The Impact of Unconditional Cash Transfers on Nutrition: The South African Child Support Grant*

Jorge M. Agüero†
University of California, Riverside

Michael R. Carter
University of California, Davis

Ingrid Woolard
University of Cape Town

August 2010

*We thank DfID and USAID for financial support. We would like to especially thank Laura Schechter, Steven Helfand and seminar participants at the University of Wisconsin-Madison, the American Agricultural Economics Association, University of Colorado-Denver, University of Florida-Gainesville, UC Riverside, Centre for the Study of African Economies, PEGNet and the North American Summer Meetings of the Econometric Society.

†Corresponding author. Contact information: 4108 Sproul Hall, Riverside CA 92521, Phone: 951.827.4108 Fax: 951.827.5685, E-mail: jorge.aguero@ucr.edu.

Abstract

This paper analyzes the impacts of South Africa's unconditional Child Support Grant (CSG) in which monthly cash grants are targeted at women, but no strings are attached. Exploiting spatial and temporal variation in program rollout that generates exogenous variation in the duration of CSG treatment, this paper employs continuous treatment methods to identify the impact of grants on child height-for-age. The analysis finds significant impacts, which predict large future economic returns, suggesting reconsideration of the presumption that costly conditionalities attached to transfer program elsewhere are necessary.

JEL: I12, H53, C21, D13.

Keywords: Nutrition, cash transfers, continuous treatment estimator, South Africa.

Despite improvements over the last two decades, child malnutrition remains a serious health problem in developing countries, and is the main contributor to child mortality (World Bank 2006). For those that survive, early childhood malnutrition contributes to the intergenerational transmission of poverty. In a striking study of Ecuadorian children, Christina Paxson & Norbert Schady (2007) show that by age 5, children who were likely inadequately nourished have already fallen well behind the cognitive development of their better nourished peers. Assuming that these disparities persist, the malnourished children will likely do less well in school, accumulate less human capital and enjoy lower adult earnings than their peers. We would expect that their own children would in turn repeat this inequitable cycle.

In an effort to break this intergenerational transmission of poverty (poverty which "lays its own eggs," in the words of an informant quoted in the Chronic Poverty Research Centre (2004)), some middle income countries have adopted cash transfer programs designed to bolster the nutrition, health and education of the children of poor families. At first glance, reliance on cash transfers to achieve these goals may seem somewhat surprising in light of the earlier debate about whether nutrition responds at all to income increases amongst poor families. For example, Jere R Behrman & Anil B Deolalikar (1987) find an income elasticity close to zero for a sample of families in south India. While other studies find a positive elasticity, the issue is far from solved.[1]

There are, however, several critical differences between the new generation of cash transfer programs and the market-generated income difference used to identify nutritional elasticities in the earlier literature. Mexico's *Progresa* program (now *Oportunidades*), the best known of these cash transfer programs, has two key design features that may mediate its nutritional impacts.[2] First, cash transfers are conditional on the household meeting certain required

[1]For example, see the studies by Shankar Subramanian & Angus Deaton (1996), Howarth E. Bouis & Lawrence J. Haddad (1992) and Jere Behrman, Mark Rosenzweig & Andrew D. Foster (1994). Lawrence Haddad, Harold Alderman, Simon Appleton, Lina Song & Yisehac Yohannes (2003) provides a review of the literature over the last 20 years.

[2]Other Latin American cash transfer programs include the Bolsa Escola program in Brazil (Mary Arends-Kuenning, Ana Fava, Ana Lucia Kassouf & Alexandre de Almeida 2005), the Social Protection Network program in Nicaragua (John A. Maluccio & Rafael Flores 2005) and PRAF in Honduras (Paul Glewwe & Pedro Olinto 2005).

behaviors: Older children must attend school; and, younger children must visit clinics for regular medical check-ups and nutritional monitoring (where among other things they are given nutritional supplements).

In addition to these conditionalities, *Progresa* cash transfers are also assigned to women. Unlike market driven income increases which may have been generated by increases in returns to assets owned by men, these targeted cash transfers have been designed to bolster the bargaining power of women with the idea of giving more weight to their preferences which are presumed to be more child-centric.[3]

While much of the impact evaluation literature on *Progresa* has focussed on schooling outcomes, there is evidence that *Progresa* has boosted child nutritional status (Jere R. Behrman & John Hoddinott 2005). However, because *Progresa* transfers were conditional (and included in-kind nutritional supplements), it is not clear whether these findings indicate a non-zero income elasticity of nutrition, or simply the impact of transfer conditionality. Unpacking the reasons behind this response is of more than academic interest. Aid agencies have noted that the heavy administrative burdens implied by transfer conditionality limit the ability of lower income African economies to implement programs modeled on *Progresa*. Given that it is precisely these economies where malnutrition is most severe, understanding the importance of costly conditionality is important.[4]

In their recent book, Ariel Fizbein & Norbert Schady (2009) note that paternalistic conditionalities intended to boost the elasticity of expenditures on nutrition, health and education are justified when parents have "persistently misguided beliefs" that lead them to underinvest in the human capital of their children.[5] But do poor parents have such

[3] A number of the key studies on the income elasticity of nutrition drew on data generated as part of the International Food Policy Research Institute (IFPRI) commercialization of agriculture studies. In most cases, the IFPRI studies concerned communities where technological and other changes enhanced returns to male-owned assets, raising the issue as to whether low estimated nutritional elasticities represented immutable family preferences or simply a re-weighting of child nutrition unfriendly male preferences (see the discussion in the summary volume by Joachim von Braun & Eileen T. Kennedy (1994).

[4] On the basis of their ex ante analysis, N. Kakwani, F. Soares & H. Son (2005) suggest that without school attendance conditionalities, cash transfers in Africa will not increase school attendance. They do not, however, speak to the question of the necessity of conditionalities to boost nutrition.

[5] When parents' beliefs are misguided in this way, a second benefit of conditionalities is that they bolster

misguided beliefs? In their aptly titled book, *Just Give Money to the Poor*, Joseph Hanlon, Armando Barrientos & David Hulme (2010) argue strenuously that they do not, and that the distortions and extra administrative costs associated with conditionalities are unnecessary.[6]

Despite its importance, the direct evidence to date on this question is slender. Maria Caridad Araujo & Norbert Schady (2006) try to explore this question by exploiting the fact that some recipients of cash transfers thought that the transfers were unconditional, while others did not. While they found that impacts were greater amongst those who believed the transfers were conditional, their results potentially misidentify the impacts of conditionalities with the characteristics of those who (correctly) understood the nature of the program. In a study designed to directly study the question of conditionalities, Sarah Baird, Craig Mcintosh & Berk Ozler (2010) find that impacts of cash transfer on school attendance in Malawi were the same irrespective of whether or not the cash transfers were conditional.

This paper aims to contribute to knowledge in this area by studying the impact of the unconditional South African Child Support Grant (CSG), which was first rolled out in 1998. Like *Progresa* and its sister programs in Latin America, CSG cash transfers are targeted at women. Unlike those programs, CSG transfers are unconditional, and come with no strings attached, nor with any in-kind transfers. Analysis of this program thus promises a sharper look at the unconditional (natural) income elasticity of nutrition, at least in the context of income transfers targeted at women.

However, in comparison to the *Progresa* program, evaluation of the CSG presents a particular methodological challenge. By randomly selecting rural areas to receive the cash transfer treatment, *Progresa* quickly became a showpiece for impact evaluation. In contrast, the CSG was a rolled out as a single, national program, depriving analysts of purposefully randomized treatment and control groups. Alternative methods are thus needed to estimate the impact of the South African CSG.

---

support for redistribution amongst the tax paying public (see Fizbein & Schady 2009, Jonah B. Gelbach & Lant Pritchett 2002).

[6]Emmanuel Skoufias (2005) calculates that roughly one third of the administrative costs associated with Mexico's Progresa program are the result of conditionalities.

One method would be to follow the current literature and use matching methods to evaluate CSG support as a binary treatment. Statistical problems aside, this approach would overlook the fact that extent of CSG treatment (the 'dosage') varies significantly across the treated population. During the nutritionally critical first 36 months of life (which will be the focus of the analysis here), some children received CSG support nearly 100 percent of the time, while others received only a month or two of support. The nutritional impact of the latter is likely negligible, while the impact of the former could be substantial. The analysis here will thus use variation in the extent of treatment to identify the impact of the CSG.[7]

While the continuous treatment estimator of Keisuke Hirano & Guido W Imbens (2004) opens the door to this kind of analysis of the extent of treatment among the treated, it still requires strong orthogonality assumptions. The critical statistical identification assumption is that extent of treatment is unrelated to unobserved factors that themselves affect child health and nutrition. While this assumption may seem hard to sustain for a voluntary program that had a single national eligibility date, we show how information on effective program rollout can be used to derive a caregiver "eagerness"[8] measure that can be used to control for otherwise unobservable, and confounding, characteristics. Conditional on eagerness and other covariates, we argue that exposure to the CSG depends (randomly) on the interaction between the child's birthdate and the effective program roll-out for the child's locational and temporal cohort.

Exploiting this eagerness variable, our analysis of the impacts of early-life CSG as a continuous treatment case uncovers economically and statistically significant effects for large dosages of CSG support. These estimates show that effects are insignificant for children

[7] Continuous treatment estimators that exploit the duration of treatment are especially important in cases like ours where longer-term impacts cannot be reliably estimated by simply projecting out short-term rates of impact. In an analysis of the Mexican cash transfer program, Jere R Behrman, Piyali Sengupta & Petra Todd (2005) try to solve this problem by assuming impacts on an initially older cohorts predict the long term impacts on initially younger cohorts. In an analysis methodologically similar to this paper, Jere R. Behrman, Susan W. Parker & Petra E. Todd (2009) revisit the Mexican program exploiting later data that allows them to exploit (randomly generated) differences in the intensity of treatment and to more reliably trace out the longer-term impacts of the cash transfer program.

