UNITED NATIONS
UNIVERSITY
UNU-WIDER

WIDER Working Paper 2020/154

Orphans and stunted growth

Investigating the potential of spatial network effects in service delivery for reducing stunting in orphans

Grace Bridgman*

November 2020

Abstract: Stunted growth in early life has serious implications for children and is a well-established constraint to productivity, life expectancy, and cognitive development. This paper evaluates the relative contributions of household resources and public service delivery in reducing the orphan-stunting penalty—the higher likelihood of orphans to be stunted. The results indicate that the effects of access to sanitation significantly reduce this penalty. Moreover, sanitation is at least as important as private wealth in reducing the penalty while access to other services is not. The results indicate that co-ordination of services and economies of scale, as well as regional herd immunity in stopping diarrhoea among children, could play an important role in providing lasting solutions to stunting more generally and orphan vulnerability in particular. This evidence is crucial as a case study in understanding how best to care for vulnerable children in South Africa but also for developing countries.

Key words: orphan, vulnerable children, stunting, sanitation, South Africa

JEL classification: I14, I15, I38, C21

Acknowledgements: I would like to thank Dieter von Fintel, my supervisor, for excellent comments, guidance, and encouragement in writing this paper. I would also like to thank Dirk Brand, Cobus Burger, and Jonathan Hulley for assistance in creating the dataset used in this analysis.

---

* Stellenbosch University, Stellenbosch, South Africa; gmmbridgman@gmail.com

This study has been prepared within the UNU-WIDER project Southern Africa—Towards Inclusive Economic Development (SA-TIED).

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9256-911-2

https://doi.org/10.35188/UNU-WIDER/2020/911-2

Typescript prepared by Mary Lukkonen.

United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Finland, Sweden, and the United Kingdom as well as earmarked contributions for specific projects from a variety of donors.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

1 Introduction

Stunted growth in early life is a well-established constraint to productivity and life expectancy and has serious implications for cognitive development in children (May and Timæus 2014; Zere and McIntyre 2003). Internationally, the evidence that *orphans* are more likely to be stunted is mixed, indicating that in some countries orphans have adequate support that keeps them from falling behind (Ali et al. 2018; Finlay et al. 2016), but more evidence is required to properly understand these support mechanisms. Studies from developing countries such as Ghana, Indonesia, Guatemala, and Bangladesh (Beal et al. 2018; Darteh, Acquah, and Kumi-Kyereme 2014; Dewey and Begum 2011; Quisumbing and Maluccio 2003), as well as regional studies from South Asia, Sub-Saharan Africa, and Latin America (Aguayo and Menon 2016; Bernal et al. 2014), document that orphans face more general socio-economic disadvantages compared to children who live with their parents. In South Africa, orphaned children are more vulnerable to lower educational outcomes (Ardington and Leibbrandt 2010); however, there is insufficient evidence on the vulnerability of orphans to lower health outcomes—and in particular to stunting.

Where orphans experience a higher propensity to exhibit stunted growth, our understanding of the mechanisms that led to this situation, and the factors that could mitigate the problem, remains limited. While household living arrangements—such as the presence of caregivers from extended families—can reduce orphans' vulnerability to stunting (Ali et al. 2018; Finlay et al. 2016), the role of public services in supporting orphans has not yet been investigated. While increased access to adequate sanitation plays a significant role in reducing the likelihood of stunting in children more generally, there is no evidence that shows whether this could also help in reducing nutrition gaps between orphans and other children. If orphaned children tend to live in regions with poor service delivery, it potentially amplifies the vulnerabilities caused by losing a parent.

In South Africa, orphaned children tend to live in poorer areas of the country with lower levels of access to sanitation and other service delivery. The evidence suggests that children who live in these types of areas are more likely to experience chronic diarrheal illness, which can lead to stunting. This implies that orphans are more likely than other children to experience stunted growth. In order to investigate these relationships, a novel dataset has been constructed that allows for the estimation of the effects of service provision on child health outcomes, specifically stunted growth, and investigates the orphan-stunting penalty—that is, the higher likelihood that orphaned children will experience stunted growth.

This paper evaluates the relative contributions of household resources and public service delivery in reducing the orphan-stunting penalty in South Africa. The results indicate that sanitation is at least as important as private wealth in reducing this penalty, while access to other services is not. Importantly, the results indicate that public infrastructure affects stunting with spatial network effects. In areas where there are spatial agglomerations of sanitation, the mitigating effects are stronger. Coordination of services and economies of scale over large regions, as well as regional herd immunity in stopping the spread of diarrhoeal illness among children, could play an important role in providing lasting solutions to stunting more generally and orphan vulnerability in particular. By contrast, proximity to other social services, such as health facilities, cash transfer payment centres, and early childhood development centres, does not alter the stunting profile of orphans and other children to the same extent.

The implication is that policy should increase the capacity of households to pursue better nutrition outcomes of (orphaned) children and that long-run health gains would more likely be achieved through prioritizing infrastructure investment rather than health and social services directly. This evidence is crucial as a case study in understanding how best to care for vulnerable children, such as orphans in South Africa, but also for developing countries in which caring for orphaned and vulnerable children (OVC) is a pertinent policy concern.

To begin, Section 2 briefly outlines the literature on OVC, stunting, and service delivery in South Africa and the context into which this research speaks. Section 3 provides a description of the novel dataset used in this analysis, and Section 4 goes on to provide descriptive statistics and spatial patterns as seen in the data. Section 5 presents the results of a baseline ordinary least squares (OLS) model and the Spatial Autocorrelation Regression (SAR) model used to arrive at the causal estimates of interest, and Section 6 concludes.

2 Context and literature review

2.1 Orphanhood in South Africa

This paper focuses on the increased likelihood of stunted growth in orphans, making it necessary to first review existing studies of welfare outcomes for orphans. Thereafter, it is also important to examine literature on stunted growth in children more specifically. Evidence from Sub-Saharan Africa suggests that orphaned children in the region are less likely to have access to health services and adequate nutrition and are considerably less likely to complete schooling than non-orphaned children (Lombe and Ochumbo 2008). While evidence from 49 lower- and middle-income countries finds that OVC are at a higher risk of adverse health outcomes (Finlay et al. 2016), other international studies suggest that OVC are no more likely to be stunted than other children once socio-economic status has been accounted for (Ali et al. 2018).

There are some international reviews on the importance of the type of living arrangements for orphans. Results suggest that the living arrangement is critical in predicting outcomes for orphans, where orphans are more likely to be better off if they live within an extended family network (Li, Chng, and Chu 2017). There is little evidence regarding the importance of public service provision for supporting orphans, however. The only study that indicates the importance of adequate access to services for orphaned children investigates the adverse effects of poorly resourced orphanages on child health. This study finds that this type of institution has a severe negative impact on orphans, both physically and cognitively (Carr, Duff, and Craddock 2018). Therefore, there is a need for more robust evidence on the role that state service provision may play in mitigating the negative effects of orphanhood on child health outcomes. This study works towards filling this gap and can be seen as a case study of the role that public service provision plays in reducing the orphan-stunting penalty.

There is a substantial but dated body of literature on OVC in South Africa that focuses largely on the increase in the prevalence of orphanhood caused by HIV/AIDS (Ardington and Leibbrandt 2010; Boler 2007; Richter and Desmond 2008). Between 2000 and 2010, a wealth of empirical investigation on the educational outcomes of OVC was undertaken and called to light both the growing crisis of orphanhood in South Africa (Chuong and Operario 2012; Meintjes et al. 2010; Parikh et al. 2007; Richter and Desmond 2008; Saloojee et al. 2007; Skinner et al. 2006) and the consistently lower welfare outcomes of OVC in comparison to non-orphans (Case and Ardington 2006; Case, Paxson, and Ableidinger 2004; Evans and Miguel 2007). In line with this literature, the dominant convention is adopted in this paper, defining orphans as follows:

children whose parents are both deceased are defined as double orphans, and children with only one living parent are defined as single orphans.[1]

Ardington and Leibbrandt (2010) investigated the extent to which double orphans are at risk of lower educational outcomes and found that OVC are more vulnerable to worse educational outcomes than non-orphans at every level of education. A further study that uses a similar dataset finds that orphaned children are no more likely to be behind in school than non-orphaned children once household characteristics such as socio-economic status, living arrangements, and household size are taken into consideration (Chuong and Operario 2012). Ardington and Leibbrandt (2010) also assert that these children are absorbed into extended family networks and especially into the households of grandparents.

The majority of research on OVC in South Africa has placed a focus on the increase in orphans in the country caused by HIV/AIDS and asserts that the number of orphans was expected to increase. It should be mentioned, however, that the present analysis does not distinguish between orphans who have lost their parents because of HIV/AIDS and those who have lost their parents for other reasons.

2.2 Stunting trends in South Africa

This analysis uses the mean prevalence of stunting in children under age 5 by electoral ward in South Africa as the primary outcome variable to measure health outcomes. By way of definition, stunting is used to quantify child growth failure (CGF), 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). Stunting is considered a consequence of chronic diarrheal illness and malnutrition (Delpeuch et al. 2000).

This measure has been used in the economic analysis of the state of health outcomes for young children and is seen to quantify vulnerability (Duflo 2000, 2003; Osgood-Zimmerman et al. 2018; Prendergast and Humphrey 2014; WHO 1995). However, it should be emphasized that food, or the lack thereof, is not the only primary need of children. Children are also in need of other forms of emotional support and cognitive stimulation. However, because this paper is confined to quantifiable data, this anthropomorphic measure will be used as a proxy for the wellbeing of children, with the understanding that meeting a child's physical needs forms only part of the care that a child needs.

