Stunting, double orphanhood and unequal access to public services in democratic South Africa

GRACE BRIDGMAN
DIETER VON FINTEL

Stellenbosch Economic Working Papers: WP09/2021

www.ekon.sun.ac.za/wpapers/2021/wp092021

November 2021

KEYWORDS: Stunting, double orphanhood, spatial inequality, WASH infrastructure, service delivery, spatial econometrics, South Africa
JEL: I14, J13, H4, R1

DEPARTMENT OF ECONOMICS
UNIVERSITY OF STELLENBOSCH
SOUTH AFRICA

A WORKING PAPER OF THE DEPARTMENT OF ECONOMICS AND THE
BUREAU FOR ECONOMIC RESEARCH AT THE UNIVERSITY OF STELLENBOSCH

www.ekon.sun.ac.za/wpapers

Stunting, double orphanhood and unequal access to public services in democratic South Africa[1]

Grace Bridgman[2] and Dieter von Fintel[3]

ABSTRACT

Orphans who lack household or community support face significant socio-economic disadvantages. In particular, they are at greater risk of malnutrition and stunting in developing countries. Children who have no living parents, also called double orphans, are most likely to require support from extended families or public institutions. This paper explores how WASH infrastructure, and public health and social services relate to stunting. It is one of the first studies to analyse these factors with a specific focus on double orphans, who tend to live in under-serviced areas with high stunting rates and poor access to public resources. We collate a cross sectional spatial dataset with local child stunting rates from 2013, rates of double orphanhood, private household resources, and public services from 2011 for South Africa, a country where the HIV/AIDS pandemic has led to high rates of double orphanhood. We estimate spatial econometric models that account for unobserved regional shocks and measurement bias, but which do not address other biases. Our results show that high stunting rates, particularly in areas with high proportions of double orphans, overlap strongly with poor provision of WASH and the availability of household resources. By contrast, other softer services accessed outside the home, such as access to health, social welfare and early childhood development facilities are not correlated with stunting in the same way. WASH is more strongly related to reduced stunting when infrastructure covers larger geographic areas and with the combined use of services from adjacent areas. This occurs because of economies of scale in provision and preventing transmission of disease across regions. Policy makers can explore the option to reduce stunting by expanding geographic networks of WASH service delivery into under-serviced areas where double orphans tend to locate.

---

[1] Grace Bridgman gratefully acknowledges funding from the SA-TIED programme, a collaboration funded by the National Treasury of South Africa, the Department of Trade and Industry of South Africa, the Delegation of the European Union to South Africa, IFPRI and UNU-WIDER. The authors are grateful for helpful comments from Jukka Pirttilä, Servaas van der Berg, Murray Leibbrandt and seminar participants at Stellenbosch University. We are also grateful for data assistance from Dirk Brand, Cobus Burger, Jonathan Hulley, Kath Hall and Aislinn Delany. Errors remain our own.

[2] Department of Economics and Research on Socioeconomic Policy (ReSEP), Stellenbosch University, Private Bag X01, Matieland, 7602, South Africa gmmbridgman@gmail.com

[3] Department of Economics and Research on Socioeconomic Policy (ReSEP), Stellenbosch University, dieter2@sun.ac.za

1. Introduction

Orphans in developing countries face various socio-economic disadvantages relative to children who live with both parents (Beal *et al.*, 2018; Darteh, Acquah & Kumi-Kyereme, 2014; Dewey & Begum, 2011; Quisumbing & Maluccio, 2003). In particular, they are at greater risk of malnutrition and stunting (Finlay et al., 2016). Parental loss disrupts children's socio-economic and socio-emotional stability, and this affects their nutritional status and their subsequent development path negatively. Stunted growth in early life is a constraint to individual productivity and life expectancy and impedes cognitive development in children (May & Timæus, 2014; Zere & McIntyre, 2003).[4] Because orphans are more likely to be stunted in childhood in some countries (Singh & Saker, 2021), they are therefore also more likely to have lower socio-economic status in adulthood. Understanding how to reduce stunting among orphans, therefore, has the benefit of reducing differences in adult outcomes between orphans and non-orphans.

The disadvantage that orphans have is sometimes mitigated by adequate material and psychosocial support from other household members, a private rather than public resource (Ali *et al.*, 2018; Finlay *et al.*, 2016). Household wealth plays an important role in reducing stunting more generally. In developing countries, positive wealth shocks during childhood enable stunted boys to grow taller and partially catch up to their peers in adolescence (Duc, 2019). However, negative wealth shocks, particularly when children lose both parents, mean that orphans are exposed to resource poor environments that raise their risk for stunting. Many *single* orphans, children who have lost one parent, continue to receive care from their surviving single parent. However, unless extended families provide support to *double orphans* they are potentially at greater risk of stunting compared to single orphans – we call this hypothesised effect the "double orphan stunting premium". There is existing support for such

[4] Multiple studies link lower overall life outcomes with stunted growth in childhood, including increased morbidity, poor cognitive development, worse educational outcomes, increased risk of perinatal and neonatal death in stunted women, lower productivity and reduced lifetime earnings potential (Black *et al.*, 2013; Dewey & Begum, 2011; de Onis & Branca, 2016; Victora *et al.*, 2008).

a hypothesis in other domains. Some studies show that double orphans are more likely than single orphans to have depressive symptoms, to be forced to do chores and to be physically assaulted by others; by contrast, single orphans have outcomes similar to children whose parents are alive (Sherr, *et. al.*, 2014; Ruiz-Casares, *et. al.*, 2009). Little research has been done to establish whether these psychosocial disadvantages experienced by double orphans transfer to the material support and physical health domains.

Public services such as infrastructure and social services could play an important role in lowering stunting among this high-risk group. If double orphans are less likely to have private support, they are naturally more reliant on public services than other children. However, orphaned children tend to live in regions with poor public service delivery which amplifies the vulnerabilities caused by losing a parent (Case & Ardington, 2006). This study is one of the first to investigate whether access to *public* support services and infrastructure relates to stunting and double orphanhood in developing countries, and sheds light on approaches that could address the hypothesised "double orphan stunting premium".

We distinguish between access to "softer" public services such as health facilities, social welfare and early childhood development (ECD) and "hard" public infrastructure, such as access to water and sanitation (WASH). The relationship between stunting and "soft services" is under-studied, but is potentially important because these services directly target child and adult health and development. On the other hand, WASH has been extensively studied in relation to stunting. Access to WASH in the home is crucial to reducing child mortality and improving other health outcomes such as stunting (Headey & Palloni, 2019).[5] This is because children who live in areas with inadequate WASH are more likely to have chronic diarrhoea, which is a determinant of stunting (Badriyah & Syafiq, 2017; Dillingham & Guerrant, 2004; Fink *et al.*, 2011; Spears, 2013). While we know that increased access to sanitation plays a significant role in reducing stunting in all children (Fink *et al.*, 2011;

[5]Many other factors cause stunting in children, including chronic malnutrition, low maternal education, low birth weight, socioeconomic status, low maternal body mass index, and living in a district that is peripheral to an urban area (Beal *et al.*, 2018; Bernal *et al.,* 2014; Darteh, Acquah & Kumi-Kyereme, 2014; Delpeuch *et al.*, 2000; Zere & McIntyre, 2003).

Spears, 2013), there is no evidence whether this access could help reduce nutrition gaps between orphans and other children particularly.

We study these relationships for the case of South Africa, where the HIV/AIDS pandemic has increased the prevalence of orphanhood since 1990 (Case & Ardington, 2006:401). In particular, the HIV/AIDS pandemic catalysed a large increase in the number of *double* orphans. Their share more than doubled, from 2% of all children in 2002 to 4.6% in 2011 (the time period of our analysis). In contrast, changes in the rates of single orphanhood were negligible (Hall, 2019). The trend reversed a number of years after publicly funded anti-retroviral treatment was rolled out for the first time in 2004. However, we study a period when the rate of double orphanhood and the impact of the HIV/AIDS pandemic on children was near its peak. This was also a period when public sector support for orphans came under severe pressure. Relatives usually care for orphans in South Africa. Publicly funded cash transfers (Foster Care Grants) were historically targeted towards children who were placed in foster care, but were extended to relatives of orphans in 2002 (Hall & Proudlock, 2011). Under these circumstances orphans – including double orphans – should not have significantly different material circumstances compared to other children who live in households that qualify for other cash transfers (Child Support Grants). However, during the first decade of the 2000s, the rising demand for Foster Care Grants crippled the system, and many eligible recipients – including the rising number of double orphans – received inadequate support (Hall & Proudlock, 2011). We therefore study a period when double orphanhood was on the rise, and public support to households that cared for these double orphans was complicated.

We build a new regional dataset to establish the relationships of interest. We first show that double orphans live in regions with higher stunting rates. These patterns are consistent with the "double orphan stunting premium" hypothesis. Our spatial econometric models illustrate that regional inequalities in private household resources and access to WASH overlap strongly with the higher risk of stunting observed in regions where double orphanhood is also most likely. We also compare whether stunting has a stronger relationship with access to WASH infrastructure or with other softer social services. Our analysis considers whether the location of WASH or soft services correlates more closely to orphanhood, and whether this determines the strength of the nature of the relationship between double orphanhood and stunting rates.

In effect, we generate new insights on the potential role that private resources and public WASH infrastructure could play in reducing stunting rates, particularly in impoverished areas where double orphans are most likely to live. However, because we do not use data collected at the level of the individual and because we estimate partial correlations, our results cannot assert that these relationships are causal. However, our use of regional data allows us to assess whether these relationships are stronger when various services are concentrated in larger geographic networks.

Our most parsimonious models show that local orphan and stunting rates are significantly negatively correlated. Controlling for access to piped water and sanitation networks in our models has the same effect as controlling for private household resources: orphanhood and stunting no longer have a significant relationship. Access to other "softer" public services such as clinics, early childhood development facilities, and social welfare offices, does not have a similar statistical effect. Together, these results are consistent with the hypothesis that household resources and WASH are more important at reducing stunting in areas where double orphans live, while other public services potentially play a lesser role. The results show that public infrastructure is related to stunting with spatial network effects: in areas where there are spatial agglomerations of WASH, the negative relationship becomes stronger. This evidence is crucial as a case study in understanding how best to care for vulnerable children in South Africa, and for other developing countries in which caring for orphaned and vulnerable children is a policy concern.

To begin, section two outlines the relevant literature, while section three provides a description of the dataset used in this analysis. Section four discusses the spatial econometric methodology and how this addresses biases in estimating the relationships of interest. Section five provides descriptive statistics and maps to show the spatial patterns of the key variables.

The section proceeds to show the core results from Ordinary Least Squares (OLS) and Spatial Autocorrelation Regression (SAR) models, while section six concludes.

2. Context and Literature Review

2.1. Stunting, WASH and orphanhood

This analysis uses the prevalence of stunting in children under five as the primary outcome variable to measure health outcomes. Stunting quantifies child growth failure, where children are too short for their age. Specifically, stunting is quantitatively defined as "children with a height-for-age... z score that is more than two standard deviations below the 2006 World Health Organization (WHO) growth reference population" (Osgood-Zimmerman *et al.*, 2018). This measure has been used in economic analysis of child health outcomes and quantifies their vulnerability (Duflo, 2000, 2003; Osgood-Zimmerman *et al.*, 2018; Prendergast & Humphrey, 2014; WHO, 1995).