[8] An eager caregiver is one who enrolls her child early in the CSG program relative to other caregivers in the same time and place.

who received CSG support for less than 50 percent of their 36 month window. These results are robust to the inclusion of cluster variables (village fixed-effects) meant to control for differences in the supply of health-related public goods and other locational differences. We also show that there is no independent time trend in child health that might confound our results.

Finally, in an effort to get an understanding of the possible economic value of these nutritional gains, we project forward in time using best estimates from the literature concerning the impact of adult height on wages. Adaptation of these estimates to the South African reality suggest that the present value of early CSG support is 1.39 times as large as the direct cost of that support. These findings in no way imply that there are not further gains from CSG support later in childhood. Indeed, these results suggest that such further gains are quite likely. However, within the confines of this study, we have not been able to estimate their magnitude.

The remainder of this paper is organized as follows. Section 1 provides background description of the South Africa Child Support Grant. Section 2 presents descriptive statistics on the program, discussing key measures and identification strategies that are available. Section 4 presents the methods and results from using continuous treatment effects. Section 5 attempts to infer the lifetime economic value of the nutritional impact identified in Section 4, and Section 6 concludes the paper.

1 The South African child support grant

The South African system of state welfare transfers changed little in terms of its basic structure up to the 1990s. The system remained dominated by means-tested, non-contributory old age pensions and disability pensions with discrimination between different racially defined population groups in terms of access to the grants and the levels of benefits. Substantial progress was, however, made during the early 1990s in removing racial discrimination from these two programs. This equalization resulted in increased access and real benefit levels for Africans and reduced real levels of benefits for white pensioners. Fiscal costs expanded significantly. Despite its apartheid past, South Africa now possesses a substantial system of state provided, cash social assistance with wide coverage of the population, most notably of the formerly disenfranchised African majority. This important role of cash social assistance is fairly exceptional compared to most other middle-income countries (see F Lund 1993, Frances Lund 2001, Servaas Van der Berg 1997, Anne Case & Angus Deaton 1997).

The reforms of the early 1990s did not extend to the third most important component of state transfer payments, namely State Maintenance Grants (SMGs). These grants, payable in the form of parent and child grants were means-tested benefits payable to a natural parent who could not, for a number of reasons, rely on the support of the second parent. If the second parent was alive, it was necessary to apply for a private maintenance order through the courts and only if this failed (or the amount awarded was very low) was the child eligible for the grant. This bureaucratic hurdle in conjunction with very low awareness of the grant was effective in excluding many eligible children from accessing the grant. In 1990, only 0.2 percent of African children were in receipt of the SMG, while 1.5 percent of white children, 4.0 percent of Indian children and 4.8 percent of Coloured children received benefits (John Kruger 1998). It became apparent in the mid-1990's that providing equal access to the SMG would have severe fiscal implications given poverty levels and household structures, with simulations based on household survey data predicting a more than twenty-fold increase in expenditures (Claudia Haarman & Dirk Haarman 1998).

In December of 1995, the new democratic government of South Africa established the Lund Committee to evaluate the existing system of state support and to explore new alternative policy options targeting children and families. In 1998 the Child Support Grant (CSG) replaced the existing SMG. The benefit was initially limited to children under seven (unlike the SMG which covered children up to age 18). In proposing the CSG, the Lund Committee emphasized that the grant must “follow the child”, meaning that the benefit should be independent of the child's family structure. This approach represented a move from a family-based benefit to a child-focused one. Legally, however, the grant must be paid over to an adult and it is the intention that the person to whom the grant is paid is the "primary care giver" of the child for whom the benefit is intended. In cases where the applicant is not the biological parent of the child, a sworn affidavit from the parents or guardians is required to confirm that the applicant is indeed the primary care giver. In practice, the designation of the primary care giver as the grant recipient has effectively targeted women. In the data used here, 98 percent of designated primary care givers are female.

When the Child Support Grant (CSG) was introduced it was intended to cover the poorest 30 percent of children and was means-tested, i.e. the child had to be residing in a household with a household income below a certain threshold. The threshold was set at R800 (approximately US$110) for children living in urban areas and at R1100 (US$150) for those living in rural areas or in informal settlements. In 1999, due to a low take-up rate, the Department of Welfare altered the income test from a household based measure to one which considered only the income of the primary caregiver plus that of his/her spouse (net of other state transfers). The means test has remained unchanged in nominal terms since 1998, despite the fact that the Consumer Price Index rose 40 percent between April 1998 and September 2004. Despite this increasingly strigent means test, about half of age-eligible children were in receipt of the grant by this latter date.[9]

The government has increased the age limit for eligibility in recent years. In April 2003 the age limit was raised to nine years old and a year later this was increased to eleven years. In April 2005 the age limit was raised to fourteen. The amount granted has also changed since 1998 and the increases have outstripped inflation. While the initial monthly benefit was R100 in 1998, it is currently R180. During the time of the survey which we discuss, the monthly benefit was R170 which equates to approximately to US$25 using the market exchange rate (or, PPP US$50).

[9] In September 2004 there were slightly over 7 million children aged 0 to 6 in South Africa. Administrative data from the Department of Social Development indicates that 3.54 million of them were receiving the CSG.

2 The KwaZulu-Natal data and the measurement of nutrition

The data for this study comes from the KwaZulu-Natal Income Dynamics Study (KIDS).[10] The KwaZulu-Natal province is home to approximately 20 percent of South Africa's population of 40 million and was formed in 1994 by combining the former Zulu homeland with the old Natal province. Although KwaZulu-Natal is not the poorest province in South Africa, it arguably has the highest incidence of deprivation in terms of access to services and perceived well-being (S. Klasen 1997, M. Leibbrandt & I. Woolard 1999). KwaZulu-Natal is also home to most of South Africa's ethnically Indian people who constitute 12 percent of the province's population. Africans comprise about 85 percent of the province's population, with people of European descent (largely British) comprising most of the remainder.

As explained in greater detail by Julian May, Jorge Agüero, Michael Carter & Ian Timaeus (2007), respondents to the KIDS study were first interviewed in 1993 as part of a nationwide living standards survey. In 1998, study reinterviewed households from the 1993 survey that were located in the KwaZulu-Natal province. A third round of surveys was undertaken in 2004. While the 1993 survey constructed households based on the residents of randomly selected dwellings, the 1998 and 2004 studies focused on reinterviewing designated 'core people' from the 1993 surveys. A household member was designated as a Core person if s/he satisfied any of the following criteria:

- A self-declared head of household (from 1993).
- Spouse/partner of self-declared head of household (from 1993).
- Lived in a three generation household and all of the following were true: Child, child-in-law, or niece/nephew of self-declared head at least 30 years old have at least one child living in household.

[10] The KIDS data set can be downloaded from `http://sds.ukzn.ac.za`.

- Spouse/partner of person satisfying criterion.

Thus all heads of households and spouses of heads are automatically included and in some three-generation households, adult children of household heads are included. The 2004 survey was able to interview at least one core person from 71 percent of the original 1993 households. In addition, children of core people who had had their own children and established new, separate residences were also interviewed in 2004. The analysis to follow will consider children resident with core people as well as grandchildren of core people now living separately. Note that children who are resident with core people includes grandchildren of the cores whose parents have not established independent residences.

2.1 Height as an *ex post* indicator of early childhood nutrition

The CSG program issues monthly payments to the care givers of eligible children. Evaluating the impact of such payment flows requires an indicator whose *ex post*, measurable value reflects the cumulative effects of those flows. School attainment, or amount learned are the kind of after the fact observable stock measures whose values reflect earlier inflows of educational inputs. Similarly, a child's height-for-age z-score (HAZ) can serve as an ex post indicator of nutritional inputs, especially for inputs received during the first three years of life.[11]

Stunting, or short height relative to standards established for healthy populations, is an indicator of long-term malnutrition. Stunting is an indicator of past growth failure. It is associated with a number of long-term factors, including chronic insufficient protein and energy intake, frequent infection, sustained inappropriate feeding practices and poverty. In children under 3 years of age, the effects of these long-term factors may not be reversible (UNICEF 1998, p. 21-23). Put differently, children under 3 are particularly vulnerable to nutritional shortfalls, and the impacts of poor nutrition during the first three years of life

[11] Z-scores (z) are defined as $z = \frac{h-\bar{h}}{\sigma_h}$ where h is height, $\bar{h}$ and σ_h are, respectively, the mean and standard deviation of height given the age. See Jere R Behrman & Anil B Deolalikar (1988) and John Strauss & Duncan Thomas (2008) for reviews of the role of height on children's nutrition.

are likely to leave a permanent mark on the child's z-score. We should thus be able to *ex post* evaluate the nutritional impact of the CSG by looking at the impact of CSG payments received during the child's 0-3 years of age "window of nutritional vulnerability."

While z-scores offer a promising way to examine the impact of income transfers on nutrition, there are two important differences between z-scores and the household food use or individual food ingestion measures used in much of the nutritional elasticity literature. First, behavioral changes potentially induced by income increases (*e.g.*, the purchase of more food) may not increase height if the child's body is unable to process or effectively use additional nutrients. Thus, a failure to find a response of HAZ to an income increase could either reflect the lack of a behavioral response (*i.e.*, the household purchased no more food), or the inability of individuals to physically transform increased nutritional inputs into improved nutritional status.[12]

A second difference between a HAZ and food use or intake measures of nutrition concerns the likely extent of measurement error. HAZ simply requires measurement of child height and age. In contrast, measurement error is an important factor when measuring food consumption because in most cases the data is collected retrospectively and the most informed person about food purchases is not always present to answer the survey questions (John Strauss & Duncan Thomas 1998).[13] Failure to find significant nutritional elasticities in studies that use food use or intake measures could thus be a problem of noisy data.