Stunting is therefore an appropriate outcome measure for econometric analysis; however, the many known causes of stunted growth in children should also be considered when modelling the relationship between orphanhood and stunting. Well-established factors that contribute towards stunted growth include chronic malnutrition, low maternal education, low birth weight, and chronic illness caused by poor sanitation (Beal et al. 2018; Bernal et al. 2014; Darteh, Acquah, and Kumi-Kyereme 2014; Delpeuch et al. 2000; Zere and McIntyre 2003). While poor access to sanitation and malnutrition are both factors that are influenced by poverty, they each separately reduce the ability of a child to grow to their full potential (May and Timæus 2014).

This analysis will focus primarily on the link between stunting prevalence and access to sanitation for two reasons. The primary reason is because of the nature of the data used in this analysis,

[1] Children whose mother is deceased and whose father is known to be alive are defined as maternal orphans, and children whose father is deceased and whose mother is known to be alive are defined as paternal orphans.

where access to sanitation is more easily and reliably measured. Secondly, the provision of adequate sanitation and adequate nutrition for vulnerable children are both pressing policy concerns and deserve proper attention. Therefore, a brief discussion of service and sanitation provision is necessary before proceeding, as provided below.

2.3 Sanitation and public service provision in South Africa

The provision of sanitation, and service provision more generally, is fundamental in addressing inequality in South Africa, not least in addressing poor child health outcomes and stunting (Aguayo and Menon 2016; Beal et al. 2018; Bernal et al. 2014; Jinabhai, Taylor, and Sullivan 2006; May and Timæus 2014; Zere and McIntyre 2003). It is well understood that sanitation reduces stunting prevalence; however, this paper seeks to understand whether the lack of access to this public good can mitigate the orphan-stunting penalty as well.

Unfortunately, access to sanitation displays a distinct spatial dimension in which the geography of apartheid is still evident (Sutherland et al. 2014: 470), and equal access to services such as these is emphasized in the South African Constitution, the Reconstruction and Development Programme (RDP), and the National Development Plan (NDP) to address this. Considerable progress has been made in providing all South Africans with proper sanitation since 1994, but even so, the 2016 StatsSA Community Survey suggests that only 59 per cent of South African households have access to flush toilets, and 31 per cent of households still use pit latrines (StatsSA 2016).[2] Access to sanitation can also be seen as a distinct public service that is more personal than access to healthcare, for example. Additionally, while the provision of sanitation benefits individuals personally, its provision is a public good since it contributes to the overall removal of harmful matter in an area.

Therefore, proper sanitation is of great concern for the health of children. Lewin et al. (2007) estimate that the burden of disease in children under age five caused by inadequate sanitation and hygiene was 9.3 per cent in 2000. More recent studies estimate widespread lack of sanitation in some areas to be in the top three causes of death for children under age five in South Africa (Gray and Vawda 2018). Moreover, while many do not die from poor sanitation, early childhood stunting can still cause lasting damage. There are multiple studies that link lower-life outcomes with stunted growth in childhood, including increased morbidity, poor cognitive development and worse educational outcomes, increased risk of perinatal and neonatal death in stunted women, lower productivity, and reduced earning potential (Black et al. 2013; Dewey and Begum 2011; de Onis and Branca 2016; Victora et al. 2008).

In the NDP, the president of South Africa undertakes providing the 'hard infrastructure' required for proper sanitation for all South Africans. This promise is yet to materialize in some regions of the country, however. This fact is crucial in understanding patterns of stunting in vulnerable children and will be discussed further in the remainder of this paper.

3 Data in brief

This analysis uses a novel dataset that brings information together from multiple publicly available resources focused on OVC 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

[2] The remaining 10 per cent of households are supplied with either chemical toilets or 'other' sanitation.

South Africa (OVCSA). Because this dataset is novel, it is necessary to briefly highlight the nature of pertinent variables.

In compiling the data, each electoral ward in South Africa was taken as the unit of observation.[3] This GIS data collection technique has allowed for multiple sources to be linked to allow for a unique spatial analysis of orphanhood, child health outcomes and service provision, and the inequalities in these measures. The sources that have been used to create the OCVSA 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 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.

3.1 Measures of CGF: primary outcome variable of interest

The primary outcome variable used to measure health outcomes for OVC in South Africa is the mean stunting prevalence. This variable was acquired from datasets that were estimated by Osgood-Zimmerman et al. (2018).[4] These data estimate the prevalence of stunting, wasting, and underweight children under age five for the years 2000 to 2018 at the 5-km x 5-km level across Sub-Saharan Africa. For the purpose of this dataset, stunting prevalence at the 5-km x 5-km level was averaged in each electoral ward of South Africa to create a variable that measures the mean stunting prevalence at the electoral ward level.

Importantly, the measures of CGF are created using the Demographic Household Survey (DHS) data for Sub-Saharan Africa and apply Bayesian geo-statistical methods to estimate the prevalence of child stunting at the 5-km x 5-km level (Osgood-Zimmerman et al. 2018). To ensure anonymity, a random displacement vector is added to each DHS observation, where urban clusters are randomly displaced by 2 km, and rural clusters are displaced by up to 5 km. These displacements are constrained to remain within the district municipality, however. Because of this data feature, it is crucial to account for the random displacement vector as well as the Bayesian estimation technique when modelling the data. This feature of the outcome variable makes SAR modelling particularly appropriate.[5]

The measures used are estimates and are not true prevalence statistics, however, 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 systematic review of stunting trends in South Africa and report stunting levels that are comparable to those used in the present analysis. Specifically, Said-Mohamed et al. (2015: 7) report a national stunting prevalence of 26.9 per cent in boys and 25.9 per cent 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).

---

[3] There are 4,277 electoral wards in South Africa, but information was not recorded in the 2011 census in one ward. Therefore, there are 4,276 observations in this dataset.

[4] These data are open access and have been made available by the authors as a free download.

[5] SAR modelling is described in detail in Appendix C.

3.2 Data from other sources

While there are multiple variables in the OVCSA data taken from the 2011 census, there are three of particular importance: the proportion of children between 0 and 4 years who are double orphans in each ward; the proportion of households with access to sanitation; and the average private asset ownership in each ward. The first divides the number of children aged 0 to 4 for whom it was reported that neither parent is alive by the total number of children aged 0 to 4. The second divides the number of households who reported having access to a flush toilet by the total number of households in each ward.[6] A similar variable has also been constructed that measures the average access to piped water in a ward. Each of these variables are measured at the ward level and where obtained using the full census of 2011 via the SuperCross website (StatsSA 2011).

The asset index, which measures private household asset ownership, was constructed in multiple steps. First, 128 combinations of household asset ownership were created by recording the proportion of houses by electoral ward who owned a fridge, electric/gas stove, washing machine, computer, telephone, TV, and/or radio. This created a [4,267 x 128] 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 were taken to produce an asset index that 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), which was estimated to be as high as 0.57 in 2011 but 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 and Vollmer 2011) and in the context of controlling for wealth when measuring child health outcomes (Sahn and Stifel 2000).

The other pertinent variables in the following analysis include the number of ECD centres, health facilities (including clinics and hospitals), police stations, SASSA points, and DSD offices, hereafter referred to as soft services or soft infrastructure.[7] Table 1 presents descriptive statistics for the variables mentioned in this section. Importantly, there are no missing data in the OVCSA dataset.[8] While it would be ideal to use the log transformation of the variables that record the number of soft services per ward, it is not possible to do so because the log of zero is undefined. Therefore, to retain the observations where there are zero soft service centres in a ward, the inverse hyperbolic sine transformation is applied. This transformation allows the distribution of these variables to more closely resemble a normal distribution while retaining true-zero values in the dataset (Bellemare and Wichman 2020).

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

[7] There were 38 health facilities, eight police departments, and 760 ECD facilities that could not be mapped to an electoral ward because of missing data. This is 0.013 per cent, 0.007 per cent, and 4 per cent of observations, respectively.

[8] This is required for SAR modelling using a spatial weighting matrix.

Table 1: Descriptive statistics for pertinent variables

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.	4,276	0.01	0.008	0	0.096
Asset Index	The average private asset ownership in the ward.	4,276	0.31	0.19	0.008	0.936
Piped Water	The proportion of households with access to piped water in the ward.	4,276	0.71	0.31	0	1
Flush Toilets	The proportion of households with access to flush toilets in the ward.	4,276	0.48	0.41	0	1
ECD Centres*	The number of ECD centres (government and private) in the ward.	4,276	3.46	4.88	0	95
Health Facilities*	The number of health facilities (clinics and hospitals) in the ward.	4,276	1.45	1.84	0	22
Police Departments*	The number of police departments in the ward.	4,276	0.26	0.49	0	4
SASSA Centres*	The number of SASSA centres in the ward.	4,276	0.09	0.29	0	2
DSD Offices*	The number of DSD offices in the ward.	4,276	0.15	0.69	0	24

Note: *in the estimations in Section 6, these variables enter the model transformed by the inverse hyperbolic sine function.

Source: author's calculations based on data from the OVCSA dataset.