In 2015, over a quarter of all children under the age of five in South Africa were found to have stunted growth (Osgood-Zimmerman *et al.*, 2018). Additionally, Said-Mohamed *et al.* (2015) demonstrate that the stunting prevalence in South Africa has remained persistently high since the 1990s, and that stunting is worse for children in rural areas.

Many studies have investigated the mechanisms by which increased access to WASH reduces child mortality and improves other child health outcomes such as stunting. Human faecal matter is a primary carrier of pathogenic bacteria and soil-transmitted helminths. These can cause multiple illnesses such as diarrhoea, cholera or dysentery. Pathogens can also cause environmental enteric disorder, trachoma, intestinal parasites and wasting in children (Mara, Lane, Scott & Trouba, 2010; Omiat & Shively, 2020). These outcomes are known to increase the likelihood of malnutrition and stunting in children. Furthermore, Mara *et al.* (2010) find that the unsafe disposal of human faecal matter increases the risk of diarrhoea by 23%. Spears (2013:2) argues that variation in the practice of open defecation in developing countries can significantly explain the variation in child stunting. These studies emphasise the need for safe human waste disposal to reduce harmful pathogens that contribute to illness in young children. Adequate WASH infrastructure therefore plays an important role in curbing high stunting rates.

In a review of 172 Demographic and Health Survey (DHS) datasets, Gunther and Fink (2010) find a robust association between improved access to WASH and lower child mortality in developing country contexts. They also find a significant relationship between improved WASH and lower child stunting in a further study (Gunther, Fink & Hill, 2010). Headley and Palloni (2019) also investigate the link between WASH and stunting using DHS data from 59 developing countries and find a significant association between improved access to WASH and lower stunting prevalence. Their results further suggest that access to improved WASH infrastructure in the home is a critical factor, as opposed to improved access to water in the community. A recent study illustrates that children who live close to refugee camps – where limited water sources are shared by many people – have long-run height disadvantages in adulthood that can be explained by the spread of disease (Nsababera, 2020).

In developing country contexts access to WASH is not yet universal, with impoverished areas lagging most severely (Aguayo & Menon, 2016; Fink *et al.*, 2011; Penner, 2010). Children who live in poor areas are therefore more likely to experience chronic diarrhoea and therefore have higher rates of stunting. Furthermore, orphaned children in sub-Saharan Africa are less likely than non-orphaned children to have access to health services and adequate nutrition, and are considerably less likely to complete schooling (Lombe & Ochumbo, 2008).

A review by Li Chng and Chu (2017) finds that orphans face fewer disadvantages if they live within an extended family network. There is little evidence regarding the importance of public service provision for supporting orphans, however. The only study which indicates the importance of adequate access to services for orphaned children investigates the adverse effects of poorly resourced orphanages on child health, and not the impact of infrastructure on orphans who continue to live in private homes in their communities. Living in an orphanage has a severe negative impact on children's physical and cognitive outcomes (Carr, Duff & Craddock, 2018). Therefore, there is a need for more robust evidence on the role that public service provision may play in mitigating the negative effect of orphanhood on child health outcomes, especially in private homes.

2.2. WASH and orphanhood in South Africa

2.2.1 WASH in South Africa

The provision of sanitation is fundamental to addressing inequality in South Africa, not least in addressing poor child health outcomes and stunting (Aguayo & Menon, 2016; Beal *et al.*, 2018; Bernal *et al.*, 2014; Jinabhai, Taylor & Sullivan, 2006; May & Timæus, 2014; Zere & McIntyre, 2003). Access to sanitation in households in South Africa has a distinct spatial dimension that reflects the geography of *apartheid* (Sutherland *et al.*, 2014:470). The *apartheid* government created 'homelands' by a series of laws enacted from 1951 to entrench its policy of separate development. The majority of South Africa's black population was limited to living in rural 'homelands', and about 3.5 million people were forcefully resettled to live in these areas. This contributed to overcrowding, competition for resources, and to economic stagnation (Abel, 2019; von Fintel & Fourie, 2019). While these areas were re-integrated into South Africa's provinces at democracy in 1994, spatial inequality continues to be defined by these demarcations (von Fintel, 2018). This includes access to health-promoting WASH infrastructure and nutrition indicators. Households in the homelands continued to experience significantly higher levels of hunger in the years following the transition to democracy, though access to cash transfers narrowed this gap in later years (Pienaar & von Fintel, 2014).

Before the advent of democracy in 1994, water and sanitation provision was relegated to local authorities, with little support or guidance from the national government (Hattingh *et. al.*, 2007). Cross-regional co-ordination was limited and provision across areas was not standardised. This hindered the establishment of effective regional WASH infrastructure networks. This scattered approach laid the basis for large spatial inequalities in WASH. Rural areas were severely under-serviced, while urban areas had better access. Because the homelands were considered to be separate states by the *apartheid* government, investment in WASH infrastructure lagged significantly behind the rest of the country. By 1994, more than half of the population could not access adequate sanitation, and more than a third had no access to clean drinking water (Hattingh *et. al.*, 2007).

The newly elected democratic government had to address these severe historical backlogs in rural infrastructure. Equal access to services is emphasised in democratic South Africa's constitution, the Reconstruction and Development Programme (RDP) as well as the National Development Plan (NDP). The expansion of WASH was centralised to a dedicated Department of Water Affairs and Forestry in 1998. Considerable progress has been made in providing all South Africans with proper sanitation, but even so, the 2016 Statistics South Africa Community Survey suggests that only 59% of South African households have access to flushing toilets, and 31% of households still use pit latrines (Stats SA, 2016). Burger *et. al.* (2017) estimate that household deprivation in sanitation and access to water declined by 34% and 26% respectively between 1996 and 2011. Progress in water provision was most pronounced in provinces that contained former homelands, ranging from a 19.9% reduction in deprivation in Limpopo to 32% in the Eastern Cape. That said, deprivation levels in the Eastern Cape in 2011 remained 14% higher than the national average 15 years earlier in 1996. While the provision of sanitation progressed well, improvements have been uneven. The commercial centres of the country – the Western Cape and Gauteng – experienced 30-40% reductions in deprivation compared to only 20% in rural Limpopo. Households in Limpopo had 37% higher deprivation levels in 2011 compared to the national average in 1996. All other parts of the country surpassed the 1996 national baseline. Household access to WASH expanded dramatically in the first 15 years of democracy, but many rural areas continue to face significant disadvantages in this regard.

2.2.2 Orphanhood in South Africa

This paper is one of the first to focus on the relationship between stunted growth and orphanhood in South Africa. We also review existing studies that describe other outcomes. There is a substantial but dated body of literature on orphans in South Africa which focuses largely on the increase in the prevalence of orphanhood due to HIV/AIDS among parents

(Ardington & Leibbrandt, 2010; Boler, 2007; Richter & Desmond, 2008). Consequently, it is well-documented how parental health influences the orphan status of children, but little is known about how orphans' own health situation is affected by the loss of their parents. Between 2000 and 2010, a wealth of empirical investigation on educational outcomes brought to light both the growing crisis of orphanhood in South Africa (Chuong & Operario, 2012; Meintjes *et al.*, 2010; Parikh *et al.*, 2007; Richter & Desmond, 2008; Saloojee *et al.*, 2007; Skinner *et al.*, 2006), as well as the consistently lower welfare outcomes of orphans in comparison with non-orphans (Case & Ardington, 2006; Case, Paxson & Ableidinger, 2004; Evans & Miguel, 2007).

Ardington and Leibbrandt (2010) find that orphans in South Africa have worse educational outcomes than non-orphans. A further study finds that orphaned children in South Africa are no more likely to drop out or stay behind in school than non-orphaned children once household characteristics such as socioeconomic status of the household, living arrangements, and household size are taken into consideration (Chuong & Operario, 2012). This concurs with international findings on stunting, where household socioeconomic differences can account for higher stunting rates among orphans (Ali *et. al.*, 2018). This analysis shifts the focus to stunting and away from educational outcomes.

3. Data on Orphaned and Vulnerable Children in South Africa

This analysis brings information together from multiple publicly available data sources, focused on orphaned children and service provision in South Africa. For the remainder of the paper, this dataset will be referred to as the data on Orphaned and Vulnerable Children in South Africa (OVCSA).[6] We used electoral wards as the unit of observation, and aggregated data from various sources to create a spatial cross section.[7]

[6] A full description of the data can be found in online Appendix C: data in brief.

[7] There are 4277 electoral wards in South Africa, but information was not recorded in the 2011 census in one ward. Therefore, there are 4276 observations in this dataset.

The sources used to create the OVCSA dataset include the 2011 South African Community Census, the 2013/14 Audit of Early Childhood Development (ECD) facilities in South Africa, a comprehensive geospatial estimation of child growth failure in Africa by Osgood-Zimmerman *et al.,* (2018), and publicly available lists of government facilities. The government facilities include the locations of health facilities, police stations, South African Social Security Agency (SASSA) offices and Department of Social Development (DSD) service points.

3.1. Measures of CGF: primary outcome variable of interest

The primary outcome variable used to measure health outcomes for orphans in South Africa is the stunting prevalence. This variable was obtained from datasets created by Osgood-Zimmerman *et al.* (2018)[8]. This data contains estimates of the prevalence of stunting, wasting and underweight in children under five in 5km x 5km gridded cells across Sub-Saharan Africa. For the purpose of this dataset, grid-level stunting prevalence was averaged to create stunting prevalence in electoral wards. Our outcome variable is therefore the stunting rate of children under five in 2013. Many of the other variables, including maps of soft infrastructure facilities, are also taken from 2013.

The stunting measures are small area estimates. The original data also contains confidence bands of these estimates. Additional results in the online appendix use interval regressions with the confidence bands as dependent variables to stress test the core results to using an estimated dependent variable. However, basic checks suggest that the data are reliable. The national and provincial stunting prevalence reported by Said-Mohamed *et al.* (2015) is consistent with that estimated by Osgood-Zimmerman *et al.* (2018).

3.2. Data from other sources

While multiple variables in the OVCSA data are taken from the 2011 Census, five are of particular importance: (i) the proportion of children between 0 and 4 years in each ward who are double orphans; (ii) the proportion of households with access to flushing toilets in

[8] This data is open access and has been made available by the authors for free download in raster format.

the home[9] and (iii) piped water in the home or via a tap in the garden; (iv) the average private household asset ownership in each ward; and lastly, (v) the proportion of households that reported living in an urban area in each ward. Ward-level aggregates were obtained from Statistics South Africa's SuperCross website (StatsSA, 2011).

We constructed an asset index to measure private household socio-economic status. First, we calculated the proportion of households in each electoral ward that owned all 128 possible combinations of a fridge, electric/gas stove, washing machine, computer, telephone, TV and/or a radio. Second, each of these combinations was assigned a wealth ranking: owning all of the assets was given a weight of one, and owning zero assets received a weight of zero. We assumed that owning more assets made a household wealthier. The regional asset index varies between 0 and 1[10]. The Gini coefficient across wards is 0.34, which is lower than the national income inequality estimate of 0.57 across households in 2011 (StatsSA, 2019). Our estimates therefore smooth over substantial variations within wards. This type of asset index is appropriate to proxy for household income in South Africa (Wittenberg & Leibbrandt, 2017), has been used to approximate household wealth in other developing countries using similar data (Harttgen & Vollmer, 2011), and has also been used to control for wealth when measuring child health outcomes (Sahn & Stifel, 2000).