While HAZ measures are not free of such errors (Behrman & Deolalikar 1988), the data used here come from a survey where measures of height were taken at least twice and should

[12]Nutritionists discuss how many nutrients the human body can absorb from food compared to preformed nutrients (i.e., fortified foods or pharmaceutical supplements). For example, Kristina L Penniston & Sherry A Tanumihardjo (2006) discuss the case of vitamin A (or β-carotene) where the absorbtion rate is 70-90 percent from preformed vitamin A but only between 20-50 percent from regular food. For the latter, the rate of absorbtion depends on dietary and non dietary factors.

[13]In addition, in many studies, the constructed variable is not "nutrient intake" but rather "nutrients available" as acknowledge by Subramanian & Deaton (1996) among others. The main problem here is that a non negligible part of the purchased (or cooked) food is lost or wasted, leading to an overstatement of intake. Also, obtaining nutrient availability from food consumption is more complicated for items such as food away from home, affecting especially households with higher income or those in urban areas. Finally, this transformation also is affected by issues such as the quality of the food purchased (Behrman & Deolalikar 1988). Food quality is rarely included in household surveys.

therefore be quite reliable. As described in the KIDS fieldworker manual (available from the data website), a child's height was taken twice and the enumerators needed to compare both measures to make sure there the difference never exceed 0.5cm. In Table 1 we show that the mean absolute difference of the two measures is 0.015. The median difference is zero for all cases and only two children have measures that differ in more than 0.5cm. In addition, whenever possible, the child's age was taken from the child's public health card. These procedures support the idea that measuring nutrition using HAZ is subject to a minimal level of measurement errors.

Before turning to the analysis, one final comment is warranted. While the analysis here will attempt to measure the impact of CSG support received during the critical first three years of the child's life, this analytical choice does not mean that CSG support outside this three year window is unimportant. Indeed, it might be quite critical. However, it is the likely irreversibility of early nutritional effects that make it more likely that we can with greater confidence and accuracy measure the impact of the CSG using only information on early treatment.

2.2 Descriptive statistics

Table 2 presents descriptive statistics regarding children, their caregivers and the child support grant using the KIDS data. Reported in Table 2 are data on all age-eligible children. Children are grouped according to whether or not they received CSG support during the critical 3 year window from birth to 36 months of age. Had the CSG program been experimentally rolled out as was Mexico's *Progresa* program (see Behrman & Hoddinott 2005), then evaluation of treatment effects by comparing the treated and the not-treated would be relatively simple. However, the South African CSG was not implemented with an experimental design. As shown in the table, non-treated children can be grouped into three categories: those that received child support grant only after they were 3 years old (321 children, column 4); those who had applied for CSG support, but who had their applications rejected or had not yet received benefits by the time of the survey (154 children, column 5); and, those for whom CSG applications were never made (886 children, column 6).

The latter group of non-applicants is clearly suspect as an adequate control group given that the CSG is a means-tested program. As can be seen, household per-capita expenditures are 40 percent higher on average for this group compared to all other groupings in the table.

However, the other two groups of non-treated children appear more promising as a comparator group around which to build an analysis. Household expenditures are quite similar between these groups and treated children. Indeed, other studies (see for example Joshua D. Angrist 1998) have used rejected applicants as relevant control group.[14]

Looking at Table 2, the descriptive statistics hint at what such a comparison might reveal. Child z-scores are higher for treated children (-0.84) versus the beneficiary and applicant groups of non-treated children (-0.91 and -1.08, respectively). However, a closer look at the data reveal that fewer than 10 of the 154 non-beneficiary applicants are actually rejected applicants. The others are still in process and are perhaps better described as tardy applicants. As reported in Table 2, the average caregiver for this group of children delayed application by nearly 1450 days after her child became eligible. Note that this figure is nearly four-times higher than the average delay for beneficiary children. A similar observation applies to beneficiary children who did not receive CSG support during the first three years of their lives.

While the height-for-age z-scores for applicant and non-treated beneficiary children are lower than those for treated children, it is unclear whether this difference is the result of the CSG treatment received by the latter, or whether the long delay in application by the non-beneficiary applicant signals something about the caregivers of the former group (*e.g.*, their preferences, family organization or childrearing skills). This observation questions the adequacy of these individuals as a control group for the purposes of impact analysis. An alternative approach is to exploit the variation in the extent of treatment to identify the

[14]The key idea idea here is that the selection is based on observables. The information used to select candidates allows the researcher to identify the set of covariates that explain participation in a "program".

impact of CSG cash transfers.

3 Continuous treatment impact evaluation strategy

It is common in the program evaluation literature to measure the treatment as a binary variable (see for example Richard Blundell & Monica Costa-Dias 2002, Guido W. Imbens 2004, James J Heckman, Hidehiko Ichimura & Petra E Todd 1998).[15] This approach makes sense when there is a randomized design in which all treated individuals or villages receive the same treatment or dosage. However, the lack of a randomized design in the CSG means that selection into treatment is not random and that that the dosage received by the treated is not uniform. A conventional binary treatment approach would identically classify all treated beneficiaries, despite the fact that some children have received CSG support for nearly 100 percent of their life under 36 months, while others have received benefits for a small fraction of their lives.

As the descriptive statistics in Table 2 show, the extent of treatment received by children varies substantially among the treated group. Splitting the treated around the median treatment level, we see that the low treatment group has averaged 18 percent of their early life covered by CSG support, whereas the average treatment level is 61 percent for the high treatment group. One might expect that the cumulative impact of CSG support for the latter group should be larger, and indeed, the z-scores for this group are -0.75, as compared to -0.93 for the low treatment cohort. Treating these two groups as the same, as a binary approach would do, thus seems likely to understate the potential effect of full CSG treatment.[16]

While there are solid analytical reasons for exploiting the duration of treatment, there can be no presumption that variation in treatment level has been randomly generated. Returning

[15]Given the available panel data, another option would be to use anthropometric measures from earlier rounds of the KIDS data to underwrite a difference-in difference analysis. Unfortunately, there are only 87 beneficiary children for whom we have anthropometric data in both 1998 and 2004, and 323 non-beneficiary children. Making matters worse, children in this panel sub-sample are by definition older and very few of them (only 4) received any CSG support during the critical first three years of their lives.

[16]In fact when binary methods were used we were not able to identify a significant treatment effect of the CSG (full results are available upon request from the authors).

again to Table 2, we see that that while per-capita expenditures levels are quite similar for households with high versus low levels of treatment, the application delay between the groups is large. While not surprising, we must again worry whether the z-score difference between children in these households reflects differences in the extent of treatment *per se*, or differences in household attitudes and preferences.

3.1 CSG rollout and caregiver eagerness

There are at least two forces at work shaping the application delay for any child. The first is the characteristics of the child's caregiver and the family environment. The second is the effective rollout of the program in the child's community relative to the child's birth/CSG eligibility date.

While the CSG program was announced nationally in mid-1998, it took time for information on the program to filter down to all communities. The average delay in application was initially high, dropping off to a lower level as the program became better known (Anne Case, V. Hosegood & Frances Lund 2003). While the program was not purposefully rolled out differently across communities (unlike the Brazilian Bolsa Escola program discussed by Arends-Kuenning et al. (2005)), we might anticipate that the program uptake might have been quicker in urban and less-isolated rural areas.

These observations suggest that a better indicator of latent caregiver characteristics might be not the gross delay in CSG application, but deviation from the average delay for children in the same age and locational cohort. Figure 1 displays the results of such an analysis. The horizontal axis displays the date the child became eligible for the CSG grant (measured as number of days since the 1998 creation of the program). An eligibility date of zero means that the child was already born (and under age 7) when the program was announced. An eligibility date of 1500 means that the child was born 1500 days after the program was created. Projected onto the figure are the actual data points for treated children in the KIDS dataset.

The vertical axis in Figure 1 shows the delay in application. The non-parametric fit (displayed as a solid line) shows that as expected, average application delay dropped off sharply as eligibility date increased. For children immediately eligible for the program when it was announced, it took an average 1500 days (roughly 4 years) before caregivers applied for the grant. By the time of the survey in 2004, newly born children were on average being enrolled in the program in less than a year.

However, this very short application delay of the youngest children likely understates the eventual application delay as there were younger children yet to be enrolled at the date of the survey. To control for this problem, we furhter explore the application delay using only data on children born two years or more before the survey date (those data points to the left of the vertical line in Figure 1). This two year cut-off is consistent with the study by Case, Hosegood & Lund (2003) who show that by 2002, the CSG take-up rate for an age cohort leveled off at 40 percent within two years of birth (recall that the CSG program was originally designed to service 30 percent of the population). Cohort enrollment rates are somewhat higher in the KIDS data. By the survey date in the second quarter of 2004, roughly 60 percent of the cohorts born in 1998, 1999, 2000, and 2001 had applied for the CSG. The take-up rate for the 2002 cohort was slightly lower at 55 percent.

Letting t_i denote the application delay of child i, we write the application delay parametric regression function as:

$$t_i = u_i(\gamma_0^u + \gamma_1^u \ell_i + \gamma_2^u \ell_i^2) + (1 - u_i)(\gamma_0^r + \gamma_1^r \ell_i + \gamma_2^r \ell_i^2) + \nu_i$$

where u_i is a binary indicator variable if the child is resident in an urban area, ℓ_i measures the date when the child became eligible for the CSG, and ν_i is the error term which we assume has the usual properties. The estimated parameters of this regression function indicate that the application delay levels off at about 255 days in both urban and rural areas. This long-run expected application delay level is estimated to have been reached 1.5 years earlier in urban areas than in rural areas. Figure 1 illustrates the expected application delay for children in urban areas (the dashed line), as well as the horizontal long-run expected delay line. The expected delay in rural areas is nearly parallel to the urban line, but shifted to the northeast.