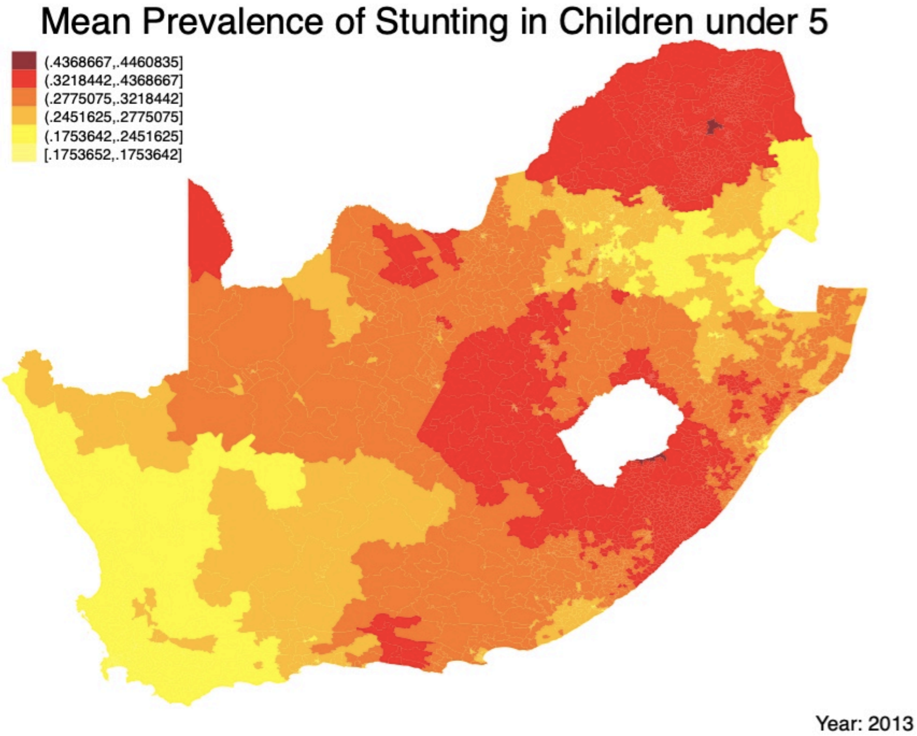
4 Descriptive statistics: spatial and bivariate analysis

4.1 Analysis of spatial distributions

Given that this analysis exploits the variation in spatial distribution of different characteristics across South Africa, it is both necessary and instructive to analyse the patterns of spatial distribution in key measures, specifically in mean stunting prevalence, the distribution of private assets and sanitation, and the distribution of orphans across South Africa. Figure 1 shows the mean prevalence of stunting in children under age 5 by electoral ward in South Africa. This distribution loosely follows municipal lines, especially in the northeast part of the country.

Figure 1 shows that stunting can be considered an extreme issue in some areas of the country, where as many as 45 per cent of children are stunted in some wards. Stunting is particularly prevalent in KwaZulu Natal, the North-West Province, and Limpopo, as well as in some areas of the Eastern Cape and the Northern Cape. The mean stunting prevalence across the entire country in 2013 was 28 per cent, which implies that over one-quarter of South African children under age 5 were stunted. However, if electoral wards in South Africa are divided into quintiles based on the proportion of double orphans in the ward, the mean stunting prevalence of the 855 wards in the top quintile is 31 per cent. This suggests that in wards with a higher prevalence of orphans, there is more stunting, and therefore, orphans may be more likely to experience stunting.

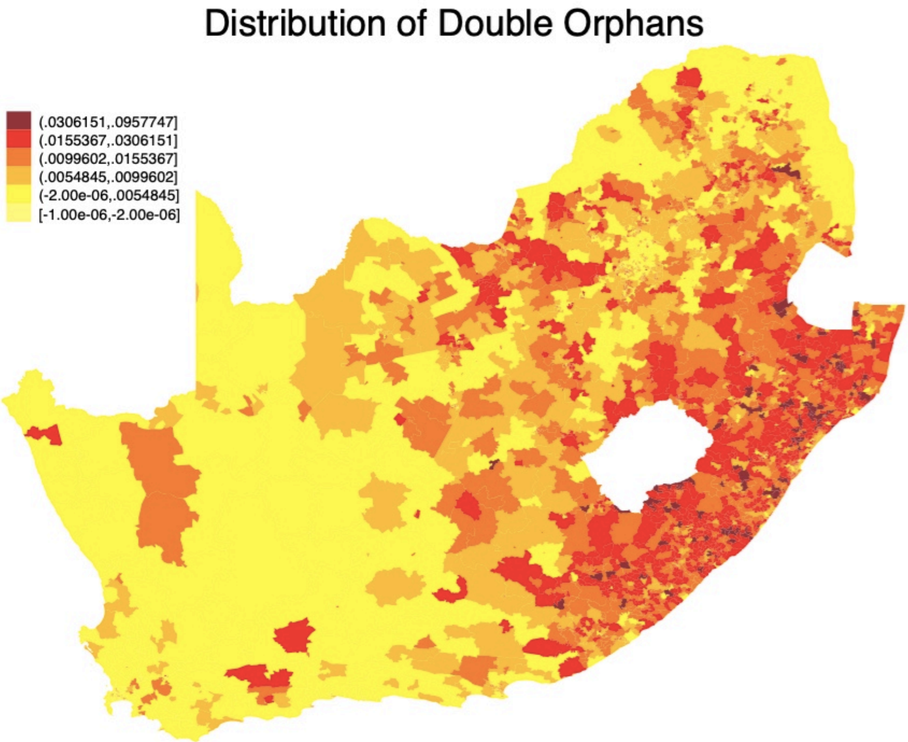
Figure 1: Mean stunting prevalence of all children under age 5 by electoral ward in South Africa



Source: author’s illustration based on data from the OVCSA dataset.

Figure 2 shows the distribution of double orphans under age 4 by electoral wards and demonstrates that this distribution follows that of the prevalence of stunting, except in the northeast part of the country. Stunting prevalence and the proportion of double orphans under age 4 have a correlation coefficient of 0.27, which is lower than may be expected from the distributions in Figures 1 and 2. While the proportion of double orphans does not exceed 10 per cent of children anywhere, it is important to note that these children are extremely young and that rates of orphanhood increase as children age. Therefore, the proportion of orphans in some places of more than 3 per cent can still be taken as a true indication of an area with a higher prevalence of vulnerable children.

Figure 2: Proportion of children aged 0 to 4 who are double orphans, by electoral ward



Source: author's illustration based on data from the OVCSA dataset.

The third spatial distribution that is produced in Figure 3 is the distribution of the constructed asset index across electoral wards in South Africa. Because this asset index, described in Section 4, necessarily averages across the average asset ownership in each ward, the distribution will be less varied than the true distribution of asset ownership across households in South Africa.

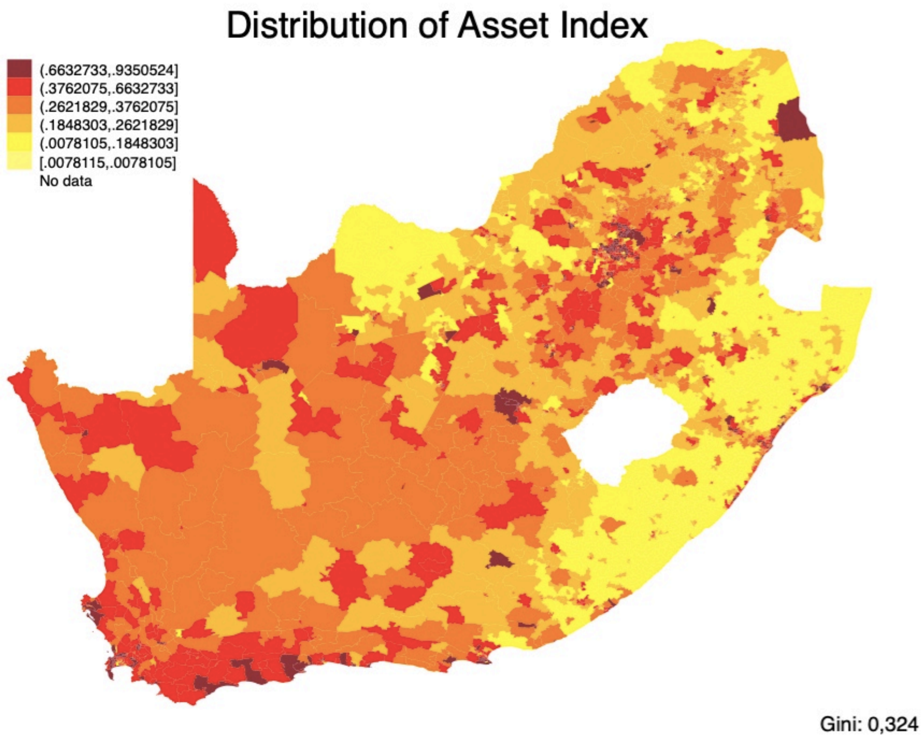
From a statistical perspective, fortunately, electoral wards in South Africa are small enough that meaningful variation in private asset ownership is still maintained. Moreover, this distribution of assets is consistent with other multidimensional analyses of poverty and inequality in South Africa (David et al. 2018). The pattern of asset ownership still maps closely the distribution that would be expected. Urban areas such as the municipality of Cape Town, Durban, and Johannesburg display greater levels of asset ownership, and there are pockets of higher asset ownership around Port Elizabeth, East London, and Bloemfontein. At the provincial level, the Western Cape and Gauteng have higher asset ownership, which is confirmed by Table 2. Lastly, it should be noted that asset ownership and stunting prevalence are strongly negatively correlated (correlation coefficient of -0.45), while mean asset ownership and the proportion of households with access to sanitation are strongly positively correlated (correlation coefficient of 0.68).

Table 2: Distribution of mean private asset ownership and sanitation by province

Province	**Private assets**		**Sanitation**	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**
Eastern Cape	0.21	0.16	0.33	0.39
Free State	0.34	0.12	0.71	0.32
Gauteng	0.44	0.22	0.86	0.21
KwaZulu Natal	0.24	0.17	0.3	0.36
Limpopo	0.27	0.11	0.17	0.31
Mpumalanga	0.33	0.16	0.46	0.4
North West	0.29	0.15	0.43	0.4
Northern Cape	0.35	0.15	0.67	0.31
Western Cape	0.47	0.21	0.9	0.12

Source: author's calculations based on data from the OVCSA dataset.