The other pertinent variables in the analysis include the number of early childhood development centres, public health facilities (including clinics and hospitals) [11], police stations, South African Social Security Agency service points and Department of Social Development offices per 100000 inhabitants, hereafter referred to as "soft services". Most of these public services are accessed outside of the home, unlike WASH infrastructure. ECD centres are generally privately owned, but receive substantial support from the state. These variables are included as a measurement of direct government support for health and child

[9] The toilet could either be attached to the municipal waste system, or a septic tank.

[10] As a robustness check, this index is compared with a data-driven PCA approach that is shown in online appendix C. While the index produced via PCA is helpful, it did not capture as much variation in asset ownership (Gini coefficient across wards of 0.27) as our approach and was therefore not used.

[11] We ran separate models that included private health facilities. However, only 33% of listed private facilities have usable location data to link them to wards. The coefficients on private health facilities were insignificant, likely due to mismeasurement error, and we therefore limited our focus to estimating associations with public health facilities. Only 22% of public health facilities, 0.007% of police departments and 4% of ECD facilities had insufficient location data to link them to a ward.

welfare in an area which may reduce stunting, especially in areas with higher proportions of orphans. Because many wards have no such services, we accommodate large numbers of zeroes and skewness in these variables' distributions by applying the inverse hyperbolic sine transformation (Bellemare & Wichman, 2020). We multiply these particular coefficients by 100, so that they can be interpreted as a percentage point change in stunting rates for a percentage change in access to these facilities.

4. Methodology

Our estimation strategy relies on the spatial properties of the data. Spatial econometric analysis draws heavily on the first law of geography, formulated by Tobler (1970:236), who noted that "everything is related to everything else, but near things are more related than distant things". In our data, outcomes in wards that are closer together are more closely related to each other. The stunting rate in one ward is related to the stunting rate in neighbouring areas, because correlated shocks affect and propagate across larger regions. Where there is an increase in child malnutrition or incidence of diarrhoea in one area, the effect is likely to spread to a number of wards around the epicentre of the shock. Disease spreads across areas, and local market shocks affect prices in surrounding markets. Correlated shocks such as food prices, weather or health shocks affect the nutritional status of children in large regions beyond the initial source of the shock.

We therefore introduce an autoregressive spatial error term to our models. It operates as a regional fixed effect, and takes into account that unobservable shocks in one area also affect the current stunting in adjacent areas. Spatial error terms reduce coefficient bias from omitted regional shocks, but do not address all biases that continue to prevent causal interpretation. Furthermore, this approach accounts for the dependent variable being constructed using spatial interpolation (Osgood-Zimmerman *et. al.*, 2018). Prior modelling introduced high spatial correlation in the dependent variable that does not necessarily reflect more abrupt changes in actual stunting rates across wards. We control for spatial persistence introduced by this process with the spatial error term, and this mitigates against this specific type of measurement bias.

Access to infrastructure is also likely to have "spill-over" effects, or spatial network effects. If there are multiple healthcare facilities in one area, residents of neighbouring areas are also likely to have improved access to healthcare, albeit not in their own wards. Our model assumes that the control variables $\boldsymbol{x}_i$ have an own-region direct relationship with stunting, and an indirect relationship that comes from larger geographic networks. Infrastructure operates with network effects: if adjacent regions are part of a WASH network, economies of scale operate, and the connected parts of the geographically spread network operate together rather than inside isolated wards. Arguably, this is more important for physical networks than networks of softer service provision. We therefore do not include spatial lags of household variables, such as the asset index and orphan rates, as we do not expect indirect impacts of socio-economic conditions other than through shared services between households in neighbouring wards and other local unobservable factors.

A spatial weighting matrix W is used to model both spatial dependence and spill-over in covariates. The elements w_{ij} indicate relationships between ward i and ward j. We use a rook contiguity matrix whose elements w_{ij} equal 1 if ward i and ward j are neighbours, and 0 if they are not (Anselin, Gallo & Jayet, 2008). If the spill-over effect of ward i shocks or characteristics are diluted as the number of neighbours grows, it is common to row-normalise W, so that weights sum to 1 in each row. For example, if ward i has many health facilities, the ward's neighbours experience increased access to services. However, those spill-over effects are measured to be smaller if more neighbours share the same health facilities among many people. We therefore row-normalise our spatial weighting matrix.

We run the Moran test for spatial dependence in the residuals of our baseline OLS models to motivate the estimation of spatial econometric specifications. The most general form of the equation we estimate is:

$$stunting\ rate_i = \beta_0 + \beta_1 double\ orphan\ rate_i + \boldsymbol{\delta}'_i \boldsymbol{x}_i + \boldsymbol{\beta}'_2 W \boldsymbol{x}_i + (I - \rho W)^{-1} u_i. \tag{1}$$

where *i* indexes wards and $\boldsymbol{x}$ is a vector of regional household resources and public services. The simplest models estimate β_1, the coefficient of interest, without controls or autoregressive errors (by setting $\boldsymbol{\delta} = \boldsymbol{\beta_2} = \mathbf{0}$ and $\rho = 0$) by OLS. The first spatial extension introduces spatial autocorrelation in the errors by allowing $\rho \neq 0$. Full specifications include the control vector $\boldsymbol{x}_i$ with $\boldsymbol{\delta} \neq \mathbf{0}$ and covariate spill-over, with $\boldsymbol{\beta_2} \neq \mathbf{0}$. In the case of spill-overs, we estimate direct effects, indirect effects, as well as the total effects of the various services on stunting rates[12].

5. Results

5.1. Descriptive Statistics and spatial distributions

Table 1 presents descriptive statistics for the variables of interest. In the average ward, about 1% of children are double orphans, with some wards experiencing rates as high as 9.6%. Asset ownership is low on average, but some wards are close to having universal access to all household goods. Access to sanitation is low (48% of households in the average ward), while access to piped water in the home is significantly higher (71% of households in the average ward). Access to softer services, such as ECD centres and health facilities is more widespread. South Africa's democratic governments have therefore been successful at addressing the geographic spread of services outside the home, but access to WASH infrastructure within the home is not yet universally available.

We also show the spatial distributions of stunting, private assets, sanitation and of orphanhood. Figure 1(a) shows the prevalence of stunting in children under five by electoral ward in South Africa, where stunting rates are as high as 45% in some wards. Stunting is particularly prevalent in provinces with large rural populations (KwaZulu Natal, the North-West Province and Limpopo, as well as some areas of the Eastern Cape and the Northern Cape). The national stunting rate was 28% in 2013. Stunting is more severe in the 10% of wards where double orphanhood is most common (31%), while it is significantly lower (p<0.001) in the 10% of wards with the lowest rates of orphanhood (25.7%). Wards with a higher concentration of double orphans therefore also have more child stunting.

[12] The term "effect" is used widely in the spatial econometric literature, but does not necessarily imply causality.

Table 1: Descriptive statistics

Variable Name	Description	N	Mean	Standard Deviation	Min.	Max.
Proportion Double Orphans†	The proportion of children aged 0 to 4 who are double orphans in the ward.	4276	0.01	0.008	0	0.096
Asset Index†	The average private asset ownership in the ward.	4276	0.31	0.19	0.008	0.936
Piped Water†	The proportion of households with access to piped water in the ward.	4276	0.71	0.31	0	1
Flush Toilets†	The proportion of households with access to flush toilets in the ward.	4276	0.48	0.41	0	1
Urban†	The proportion of households in the ward which reported being urban.	4276	0.52	0.46	0	1
ECD Centres*	The number of ECD Centres (Govt. and private) in the ward.	4276	3.46	4.88	0	95
Health Facilities*	The number of public health facilities (clinics and hospitals) in the ward.	4276	1.34	1.58	0	16
Police Departments*	The number of police departments in the ward.	4276	0.26	0.49	0	4
SASSA Centres*	The number of South African Social Security Agency Centres in the ward.	4276	0.09	0.29	0	2
DSD Offices*	The number Department of Social Development Offices in the ward.	4276	0.15	0.69	0	24

NOTES: Own calculations. †Ward-level aggregates from the 2011 census (StatsSA, 2011); *Ward-level totals from publicly available geolocation data; Descriptions of variables in the table; the asset index ranges from 0 = all households in the ward have no assets to 1 = if all households in a ward possess all asset combinations in the census.

This correspondence is reflected in Figure 1(b), which shows the spatial distribution of double orphans under the age of 4 by electoral wards[13]. The distribution closely follows that of the prevalence of stunting, except in the north-east of the country, with high rates of double orphanhood in KwaZulu-Natal and the eastern part of the Eastern Cape. Stunting rates and the prevalence of double orphanhood have a strong positive relationship, with a ward-level correlation coefficient of 0.27.

Figure 1(c) shows the geographic distribution of the asset index across electoral wards in South Africa. This distribution of assets is consistent with other multidimensional analyses

[13] The number of double orphans in a ward includes all double orphaned children under 4, while the stunting prevalence in the ward includes all children under 5. This inconsistency is unavoidable due to the different age-groupings in the two different datasets used to compile the OVCSA data.

of poverty and inequality in South Africa (David *et al.*, 2018). Asset ownership is highly concentrated in metropolitan areas such as Cape Town, eThekwini, Tshwane and Johannesburg, and there are pockets of higher asset ownership around the metropoles of Nelson Mandela Bay, Buffalo City and Mangaung. In contrast, asset ownership declines significantly in non-metropolitan rural areas. Asset ownership and stunting prevalence are negatively related (correlation coefficient of -0.45), while asset ownership and access to sanitation are positively related (correlation coefficient of 0.68). Private assets are therefore strongly related to lower stunting. By implication, there is a strong urban bias in these relationships. Stunting is low and household socio-economic conditions are favourable in urban areas.

Figure 1(d) shows rates of household access to flushing toilets in each ward. There are pockets of extremely low access to sanitation: as few as 3% of households in some wards had a flushing toilet in 2011. Again, there is a strong urban bias that is dominated by low access to sanitation in the mainly rural former *apartheid* 'homelands'. Access to sanitation and the prevalence of double orphanhood is negatively correlated (correlation coefficient -0.36).

Figure 1 has highlighted the strong relationships between stunting rates and local socio-economic conditions. Stunting corresponds closely to double orphanhood and conditions in the home. The spatial distributions show a clear rural disadvantage in the provision of hard sanitation infrastructure; this overlaps with the stunting disadvantage in rural areas where many double orphans are also located.

Figure 2 shows the number of soft service points per 100 000 inhabitants in each ward, including ECD centres, public health facilities (clinics and hospitals), police stations, South African Social Security Agency (SASSA) service points, and Department of Social Development (DSD) offices[14]. In contrast to the provision of sanitation infrastructure, soft services have expanded widely to reach remote parts of the country. They are well-represented in rural areas, and emphasise government's efforts to provide services to citizens

[14] SASSA service points are centralized offices where recipients of social grants in South Africa can apply for grants in person, receive grants in person and query or update their grant recipient details. Access to these centralized offices is likely to increase the support that is accessible to individuals through the cash transfer (or social grant) system in South Africa, which is a substantial safety net for the most vulnerable individuals in the country.

across all parts of the country. McLaren *et. al.* (2014) show that a majority of South Africans live within 2km of a public clinic: proximity is less of a constraint to healthcare than high transport costs and poor quality of care. These distributions are, however, not strongly correlated with urbanisation (see Table 2) and also not with stunting. Soft services, including health facilities, are accessible to most citizens. But this descriptive evidence suggests that their likely impact on stunting is weak.