Using the estimated coefficients from the application delay regression, we calculated the expected delay for each child in the sample[17]. Denoting this conditional expectation as $\hat{t}(\ell_i, u_i)$, we defined caregiver eagerness as:

$$e_i = \frac{\hat{t}(\ell_i, u_i) - t_i}{\hat{t}(\ell_i, u_i)}$$

Positive eagerness values mean extra-eager, early applicant caregivers, while negative values indicate less-eager or tardy applicant caregivers.

Table 2 reports information on these standardized eagerness measures. As can be seen, non-treated children come from less eager families than do treated children. The average treated child has an eagerness score of 29 percent (i.e., they applied 29 percent earlier than the average of their cohort). CSG beneficiaries who did not receive any CSG treatment during the first three years of life applied with the near normal delay on average (-3 percent eagerness), whereas the applicant group that had not yet received any CSG payments by the survey date are shown to be tardy, with an average eagerness score of -40 percent .

These eagerness figures thus suggest that the non-treated, potential control groups are different than the beneficiary group. It also suggests that the eagerness measure can be used control for latent caregiver and family characteristics.[18] Using this idea, we will devise an identification strategy built around the notion that conditional on eagerness, the extent of CSG treatment should be random (related only to the accidents of birth time and location) and hence orthogonal to the expected effect of the treatment.

[17]Children born after the date when eagerness fell to its minimum value were assigned the asymptotic eagerness value as their expected application delay.

[18]In a simple regression analysis of z-scores, the coefficient of the eagerness is positive and highly significant even in the presence of other covariates, including CSG treatment.

3.2 Testing for cohort effects

Before turning to the impact analysis of the CSG, this section explores the reliability of the proposed identification strategy. Conditional on eagerness (as well as other family and child characteristics), the use of program roll-out to generate variation in treatment implies that children born in later cohorts will be more likely to have more extensive treatment. Children in the lowest tercile of the treament distribution were, in 2004, 2.5 years older on average than children in the highest treatment tercile. While this is a rather modest difference (i.e., the cohorts are not very far apart in age), there could still be concerns that the later birth cohort may be systematically better off irrespective of the CSG. For example, an overall improvement in the living standards of South African households, for reasons unrelated to the roll-out of the CSG, could lead us to an overestimation of the role of the CSG. The later cohorts might also have enjoyed better access to clinics and other health facilities, creating a spurious correlation between treatment and other nutrition-promoting interventions. In this section we address this issue by both testing for trends in the z-scores of CSG ineligible children and looking at data on trends in the quality of heath care.

To test for these cohort effects, Figure 2 graphs the estimated trend between z-score and child age at the time of the KIDS 2004 survey for two sets of children whose reported family per capita expenditures make it unlikely that they meet the means test for the CSG. The solid line plots the OLS relationship between z-score and age for all ineligible children. The dashed line illustrates the same relationship for the poorest 40 percent of CSG-ineligible children. This latter group, whose living standards hover just above the CSG eligibility level is arguably a better indicator for trends that may have affected CSG-eligible children.

As can be appreciated visually in Figure 2, there is a very slight downward trend (older children have lower for both subsets of children who are ineligible for the CSG. Table 3 displays the estimated regression coefficients used to generate Figure 2. In column (1) we consider all non-eligible children and show that the apparent negative relationship is not statistically significant. Adding a quadratic term for the child's age, as in column (2), does not alter our conclusion.

Columns (3) and (4) of the table show the results when we focus this exercise on the bottom 40 percent of ineligible children. The data points projected onto Figure 2 are those for this poorer subset of CSG-ineligible children. Once again, the analysis does not support the idea that there was a significant improvement in the health status of children from later cohorts in the non-eligible sample. While the estimated slope is downward sloping, its magnitude is so slight that even if it were significant, it would imply only a 0.02 percentage increase in z-score between children in the upper tercile of the treatment distribution as compared to those in the lowest tercile of the treatment distribution. As the next section shows, the estimated CSG treatment effect is an order of magnitude larger than this amount. This positive estimated effect of the CSG on HAZ is thus unlikely to be driven by confounding factors affecting recent cohorts of South African children.

The KIDS data provide a second window to explore the possibility of a cohort effect that improves child health and height independently of the CSG. The KIDS 2004 survey included a community questionnaire in which local leaders were asked their perceptions about the quality of local services in both 1999 and 2004. Table 4, organized as a transition matrix, with 1999 perceptions defining the rows and 2004 perceptions the columns, shows a preponderance of stasis (the bolded main diagonal of the matrix) or even deterioration in the perceived quality of local health services. While there are a few communities where services may have improved (the upper triangle of the transition matrix), these data suggest if anything, that later cohorts of children may have had access to poorer quality health services. Given that these communities have been hit hard by increasing numbers HIV/AIDS-related illnesses and deaths, this apparent deterioration in health services may be the result of a health care system that has been overwhelmed by these new demands.

The community questionnaire also asked respondents to identify the main improvements in their communities since 1999. Corroborating the evidence that health services have not been generally improving over the period of the CSG grant, only 6 percent of communities identified hospitals and clinics as the most improved service since 1999, and another 4.7 percent and 6.4 percent listed them as the second and third most improved service. These numbers are much lower than for primary school (27.3 percent , 21.5 percent and 4.8 percent , respectively) and for the supply of electricity (12.1 percent , 20.0 percent and 12.7 percent , respectively). While these issues deserve further attention, at a descriptive level at least the evidence suggests if anything that complementary health care services have deteriorated for children in the later birth cohorts who have received higher levels of CSG support.

As mentioned before, South Africa has an extensive welfare system. Thus, it is possible that the possible effects associated to the CSG could confound a joint effect of the SG plus other programs, in particular, the widely studied Old Age Pension (see Anne Case & Angus Deaton 1998, Anne Case 2004, Esther Duflo 2003). In table 5 we show that this is not a concern. We report the proportion of pension recipients in the household per decile of CSG exposure. There is no systematic (and statistically significant) trend between low and high levels of exposure to the CSG and the pension grantees. The lack of a relationship is expected, in part, by to the extensive coverage of pension grant as described by Duflo (2003). Thus, children with low and high exposures to the CSG do not differ in the proportion of adults that receive pensions and this is unlikely to drive our results.

4 Continuous treatment estimates of the impact of the CSG on nutrition

In this section we evaluate the impact of the CSG when the treatment is defined as a continuous variable using an estimator proposed by by Hirano & Imbens (2004).[19] As discussed above, we achieve identification of program effects by exploiting variation of treatment amongst the subset of treated children. As with a binary treatment analysis, the key

[19] Alternative approaches for non-binary treatment include Guido Imbens (2000) for categorical treatments. See also Jere R. Behrman, Yingmei Cheng & Petra E. Todd (2004).

identifying assumption is that conditional on observables (including eagerness), variation in treatment status is the result of random factors related to child age and program roll-out.

4.1 Identification strategy

The intuition behind the Hirano and Imbens estimator is most easily explained with the empirical example used by these authors. In their study, Hirano & Imbens (2004) use their continuous treatment estimator to evaluate the impact of lottery winnings on labor supply of the "treated" population of lottery winners. The treatment dosage (size of lottery winnings) is clearly randomly distributed amongst lottery winners, satisfying a general unconfoundedness condition that treatment dosage is orthogonal to the outcomes of interest (*e.g.*, leisure-seeking individuals were no more likely to receive large winnings than were those who wanted lottery winnings to start a new business and work more hours). When this unconfoundness assumption is fulfilled, identification of treatment effects should be relatively straightforward and credible. However, in the empirical analysis of Hirano & Imbens (2004), the random distribution of treatment is disrupted by non-random survey response problems (their completed sample is biased towards winners of smaller lottery prizes). Their key contribution is to show how to integrate other covariates into the analysis when the observed treatment is not purely randomly distributed.

Formally, consider a random sample of individuals indexed by i where $i = 1, \ldots, N$. Let $d \in \mathcal{D}$ denote the dosage (in our case, the extent of CSG treatment during the child's first 36 months of life). For each i there is a set of potential outcomes, $Y_i(d)$, which capture i's response to a *dose*. In our case, $Y_i(d)$ is the treated child's HAZ score. When the dose is binary then $\mathcal{D} = \{0, 1\}$, but for the purpose of our paper we consider the continuous treatment case where $\mathcal{D}$ lies in the interval $[d_0, d_1]$. For each unit i we observe a set of covariates X_i, the level of the treatment received, $D_i \in [d_0, d_1]$, and the corresponding outcome $Y_i = Y_i(D_i)$.[20]

[20]With continuous treatment several restrictions apply to the probability space. See Hirano & Imbens

Let the average dose-response function at d be described by $\mu(d) = E[Y_i(d)]$. We are interested in estimating the average gain on height-for-age from receiving the CSG for a fraction d of the window compared to a smaller reference dose, $\tilde{d}$. Thus, our impact measure of interest is:

$$\theta(d) = \mu(d) - \mu(\tilde{d}) = E[Y_i(d)] - E[Y_i(\tilde{d})] \qquad \tilde{d}, d \in \mathcal{D}. \tag{1}$$

To keep the notation simple, we will drop the i subscript until the estimation section. The key assumption of the method suggested by Hirano & Imbens (2004) is a generalization of the unconfoundedness assumption found in the binary treatment literature. The central idea is that after adjusting for differences in a set covariates X all biases in the comparison and treatment groups are removed. The authors capture this assumption as follows (p. 74):

Assumption 1 (Gen. Weak Unconfoundedness) $Y(d) \perp D | X \; \forall d \in \mathcal{D}$.