Figure 3: Spatial distribution of private asset ownership by electoral ward[9]



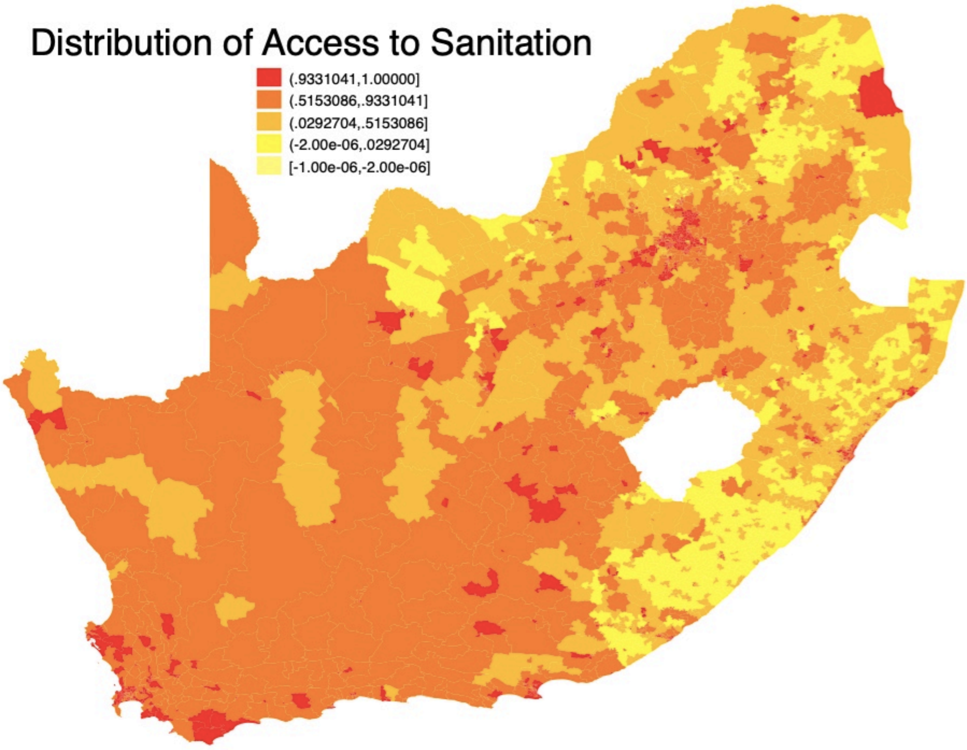
Source: author's illustration based on data from the OVCSA dataset.

The final map, Figure 4, shows the spatial distribution of the proportion of households with access to sanitation in the form of a flush toilet. It is immediately clear from this distribution that there are pockets of extremely low access to sanitation. As few as 3 per cent of households in some wards have a flush toilet. As discussed in Section 2.2, proper access to sanitation has important health consequences and is a matter of concern for the entire population of those wards without widespread access to sanitation. At a provincial level, Table 2 confirms that the

[9] The colour scale for this map is inverted in comparison to previous maps, where darker colours indicate that there is higher average private asset ownership.

average access to sanitation for households in Limpopo and KwaZulu Natal is as low as 17 per cent and 30 per cent, respectively, while that of the Western Cape and Gauteng are 90 per cent and 86 per cent, respectively. Lastly, access to sanitation and the prevalence of double orphanhood is negatively correlated at -0.36.

Figure 4: Spatial distribution of the proportion of houses with access to sanitation by electoral ward[10]



Source: author's illustration based on data from the OVCSA dataset.

4.2 Bivariate analysis

While there are multiple variables that enter the analysis in the following section, understanding three key variables and the relationships between them is crucial before moving on with any multivariate model. These variables are the mean stunting prevalence, the proportion of children who are double orphans, and access to sanitation in an electoral ward.

Figure 5 demonstrates that the mean stunting prevalence decreases systematically as the average access to sanitation in an area increases. There are two local polynomials drawn in this figure, where one demonstrates the relationship between stunting and sanitation for those areas with household asset ownership above the national average and the second shows that for areas with household asset ownership below the national average. This differentiation shows that stunting prevalence is higher for poorer wards at every level of access to sanitation, but access to sanitation dramatically reduces stunting prevalence in both wealthier and poorer wards. This figure suggests that further investigation into this relationship is warranted.

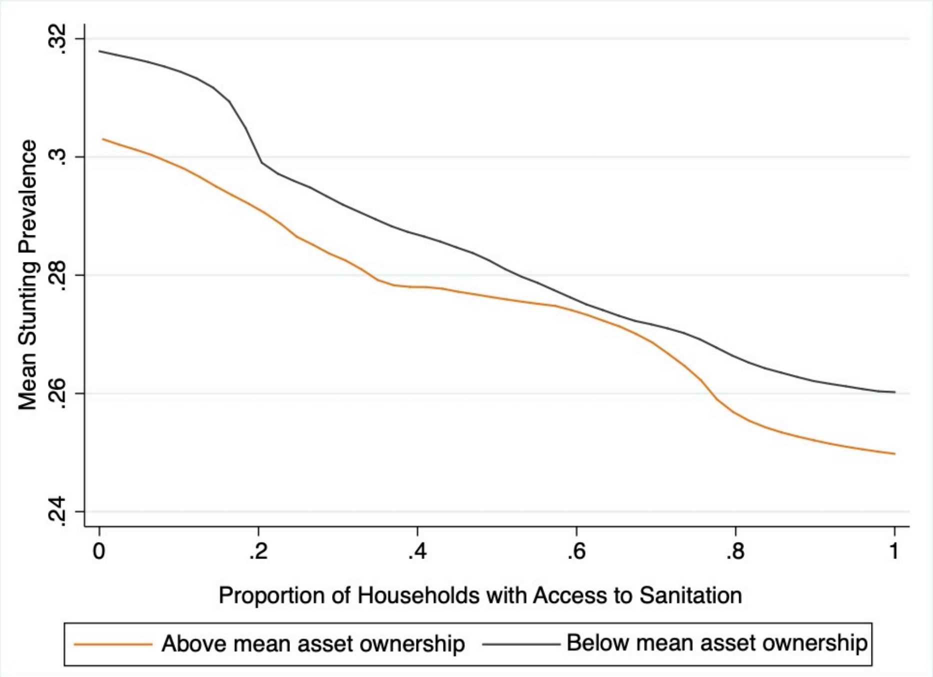
Figure 6 shows kernel density plots of stunting prevalence. In a similar fashion as Figure 5, this figure has two density plots: the first demonstrates the stunting prevalence for wards with

---

[10] The colour scale for this map is inverted, where darker colours indicate that there is higher access to sanitation.

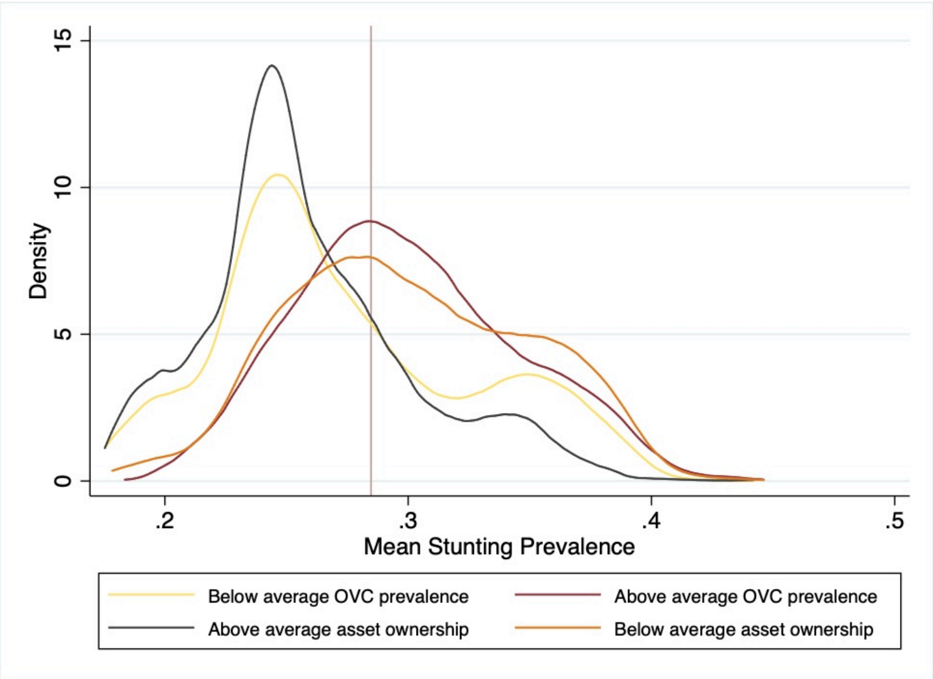
proportions of double orphans above the national average, and the second shows that figure for wards with proportions of double orphans below the national average. These plots show that stunting prevalence is consistently higher in areas where there are more double orphans.

Figure 5: Local polynomial showing the mean stunting prevalence as access to sanitation improves



Source: author's illustration based on data from the OVCSA dataset.

Figure 6: Kernel density plot of stunting prevalence with above and below average prevalence of double orphans



Source: author's illustration based on data from the OVCSA dataset.

Taken together, Figures 5 and 6 suggest that stunting prevalence is higher in areas where there is a higher concentration of double orphans and that stunting is dramatically reduced as the average access to sanitation in an area increases. With these relationships in mind, the following section takes the investigation further with a multivariate analysis and SAR modelling.[11]

5 Results

In each model below, the prevalence of stunting in children under age 5 in 2013 is used as the outcome variable, and the prevalence of double orphans aged 0 to 4 in 2011 is used as the measure of OVC in an area. Mean stunting prevalence in 2013 will be used as the outcome variable throughout the analysis for two reasons. First, the prevalence of double orphans in an area is taken for 2011, and it is therefore appropriate to use the stunting prevalence in 2013 to measure the outcomes for these children. Second, many of the other variables, including maps of soft infrastructure facilities, are taken from 2013. The analysis will begin with a baseline estimation and will proceed to estimate an SAR model that takes spatial patterns and spill-over effects into account.