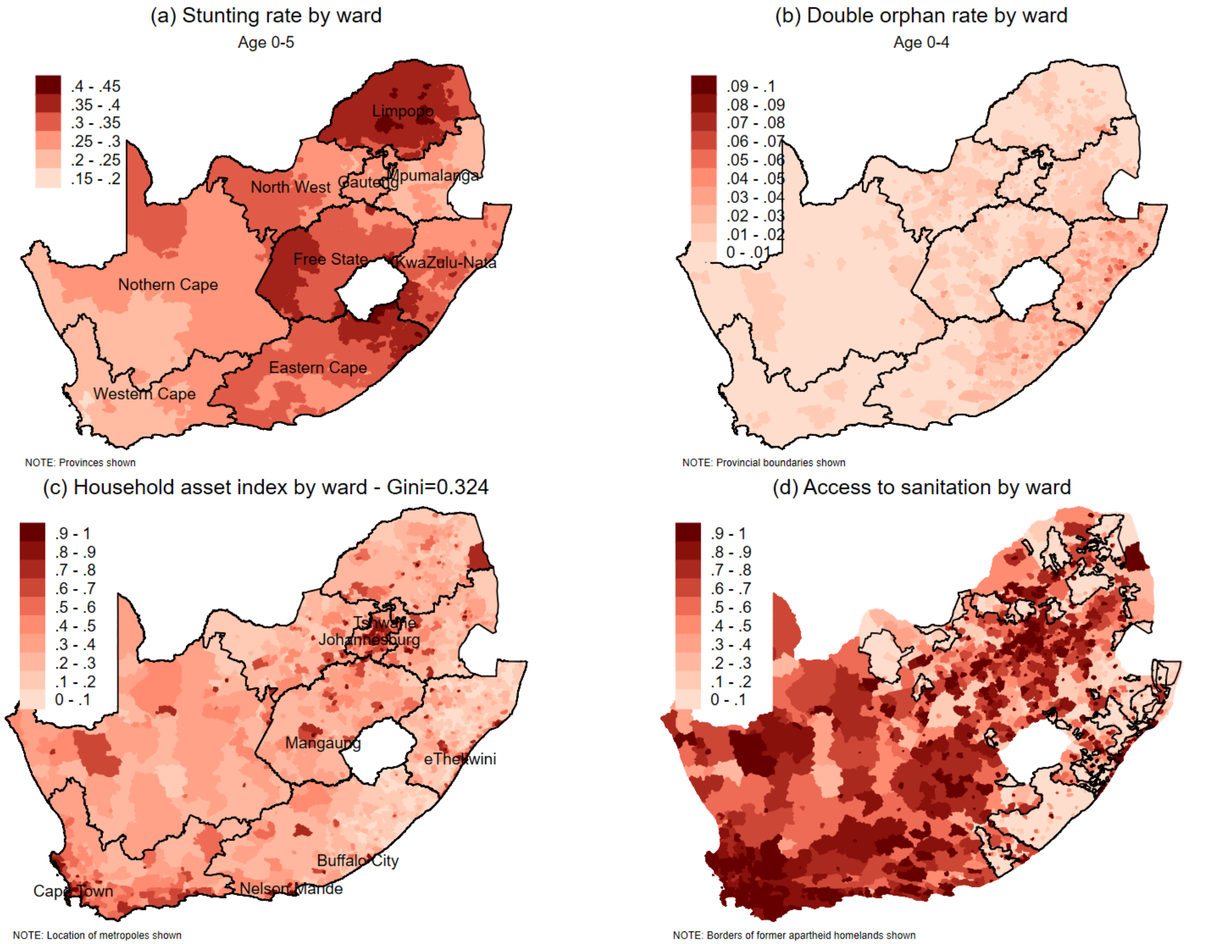
These figures emphasise that it is challenging to establish a large network of WASH infrastructure that reaches homes, and access to these services is biased towards urban areas (see Table 2). It is less complicated to provide "stand alone" services such as clinics or police stations which do not require a network of connected infrastructure such as dams, pipes, sewage networks, and access to municipal water. Further analysis of bottle-necks in service provision is critical in formulating policy which can successfully implement increased access to sanitation, but is unfortunately outside the scope of the current analysis.

Even so, the greater access to sanitation and piped water in urban areas is not surprising, given South Africa's history of separate development. Investigating the relationship between stunting and orphanhood therefore requires controlling for urbanisation in the econometric analysis in section 6.

Table 2: Correlation between service provision & ward-level urbanisation rates

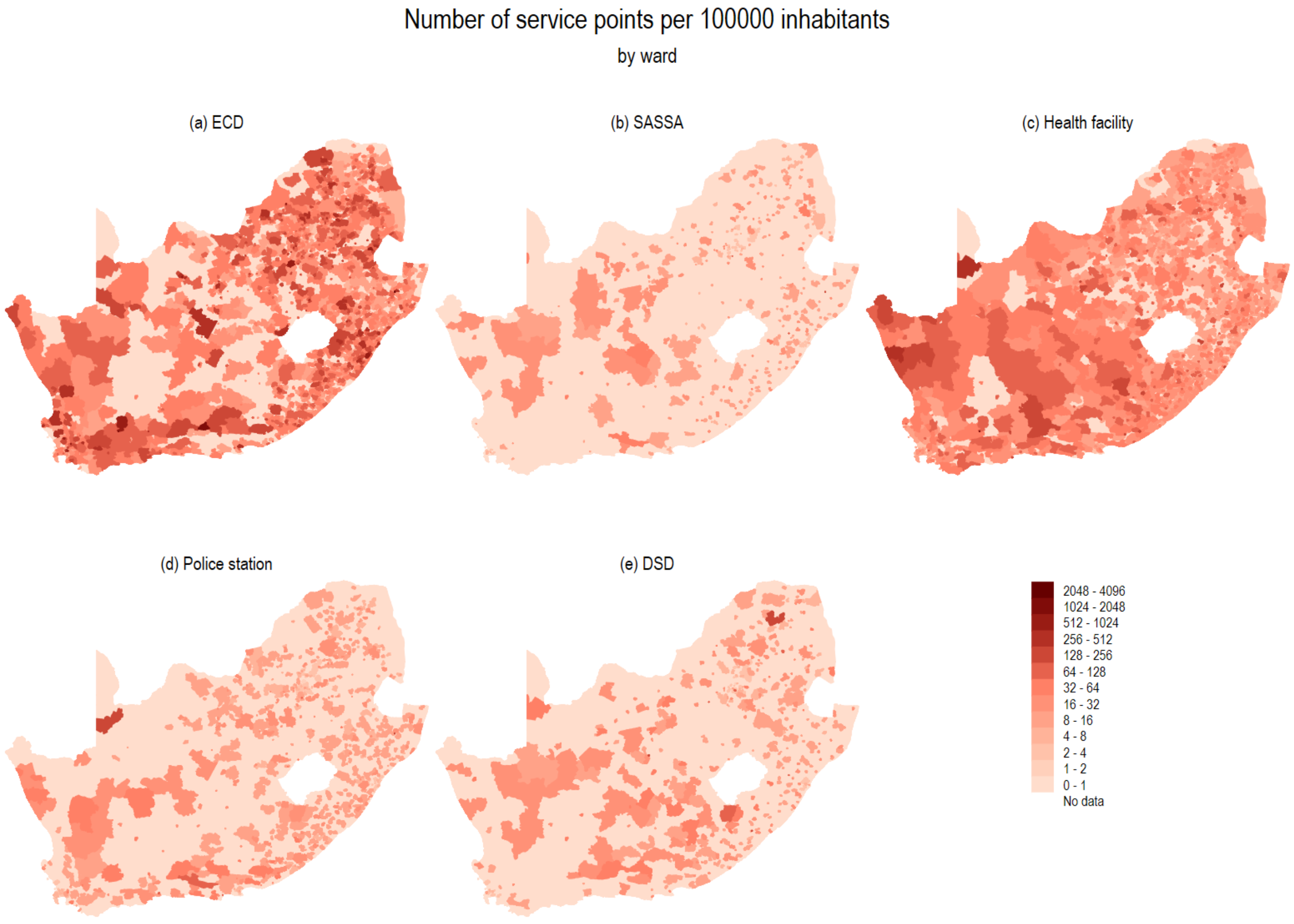
Variable	Correlation Coefficient
South African Social Security Agency (SASSA) Service Points*	0.09
Department of Social Development (DSD) Offices*	0.08
Early Childhood Development Centres*	0.08
Public Health Facilities*	0.10
Police Departments*	0.05
Proportion of Households with access to flush toilets†	0.89
Proportion of Households with access to piped water†	0.71

NOTES: Own calculations. †Ward-level aggregates from the 2011 census (StatsSA, 2011); *Ward-level totals from publicly available geolocation data.



Source: own calculations from OVCSA dataset. Figures are measured in proportions of: (a) the population aged 0-5 (b) the population aged 0-4 and (d) all households. In (c) the index ranges from 0 to 1

NOTES: (a) Ward-level stunting rates are for children aged 0 to 5 in 2013 from Osgood-Zimmerman *et al.* (2018). Province names are shows. (b) Orphanhood rates are the proportion of children aged 0 to 4 in a ward who are double orphans in 2011, from census 2011 (StatsSA, 2011) (c) The asset index ranges from 0 = all households in the ward have no assets to 1 = if all households in a ward possess all asset combinations in the census. Metropolitan areas shown (d) Proportion of households who have access to sanitation from census 2011 (StatsSA, 2011). Borders of former *apartheid* homelands shown.



NOTES: Own calculations using publicly available geolocation data. All figures are measured as the number of service points per 100000 inhabitants in each ward. (a) Early Childhood Development (b) South African Social Security Agency centres (c) Public sector health facilities (d) Police stations (e) Department of Social Development Offices.

5.2. Econometric models

Table 3 presents the core results of our study. Column (1) shows the OLS estimate of $\hat{\beta}_1$ without any controls, and before accounting for spatial autocorrelation in the error term. The correlation between double orphanhood and stunting is large and significant, and confirms that orphans are located in areas with high stunting rates. This pattern is consistent with the presence of a double orphan stunting penalty. As the proportion of double orphans in a ward increases from 0 to 1, the stunting rate grows by 167 percentage points. This coefficient is, however, upward biased because it also measures the effects of unobservable regional spill-overs, such as common price, climate and disease shocks, and because there is measurement bias from spatial smoothing in the dependent variable. The highly significant Moran statistic of 0.901 confirms this, and indicates that there is strong positive spatial autocorrelation in the residuals.

As expected, allowing for spatial dependence in the residuals ($\rho \neq 0$) in column (2) reduces the coefficient substantially. The stunting rate now grows by only 8.2 percentage points as the proportion of double orphans in the ward increases from 0 to 1. We estimate $\hat{\rho}$ at 0.943, indicating strong spatial persistence in the residuals. Our results strongly support the inclusion of spatial errors in all subsequent models.

The relationship between stunting and double orphanhood halves in size and becomes statistically insignificant when a control for ward-level household assets is introduced in column (3). The significant negative association between stunting and household resources is consistent with previous evidence (Ali *et. al.,* 2018, Duc, 2019). Controlling only for access to piped water and flushing toilets in column (4) has a similar statistical effect on the specification as controlling only for household resources in column (3). The coefficient on double orphanhood halves in size relative to column (2) and becomes insignificant.