In other words, conditioning on the covariates the extent of treatment is random, unconfounded with any unobserved factors that might effect the extent of treatment (and outcomes)

While this unconfoundedness assumption is obviously met in the case of the lottery winnings studied by Hirano & Imbens (2004), our key identification assumption here is that conditional on eagerness (and other observable characteristics), the extent of early life CSG treatment or dosage is random, depending only on the child birthdate and local program roll-out. Unless childbirth decisions were postponed by more ardent caregivers (in anticipation of the CSG), this assumption should be met. Such postponement is exceedingly unlikely, however, as the program was announced in mid-1998 and in principal available to all eligible children at that date. Put differently, anyone with immediate knowledge of the program would have been positioned to benefit from it immediately and would have had no incentive to postpone childbearing.

(2004) for details.

Following Hirano & Imbens, the next step is to define the "generalized propensity score" or GPS.

Definition 1 (Generalized Propensity Score) *Let $r(d, x)$ be the conditional density of the treatment given the covariates:*

$$r(d, x) = f_{D|X}(d, x)$$

Then the generalized propensity score is $R = r(D, X)$.

Like in the standard (binary) propensity score, the GPS has the property that within strata with the same value of the conditional density $r(d, X)$, the probability that $D = d$ does not depend on the value of the covariates X. The authors use the GPS to show that, when using the weak unconfoundedness assumption, the assignment to treatment is unconfounded given the GPS. That is, for every d, we have $f_D(d|r(d, x), Y(d)) = f_D(d|r(d, X))$. [21]

The estimation of $\theta(d)$ requires computing two functions. First, let $\beta(d, r) = E[Y|D = d, R = r]$ be the conditional expectation of the outcome as a function of the treatment level D and the score R (note that both variables are scalars.) The dose response function at a particular level of the treatment is the average of the conditional expectation over the GPS at the particular level of treatment. This is given by $\mu(d) = E[\beta(d, r(d, X))] = E[Y(d)]$. Once $\mu(d)$ is computed, we can obtain our estimate for $\theta(d)$ as defined above.[22]

4.2 Impact estimation

To estimate $\theta(d)$ for all $d \in \mathcal{D}$ we use a two-stage approach as follows. In the first stage we assume a normal distribution for the treatment given the covariates:

$$D_i|X_i \sim N(\psi' X_i, \sigma^2)$$

[21] See theorem 1 in Hirano & Imbens (2004).

[22] As the authors note, the averaging is over the score evaluated at the treatment level of interest, $r(d, X)$ and not over the GPS $R = r(D, X)$. Theorem 2 in Hirano & Imbens (2004) formally demonstrates the relation between $\beta(d, r)$ and $\mu(d)$. It also shows that $\beta(d, r) = E[Y(d)|r(d, X) = r]$

where the parameters ψ and σ^2 are estimated by maximum likelihood. This allows us to estimate the GPS as

$$\hat{R}_i = \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} \exp\big(- \frac{1}{2\hat{\sigma}^2}(D_i - \hat{\psi}' X_i)^2\big)$$

In the second stage we use a flexible function for $\beta(d, r)$. As suggested by Hirano & Imbens (2004) we use a quadratic approximation

$$\beta(D_i, R_i) = E[Y_i|D_i, R_i] = \alpha_0 + \alpha_1 D_i + \alpha_2 D_i^2 + \alpha_3 R_i + \alpha_4 R_i^2 + \alpha_5 D_i R_i$$

The set of parameters $\alpha = (\alpha_0, \ldots, \alpha_5)$ can be estimated using ordinary least squares. Given the estimated parameters we can compute the average potential outcome at the treatment level d

$$\hat{\mu}(d) = E[\hat{Y}(d)] = \frac{1}{N}\sum_{i=1}^{N}(\hat{\alpha}_0 + \hat{\alpha}_1 \cdot d_i + \hat{\alpha}_2 \cdot d_i^2 + \hat{\alpha}_3 \cdot \hat{r}(d, X_i) + \hat{\alpha}_4 \cdot \hat{r}(d, X_i)^2 + \hat{\alpha}_5 \cdot d_i \cdot \hat{r}(d, X_i))$$

We can compute $\hat{\mu}(d)$ for all levels of d. To estimate $\hat{\theta}(d)$, the gains in Y from receiving a dose of d compared to a dose of $\tilde{d}$, we proceed as follows

$$\hat{\theta}(d) = \hat{\mu}(d) - \hat{\mu}(\tilde{d}) \qquad \forall d \in \mathcal{D}$$

The authors suggest computing the confidence intervals for the estimates using bootstrap methods.

4.3 Econometric results

As described in the prior section, our strategy to satisfy the weak unconfoundedness orthogonality assumption is to include in the set of covariates a reliable indicator (eagerness) of the family preferences and values that usually remains unobserved and confound casual analysis. Conditional on these variables, including the "eagerness" to apply for the CSG, should remove most any bias that results from estimating impacts by comparing across different treatment levels[23].

Table 6 shows the estimates of the first and second stages of our core impact evaluation model. The first-stage estimates indicate the importance the eagerness variable has on explaining the treatment dose. Other covariates include the age, education, sex, marital status and employment status of the child's caregiver. The model also includes interaction terms between eagerness and caregiver sex and marital status.[24] While most of these variables have intuitive signs (children of younger, better educated, female caregivers are more likely to have larger CSG coverage), the individual coefficients are not significant at conventional levels.

4.3.1 Balance of the Core Model

Table 7 allows us to explore whether or not the GPS balances the characteristics of the observations in the different ranges of the treatment variable. The analysis here builds on the suggestions of Hirano and Imbens who propose a natural extension of the concept of balance as developed by Rosenbaum and Rubins (1993) for the binary treatment case. The first three columns of Table 7 test whether the mean value of the covariates are the same for the observations in the three different treatment terciles. For example, the first cell of the table tests whether the mean eagerness of those who receive low treatment amounts (0-39 percent of their window) is different from those that did not receive low treatment. As can be seen, low treatment children come from families that are 51 percentage points less eager than the families of children who received higher treatment levels. Overall, the raw data is fairly well balanced along most dimensions except eagerness and the marital status and sex

[23] Standard errors were robust to unknown forms of heteroskedasticity and clustered at the village level.

[24] As can be seen in Table 6 below, the raw data are unbalanced in terms of caregiver sex and marital status. Compared to children with high and low levels of treatment, children receiving intermediate levels of treatment are more likely to have male cargiver and less likely to have a married caregiver. In an effort to capture these interactions, the binary variables for caregiver sex and marital status were interacted with eagerness with the expectation that the interaction terms would carry a negative coefficient.

of the caregiver. Of the twenty-one possible comparisons between treatment groups, five are statistically different from zero, indicating no random allocation of treatment.[25]

The second set of three columns in Table 7 explore whether the covariates are better balanced once we condition on the estimated GPS. Again we follow the suggestion of Hirano and Imbens on how to test balance in the continuous treatment case. For each treatment tercile, we first calculate the estimated probability that each observation might have received the median treatment level for the tercile. In other words, letting d_t denote the median treatment level received in tercile t, we calculate $r(d_t, X_i)$ for each observation. For example, for the lowest treatment tercile, we calculate $r(23\%, X_i)$, where 23% is the median treatment level for those who actually received low levels of treatment.

Continuing with the example of the lowest treatment tercile, we then separate the observations up into the five quintiles defined by $r(23\%, X_i)$. For each of these GPS quintile blocks, we then test whether the means of the covariates for the observations that actually received low treatment is different from the means for those that did not receive low treatment. Note that if GPS successfully balanced the covariates, we should expect low and not-low treatment groups to look similar once we block or condition on the GPS.

As can be seen in the second group of columns in Table 7, the GPS score improves the balance of the data as now only three of the twenty-one comparisons are statistically significant. Differences in eagerness values between treatment groups are now less than a third of their magnitude in the raw data, and only one of the differences remains statistically significant (in the balanced data, high treatment children come from more eager families than do the children in the lower treatment terciles). Also none of the other covariates exhibit significantly different means across treatment groups.[26]

[25]The data are not balanced across treatment terciles in terms of child age in 2004 as would be expected given the rollout pattern of the CSG.

[26]The balanced data exhibit greater differences in age amongst the treatment cohorts than does the raw data. This result is expected as we achieve balance along the dimension of eagerness by effectively comparing equally eager parents whose children were born in different age cohorts. That is, we are using the rollout process to create the random variation in treatment that we need to identify the effect of a program that is in part also driven by self-selection based on parental eagerness. As discussed in the prior section, there is no evidence of an independent cohort effect on child health that would invalidate this assumption.

4.3.2 Estimated Continuous Treatment Effects

Given that our GPS procedure systemaically improves the balance of eagerness and other characteristics across treatment groups, we turn now to examine the quantitative implications of our estimates. As Hirano & Imbens (2004) argue, the parameters of the second-stage estimate in Table 6 do not have a direct meaning, so whether the treatment has a statistically significant impact on the outcome cannot be inferred directly from those parameters.

To evaluate CSG treatment effects, we estimate $\theta(d)$ over the $d = [0.01, 1.00]$ interval depicted in figure 3. Here the baseline against which we compare all treatment levels is set at reference level $\tilde{d} = 0.01$. We thus compare the gains on HAZ from receiving the CSG for different proportions of the window of nutritional opportunity against receiving a "small" dose.

Figure 3 shows on the right axis a non-parametric estimate of the distribution of CSG treatment in the KIDS data (shown as the red or dotted line). The peak of the distribution is close to receiving treatment for three-quarters of the windows and decreases quickly after that. The solid line (measured on the left axis) is our estimates for $\theta(d)$ for different values of d measured in the horizontal axis.