5.1 Baseline model

To begin, this analysis will examine the following baseline model using OLS estimation:

$$Stunting_i = \beta_0 + \beta_1 OVC_i + \delta_i \boldsymbol{X_i} + u_i \tag{1}$$

where β_1 is the impact of the proportion of children who are double orphans on stunting prevalence[12], $\boldsymbol{X_i}$ is the matrix of area-specific controls added to the model, and δ_i is the [n x 1] vector of coefficients. In Table 3, column 1 shows that β_1 is large and significant. This relationship remains significant but becomes smaller when a control for the average level of asset ownership is added in column 2, suggesting that wealth plays a significant role in reducing stunting for both orphans and non-orphans. However, when controls for the proportion of households with piped water and a flush toilet are added in column 3, β_1 becomes statistically insignificant. This implies that proper sanitation and clean water are crucial in reducing stunting in children under age 5, as the literature also suggests (Beal et al. 2018; Bernal et al. 2014; Darteh, Acquah, and Kumi-Kyereme 2014; Delpeuch et al. 2000; Prendergast and Humphrey 2014).

In column 4, controls for the provision of soft services are added. While ECD and health facilities reduce stunting, the effect sizes are not economically significant. Furthermore, there is a considerable and significant β_1 coefficient in this model, suggesting that soft services do not specifically reduce the orphan-stunting penalty.

In the final specification in row 6, all the controls are included, and the β_1 coefficient is much reduced and statistically insignificant. The patterns in columns 2–4 suggest that access to piped water and proper sanitation play the most prominent role in this reduction and are the primary reason why the β_1 coefficient is no longer significant in column 6. Therefore, these results suggest that while the roll-out of improved access to soft services may be important in other regards, it does not play a large role in protecting orphans against stunting in particular. Rather,

[11] A description of SAR modelling is included in Appendix C.

[12] Also known as the orphan-stunting penalty.

this model indicates that hard infrastructure, such as piped water and proper sanitation, are crucial in reducing the orphan-stunting penalty.

Table 3: Baseline OLS regression on the mean prevalence of stunting in 2013

	(1)	(2)	(3)	(4)	(5)
Proportion Double Orphans	1.665***	0.637**	0.373	1.606***	0.269
Asset Index		-0.107***			-0.023**
Piped Water[#]			-0.035***		-0.033***
Flush Toilet[#]			-0.048***		-0.043***
ECD Centres				-0.006***	-0.004***
Police Departments				-0.005**	-0.003**
Health Facilities				-0.001	0.000
SASSA Offices				0.003	0.009***
DSD Facilities				-0.000	0.005***
Constant	0.266***	0.311***	0.329***	0.277***	0.338***
R-squared	0.077	0.207	0.340	0.099	0.358
N	4,275.000	4,275.000	4,275.000	4,275.000	4,275.000

Note: * 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, this refers to flush toilets only. Inverse hyperbolic sine transformation of soft services applied. Standard errors are clustered at the local municipal level.

Source: author's calculations based on data from the OVCSA dataset.

This baseline OLS regression is statistically hazardous for two reasons. First, the outcome variable of interest is an estimate in itself. This implies that there is a level of uncertainty in the exact value of the measure that needs to be accounted for when calculating standard errors. Second, the outcome variable is a fraction that varies between zero and one. 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, such as the β_1 coefficient in columns 1 and 4 (Ferrari and Cribari-Neto 2004).

To address the first issue, this baseline model is estimated using an interval regression; to address the second, a fractional probit is estimated in Appendix A. These models show that effect sizes and significance levels are similarly estimated using OLS, a fractional probit, and an interval regression model, indicating that the estimates presented using OLS are robust to the nature of the outcome variable.

5.2 SAR modelling: measuring causal effects

In order to address the bias that is introduced because of spatial autocorrelation, spatial autocorrelation regression modelling can be used. For a full description of this methodology, see Appendix C. The first SAR model that will be estimated is given by the following equation:

$$Stunt_{Prev_i} = \beta_0 + \beta_1 OVC_i + \delta_i \boldsymbol{X_i} + (I - \rho \boldsymbol{W})^{-1} u_{i,} \tag{2}$$

where similar parameters remain the same as in equation (1), ρ is the spatial autocorrelation coefficient, and $\boldsymbol{W}$ is the spatial weighting matrix. This model assumes that the stunting prevalence in one ward will affect the stunting prevalence in all other wards. Because of the specification of $\boldsymbol{W}$, the stunting prevalence in areas that are closer will have a larger spill-over effects. Models 1 and 2 in Table 4 show the total effects of ward characteristics, as estimated by the SAR model with a spatial lag, which was run using the -sp- package in Stata. Spatial lags refer to wards that neighbour each other, where a first-order spatial lag includes the immediate neighbours of ward *i,* and a second-order spatial lag includes immediate neighbours of ward *i* as well as the second layer of neighbours. That is to say, the SAR model allows for the characteristics of each ward *i* to impact ward *i* to a greater or lesser extent.[13]

The estimated value of the ρ coefficient for this specification is 1.86, significant at the 0.01-percent significance level. This implies that the model solution is explosive and that a shock in stunting prevalence in one area will be amplified by the spill-over effects across wards. This result has important implications for the interpretation of this model and will be discussed explicitly below. The results in panel 1 (a) and 2 (a) of Table 4 suggest that a higher proportion of orphans is associated with a significantly higher mean stunting prevalence in a ward, even once autocorrelation in the residuals is taken into account. These results are comparable to columns 1 and 2 in Table 3. Panel 2 (a) also implies that higher average asset ownership is significantly associated with a lower stunting prevalence, although the association between orphanhood and stunting measured in Table 4 is reduced by more than half. This suggests that accounting for spatially correlated residuals has significantly reduced the bias in this model.

While this first model specification accounts for autocorrelation in the residuals, spill-over effects from other ward characteristics have not yet been included. Therefore, the second SAR model that will be estimated is given by the following equation:

$$Stunt\_Prev_i = \beta_0 + \beta_1 OVC_i + \delta_i \boldsymbol{X_i} + \beta_2 \boldsymbol{W} \boldsymbol{X}_i + (I - \rho \boldsymbol{W})^{-1} u_{i,} \tag{3}$$

where similar parameters remain the same as in equation (2), and $\beta_2 \boldsymbol{W} \boldsymbol{X}_i$ is the measure of the spill-over effects of independent control variables on $Stunt\_Prev_i$. This model assumes that the control variable $\boldsymbol{X}_i$ will have a direct effect of δ_i and an indirect effect of $\beta_2 \boldsymbol{W}$ on $Stunt\_Prev_i$. Because of the recursive nature of the spill-over effects, it is necessary to explicitly estimate the direct effect, the spill-over or 'indirect' effect, and the total effects.[14]

Models 3, 4, and 5 in Tables 4 and 5 include two different specifications, labelled (a) and (b). The (a) specifications do not include indirect spill-over effects of the independent variables and can be taken as the control specification. The (b) specification differentiates between the direct effect of ward characteristics and the indirect spill-over effect. The effects reported in this table are derivatives, otherwise interpreted as the effects of a one-percentage-point change in $\boldsymbol{x_i}$ on $Stunt\_Prev_i$.

In model 3, panels 3 (a) and 3 (b) correspond with column 3 in Table 3 and have a similar interpretation: hard infrastructure in the form of improved access to piped water and flush toilets substantially reduces the orphan-stunting penalty. SAR model 3 suggests that while this form of

[13] See Appendix C for a fuller description of spatial lags and the weighting matrix used to determine the weighting of each neighbor on ward *i.*

[14] These regression outputs are included in Appendix B for completeness but will not be interpreted or referred to explicitly.

service delivery reduces the penalty, there is still a significant relationship between orphanhood and stunting present. Importantly, the indirect effects of increased access to sanitation is large and significant. This suggests that the spill-over effects of adequate sanitation are important to consider and call to light the benefit of large-scale increases in sanitation-improving health outcomes in general.[15]

This result implies that the network effects of sanitation are two-fold: in provision and in usage. First, network effects are present in the provision of sanitation in that where there is isolated access to sanitation, it will have as large an effect. The provision of sanitation must be part of a larger sanitation network that experiences economies of scale in providing adequate services. Second, network effects are present in the use of sanitation in that the more people in the area who have adequate access to sanitation, the less likely children will become chronically ill. Thus, diarrheal illness is less likely to spread in areas that are well serviced by sanitation networks and where people use the service effectively.

Model 4 in Table 4 indicates that soft infrastructure also has an effect that reduces the orphan-stunting penalty but to a lesser extent than hard infrastructure. In panels 4 (a) and 4 (b), the relationship between stunting prevalence and the proportion of double orphans in a ward is still large and significant (0.48 at the 0.01-per-cent significance level). Encouragingly, however, the relationship between adequate provision of these services and stunting is negative, except for having access to SASSA offices. This can be interpreted as suggesting that SASSA offices are well targeted to poorer wards.

The indirect effect of these soft services is larger than the direct effect, indicating that these facilities benefit more than just those who live nearby. This suggests that there are beneficial agglomeration effects that occur in the provision of these services, where the benefit of increased soft-service provision accrues to more than just the ward in which the facility is situated. Thus, where there is a larger network of public services, they are less likely to be over-crowded and would therefore be more able to provide a higher level of care than would be the case if they were over-crowded (Hörcher et al. 2020).

Model 5 in Table 5 includes controls for asset ownership, hard infrastructure, and all soft services and corresponds with column 5 of Table 3.[16] It is important to note that the orphan-stunting penalty is small and statistically indistinguishable from zero. This implies that once all area characteristics are accounted for, orphans are no more vulnerable to stunting than other children in the area. The individual effects remain consistent with the discussion above, but taken together, these results suggest that it is hard infrastructure that is most important for reducing the effects of vulnerability on children's health outcomes.

Specifically, these results unambiguously point towards the importance of sanitation, which is at least as important as asset ownership in reducing the orphan-stunting penalty. Model 5 presents results that indicate causal relationships, specifically that the effect of hard infrastructure is to significantly reduce stunting prevalence in areas where it is widely available. Moreover, the effects measured here ultimately imply that service provision and the network of facilities included successfully support orphans.

[15] See the note in Table 3 regarding the difference between flush toilets and piped water.