These results, together with the maps in Figure 1, emphasise that there is strong regional overlap in access to WASH and household wealth. Both factors represent resources accessed in the home, and both relate to lower stunting rates. However, the two factors are not perfect substitutes: one aspect represents private wealth, and the other public infrastructure. Later models explore whether they have independent negative relationships with stunting.

The two factors also confound the relationship between stunting and double orphanhood. Because this relationship becomes insignificant, WASH and household wealth are more closely associated with stunting than double orphanhood. Orphans therefore predominantly live in poorly resourced communities; their location in poor communities describes the strong association between double orphanhood and high stunting rates. Figure 1 confirms the strong correlations between double orphanhood and WASH (correlation coefficients of -0.346 for piped water and -0.363 for sanitation respectively), and double orphanhood and household assets (correlation coefficient of -0.429). These observations are consistent with prior evidence showing that orphans have similar outcomes to other children when adequate resources are available (Ali *et. al.,* 2018). Our results therefore point to inadequate household resources and inadequate access to WASH as the proximate factors that generate a correlation between stunting and double orphanhood.

Column (5) only controls for "soft services" that do not require networks of infrastructure to operate. Areas with high concentrations of ECD facilities and social security service centres have significantly lower local stunting rates. These services focus directly on children's well-being, even if they do not directly address healthcare concerns. In contrast, more police departments in an area have no significant association with stunting rates, and the number of public health facilities is significantly positively associated with stunting. The latter result suggests that public health facilities are appropriately targeted at areas with poor health outcomes, rather than measuring the impact of facilities on stunting. More pertinently, comparing column (5) with column (2) shows that controlling for access to soft services does not remove the statistical significance of the coefficient on orphanhood. The coefficient remains large (at 0.104) and is highly significant. Furthermore, the coefficients on the various soft services are economically small – when the availability of each of the facilities increases by a percentage, stunting rates change by a small fraction of a percentage point. This is unsurprising. These public services have wide geographic reach, as seen in Figure 2, and include areas with high stunting and double orphanhood rates. WASH infrastructure is, however, more concentrated in areas with low stunting and high availability of private resources. These first results emphasise that access to healthcare facilities and other public services are less important than access to WASH for understanding the relationship between stunting and double orphanhood. The associations are consistent with a hypothesis that soft services are not as important as hard infrastructure for reducing stunting more generally, and for double orphans in particular.

In column (6) we assess which of the partial correlations remain significant when we include all covariates simultaneously. However, we also account for an additional source of bias. Increased access to these services may stand proxy for urbanisation, which is important in explaining stunting (Beal *et al.*, 2018; Bernal *et al.*, 2014; Darteh *et al.*, 2014; Delpeuch *et al.*, 2000; Prendergast & Humphrey, 2014). In column (6) we also control for the proportion of households that live in urban areas to account for this potential bias. Similarly to specifications (3) and (4) the coefficient on double orphanhood becomes statistically insignificant. Urbanisation is negatively correlated with stunting. Despite including a full set of controls, *both* household resources and access to piped water remain significantly associated with stunting levels.

Our final model (7) investigates spatial spill-overs in infrastructure provision. We include spatial lags of access to WASH and access to soft services. We distinguish "direct effects" – the *within*-ward correlation between the respective services and stunting rates – from "indirect effects" – the relationship that arises because coverage outside residents' own wards also correlate with local stunting. The indirect relationship between increased access to sanitation and stunting is negative, large and significant, while the direct relationship is negligibly small and statistically insignificant. Access to piped water has a significant direct negative relationship with stunting, but the indirect relationship is much larger. The correlation between service provision and stunting is stronger when the service is provided as part of a larger regional network.

In addition to positive regional network spill-overs, the results are consistent with the hypothesis that there are benefits to bringing piped water close to the home. However, sanitation only has a significant negative relationship with stunting when a larger network is in place. Sanitation has two hypothesised network benefits. First, networks may raise the efficiency of service delivery and WASH provision. Larger sanitation networks leverage economies of scale and network externalities in providing adequate services. Second, network effects may arise from how sanitation is used. Transmission of diarrhoea across wards is less likely in areas that are well serviced by extensive, well-functioning sanitation networks and where people use the services effectively (Badriyah & Syafiq, 2017; Dillingham & Guerrant, 2004; Fink *et al.*, 2011; Spears, 2013). Consequently, stunting is lower in areas with well-established use of sanitation across large spaces.

There are also network spill-overs in some of the soft services. The direct relationship between stunting, and proximity to ECD and social security (SASSA) centres is negative. The indirect relationship between these soft services and stunting rates is larger than the direct relationship. This model also shows a significantly positive direct and indirect relationship between stunting and proximity to health facilities and Department of Social Development centres. These facilities and centres are therefore purposefully located in areas of high poverty and local stunting, instead of causally increasing stunting.

These models are consistent with the notion that there is positive agglomeration in various forms of infrastructure. Agglomeration effects can be defined as the benefit of spill-over effects which produce a larger marginal benefit with spatial concentration of specific characteristics (Marshall, 1920). Widespread and effective provision and use of local services, especially hard infrastructure, should be of the utmost concern for policy in South Africa.

Our results obtained from a regional cross section are also consistent with the hypothesis that private household resources can mitigate children's stunting rates, including those of orphans, in South Africa. This has also been shown in other parts of the world (Duc, 2019). Evidence on the importance of household living arrangements has shown that the environment in which children develop can significantly impact their cognitive and physical development (Carr *et al.*, 2018; Meintjes *et al.*, 2010; Salifu Yendork, 2020). However, it is difficult for the state to intervene in private household composition. Nevertheless, many

South African orphans are already in the care of relatives who qualify for state-supported cash transfers (Hall & Proudlock, 2011): the public purse can influence orphan stunting indirectly by strengthening this support, and provided that the resources are available and implementation capacity are in place. Without adding conditions to cash transfers, the state cannot have direct influence on private choices to prioritise expenditure towards enhancing nutrition. But even unconditional cash transfers have positive impacts on child growth in South Africa (Coetzee, 2013).

More importantly, this analysis supports previous research highlighting the important role of WASH networks in improving sanitary circumstances in homes, and thereby creating an enabling environment where stunting rates can be reduced. However our results also suggest that because double orphans live in areas with low access to WASH, such a policy could be effective at reducing stunting among this vulnerable group in particular.

5.3. Robustness checks

Apart from spatial dependence and network effects, we consider two further econometric concerns. First, the outcome variable of interest is estimated. While measurement errors in dependent variables are less serious for coefficient bias, we re-estimated our models using interval regressions, using the upper and lower estimates of stunting provided by Osgood-Zimmerman *et al.*, (2018) as dependent variables. Second, the outcome variable is a fraction which varies between 0 and 1. In the case where the dependent variable is limited to the unit interval, OLS can predict fitted values that exceed the upper and lower bounds (Papke & Wooldridge, 1996). We show estimates using interval regressions and the fractional probit in online Appendix A. Coefficient sizes and significance levels are similar when using OLS, a fractional probit and an interval regression model, indicating that the specification is robust to the nature of the outcome variable.

Table 3: Direct, indirect and total effect of area characteristics on the stunting prevalence by electoral ward

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)		
Stunting rate							Direct	Indirect	Total
Double orphan rate[†]	1.667***	0.082***	0.041	0.041	0.104***	0.037	0.028		0.028
	(0.354)	(0.029)	(0.028)	(0.028)	(0.031)	(0.029)	(0.027)		(0.027)
Asset index[†]			-0.012***			-0.005***	-0.006***		-0.006***
			(0.002)			(0.002)	(0.002)		(0.002)
Piped Water[†]				-0.013***		-0.010***	-0.013***	-0.024***	-0.037***
				(0.001)		(0.001)	(0.001)	(0.003)	(0.003)
Flush Toilets[†]				-0.007**		-0.002	-0.001	-0.021***	-0.022***
				(0.001)		(0.001)	(0.001)	(0.002)	(0.003)
ECD Facilities[#]					-0.024**	-0.017*	-0.029**	-0.052*	-0.082**
					(0.011)	(0.010)	(0.011)	(0.030)	(0.038)
Police Depts[#]					-0.009	-0.003	-0.021	-0.056	-0.077
					(0.015)	(0.013)	(0.017)	(0.046)	(0.059)
Health Facilities[#]					0.034***	0.043***	0.054***	0.089**	0.143***
					(0.012)	(0.011)	(0.014)	(0.037)	(0.048)
SASSA Centres[#]					-0.045**	0.002	-0.051*	-0.137*	-0.188*
					(0.023)	(0.021)	(0.029)	(0.077)	(0.100)
DSD Centres[#]					-0.027	0.008	0.048*	0.204***	0.252***
					(0.020)	(0.018)	(0.025)	(0.063)	(0.084)
Urban[†]						-0.007***	-0.005***		-0.005***
						(0.001)	(0.001)		(0.001)
Constant	0.266***	0.274***	0.277***	0.289***	0.276***	0.291***	0.319***		
	(0.007)	(0.002)	(0.002)	(0.002)	(0.146)	(0.186)	(0.003)		
Spatial errors	N	Y	Y	Y	Y	Y		Y	
Spatial lag of $\boldsymbol{x}$	N	N	N	N	N	N		Y	
N	4276	4276	4276	4276	4276	4276		4276	
Moran's I[‡] / ρ[♦]	[‡]0.901***	[♦]0.943***	[♦]0.941***	[♦]0.937***	[♦]0.943***	[♦]0.937***		[♦]0.928***	
Pseudo-R^2	0.077	0.077	0.207	0.323	0.077	0.345		0.406	

NOTES: * $p<0.1$, ** $p<0.05$, *** $p<0.01$; Row-normalised rook spatial weight matrix.. [‡]Moran's I of OLS residuals. ρ[♦] is the estimated spatial autocorrelation in errors. Source: Own calculations. [†]Ward-level aggregates from the 2011 census (StatsSA, 2011); [#]Ward-level totals per 100000 inhabitants from publicly available geolocation data and inverse hyperbolic sine transformation applied, with coefficients and standard errors multiplied by 100; Ward-level stunting rates are for children aged 0 to 5 in 2013 and from (Osgood-Zimmerman *et al.*, 2018). Orphanhood rates are the proportion of children aged 0 to 4 in a ward who are double orphans in 2011. The asset index ranges from 0 = all households in the ward have no assets to 1 = if all households in a ward possess all asset combinations in the census. Direct and indirect effects are estimated using the rook spatial weight matrix.

6. Conclusion

Our study is one of the first to investigate the relationships between stunting, double orphanhood and public service provision. Orphans tend to live in areas with high rates of stunting in South Africa. This observation is consistent with the hypothesis that there is a "double orphan stunting penalty" in South Africa. The pattern also corresponds closely with the observation that orphans tend to live in rural areas with lower average private household resources and WASH infrastructure. Proximity to soft public services – which have the intention of directly affecting healthcare and child welfare – are only weakly correlated with stunting. More importantly, these factors are not as closely related to stunting and orphanhood in our regional cross section as are household resources and WASH. Access to soft services is therefore an unlikely reason for why there is a high regional correspondence between orphanhood and stunting. Rather, greater access to hard (WASH) infrastructure and private household resources have negative relationships with stunting that remain significant in our most fully specified models. Further, controlling for these factors removes the significance of the negative relationship between double orphanhood and local stunting rates. While we cannot infer causality from our analysis of a regional cross section dataset, these relationships are consistent with the hypothesis that expanding hard WASH infrastructure can play a role in breaking down the negative relationship between double orphanhood and child stunting.

