19	0.00	1.00	0.0110	(1.51)
toilet	0.21	0.41	0.00	1.00	0.29	0.45	0.00	1.00	0.0820***	(4.64)
Food expenditure (AES)	192.57	123.40	4.51	1479.61	213.34	166.13	4.45	1809.98	72.59***	(4.21)
Child's age	6.69	3.81	0.26	15.25	7.68	5.19	-0.16	15.33	0.985***	(5.43)
rural	0.72	0.45	0.00	1.00	0.67	0.47	0.00	1.00	-0.0566**	(-2.98)
Child's relationship to grant recipient	0.76	0.43	0.00	1.00	0.00	0.00	0.00	0.00	-0.759***	(-53.51)
Informal dwelling	0.09	0.29	0.00	1.00	0.10	0.30	0.00	1.00	0.00842	(0.70)
Hospital visit (No=1)	0.45	0.50	0.00	1.00	0.49	0050	0.00	1.00	0.0345	(1.66)
observations		1596				907			2503	

t statistics in parentheses * p<0.05, ** p<0.01, *** p<0.001

Figure A1: Conditional quantile effect of CSG (sample 2)



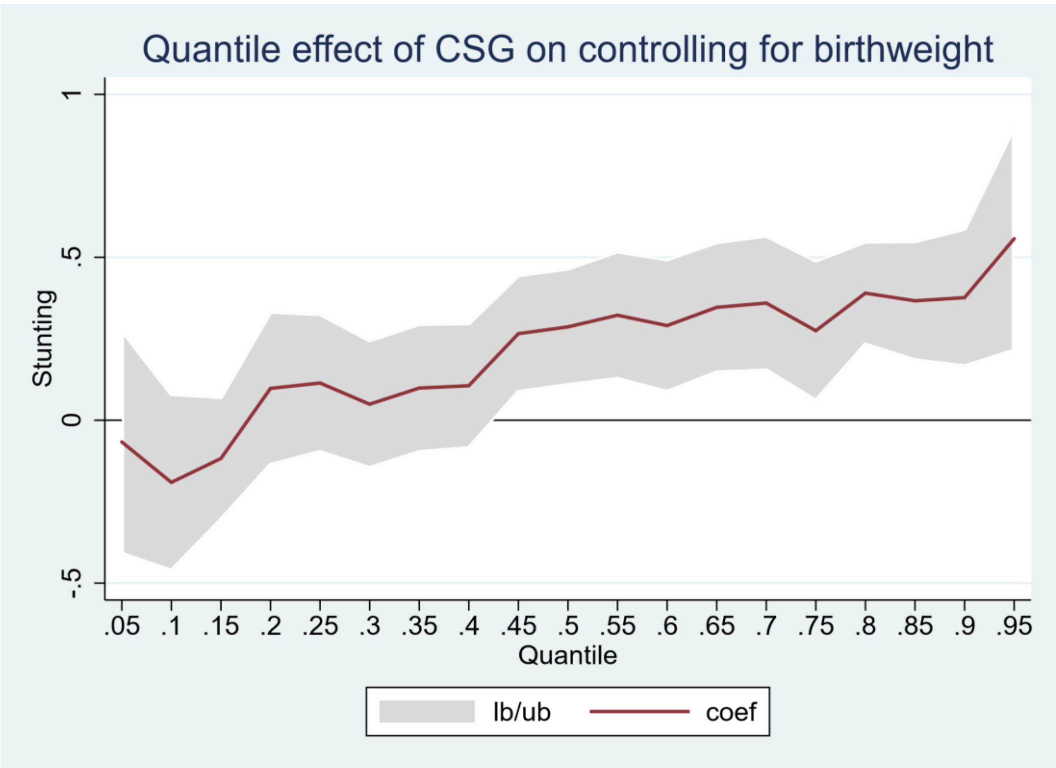


Figure A3: Cumulative distribution of birth weight