These points estimates show that for treatments covering less than 20 percent of the window we find no gains. The gains are at a maximum when the treatment covers around three-fourths of the nutritional window. A child receiving treatment for two-thirds of the windows, on average, has 0.20 more HAZ than a child with a treatment covered for only 1 percent of the window. This gain is statistically significant as the 90 percent confidence interval estimate excludes zero. The portion of the impact curve for for which we can reject the hypothesis of zero impact are demarcated with rectangles. The decline of the gains on HAZ after a dosage of 80 percent cannot be interpreted as an indicator that treatment is less effective after that level, since this decline coincides with a growth of the width of the interval estimator provoked by a small number of dosage levels beyond that point (as shown by the treatment distribution function).[27].

4.4 Robustness check using only temporal rollout variation to identify CSG impact

Identification of our core continuous treatment model depends on the assumptions that neither birthdate nor birth location directly influence a child's HAZ. While there is ample evidence that there is no independent birthdate or cohort effect, we might worry that there is an independent spatial pattern (*e.g.*, urban communities with earlier rollout may independently have higher z-scores). While it is a bit hard to know what to expect in this regard given spatial structure of the South Africa, we check the robustness of our results by re-estimating our model after effectively including a set of cluster (village) fixed effects in the first stage GPS regression. Given that the KIDS data come from 72 clusters, this is clearly a statistically expensive and conservative approach.

The results from this robustness check are shown in Table 6 and the estimated treatment effect is shown as the dashed line in Figure 3. As can be seen, the impact of exposure to the CSG from these estimates parallel the core model estimates. However, the estimated impacts are statistically significant at the 10 percent level over a slightly smaller treatment range (the 90 percent interval excludes zero for dosages ranging from 54 percent to 68 percent).

5 From cash transfer flows to human capital stocks

The nutrition of young children is of importance not only because of concern over their immediate welfare, but also because nutrition in this formative stage of life is widely perceived to have substantial, persistent impact on their physical and mental development. This in

[27]Zhong Zhao (2008) shows that in the case of binary propensity score matching, a misspecified first stage does not lead to important biases as long as the unfoundedness assumption is maintained. The paper does not explore the case of continuous treatment, however we could expect the conclusions to hold for this case as well. Our paper argues that the set of variables including eagerness are key to satisfy this (untestable) assumption.

turn affects their school success and later labor market productivity. Futoshi Yamauchi (2008), using also the data from KIDS, finds that siblings with improved height-for-age z-scores significantly start school at an earlier age, have higher grade attainment, lower grade repetition, and better learning performance in the early stage of schooling. Improving the nutritional status of malnourished infants and small children may, therefore, have important payoffs over the long term (e.g., Behrman & Hoddinott 2005, John Maluccio, John Hoddinott, Jere R. Behrman, Reynaldo Martorell & Agnes R. Quisumbing 2009). In this section we try to quantify the gains in height-for-age (z-scores) found in the previous section in terms of adult wages. We then use these monetary figures to calculate private to CSG payments. Our methodology is closely related to the work by Behrman, Sengupta & Todd (2005) where they calculate the lifetime earnings for children with and without participation in Mexico's Progresa using their estimated impact of the program on child schooling and the returns to education from the existing literature.

Consider the case of a male child who is treated before age 1 and receives CSG benefits for two-thirds of the first three years of life. The gains in the child's z-score would be estimated to be around 0.20.[28] To compute the gains in height for an adult we assume that as an adult, this child will have a z-score 0.2 higher than the current average male between 25-35 years of age. This assumption is consistent with the evidence that early childhood height losses and gains are irreversible, leading to a permanent change in the child's position in the height distribution.[29]

Using data from the 1998 wave of KIDS we obtained the z-score of adult males age 25-35, as well as their average monthly wage earnings. The average z-score for this adult group is -0.68, and a z-score gain of 0.2 translates into a 1.8 cm., or 1.1 percent , gain in adult height.[30]

[28]Our estimates transforming z-scores to children's height do not vary by gender.

[29]This assumption is consistent with the conventional pediatric wisdom that a child's height (z-score) at age 2 is an excellent predictor of adult height (z-score.)

[30]Recall that z-scores (z) are defined as $z = \frac{h-\bar{h}}{\sigma_h}$ where h is height, $\bar{h}$ and σ_h are, respectively, the mean and standard deviation of height given the age. Changes in height are computed by $\Delta h = \sigma_h \Delta z$.

A number of authors have examined the impact of adult height on wages. In the context of lower income countries, adult height is seen to be an indicator of a broad array of human capabilities, including health, cognitive development and work capacity.[31] The study by Duncan Thomas & John Strauss (1997) takes a particularly careful look at the relation between wages and health in Brazil. Controlling for achieved levels of eduction, they report for urban males an elasticity of wages with respect to height of 2.43 to 3.36. Using these elasticities, the gains in monthly South African wages from an increment in height of 1.1 percent would be between R67 to R92. Note that these calculations ignore any general equilibrium effects that would occur from having a better nourished adult population.[32] While it would be good to have similar elasticity estimates specifically for South Africa, we use the Thomas and Strauss estimate as a way to arrive at a rough evaluation of the benefits of the increased height generated by the CSG.

To compute the returns to the CSG payments, we calculate the present discounted value of a monthly flow of R62 (and R92) from age 25 to 65. Note that this calculation assumes that the individual is fully employed throughout this time period. At an annual real discount rate of 5 percent it yields a discounted present value of R3,896 (R5,380) at birth. Given the cost of 20 months of the CSG (20 $\times$ R170) and assuming administrative costs of 6.6 percent [33], our calculations show a Benefit-Cost ratio between 1.5 and 2.1. Note that a benefit-cost ratio of 1.0 would imply a real rate of return of 5 percent . A ratio in excess in 1.0 implies a real return in excess of this rate. If we more adjust our calculation and more conservatively assume individuals are unemployed 33 percent of the time (with unemployment spells randomly distributed across the life cycle), then the benefit figures are cut by a third, and the estimated benefit-cost ratios fall to between 1.0 and 1.39.

[31] In contrast, work on developed country (such as Andrew Postelwaite, N. Persico & D. Silverman (2004)) interpret greater adult as an indicator of higher status or other social processes that boost adult earnings.

[32] In principal, these general equilibrium effects could be negative (if returns diminish to increasingly plentiful human capital), or positive (if returns to human capital increase once the labor market becomes sufficiently dense in better nourished, and better educated workers).

[33] Skoufias (2005) found that the administrative costs of Progresa were about 8.9 percent and reduced to 6.6 percent without the conditionality

While these numbers need to be treated with extreme caution, and are at best only indicative of the order of magnitude of the long-term gains that might be anticipated from the CSG, there are several reasons why these impressive returns to the CSG may be conservative. First, this simple analysis ignores the impact that z-scores can have on educational attainment and progress. Second, we did not include the potential gains from receiving the grant after the window of opportunity (from age three to fourteen).

In addition to understanding the height wage elasticity in South Africa itself, future efforts to evaluate the impact of cash transfer flows on the future value of human capital might look more closely at how wages and earnings evolve over the life-cycle. The analysis here has simply assumed that wage gains are once and for all and persist over the life cycle. In addition, the simple calculations here have not considered the horrific drop in life expectancy that HIV/AIDS has brought to South Africa. The effects of this pandemic on the labor market has yet to be fully understood.

6 Conclusions

Cash transfer programs have taken on an increasingly important role in the anti-poverty programs of middle income countries. While a number of these programs have been modeled on Mexico's *Progresa* program (*e.g.*, those in Brazil, Honduras and Nicaragua), South Africa's Child Support Grant (CSG) has followed its own logic. Implemented at the same time as *Progresa* in 1998, the CSG targets child support payments to children's caregivers (almost exclusively women). Unlike *Progresa*, receipt of the CSG is not conditional on particular child behavior (school attendance and regular medical check-ups). Indeed as originally implemented, the CSG was limited to children under seven years of age.

In this context, this paper has shown that these targeted, unconditional CSG payments have bolstered early childhood nutrition as signalled by child height-for-age. While it is of course possible that conditioning CSG payments on, say, medical check-ups would have further increased program effects, we do find robust effects even in the absence of such conditioning. In contrast to the literature on the elasticity of nutrition with respect to income, income and nutrition appear to be more tightly wedded in the case of CSG payments, perhaps because the income increases are assigned to women.

While income transfers such as those of the CSG or the *Progresa* program in Mexico should help immediately redress contemporaneous poverty, the deeper question is whether they help facilitate a longer-term (inter-generational) pathway from poverty. One way that they might contribute to this goal is by enhancing the durable human capital stock of the next generation. Augmenting our estimates of the nutritional effects of the CSG with best estimates from the literature on the elasticity of wages with respect to adult height, we calculate that the discounted present value of increased future earnings are perhaps 39 percent higher than the cost of early-life CSG support. While these estimates are crude, first attempts, they do point us toward the sort of longer term analysis needed to determine when short term cash transfers translate into the long run asset increases needed to sustainably reduce poverty in the future.

References

Angrist, Joshua D. 1998. "Estimating the Labor Market Impact of Voluntary Military Service Using Social Security Data on Military Applicants." *Econometrica*, 66(2): 249–288.

Araujo, Maria Caridad, and Norbert Schady. 2006. "Cash Transfers, Conditions, School Enrollment, and Child Work: Evidence from a Randomized Experiment in Ecuador." The World Bank Policy Research Working Paper 3930.

Arends-Kuenning, Mary, Ana Fava, Ana Lucia Kassouf, and Alexandre de Almeida. 2005. "The Impact of School Quality and School Incentive Programs on Childrens Schooling and Work in Brazil." University of Illinois Working Paper.

Baird, Sarah, Craig Mcintosh, and Berk Ozler. 2010. "Cash or condition ? evidence from a randomized cash transfer program." The World Bank Policy Research Working Paper Series 5259.

Behrman, Jere, Mark Rosenzweig, and Andrew D. Foster. 1994. "The Dynamics of Agricultural Production and the Calorie-Income Relationship." University of Pennsylvania Mimeo.