[16] Importantly, this model accounts for bias caused by an autocorrelation in the residuals as well as the spill-over effects of ward characteristics.

Positive agglomeration effects are seen in both hard and soft infrastructure in these models. Agglomeration effects can be defined as the benefit of spill-over effects that produce a larger marginal benefit than would otherwise be expected because of the spatial concentration of specific characteristics (Marshall 1920). There is a beneficial spill-over effect of greater access to hard and soft services in an area. Another indication of positive agglomeration effects is the ρ coefficient, which is larger than 1 in each of the specifications and increases substantially from model 1 to model 5. This suggests that the solution to this model is explosive in the sense that any shock to mean stunting prevalence will be amplified as the effects ripple over space. This implies that if stunting were to become worse in one ward, there would be a negative effects on stunting in the nearby areas and that the effects of worsened stunting would become stronger the further it spread from the initial shock. It should be assumed, however, that these effects do not occur indiscriminately, where the ripple effects are likely to be interrupted in areas with higher levels of access to sanitation. This implies that widespread and effective provision of local services, especially hard infrastructure, should be of the utmost concern for policy in South Africa. For example, if an adverse weather shock, such as flooding, were to cause a greater burden of diarrheal illness in ward *i*, it is likely that neighbouring wards will also become impacted by higher levels of diarrheal illness in children as children cross ward boundaries. It is likely that this pattern would continue until it is interrupted by an area with adequate sanitation services to properly mitigate the effects of a high burden of diarrheal illness in the area.

6 Conclusion

This analysis investigates the influence of public service provision in reducing the higher propensity of orphans to experience stunted growth relative to non-orphans. Orphans tend to live in areas with a high prevalence of stunting—a reality that is partially explained by the fact that orphans also tend to live in areas with lower average private wealth. From Tables 4 and 5, it is clear that greater access to public services, and specifically sanitation, is at least as important as private wealth in accounting for the orphan-stunting penalty. Soft services remain important in reducing stunting; however, these results point to the critical role that the hard infrastructure provision plays in mitigating the effects of orphanhood on health outcomes.

Moreover, positive spill-over effects of hard infrastructure call to light the importance of comprehensive and proper access to sanitation. Therefore, widespread provision of sanitation should be seen as a pressing policy concern which will work towards reducing stunting in South Africa, and specifically the orphan stunting penalty. This finding does not negate the importance of soft service provision but indicates that the direct effects of these services in reducing the orphan stunting penalty is small. The nature of hard infrastructure such as sanitation is pertinent in that it is a public good, and the benefits of widespread use of sanitation increase as more people use it. It is unlikely that any institution other than the state would be able to provide widespread access to its use, however.

These findings have critical implications for policy in South Africa, but also for developing countries with large populations of OVC. Thus, these results point to the fact that the widespread installation of adequate sanitation infrastructure is likely to significantly mitigate the orphan stunting penalty. While it is not feasible for policy makers to alter the private living arrangements of individuals, these results imply that it is possible for the government to tangibly address orphan stunting, thereby supporting some of the most vulnerable South Africans. To the author's knowledge, there is no other study which suggests such a tangible manner to support orphans, and to reduce the increased likelihood of poor health outcomes that they face due to their increased vulnerability. Finally, this analysis can be interpreted as a case study, where the findings and conclusions are likely to be applicable in other developing countries.

Table 4: Direct, indirect, and total effects of area characteristics on the mean stunting prevalence by electoral ward

	Model 1 Effects	**Model 2 Effects**	**Model 3 Effects**				**Model 4 Effects**			
	(a)	**(a)**	**(a)**		**(b)**		**(a)**		**(b)**	
	Total	**Total**	**Total**	**Direct**	**Indirect**[17]	**Total**	**Total**	**Direct**	**Indirect**	**Total**
Proportion Double Orphans	0.5***	0.28***	0.26***	0.17***		0.17***	0.48***	0.48***		0.48***
Asset Index		0.04***								
Piped Water			-0.03***	-0.02***	0.05**	0.03				
Flush Toilets			-0.01***	0.001	-0.25***	-0.25***				
ECD Facilities							-0.002***	-0.001***	-0.14***	-0.14***
Police Departments							-0.002	-0.003***	-0.2***	-0.2***
Health Facilities							0.000	0.000	-0.02**	-0.02*
SASSA Centres							-0.003**	-0.001	0.19***	0.2***
DSD Centres							-0.001	-0.003**	-0.26***	-0.26***
Chi2	30,932.36***	29,418.66***	54,718.50		16495.25***		30,195.85***		31,738.15***	
Rho Coefficient:	1.86***	1.86***	1.895106***		2.796763***		1.86***		2.64***	
Pseudo R Squared	0.0774	0.2073	0.3224		0.4789		0.0924		0.2775	

Note: * z<0.1, ** z<0.05, *** z<0.01. Row-normalized spatial weighting matrix applied. Inverse hyperbolic sine transformation of soft services applied.

Source: author's calculations based on data from the OVCSA dataset.

[17] The indirect effect should be interpreted as the average across-ward spill-over effect of a 1-percentage-point change in the x_i variable on $\overline{Stunt\ Prev}$.

Table 5: Direct, indirect, and total effects of area characteristics on the mean stunting prevalence by electoral ward

	Model 5 Effects			
	(a)	(b)		
	Total	Direct	Indirect	Total
Proportion Double Orphans	0.23***	0.07		0.07
Asset Index	-0.13***	-0.013***		-0.13***
Piped Water	-0.03***	-0.02***	0.07***	0.05**
Flush Toilets	-0.009***	0.005**	-0.19***	-0.19***
ECD Facilities	-0.002***	-0.001***	-0.07***	-0.07***
Police Departments	-0.001	-0.002**	-0.07***	-0.08***
Health Facilities	0.000	-0.000	-0.07***	-0.07***
SASSA Centres	0.000	-0.002**	-0.03	-0.03
DSD Centres	0.001	-0.001	-0.18***	-0.18***
Chi2	55,941.31***		5,695.48***	
Rho Coefficient:	1.9***		3.26***	
Pseudo R Squared	0.3360		0.5068	

Note: * z<0.1, ** z<0.05, *** z<0.01. Row-normalized spatial weighting matrix applied. Inverse hyperbolic sine transformation of soft services applied.

Source: author's calculations based on data from the OVCSA dataset.

References

Aguayo, V.M., and P. Menon (2016). 'Stop Stunting: Improving Child Feeding, Women's Nutrition and Household Sanitation in South Asia'. *Maternal and Child Nutrition,* 12(S1): 3–11. https://doi.org/10.1111/mcn.12283

Ali, Z., N. Abu, I.A. Ankamah, E.A. Gyinde, A.S. Seidu, and A.R. Abizari (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. https://doi.org/10.1186/s40795-018-0240-0

Anselin, L. (1999). 'Spatial Econometrics'. Working Paper Series. Dallas: University of Texas at Dallas.

Anselin, L., J. Le Gallo, and H. Jayet (2008). 'Spatial Panel Econometrics'. In L. Màtyàs and P. Sevestre (eds), *The Econometrics of Panel Data*. Berlin: Springer, 625–60. https://doi.org/10.1007/978-3-540-75892-1_19

Ardington, C., and M. Leibbrandt (2010). 'Orphanhood and Schooling in South Africa: Trends in the Vulnerability of Orphans between 1993 and 2005'. *Economic Development and Cultural Change,* 58(3): 507–36. https://doi.org/10.1086/650414

Beal, T., A. Tumilowicz, A. Sutrisna, D. Izwardy, and L.M. Neufeld (2018). 'A Review of Child Stunting Determinants in Indonesia'. *Maternal and Child Nutrition*, 14(4): 1–10. https://doi.org/10.1111/mcn.12617

Bellemare, M.F., and C.J. Wichman (2020). 'Elasticities and the Inverse Hyperbolic Sine Transformation'. *Oxford Bulletin of Economics and Statistics,* 82(1): 50–61. https://doi.org/10.1111/obes.12325

Bernal, J., E.A. Frongillo, A. Herrera, and J.A. Rivera (2014). 'Food Insecurity in Children But Not in Their Mothers Is Associated with Altered Activities'. *The Journal of Nutrition*, 144(24): 1619–26. https://doi.org/10.3945/jn.113.189985

Black, R.E., C.G. Victora, S.P. Walker, Z.A. Bhutta, P. Christian, M. De Onis, M. Ezzati, S. Grantham-Mcgregor, et al. (2013). 'Maternal and Child Undernutrition and Overweight in Low-Income and Middle-Income Countries'. *The Lancet*, 382(9890): 427–51. https://doi.org/10.1016/S0140-6736(13)60937-X

Boler, T. (2007). *Facing the Consequenses of AIDS: Orphans Educational Outcomes and Cash Grants in South Africa*. London: London School of Hygiene and Tropical Medicine.

Carr, A., H. Duff, and F. Craddock (2018). 'A Systematic Review of Reviews of the Outcome of Severe Neglect in Underresourced Childcare Institutions'. *Trauma, Violence, and Abuse,* 21(3): 484–97. https://doi.org/10.1177/1524838018777788

Case, A., and C. Ardington (2006). 'The Impact of Parental Death on School Outcomes: Longitudinal Evidence from South Africa'. *Demography*, 43(3): 401–20. https://doi.org/10.1353/dem.2006.0022

Case, A., C. Paxson, and J. Ableidinger (2004). 'Orphans in Africa: Parental Death, Poverty, and School Enrollment'. *Demography*, 41(3): 483–508. https://doi.org/10.1353/dem.2004.0019

Chuong, C., and D. Operario (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. https://doi.org/10.1080/17441692.2011.574147

Darteh, E.K.M., E. Acquah, and A. Kumi-Kyereme (2014). 'Correlates of Stunting among Children in Ghana'. *BMC Public Health,* 14(1). https://doi.org/10.1186/1471-2458-14-504

David, A., N. Guilbert, N. Hamaguchi, Y. Higashi, H. Hino, M. Leibbrandt, and M. Shifa (2018). *Spatial Poverty and Inequality in South Africa: A Municipality Level Analysis*. SALDRU Working Paper Number 221. Cape Town: SALDRU, UCT.