Moreover, positive location spill-over of WASH infrastructure emphasises the importance of comprehensive and proper access to sanitation and piped water in the home. Access across adjacent regions strengthens the correlations we observe. Widespread and co-ordinated provision of sanitation and water should be recognised as a pressing policy concern which will work towards reducing stunting in South Africa (Momberg, Mahlangu, Ngandu, May, Norris & Said-Mohamed, 2020). Prioritising WASH particularly presents as an option to address the higher stunting rates in areas where double orphans are located. Network effects in provision raise supply-side efficiency and widespread geographic use also introduces protection against the spread of disease. WASH infrastructure is a public good, but is best accessed privately in the home. Moreover, the benefits of widespread use of sanitation increase as more people access these services. It is unlikely that any institution other than the state would be able to provide widespread access to its use, however. Co-ordination of services and economies of scale over large regions, as well as regional population immunity in stopping the spread of illness (such as diarrhoea) among children, present as an option to provide lasting solutions to stunting more generally and orphan vulnerability in particular.

These results may explain why the “orphan stunting penalty” is absent in many countries where WASH provision is more equally distributed, and where household resources mitigate the impacts of losing a parent on children’s stunting. While it is not feasible for policy makers to alter the private living arrangements of households, these results imply that targeted investment in health-promoting infrastructure can benefit double orphans who continue to live in disadvantaged regions.

References

Abel, M. 2019. Long-run effects of forced resettlement: evidence from apartheid South Africa. *The Journal of Economic History*. 79(4):915-953.

Aguayo, V.M. & Menon, P. 2016. Stop stunting: Improving child feeding, women's nutrition and household sanitation in South Asia. *Maternal and Child Nutrition*. 12:3–11.

Ali, Z., Abu, N., Ankamah, I.A., Gyinde, E.A., Seidu, A.S. & Abizari, A.R. 2018. Nutritional status and dietary diversity of orphan and non - Orphan children under five years: A comparative study in the Brong Ahafo region of Ghana. *BMC Nutrition*. 4(1):1–8.

Anselin, L., Gallo, J. Le & Jayet, H. 2008. Spatial Panel Econometrics. In L. Màtyàs & P. Sevestre (eds.). Berlin: Springer *The econometrics of panel data*. 625–660.

Ardington, C. & Leibbrandt, M. 2010. Vulnerability of Orphans Between 1993 and 2005. *Economic Development and Cultural Change*. 58(3):507–536.

Badriyah, L. & Syafiq, A. 2017. The Association Between Sanitation, Hygiene, and Stunting in Children Under Two-Years (An Analysis of Indonesia's Basic Health Research, 2013). *Makara Journal of Health Research*. 21(2).

Beal, T., Tumilowicz, A., Sutrisna, A., Izwardy, D. & Neufeld, L.M. 2018. A review of child stunting determinants in Indonesia. *Maternal and Child Nutrition*. 14(4):1–10.

Bellemare, M.F. & Wichman, C.J. (2020), Elasticities and the Inverse Hyperbolic Sine Transformation, *Oxford Bulletin of Economics and Statistics*, 82: 50-61.

Bernal, J., Frongillo, E.A., Herrera, A. & Rivera, J.A. 2014. Food Insecurity in Children but Not in Their Mothers Is Associated with Altered Activities. *The Journal of Nutrition*. 144(24):1619–1626.

Black, R.E., Victora, C.G., Walker, S.P., Bhutta, Z.A., Christian, P., De Onis, M., Ezzati, M., Grantham-Mcgregor, S., et al. 2013. Maternal and child undernutrition and overweight in low-income and middle-income countries. *The Lancet*. 382:427–451.

Boler, T. 2007. Facing the consequenses of AIDS: Orphans educational outcomes and cash grants in South Africa. London School of Hygiene and Tropical Medicine.

Burger, R., van der Berg, S., van der Walt, S. & Yu, D. 2017. The Long Walk: Considering the Enduring Spatial and Racial Dimensions of Deprivation Two Decades After the Fall of Apartheid. *Social Indicators Research*. 130(3):1101–1123.

Carr, A., Duff, H. & Craddock, F. 2018. A Systematic Review of Reviews of the Outcome of Severe Neglect in Underresourced Childcare Institutions. *Trauma, Violence, and Abuse*. 21(3):484–497.

Case, A. & Ardington, C. 2006. The impact of parental death on school outcomes: Longitudinal evidence from South Africa. *Demography*. 43(3):401–420.

Case, A., Paxson, C. & Ableidinger, J. 2004. Orphans in Africa: Parental death, poverty, and school enrollment. *Demography*. 41(3):483–508.

Chuong, C. & Operario, D. 2012. Challenging household dynamics: Impact of orphanhood, parental absence, and children's living arrangements on education in South Africa. *Global Public Health*. 7(1):42–57.

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.

Darteh, E.K.M., Acquah, E. & Kumi-Kyereme, A. 2014. Correlates of stunting among children in Ghana. *BMC Public Health*. 14(1).

David, A., Guilbert, N., Hamaguchi, N., Higashi, Y., Hino, H., Leibbrandt, M. & Shifa, M. 2018. *Spatial poverty and inequality in South Africa: A municipality level analysis*. (221). [Online], Available: http://www.opensaldru.uct.ac.za/handle/11090/902%0Ahttp://www.opensaldru.uct.ac.za/bitstream/handle/11090/902/2018_221_Saldruwp.pdf?sequence=1.

Delpeuch, F., Traissac, P., Martin-Prével, Y., Massamba, J.P. & Maire, B. 2000. Economic crisis and malnutrition: Socioeconomic determinants of anthropometric status of preschool children and their mothers in an African urban area. *Public Health Nutrition*. 3(1):39–47.

Dewey, K.G. & Begum, K. 2011. Long-term consequences of stunting in early life. *Maternal and Child Nutrition*. 7(SUPPL. 3):5–18.

Dillingham, R. & Guerrant, R.L. 2004. Childhood stunting: measuring and stemming the staggering costs of inadequate water and sanitation. *The Lancet*. 363(9403):94–94.

Duc, L.T. 2019. Household wealth and gender gap widening: evidence from adolescents in Ethiopia, India, Peru and Vietnam. *Economics and Human Biology*, 34:208-215.

Duflo, E. 2000. Child Health and Household Resources in South Africa: Evidence from the Old Age Pension Program. *AEA Papers and Proceedings*. (May):393–398. [Online], Available: https://pubs.aeaweb.org/doi/pdf/10.1257/aer.90.2.393.

Duflo, E. 2003. Grandmothers and granddaughters: Old-age pensions and intrahousehold allocation in South Africa. *World Bank Economic Review*. 17(1):1–25.

Evans, D.K. & Miguel, E. 2007. Orphans and schooling in Africa: A longitudinal analysis. *Demography*. 44(1):35–57.

Fink, G., Günther, I. & Hill, K. 2011. The effect of water and sanitation on child health: Evidence from the demographic and health surveys 1986-2007. *International Journal of Epidemiology*. 40(5):1196–1204.

Günther, I & Fink, G. 2010. *Water, Sanitation and Children's Health, evidence from 171 DHS Surveys*. World Bank Policy Research Working Paper 5275, Washington DC, The World Bank.

Hall, K. & Proudlock, P. 2011. *Orphaning and the Foster Child Grant: A return to the 'care or cash' debate.* Cape Town: Children's Institute, University of Cape Town.

Hall, K. 2019. Demography of South Africa's children. In M. Shung-King, L. Lake, D. Sanders & M. Hendricks (eds.). Cape Town: University of Cape Town *Child and adolescent health – leave no one behind, South African Child Gauge 2019*. 216-220.

Hattingh, J., Maree, G.A., Ashton, P.J., Leaner, J.J., Rascher, J. & Turton, A.R. 2007. Introduction to ecosystem governance: focusing on Africa. *Water Policy*. 9(S2):5-10.

Finlay, J.E., Fink, G., McCoy, D.C., Tavárez, L.C., Chai, J., Danaei, G., Ezzati, M., Fawzi, W., et al. 2016. Stunting risk of orphans by caregiver and living arrangement in low-income and middle-income countries. *Journal of Epidemiology and Community Health*. 70(8):784–790.

Harttgen, K. & Vollmer, S. 2011. Inequality Decomposition without Income or Expenditure Data: Using an Asset Index to Simulate Household Income. *United Nations Development Programme*. 1–18.

Headey, D. & Palloni, G. 2019. Water, Sanitation, and Child Health: Evidence From Subnational Panel Data in 59 Countries. *Demography*. 56(2):729–752.

Heltberg, R. 2009. Malnutrition, poverty, and economic growth. *Health Economics*. 18(SUPPL. 1):1–22.

Hörcher, D., Borger, B. De, Seifu, W. & Graham, D.J. 2020. Public transport provision under agglomeration economies. *Regional Science and Urban Economics*. 81(103503).

Jinabhai, C.C., Taylor, M. & Sullivan, K. 2006. Persistent undernutrition amongst preschool children a decade after democracy. *Southern African Journal of Epidemiology and Infection*. 21(1):26–30.

Lombe, M. & Ochumbo, A. 2008. Sub-Saharan Africa's orphan crisis: Challenges and opportunities. *International Social Work*. 51(5):682–698.

Mara, D., Lane, J., Scott, B. & Trouba, D. 2010. Sanitation and Health. *PLOS Medicine*. 7(11). [Online], Available: https://doi.org/10.1371/journal.pmed.1000363.

Marshall, A. 1920. *Principles of Economics*. London: Macmillan.

May, J. & Timæus, I.M. 2014. Inequities in under-five child nutritional status in South Africa: What progress has been made? *Development Southern Africa*. 31(6):761–774.

McLaren, Z.M., Ardington, C. & Leibbrandt, M. 2014. Distance decay and persistent health care disparities in South Africa. *BMC Health Services Research* 14(541)

Meintjes, H., Hall, K., Marera, D.H. & Boulle, A. 2010. Orphans of the AIDS epidemic? the extent, nature and circumstances of child-headed households in South Africa. *AIDS Care - Psychological and Socio-Medical Aspects of AIDS/HIV*. 22(1):40–49.

Momberg, D.J., Mahlangu, P., Ngandu, B.C., May, J., Norris, S.A. & Said-Mohamed, R. 2020. Intersectoral (in)activity: Towards an understanding of public sector department links between water, sanitation and hygiene (WASH) and childhood undernutrition in South Africa. *Health Policy and Planning*. 35(7):829–841.

Nsabarera, O. 2020. Refugee camps – a lasting legacy? Evidence on long-term health impact. *Economics and Human Biology*. 39: 1-16.

Omiat, G. & Shively, G. 2020. Rainfall and child weight in Uganda. *Economics and Human Biology*, 38:1-18

de Onis, M. & Branca, F. 2016. Childhood stunting: A global perspective. *Maternal and Child Nutrition*. 12:12–26.

Osgood-Zimmerman, A., Millear, A.I., Stubbs, R.W., Shields, C., Pickering, B. V., Earl, L., Graetz, N., Kinyoki, D.K., et al. 2018. Mapping child growth failure in Africa between 2000 and 2015. *Nature*. 555(7694):41–47.