Behrman, Jere R, and Anil B Deolalikar. 1987. "Will Developing Country Nutrition Improve with Income? A Case Study for Rural South India." *Journal of Political Economy*, 95(3): 492–507.

Behrman, Jere R, and Anil B Deolalikar. 1988. "Health and Nutrition." Vol. I, 631–711. Amsterdam:North-Holland.

Behrman, Jere R., and John Hoddinott. 2005. "Programme Evaluation with Unobserved Heterogeneity and Selective Implementation: The Mexican "PROGRESA" Impact on Child Nutrition." *Oxford Bulletin of Economics and Statistics*, 67(4): 547–569.

Behrman, Jere R, Piyali Sengupta, and Petra Todd. 2005. "Progressing through PROGRESA: An Impact Assessment of a School Subsidy Experiment in Rural Mexico." *Economic Development and Cultural Change*, 54(1): 237–75.

Behrman, Jere R., Susan W. Parker, and Petra E. Todd. 2009. "Schooling Impacts of Conditional Cash Transfers on Young Children: Evidence from Mexico." *Economic Development and Cultural Change*, 57(3): 439–477.

Behrman, Jere R., Yingmei Cheng, and Petra E. Todd. 2004. "Evaluating Preschool Programs When Length of Exposure to the Program Varies: A Nonparametric Approach." *The Review of Economics and Statistics*, 86(1): 108–132.

Blundell, Richard, and Monica Costa-Dias. 2002. "Alternative Approaches to Evaluation in Empirical Microeconomics." Institute of Fiscal Studies Cemmap Working Papers 10/02, London.

Bouis, Howarth E., and Lawrence J. Haddad. 1992. "Are Estimates of Calorie-income Elasticities Too High?: A Recalibration of the Plausible Range." *Journal of Development Economics*, 39(2): 333–364.

Case, Anne. 2004. "Does Money Protect Health Status? Evidence from South African Pensions." In *Perspectives on the Economics of Aging. NBER Chapters*, 287–312. National Bureau of Economic Research, Inc.

Case, Anne, and Angus Deaton. 1997. "Large Cash Transfers to the Elderly in South Africa." Princeton University Discussion Paper 176.

Case, Anne, and Angus Deaton. 1998. "Large Cash Transfers to the Elderly in South Africa." *Economic Journal*, 108(450): 1330–61.

Case, Anne, V. Hosegood, and Frances Lund. 2003. "The Reach of the South African Child Support Grant: Evidence from KwaZulu-Natal." CSDS Working Paper 38.

Chronic Poverty Research Centre. 2004. "Chronic Poverty Report 2004-2005." Chronic Poverty Research Centre.

Duflo, Esther. 2003. "Grandmothers and Granddaughters: Old-Age Pensions and Intra-household Allocation in South Africa." *World Bank Economic Review*, 17(1): 1–25.

Fizbein, Ariel, and Norbert Schady. 2009. *Conditional cash transfers : reducing present and future poverty.* Washington D.C.:The World Bank.

Gelbach, Jonah B., and Lant Pritchett. 2002. "Is More for the Poor Less for the Poor? The Politics of Means-Tested Targeting." *The B.E. Journal of Economic Analysis & Policy*, 0(1).

Glewwe, Paul, and Pedro Olinto. 2005. "Evaluating of the Impact of Conditional Cash Transfers of Schooling: An Experimental Analysis of Honduras PRAF Program." University of Minnesota Mimeo.

Haarman, Claudia, and Dirk Haarman. 1998. "A Contribution Towards a New Family Support System in South Africa." *Report for the Lund Committee on Child and Family Support.*

Haddad, Lawrence, Harold Alderman, Simon Appleton, Lina Song, and Yisehac Yohannes. 2003. "Reducing Child Malnutrition: How Far Does Income Growth Take Us?" *World Bank Economic Review*, 17(1): 107–131.

Hanlon, Joseph, Armando Barrientos, and David Hulme. 2010. *Just Give Money to the Poor: The Development Revolution from the Global South.* Kumarian Press.

Heckman, James J, Hidehiko Ichimura, and Petra E Todd. 1998. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *Review of Economic Studies*, 64(4): 605–54.

Hirano, Keisuke, and Guido W Imbens. 2004. "The Propensity Score with Continuous Treatments." 73–84. New York:Wiley.

Imbens, Guido. 2000. "The Role of the Propensity Score in Estimating Dose-Response Functions." *Biometrika*, 87(3): 706–710.

Imbens, Guido W. 2004. "Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review." *The Review of Economics and Statistics*, 86(1): 4–29.

Kakwani, N., F. Soares, and H. Son. 2005. "Conditional Cash Transfers in Africa." UNDP International Poverty Centre Working Paper 9.

Klasen, S. 1997. "Poverty, Inequality and Deprivation in South Africa: An Analysis of the 1993 SALDRU Survey." *Social Indicator Research*, 41: 51–94.

Kruger, John. 1998. "From Single Parents to Poor Children: Refocusing South Africa's Transfers to Poor Households with Children."

Leibbrandt, M., and I. Woolard. 1999. "A Comparison of Poverty in South Africa's Nine Provinces." *Development Southern Africa*, 16(1): 37–54.

Lund, F. 1993. "State Social Benefits in South Africa." *International Social Security Review*, 46(1): 5–23.

Lund, Frances. 2001. "Child Support Grant." University of Natal Mimeo.

Maluccio, John A., and Rafael Flores. 2005. "Impact Evaluation of a Conditional Cash Transfer Program. The Nicaraguan Red de Protección Social." IFPRI Research Report 141, Washington D.C.

Maluccio, John, John Hoddinott, Jere R. Behrman, Reynaldo Martorell, and Agnes R. Quisumbing. 2009. "The Impact of Nutrition during Early Childhood on Education among Guatemalan Adults." *Economic Journal*.

May, Julian, Jorge Agüero, Michael Carter, and Ian Timaeus. 2007. "The KwaZulu-Natal Income Dynamics Study (KIDS) 3rd Wave: Methodology, First Findings and an Agenda for Future Research." *Development Southern Africa*, 24(5): 629–648.

Paxson, Christina, and Norbert Schady. 2007. "Cognitive Development among Young Children in Ecuador: The Roles of Wealth, Health, and Parenting." *Journal Human Resources*, XLII(1): 49–84.

Penniston, Kristina L, and Sherry A Tanumihardjo. 2006. "The Acute and Chronic Toxic Effects of Vitamin A." *The American Journal of Clinical Nutrition*, 83: 191–201.

Postelwaite, Andrew, N. Persico, and D. Silverman. 2004. "The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height." *Journal of Political Economy*, 112: 1019–1053.

Skoufias, Emmanuel. 2005. "PROGRESA and its impacts on the welfare of rural households in Mexico:."

Strauss, John, and Duncan Thomas. 1998. "Health, Nutrition, and Economic Development." *Journal of Economic Literature*, 36(2): 766–817.

Strauss, John, and Duncan Thomas. 2008. "Health over the Life Course." Vol. 4 of *Handbook of Development Economics*, 3375–3474. Elsevier.

Subramanian, Shankar, and Angus Deaton. 1996. "The Demand for Food and Calories." *Journal of Political Economy*, 104(1): 133–62.

Thomas, Duncan, and John Strauss. 1997. "Health and Wages: Evidence on Men and Women in Urban Brazil." *Journal of Econometrics*, 77(1): 159–185.

UNICEF. 1998. *The State of the World's Children.* New York:Oxford University Press.

Van der Berg, Servaas. 1997. "South Africa After Apartheid: A Welfare State in the Making?" Montreal.

von Braun, Joachim, and Eileen T. Kennedy. 1994. *Agricultural Commercialization, Economic Development, and Nutrition.* Baltimore:The Johns Hopkins University Press.

World Bank. 2006. *Repositioning Nutrition as Central to Development: A Strategy for Large-Scale Action. Directions in Development*, The World Bank.

Yamauchi, Futoshi. 2008. "Early Childhood Nutrition, Schooling, and Sibling Inequality in a Dynamic Context: Evidence from South Africa." *Economic Development and Cultural Change*, 56: 657–682.

Zhao, Zhong. 2008. "Sensitivity of Propensity Score Methods to the Specifications." *Economics Letters*, 98(3): 309–319.