Delpeuch, F., P. Traissac, Y. Martin-Prével, J.P. Massamba, and B. Maire (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. https://doi.org/10.1017/S1368980000000069

Dewey, K.G., and K. Begum (2011). 'Long-Term Consequences of Stunting in Early Life'. *Maternal and Child Nutrition,* 7(s3): 5–18. https://doi.org/10.1111/j.1740-8709.2011.00349.x

Duflo, E. (2000). 'Child Health and Household Resources in South Africa: Evidence from the Old Age Pension Program'. *AEA Papers and Proceedings,* (May): 393–8. https://doi.org/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. https://doi.org/10.1093/wber/lhg013

Evans, D.K., and E. Miguel (2007). 'Orphans and Schooling in Africa: A Longitudinal Analysis'. *Demography*, 44(1): 35–57. https://doi.org/10.1353/dem.2007.0002

Ferrari, S.L.P., and F. Cribari-Neto (2004). 'Beta Regression for Modelling Rates and Proportions'. *Journal of Applied Statistics,* 31(7): 799–815. https://doi.org/10.1080/0266476042000214501

Finlay, J.E., G. Fink, D.C. McCoy, L.C. Tavárez, J. Chai, G. Danaei, M. Ezzati, W. Fawzi, 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): 784–90. https://doi.org/10.1136/jech-2015-206346

Gray, A., and Y. Vawda (2018). *South African Health Review 2018*. L.C. Rispel and A. Padarath (eds). Durban: Health Systems Trust.

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

Hörcher, D., B. De Borger, W. Seifu, and D.J. Graham (2020). 'Public Transport Provision Under Agglomeration Economies'. *Regional Science and Urban Economics*, 81(103503). https://doi.org/10.1016/j.regsciurbeco.2019.103503

Jinabhai, C.C., M. Taylor, and K. Sullivan (2006). 'Persistent Undernutrition Amongst Preschool Children a Decade After Democracy'. *Southern African Journal of Epidemiology and Infection,* 21(1): 26–30. https://doi.org/10.1080/10158782.2006.11441258

Lewin, S., R. Norman, N. Nannan, E. Thomas, and D. Bradshaw (2007). 'Estimating the Burden of Disease Attributable to Unsafe Water and Lack of Sanitation and Hygiene in South Africa in 2000'. *South African Medical Journal*, 97(7): 755–62.

Li, D., G.S. Chng, and C.M. Chu (2017). 'Comparing Long-Term Placement Outcomes of Residential and Family Foster Care: A Meta-Analysis'. *Trauma, Violence, and Abuse,* 20(5): 653–64. https://doi.org/10.1177/1524838017726427

Lombe, M., and A. Ochumbo (2008). 'Sub-Saharan Africa's Orphan Crisis: Challenges and Opportunities'. *International Social Work,* 51(5): 682–98. https://doi.org/10.1177/0020872808093345

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

May, J., and I.M. Timæus (2014). 'Inequities in Under-Five Child Nutritional Status in South Africa: What Progress Has Been Made?' *Development Southern Africa,* 31(6): 761–74. https://doi.org/10.1080/0376835X.2014.952896

Meintjes, H., K. Hall, D.H. Marera, and A. Boulle (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): 40–9. https://doi.org/10.1080/09540120903033029

de Onis, M., and F. Branca (2016). 'Childhood Stunting: A Global Perspective'. *Maternal and Child Nutrition*, 12(S1): 12–26. https://doi.org/10.1111/mcn.12231

Osgood-Zimmerman, A., A.I. Millear, R.W. Stubbs, C. Shields, B.V. Pickering, L. Earl, N. Graetz, D.K. Kinyoki, et al. (2018). 'Mapping Child Growth Failure in Africa Between 2000 and 2015'. *Nature,* 555(7694): 41–7. https://doi.org/10.1038/nature25760

Papke, L.E., and J.M. Wooldridge (1996). 'Econometric Methods for Fractional Response Variables with an Application to 401 (k) Plan Participation Rates'. *Journal of Applied Econometrics,* 11(6): 619–32. https://doi.org/10.1002/(SICI)1099-1255(199611)11:6<619::AID-JAE418>3.0.CO;2-1

Parikh, A., M.B. DeSilva, M. Cakwe, T. Quinlan, J.L. Simon, A. Skalicky, and T. Zhuwau (2007). 'Exploring the Cinderella Myth: Intrahousehold Differences in Child Wellbeing Between Orphans and Non-Orphans in Amajuba District, South Africa'. *Aids,* 21(SUPPL. 7). https://doi.org/10.1097/01.aids.0000300540.12849.86

Prendergast, A.J., and J.H. Humphrey (2014). 'The Stunting Syndrome in Developing Countries'. *Paediatrics and International Child Health,* 34(4): 250–65. https://doi.org/10.1179/2046905514Y.0000000158

Quisumbing, A.R., and J.A. Maluccio (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. https://doi.org/10.1111/1468-0084.t01-1-00052

Richter, L.M., and C. Desmond (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): 1019–28. https://doi.org/10.1080/09540120701842738

Sahn, D.E., and D.C. Stifel (2000). 'Poverty Comparisons Over Time and Across Countries in Africa'. *World Development*, 28(12): 2123–55. https://doi.org/10.1016/S0305-750X(00)00075-9

Said-Mohamed, R., L.K. Micklesfield, J.M. Pettifor, and S.A. Norris (2015). 'Has the Prevalence of Stunting in South African Children Changed in 40 Years? A Systematic Review'. *BMC Public Health,* 15(534). https://doi.org/10.1186/s12889-015-1844-9

Saloojee, H., T. De Maayer, M.L. Garenne, and K. Kahn (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. https://doi.org/10.1080%2F14034950701356435

Skinner, D., N. Tsheko, S. Mtero-Munyati, M. Segwabe, P. Chibatamoto, S. Mfecane, B. Chandiwana, N. Nkomo, et al. (2006). 'Towards a Definition of Orphaned and Vulnerable Children'. *AIDS and Behavior*, 10(6): 619–26. https://doi.org/10.1007/s10461-006-9086-6

StataCorp (2017). *Stata Spatial Autoregressive Models Reference Manual: Release 15*. College Station, Texas: StataCorp. Available at: https://www.stata.com/manuals/sp.pdf (accessed 19 November 2020).

StatsSA (2011). *Statistics SA Data*. Available at: https://superweb.statssa.gov.za/webapi/jsf/login.xhtml (accessed 19 November 2020).

StatsSA (2016). *Community Survey 2016*. Dataset. Pretoria: Statistics South Africa.

StatsSA (2019). 'Inequality Trends in South Africa: A Multidimensional Diagnostic of Inequality'. GeoCurrents. Available at: http://www.geocurrents.info/economic-geography/inequality-trends-in-south-africa (accessed 19 November 2020).

Stewart, M.B. (1983). 'On Least Squares Estimation When the Dependent Variable is Grouped'. *The Review of Economic Studies*, 50(4): 737. https://doi.org/10.2307/2297773

Sutherland, C., M. Hordijk, B. Lewis, C. Meyer, and S. Buthelezi (2014). 'Water and Sanitation Provision in eThekwini Municipality: A Spatially Differentiated Approach'. *Environment and Urbanization*. 26(2): 469–88. https://doi.org/10.1177/0956247814544871

Tobler, W.R. (1970). 'A Computer Movie Simulating Urban Growth in the Detroit Region'. *Economic Geography,* 46(June): 234–40. https://doi.org/10.2307/143141

Victora, C.G., L. Adair, C. Fall, P.C. Hallal, R. Martorell, L. Richter, and H.S. Sachdev (2008). 'Maternal and Child Undernutrition: Consequences for Adult Health and Human Capital'. *The Lancet,* 371(9609): 340–57. https://doi.org/10.1016/S0140-6736(07)61692-4

WHO (1995). *Physical Status: The Use Of and Interpretation of Anthropometry, Report of a WHO Expert Committee.* Geneva: WHO.

Xu, J., and J.S. Long (2005). 'Confidence Intervals for Predicted Outcomes in Regression Models for Categorical Outcomes'. *Stata Journal,* 5(4): 537–59. https://doi.org/10.1177/1536867X0500500405

Zere, E., and D. McIntyre (2003). 'Inequities in Under-Five Child Malnutrition in South Africa'. *International Journal for Equity in Health,* 2: 1–10. https://doi.org/10.1186/1475-9276-2-7

Appendix

A Robustness tests of OLS results

Interval regression

Because the mean stunting prevalence in each area in 2013 is an estimated variable, the confidence intervals calculated by an OLS model may overestimate the certainty of the statistical relationships under scrutiny (Xu and Long 2005). To overcome this issue, it is necessary to use a technique that can estimate the linear parameter using the observed value of the dependent variable, while the true value of the dependent variable remains unobserved but is known to lie within a certain interval (Stewart 1983). In this case, 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 5.1, taking this uncertainty into account when calculating the confidence intervals of the coefficients in question.

Tables A1 and A2 show the results from the interval regression of the baseline OLS model presented in Section 5.1. 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 5.1 are robust to the uncertainty of the mean stunting prevalence and do not present a misleading level of certainty, even though the dependent variable is an estimated value.

Table A1: Interval regression on the upper and mean estimate of prevalence of stunting in 2013

	(1.1)	(1.2)	(1.3)	(1.4)	(1.5)
Proportion Double Orphans	1.794***	0.703**	0.429	2.600*	0.631
Asset Index		-0.113***			0.014
Piped Water			-0.037***		0.029
Flush Toilet			-0.051***		-0.091***
ECD Centres				-0.021	-0.017
Police Departments				0.003	0.009
Health Facilities				-0.009	-0.000
SASSA Offices				0.035	0.029*
DSD Centres				-0.026	-0.015
Constant	0.301***	0.348***	0.367***	0.334***	0.347***
Sigma	-3.052***	-3.146***	-3.257***	-3.757***	-4.211***
N	4,275.000	4,275.000	4,275.000	27.000	27.000

Note: * 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, this refers to flush toilets only. Inverse hyperbolic sine transformation of soft services applied. Standard errors are clustered at the local municipal level.