Papke, L.E. & Wooldridge, J.M. 1996. Econometric methods for fractional response variables with an application to 401 (k) plan participation rates. *Journal of Applied Econometrics*. 11(6):619–632.

Parikh, A., DeSilva, M.B., Cakwe, M., Quinlan, T., Simon, J.L., Skalicky, A. & Zhuwau, T. 2007. Exploring the Cinderella myth: Intrahousehold differences in child wellbeing between orphans and non-orphans in Amajuba District, South Africa. *Aids*. 21(SUPPL.

7).

Penner, B. 2010. *Flush with Inequality: Sanitation in South Africa*.

Pienaar, P.L. & von Fintel, D.P. 2014. Hunger in the former apartheid homelands: determinants of convergence one century after the 1913 Land Act. *Agrekon*. 53(4):38-67.

Prendergast, A.J. & Humphrey, J.H. 2014. The stunting syndrome in developing countries. *Paediatrics and International Child Health*. 34(4):250–265.

Quisumbing, A.R. & Maluccio, J.A. 2003. Resources at marriage and intrahousehold allocation: Evidence from Bangladesh, Ethiopia, Indonesia, and South Africa. *Oxford Bulletin of Economics and Statistics*. 65(3):283–327.

Richter, L.M. & Desmond, C. 2008. Targeting AIDS orphans and child-headed households? a perspective from national surveys in South Africa, 1995-2005. *AIDS Care - Psychological and Socio-Medical Aspects of AIDS/HIV*. 20(9):1019–1028.

Ruiz-Casares, M., Thombs, B. D., & Rousseau, C. 2009. The association of single and double orphanhood with symptoms of depression among children and adolescents in Namibia. *European child & adolescent psychiatry*, *18*(6), 369–376.

Sahn, D.E. & Stifel, D.C. 2000. Poverty comparisons over time and across countries in Africa. *World Development*. 28(12):2123–2155.

Said-Mohamed, R., Micklesfield, L.K., Pettifor, J.M. & Norris, S.A. 2015. Has the prevalence of stunting in South African children changed in 40 years? A systematic review. *BMC Public Health*. 15(534).

Salifu Yendork, J. 2020. Vulnerabilities in Ghanaian orphans: Using the ecological systems theory as a lens. *New Ideas in Psychology*. 59(April):100811.

Saloojee, H., De Maayer, T., Garenne, M.L. & Kahn, K. 2007. What's new? Investigating risk factors for severe childhood malnutrition in a high HIV prevalence South African setting. *Scandinavian Journal of Public Health*. 35(SUPPL. 69):96–106.

Sherr, L., Croome, N., Clucas, C. & Brown, E. 2014. Differential effects of single and double parental death on child emotional functioning and daily life in South Africa. Child Welfare. 93(1):149-72.

Singh, A. & Sekher, T.V. 2021. Orphans and their living arrangement in Indian households: Understanding their educational and nutritional status. *Children and Youth Services Review*. 21

Skinner, D., Tsheko, N., Mtero-Munyati, S., Segwabe, M., Chibatamoto, P., Mfecane, S., Chandiwana, B., Nkomo, N., et al. 2006. Towards a definition of orphaned and vulnerable children. *AIDS and Behavior*. 10(6):619–626.

Spears, D. 2013. *How Much International Variation in Child Height Can Sanitation Explain?* (6351). [Online], Available: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2212559%0Ahttp://nutritionatthecenter.care2share.wikispaces.net/file/view/Spears et al_Intl Variation in Child Height %26 Sanitation.pdf/480266206/Spears et al_Intl Variation in Child Height %26 Sanitatio.

StatsSA. 2011. *Statistics SA Data*. [Online], Available: http://superweb.statssa.gov.za/.

StatsSA. 2016. *Community Survey 2016 Statistical Release.* Pretoria. [Online] Available: http://cs2016.statssa.gov.za/wp-content/uploads/2016/07/NT-30-06-2016-RELEASE-for-CS-2016-_Statistical-releas_1-July-2016.pdf

StatsSA. 2019. *Inequality trends in South Africa: A multidimensional diagnostic of inequality*. Pretoria. [Online], Available: http://www.geocurrents.info/economic-geography/inequality-trends-in-south-africa.

Sutherland, C., Hordijk, M., Lewis, B., Meyer, C. & Buthelezi, S. 2014. Water and sanitation provision in eThekwini Municipality: a spatially differentiated approach. *Environment and Urbanization*. 26(2):469–488.

Tobler, W.R. 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*. 46(June):234–240.

Victora, C.G., Adair, L., Fall, C., Hallal, P.C., Martorell, R., Richter, L. & Sachdev, H.S. 2008. Maternal and child undernutrition: consequences for adult health and human capital. *The Lancet*. 371:340–357.

von Fintel, D.P. 2007. Dealing with earnings bracket responses in household surveys – how sharp are midpoint imputations. *South African Journal of Economics*. 75:293-312

von Fintel, D.P. 2018. Long-run spatial inequality in South Africa: early settlement patterns and separate development. *Studies in Economics and Econometrics*. 42(2):81-102.

von Fintel, D.P. & Fourie, J. 2019. The great divergence in South Africa: population and wealth dynamics over two centuries. *Journal of Comparative Economics*. 47(4):759-773.

WHO. 1995. *Physical status: the use of and interpretation of anthropometry, report of a WHO expert committee.* Geneva.

Wittenberg, M. & Leibbrandt, M. 2017. Measuring Inequality by Asset Indices: A General Approach with Application to South Africa. *Review of Income and Wealth*. 63(4):706–730.

Xu, J. & Long, J.S. 2005. Confidence intervals for predicted outcomes in regression models for categorical outcomes. *Stata Journal*. 5(4):537–559.

Zere, E. & McIntyre, D. 2003. Inequities in under-five child malnutrition in South Africa. *International Journal for Equity in Health*. 2:1–10.

Appendices

Appendix A: robustness tests of OLS results

Table A1: Baseline OLS regression on the prevalence of stunting in 2013

	(1)	(2)	(3)	(4)	(5)	(6)
Proportion Double Orphans	1.665***	0.637**	0.373	0.389	1.606***	0.274
	(0.353)	(0.284)	(0.248)	(0.244)	(0.343)	(0.225)
Asset Index		-0.107***				-0.026***
		(0.01)				(0.009)
Piped Water#			-0.035***	-0.031***		-0.028***
			(0.007)	(0.008)		(0.008)
Flush Toilet#			-0.048***	-0.038***		-0.030***
			(0.008)	(0.007)		(0.008)
Urban				-0.012		-0.015**
				(0.007)		(0.007)
ECD Centres					-0.006***	-0.004***
					(0.001)	(0.001)
Police Departments					-0.005**	-0.003**
					(0.002)	(0.002)
Health Facilities					-0.001	0.000
					(0.001)	(0.001)
SASSA Offices					0.003	0.009***
					(0.002)	(0.002)
DSD Facilities					-0.000	0.006***
					(0.002)	(0.002)
Constant	0.266***	0.311***	0.329***	0.327***	0.277***	0.337***
	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)
R-squared	0.077	0.207	0.340	0.342	0.099	0.361
N	4275	4275	4275	4275	4275	4275
Moran Chi Squared	80129.85	61761.37	50561.18	50755.78	76212.88	47827.93
	(P=0,000)	(P=0,000)	(P=0,000)	(P=0,000)	(P=0,00)	(P=0,000)

NOTES: * $p<0.1$, ** $p<0.05$, *** $p<0.01$, #Access to flush toilets and piped water enter the equation as two separate variables. When the phrase "sanitation" is used below, this refers to flush toilets only. Inverse hyperbolic sine transformation of soft services applied. Standard errors are clustered at local municipal level. Source: own calculations from OVCSA data.

Interval Regression

Due to the fact that the stunting prevalence in each area in 2013 is an estimated variable, the confidence intervals calculated by an OLS model may over-estimate the certainty of the statistical relationships under scrutiny (Xu & Long, 2005). In order to overcome this issue, it is necessary to use an estimates parameters, knowing that the dependent variable lies within a certain interval (von Fintel, 2007). In order to do so, it is possible to run an interval regression using the upper and lower bound estimates of stunting prevalence which were generated by Osgood-Zimmerman *et al.* (2018). The unobserved value of stunting prevalence could lie between the mean and the upper-bound estimate, it could lie between the mean and the lower-bound estimate, or it could lie somewhere between the lower and upper bound estimates other than the mean. Using an interval regression, it is possible to estimate the same baseline model presented in section 6.1, taking this uncertainty into account when calculating the confidence intervals of the coefficients in question.

Table A2: Interval regression on the Upper and Mean estimate of Stunting in 2013.

	(1.1)	(1.2)	(1.3)	(1.4)	(1.5)
Proportion double orphans	1,794***	0,703**	0,445*	1,729***	0,314
	-(0,375)	(0,305)	(0,259)	(0,363)	(0,239)
Asset Index		-0,113***			-0,029***
		(0,011)			(0,009)
Piped Water			-0,033***		-0,030***
			(0,009)		(0,008)
Flush toilet			-0,041***		-0,031***
			(0,008)		(0,008)
Urban			-0,012		-0,015*
			(0,009)		(0,008)
ECD Centres				-0,007***	-0,005***
				(0,001)	(0,001)
Police Departments				-0,005**	-0,003*
				(0,002)	(0,002)
Health Facilities				-0,001	0
				(0,001)	(0,001)
SASSA Offices				0,003	0,011***
				(0,003)	(0,003)
DSD Centers				0	0,006***
				(0,002)	(0,002)
Constant	0,301***	0,348***	0,366***	0,313***	0,376***
	(0,008)	(0,008)	(0,007)	(0,007)	(0,006)
Sigma	-3.052***	-3.146***	-3.257***	-3.757***	-4.211***
N	4275.000	4275.000	4275.000	27.000	27.000

NOTES: * p<0.1, ** p<0.05, *** p<0.01, #Access to flush toilets and piped water enter the equation as two separate variables. When the phrase "sanitation" is used below, this refers to flush toilets only. Inverse hyperbolic sine transformation of soft services applied. Standard errors are clustered at local municipal level. Source: own calculations from OVCSA data.

Table A1 and A2 below show the results from the interval regression of the baseline OLS model presented in section 6.1. When comparing these interval regressions to the OLS model, the relative size of the coefficients, as well as the pattern of statistical significance is identical, or the interval regressions calculate a greater level of statistical significance. The only deviation is that the β_1 coefficient in column 2.3 remains significant. These results show that the estimates presented using the OLS estimation in section 6.1 is robust to the uncertainty of the stunting prevalence estimates, and does not present a misleading level of certainty, even though the dependent variable is an estimated value.

Table A2: Interval regression on the Mean and Lower estimate of Stunting in 2013.