Figure 1: Standardized Eagerness

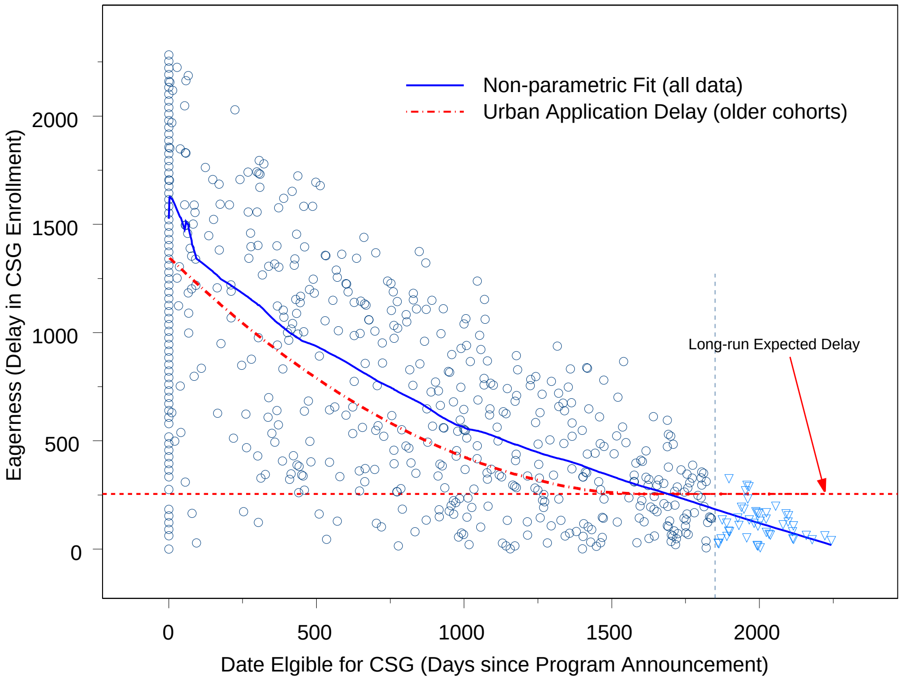

Figure 2: Trends in Height-for-Age: CSG ineligible children

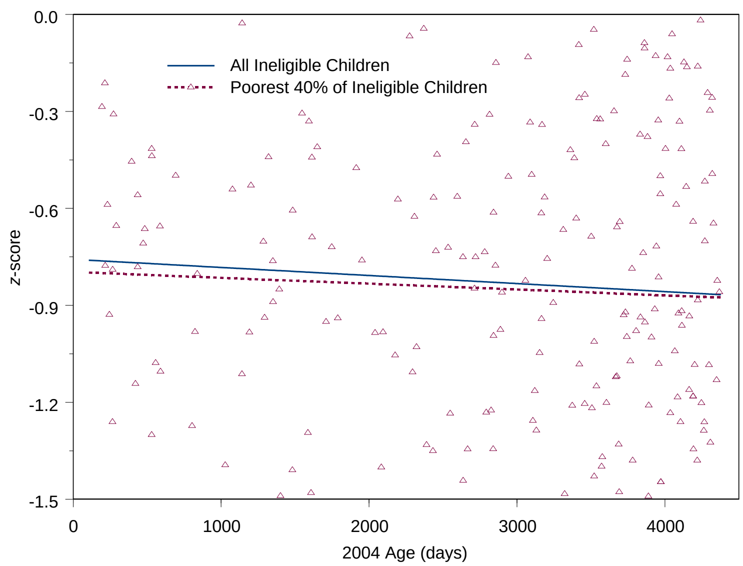

Figure 3: Gains in Height-for-Age from Child Support Grant

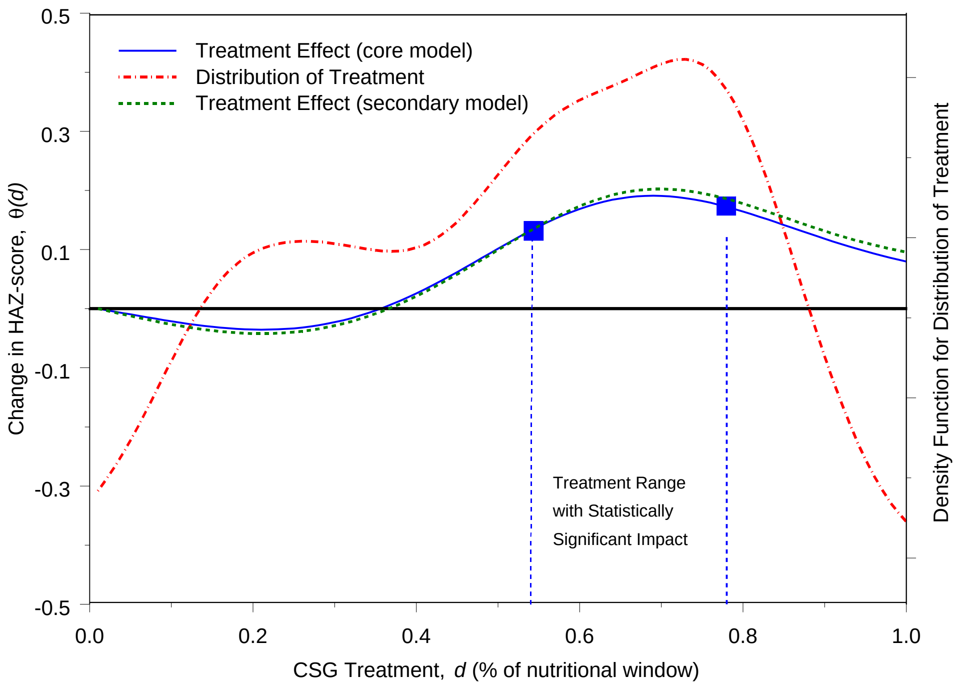

Table 1: Average differences in children's height measures, by year of birth and gender

Year of birth	Boys	Girls	All children
1997	.023	.008	.016
1998	.011	.031	.021
1999	.015	.019	.017
2000	.018	.013	.015
2001	.011	.013	.012
2002	.014	.011	.013
2003	.016	.011	.013
All years	.015	.016	.015

Note: Absolute differences are expressed in centimeters.
Source: Author's calculations using KIDS dataset.

Table 2: Descriptive Statistics by CSG Treatment Status

	Treated in 3 Year Window			**Not Treated**		
		Coverage				
	All	*Low*	*High*	*Beneficiaries*	*Applicants*	*Non-applicants*
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment						
Exposure (percent life)	51	34	69	24	0	0
Window (percent life under 3)	40	18	61	0	0	0
Child Characteristics						
HAZ 2004	-0.84	-0.93	-0.75	-0.91	-1.08	-0.83
Age in 2004	2.5	2.4	2.7	6.2	6.1	6.9
Caregiver Characteristics						
Per-capita Household Expenditure†	330	321	340	335	323	461
Application Delay (days)	329	456	194	1369	1448	-
Eagerness (percent deviation average delay)	29	0.2	60	-3	-39.9	-
Female (percent)	94.5	95.5	93.6	96.2	96.0	91.8
Age	38	39	37	41	48	46
Unemployed (percent)	50.1	48.8	52.5	42.7	25.3	29.8
Education (yrs.)	7.7	7.4	7.9	6.7	5.3	6.3
Number of Observations	245	123	122	321	154	886

†Rand per-month

Source: Author's calculations using KIDS dataset.

Table 3: Age effects in Height-for-Age for non-eligible children

Model:	All non-eligible			Bottom 40%
	(1)	(2)	(3)	(4)
Age in days (x100)	-0.025	0.187	-0.018	-0.014
	(0.030)	(0.139)	(0.045)	(0.199)
Age squared		-0.004		-0.000
		(0.003)		(0.004)
Observations	886	886	392	392

Note: Robust and clustered standard errors at the village level in parenthesis. The bottom 40 percent corresponds to children in the bottom 40 percent of ineligible households based on per capita expenditure.
Source: Author's calculations using KIDS.

Table 4: Distribution of the perceptions about hospitals and clinics services: 1999 and 2004

	2004				
1999	V. unhappy	Unhappy	Neutral	Happy	V. happy
Very unhappy (10 communities)	**40.0**	10.0	20.0	10.0	20.0
Unhappy (11 communities)	0.0	**54.6**	18.2	27.3	0.0
Neutral (6 communities)	33.3	33.3	**16.7**	16.7	0.0
Happy (23 communities)	34.8	43.5	4.4	**17.4**	0.0
Very happy (2 communities)	50.0	50.0	0.0	0.0	**0.0**

Source: Author's calculations using KIDS community questionnaire.

Table 5: Proportion of Old Age Pension recipients per deciles of exposure to CSG

CSG deciles	Old age pension recipients	
	Proportion	Std. errors
Lowest	0.098	0.016
2	0.078	0.013
3	0.076	0.013
4	0.091	0.015
5	0.067	0.012
6	0.069	0.012
7	0.104	0.017
8	0.106	0.018
9	0.074	0.013
Highest	0.078	0.013

Source: Author's calculations using KIDS.

Table 6: First and Second stages of GPS estimation

	Core Model		Secondary Model	
Variable	coefficient	s.e.	coefficient	s.e.
	First Stage: Maximum likelihood estimates			
Eagerness	0.003	0.0003	0.002	0.0004
Boy (=1)	-0.01	0.02	-0.01	0.026
Caregiver's age	-0.001	0.001	-0.004	0.001
Caregiver' educ.	-0.004	0.003	-0.006	0.004
Cargiver married (=1)	-0.03	0.03	-0.03	0.03
Married × Eager	-0.0005	0.001	-0.001	0.001
Caregiver's sex	-0.005	0.03	-0.003	0.03
Sex × Eager	-0.001	0.001	-0.0003	0.001
Caregiver works (=1)	0.002	0.02	-0.02	0.03
Village fixed effects	No		Yes	
	Second Stage: OLS estimates			
α_0	-0.84	0.40	-0.72	0.62
α_1	0.85	1.61	0.40	1.64
α_2	-177	164	-120	141
α_3	-0.58	0.63	-0.69	0.81
α_4	3.91	29	14.84	30
α_5	1.1	0.63	0.85	0.57

Note: Robust standard errors (s.e.) are clustered at village level.
Source: Author's calculations using KIDS dataset.

Table 7: T-tests for Equality of Means between Treatment Groups

	Raw Data Treatment Terciles			Data Adjusted by GPS Treatment Terciles		
	Low [0-39]	Medium [39-68]	High [58-100]	Low [0-39]	Medium [39-68]	High [58-100]
Eagerness	-50.87*	-10.93*	60.55*	-9.04	-3.52	26.42*
Boy	-0.05	0.00	0.05	-0.02	0.04	0.04
Caregiver's age	1.85	-1.00	-0.83	5.47*	-0.97	-2.09
Caregiver' educ.	-0.71	0.49	0.21	-0.64	0.83	-0.55
Cargiver married (=1)	0.07	-0.13*	0.06	0.11	-0.14	0.14
Caregiver's sex	0.03	-0.05*	0.02	0.04	-0.06*	0.04
Caregiver works (=1)	-0.08	0.06	0.02	-0.08	0.08	-0.01

Note: * denotes significance at 10 percent .
Source: Author's calculations using KIDS dataset.