	(2.1)	(2.2)	(2.3)	(2.4)	(2.5)
Proportion double orphans	1,601***	0,699***	0,473**	1,550***	0,369*
	(0,314)	(0,254)	(0,213)	(0,304)	(0,197)
Asset Index		-0,094***			-0,023***
		(0,009)			(0,007)
Piped Water			-0,029***		-0,026***
			(0,007)		(0,007)
Flush toilet			-0,034***		-0,027***
			(0,006)		(0,007)
Urban			-0,01		-0,012*
			(0,007)		(0,006)
ECD Centres				-0,005***	-0,004***
				(0,001)	(0,001)
Police Departments				-0,004**	-0,003**
				(0,002)	(0,001)
Health Facilities				-0,001	0
				(0,001)	(0,001)
SASSA Offices				0,003	0,009***
				(0,002)	(0,002)
DSD Centers				0	0,005***
				(0,002)	(0,001)
Constant	0,231***	0,270***	0,285***	0,240***	0,294***
	(0,006)	(0,007)	(0,005)	(0,006)	(0,005)
Sigma	-3.262***	-3.360***	-3.482***	-3.808***	-4.251***
N	4275.000	4275.000	4275.000	27.000	27.000

NOTES: * p<0.1, ** p<0.05, *** p<0.01, #Access to flush toilets and piped water enter the equation as two separate variables. When the phrase "sanitation" is used below, this refers to flush toilets only. Inverse hyperbolic sine transformation of soft services applied. Standard errors are clustered at local municipal level. Source: own calculations from OVCSA data.

Fractional Probit

Table A3: Marginal effects estimated after fractional probit regression on the Mean Prevalence of Stunting in 2013.

	(1)	(2)	(3)	(4)	(5)
Proportion double orphans	4,844***	1,816***	1,144***	4,676***	0,801***
	(0,313)	(0,267)	(0,236)	(0,312)	(0,238)
Asset Index		-0,325***			-0,081***
		(0,011)			(0,013)
Piped Water			-0,088***		-0,077***
			(0,009)		(0,009)
Flush Toilet			-0,115***		-0,089***
			(0,010)		(0,011)
Urban			-0,036***		-0,044***
			(0,009)		(0,009)
ECD Centers				-0,018***	-0,013***
				(0,002)	(0,002)
Police Departments				-0,015***	-0,010**
				(0,006)	(0,005)
Health Facilities				-0,002	0,001
				(0,003)	(0,002)
SASSA Offices				0,008	0,029***
				(0,009)	(0,008)
DSD Centers				-0,001	0,017***
				(0,007)	(0,006)
Constant	-0,624***	-0,491***	-0,447***	-0,593***	-0,419***
	(0,004)	(0,006)	(0,006)	(0,006)	(0,007)
N	4276	4276	4276	4276	4276

NOTES: * p<0.1, ** p<0.05, *** p<0.01. Inverse hyperbolic sine transformation of soft services applied.

Source: own calculations from OVCSA data

It is also necessary to establish the robustness of this data to the fact that the mean stunting prevalence is a proportion bound between zero and one. This type of data often displays an asymmetrical distribution around the mean (non-normal distribution) and is vulnerable to heteroscedasticity in the errors. Additionally, when fitting a linear estimator to this data, fitted values often exceed the unit interval. To overcome this problem, Papke and

Wooldridge (1996) suggest using a nonlinear link function $G(.)$ on the conditional mean of the fractional outcome variable $E\ (y \mid x)$ to ensure that predictions lie between zero and one. The link function $G(.)$ can be any known function which satisfies $0 < G(z) < 1$, and propose a quasi-maximum likelihood estimation technique (Papke & Wooldridge, 1996). For the purpose of the model presented here, a probit will be used as a link function.

Table A3 presents the baseline model given in section 6.1 which has been estimated using a fractional probit. Since a link function has been used to predict the model, it is unwise to interpret the values of the coefficients themselves. Therefore, the average individual marginal effects for each control variable is given in Table A3. Note that the estimation technique does not allow for clustered standard errors. As with the interval regression, the fractional probit suggests that the results of the OLS model are robust to the fractional nature of the outcome variable.

Appendix C: Data in Brief

The purpose of this appendix is to more fully describe the data used in this analysis, which is briefly outlined in section 3. As in section 3, this dataset will be referred to as the data on Orphaned and Vulnerable Children in South Africa (OVCSA) for the remainder of this discussion. The unit of observation is the electoral ward of South Africa, of which there are 4277[15]. The sources which have been used to create the OCVSA dataset include the 2011 South African Census, the 2013/14 Audit of ECD facilities in South Africa, a comprehensive geospatial estimation of CGF in South Africa by Osgood-Zimmerman *et al.,* (2018), and multiple publicly available lists of government facilities. The government facilities include health facilities, police stations, South African Social Security Agency (SASSA) offices and Department of Social Development (DSD) service points. The GIS data collection technique allows for these multiple sources to be linked in order to allow for a unique spatial analysis of orphanhood, child health outcomes and service provision, as well as the inequalities in these measures.

The remainder of this appendix describes the data and is structured as follows. The following section describes the child health measures which where estimated by Osgood-Zimmerman *et al.* (2018)[16]. Section 3 details the data taken from the South African Census of 2011, while section 4 describes data taken from all other sources. Section 5 briefly discusses issues regarding weighting and analysis, and section 6 concludes briefly by outlining the value of this data and possibilities for further investigations.

Data on Child Growth Failure

The data on child growth failure (CGF) has been taken from a publication in *Nature* by Osgood-Zimmerman *et al.* (2018). This data includes estimates of the prevalence of stunting, wasting and underweight in children under 5 for the years 2000 to 2018 gridded at 5km x 5km points across Sub-Saharan Africa. For the purpose of the data on OVCSA, stunting, wasting and underweight prevalence at the 5km x 5km level was averaged in each electoral ward of South Africa in order to create a variable which measures the mean stunting, wasting and underweight prevalence, each at the electoral ward level.

[15] There are 4277 electoral wards in South Africa, but information was not recorded in the 2011 census in one ward. Therefore, there are 4276 observations in this dataset.

[16] This data is open access and has been made available by the authors for free download.

Importantly, the measures of CGF are created using the Demographic Household Survey (DHS) data for Sub-Saharan Africa supplemented by South Africa's National Income Dynamics Study (NIDS). In order to produce CGF estimates for South Africa, individual level height, weight and age data for children under 5 was extracted from the household survey data. Next, the micro-data were collapsed to the level of clustering, or in the case of the DHS, the areal-level of moderate stunting, wasting and underweight. In the Bayesian geo-statistical method used to generate the estimates, the centroids of each areal-level polygon are taken as the geolocations, and the point estimates are down-weighted in the following likelihood estimations in order to account for uncertainty in the precise location of the observation. Finally, the areal data were "resampled to 10,000 coordinate locations per areal observation using a population-weighted sampling scheme over the relevant area" (Osgood-Zimmerman *et al.*, 2018:48).

It should be noted that the final measures obtained are estimates, and are not true prevalence statistics, which implies that there is an added level of uncertainty in the data analysis. Even so, Said-Mohamed *et al.*, (2015) include 50 different studies and multiple datasets in a systematics review of stunting trends in South Africa, and report stunting levels which are comparable to those used in the present analysis. Specifically, Said-Mohamed *et al.*, (2015:7) report a national stunting prevalence of 26,9% in boys and 25,9% in girls aged 0 to 3 in 2013. Additionally, the stunting prevalence reported by Said-Mohamed *et al.* (2015) at the provincial level is consistent with those estimated by Osgood-Zimmerman *et al.* (2018).

Data from the South African Census of 2011

Multiple variables in the OVCSA data are taken from the 2011 Census. Each of these variables is measured at the ward level, and was obtained using the full Census of 2011 via the SuperCross website (StatsSA, 2011). While there are more variables than can be listed here, Table C1 lists variables of interest.

In general, each variable described here is included as a frequency of households or persons in each ward with the respective characteristics. Where a variable is used in analysis, it has also been converted to a proportion of households or persons in each ward by dividing by the total number of persons or households in each ward. For example, to obtain the proportion of children between 0 and 4 years who are double orphans in each ward, the number of children aged 0 to 4 for whom it was reported that neither parent is alive is divided by the total number of children aged 0 to 4. Similarly, to obtain the proportion of households with access to sanitation the number of households who reported having access to a flush toilet is divided by the total number of households in each ward[17].

Table C1: Description of variables from South Africa Census of 2011.

	Variable description	Weighting unit
Geographic data	Ward	
	Municipality	
	Province	
	Level of urbanisation	
Demographic variables	Number of households living in different dwelling types (house, shack etc.)	Household
	Number of households with male/female household head	Household
	Number of orphaned children (both single and double orphans)	Person
	Total population of ward	Person
	Total population in ward in 5 year age-cohorts	Person
	Total number of households in ward	Household
	Total number of households with age of household head in 5 year age cohorts	Household
	Total number of people of different population groups	Person
	Number of households of different sizes	Household
	Assets owned by households	Household
	Number of people who speak each South African language	Person
	Number of people with each level of education (eg. Number who completed Matric)	Person
	Constructed asset index	Household

The final variable which will be discussed is the asset index which was constructed to control for the average level of private household asset ownership. This asset index was constructed in multiple steps. First, 128 combinations of household asset ownership

[17] The toilet could either be attached to the municipal waste system, or a septic tank.

were created, by recording the proportion of households in each electoral ward who owned any combinations of a fridge, electric/gas stove, washing machine, computer, telephone, TV and/or a radio. This created a [4267 x 128] ward by asset combination matrix. Second, each of these 128 combinations received a wealth ranking, where owning all of the assets was given a weight of one, and owning zero assets received a weight of zero. It was assumed that owning more assets made a household wealthier. This produced a [128 x 1] weight vector. Finally, the dot product of the matrix with proportions and the weighing vector was taken to produce an asset index which varies between 1 and 0, and which is monotonically increasing with average private asset ownership in the ward.

This method necessarily takes an average of private asset ownership per ward. Even so, the Gini coefficient for this index across wards is 0,34. This is lower than the national Gini coefficient of private asset ownership across households for South Africa in 2011 (StatsSA, 2019), as well as the Gini coefficients for other indices estimated with South African household survey data (Wittenberg & Leibbrandt, 2017). Even so, this index still retains much of the variation in private asset ownership. This type of asset index has been used to approximate household wealth in other developing countries using similar data (Harttgen & Vollmer, 2011), and has also been used in the context of controlling for wealth when measuring child health outcomes (Sahn & Stifel, 2000).

As a robustness check, an asset index was generated using the 10% individual level sample of the 2011 census using principal component analysis (PCA). The same assets named above where analysed in order to predict an asset index for individuals, as well as weights for the relative importance of each asset. Due to the fact that there is no ward-level data available in the 10% sample, it was not possible to compare the geo-index which had been created. Therefore, the principal component weights where used to impute weights for each of the 128 combinations of assets described above. This weighting vector was then used to create an asset index at the electoral ward level. This more data-driven approach produced an index with a Gini coefficient of 0,24. Even so, the asset index created with the PCA weighing vector, and that created with the approach above are strongly correlated (correlation coefficient = 0,98). Given that the initial index captures greater variation in household asset ownership across wards (evidenced by the higher Gini coefficient) it has been used as the control for assets in the analysis.

The value of the data and its possible further use

The geographic nature of the data allows for analyses which can inform localised policy, such as the policy implications of this paper. Additionally, there are few datasets such as this which pull together multiple household surveys in South Africa by averaging across small geographic units. Therefore, this dataset also contributes to data techniques which South African researchers have at their disposal in analysing household surveys.

By way of concluding this section, it should be noted that there are many possible areas for further analysis of this data. Most obviously, the welfare outcomes of maternal and paternal orphans in South Africa could be analysed. Secondly, it is possible to exploit other measures of CGF (the prevalence of wasting and underweight) to elucidate further the health outcomes of children in South Africa. Finally, this data is well suited to answer issues which are of concern to South African social policy, especially those with a geographic dimension such as infrastructure provision.