Source: author's calculations based on data from the OVCSA dataset.

Table A2: Interval regression on the mean and lower estimate of prevalence of stunting in 2013

	(2.1)	(2.2)	(2.3)	(2.4)	(2.5)
Proportion Double Orphans	1.601***	0.699***	0.459**	2.307*	0.643
Asset Index		-0.094***			0.016
Piped Water			-0.032***		0.034*
Flush Toilet			-0.042***		-0.085***
ECD Centres				-0.018	-0.015
Police Departments				0.007	0.013
Health Facilities				-0.008	-0.003
SASSA Offices				0.045	0.041**
DSD Centres				-0.023	-0.018
Constant	0.231***	0.270***	0.287***	0.240***	0.254***
Sigma	-3.262***	-3.360***	-3.482***	-3.808***	-4.251***
N	4,275.000	4,275.000	4,275.000	27.000	27.000

Note: * 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, this refers to flush toilets only. Inverse hyperbolic sine transformation of soft services applied. Standard errors are clustered at the local municipal level.

Source: author's calculations based on data from the OVCSA dataset.

Fractional probit

It is also necessary to establish the robustness of the results because the mean stunting prevalence variable is a proportion bound between zero and one. These types of data often display an asymmetrical distribution around the mean (non-normal distribution) and are vulnerable to heteroscedasticity in the errors. Additionally, when fitting a linear estimator to these 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 that satisfies $0<G(z)<1$ and proposes a quasi-maximum likelihood estimation technique (Papke and 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 5.1, which has been estimated using a fractional probit. Because a link function has been used to predict the model, it is unwise to interpret the values of the coefficients themselves. Therefore, the marginal effects at the mean of each control variable is given in Table A3. The estimation technique does not allow for clustered standard errors.

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	1.64***	0.61***	0.37***	2.55	0.81
Asset Index		-0.11***			0.007
Piped Water			-0.033***		0.024
Flush Toilet			-0.049***		-0.09***
ECD Centres				-0.02	-0.01
Police Departments				0.004	0.009
Health Facilities				-0.007	-0.001
SASSA Offices				0.04	0.03
DSD Centres				-0.02	-0.013
Constant	-0.624***	-0.491***	-0.443***	-0.589***	-0.505***
N	4,276.000	4,276.000	4,276.000	27.000	27.000

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

Source: author's calculations based on data from the OVCSA dataset.

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.

B SAR regression outputs

Table B1: Results from SAR model

	Model 1	Model 2 (a)	Model 2 (b)	Model 3 (a)	Model 3 (b)	Model 4 (a)	Model 4 (b)	Model 5 (a)	Model 5 (b)
Proportion Double Orphans	0.498***	0.284***	0.112**	0.263***	0.168***	0.480***	0.477***	0.225***	0.070
Asset Index		-0.037***	-0.012***					-0.013***	-0.013***
Piped Water				-0.028***	-0.022***			-0.027***	-0.021***
Flush Toilet				-0.013***	0.002			-0.009***	0.005***
ECD Centres						-0.002***	-0.001***	-0.002***	-0.001***
Police Departments						-0.001	-0.003***	-0.001	-0.002**
Health Facilities						0.000	0.000	0.000	-0.001
SASSA Offices						-0.003**	-0.001	0.000	-0.002**
DSD Centres						-0.001	-0.003***	0.001	-0.001
Constant	0.272***	0.287***	0.480***	0.303***	0.390***	0.276***	0.554***	0.307***	0.548***
Weighted Effects:									
Rho Coefficient	1.861***	1.863***	2.581***	1.895***	2.797***	1.860***	2.637***	1.900***	3.263***
Asset Index			-0.614***						
Piped Water					0.047**				0.071***
Flush Toilet					-0.252***				-0.194***
ECD Centres							-0.136***		-0.071***
Police Departments							-0.201***		-0.074***
Health Facilities							-0.024**		-0.066***
SASSA Offices							0.197***		-0.032
DSD Centres							-0.257***		-0.179***
N	4,276.000	4,276.000	4,276.000	4,276.000	4,276.000	4,276.000	4,276.000	4,276.000	4,276.000

Source: author's calculations based on data from the OVCSA dataset.

C Spatial autocorrelation regression modelling

The field of spatial econometric analysis draws heavily on the first law of geography, formulated by Tobler (1970: 236), that 'everything is related to everything else, but near things are more related than distant things'. Tobler qualifies this statement that while spatially situated data display 'infinite' complexity, spatial models can assume that things that are closer together affect each other to a greater extent and should be modelled as such. As described in Section 4, the unit of observation in the OVCSA data is the electoral ward, which gives the data an explicit spatial dimension. This means that while observations remain separate from each other, observations that are closer together will be related to each other and should be modelled as such.[18] Where $Corr\ (X_i, X_j)$ is positive, areas that are close together will tend to have similar values of X; that is, observations show spatial clustering.

Ward boundary demarcations are randomly drawn, and stunting prevalence and the number of orphans will not necessarily follow these lines. This implies that the rate of stunting prevalence in one ward will be related to, and have an impact on, the rate of stunting in neighbouring areas. Where there is an increase in child malnutrition or diarrheal illness, which increases stunting in one area, the effects are likely to be measured in a number of wards, where a few wards in particular may be the epicentre of the shock.[19] Thus, correlated shocks such as food prices or weather shocks will affect the nutritional status of children in broader regions. Additionally, the way in which variables relate to each other is also likely to display a spatial dynamic, which will cause correlation in the error term. In effect, the autoregressive error term operates as a regional fixed effect and takes into account that unobservable shocks in one area also affect the current stunting in adjacent areas, reducing bias.

It is best to explain spatially autocorrelated errors using an example. Assume there are two groups of wards, group A and group B. There are seven wards in group A, which is in an urban setting and has low mean stunting and relatively few orphans. There are seven wards in group B, which is in a rural area with larger wards, fewer government service points, higher mean stunting, and more orphans. In a regression model, the estimated residuals of those wards in group A will be more like the residuals of other wards in group A—and similarly for group B. The residuals are therefore spatially correlated—a property which will have important considerations on the interpretation of spill-over effects. Furthermore, features such as access to soft infrastructure in an area and clustered demographic characteristics are also likely to have spill-over effects. So, if there are multiple healthcare facilities in one area, neighbouring areas are likely to also have improved access to healthcare, albeit to a lesser extent.

The level of spatial autocorrelation is represented in SAR models by the ρ coefficient, or the autocorrelation parameter.[20] The value of the ρ coefficient indicates the type of spatial clustering in the data and allows for the interpretation of the type of spill-over effects that are present because of autocorrelation in the residuals. This parameter varies between -1 and 1 in stationary solutions, where the spill-over effects between areas dampen the overall effects of the characteristic. Thus, if there is a shock in one ward and if the estimated ρ coefficient is between -1 and 1, the spill-over effects will decrease over space. However, where the ρ coefficient is

[18] This fact is especially true for the measure of the mean prevalence of stunting in an area because of the manner in which the measure was constructed by Osgood-Zimmerman et al. (2018).

[19] This implies that the residuals in this model will be correlated, which is explained more fully below.

[20] The ρ coefficient should not be interpreted as an effect size or a traditional 'coefficient'.

greater than this interval, the model solution is explosive. If this is the case, a negative shock is amplified over space (StataCorp 2017).

These data can be modelled using a spatial weighting matrix and spatial lags to account for spatial autocorrelation. A spatial weighting matrix W is an N x N positive matrix in which the cross-sectional unit denotes the columns and rows. Element w_{ij} of matrix W represents the strength of the relationship between ward i and ward j. In a simplified, binary version of matrix W, element w_{ij} equals 1 if ward i and ward j are neighbours and equals 0 if they are not (Anselin, Gallo, and Jayet 2008). While this representation of the effects of neighbouring wards accounts for the direct effects of neighbours on each other, any effects of non-neighbouring wards is lost. Therefore, a more nuanced weighting matrix will be calculated for this analysis, where the strength of the relationship between ward i and ward j will be defined by the inverse of the distance between them, which is also known as the distance decay (Anselin 1999). Where the distance between two observations can be measured, this type of weighting matrix is preferred, as the effects of a ward on all other wards can be quantified, and the estimation of both direct and indirect effects of ward characteristics on the outcome variable is made possible.

For the purpose of this analysis, the inverse of the distance between the GPS location of the midpoint of each electoral ward is used to define weighting matrix W. This type of weighting matrix will also conveniently account for the spatial autocorrelation inherent in the values of the mean stunting prevalence. The spatial weighting matrix models the effects that wards may have on each other and also introduces the concept of a spatial lag. This concept is acquired from time-series modelling, where independent or dependent variables are included in the model lagged by one unit of time. In spatial models, this concept is applied across space, where a spatial lag includes variables lagged by one spatial polygon. Where contiguous weighting matrices are used, only a defined number of neighbours are included, and there are a limited number of spatial lags. Where distance decay matrices are used, the spatial lag of every polygon is included in the model (Anselin 1999).

In creating a spatial weighting matrix, it is important to consider the type of normalization of the weights that is applied because it impacts the scale of estimates. That is, because spatial lags in the independent variables will enter the model as $\boldsymbol{W}$x, and autoregressive residuals are modelled using $\boldsymbol{W}$u, the spatial lag coefficient will be inversely affected by the scale of $\boldsymbol{W}$ (StataCorp 2017). There are three main normalization methods for applied work: row, spectral, and minmax normalization. Minmax and spectral normalization calculate symmetric matrices and should be used when it can be assumed that ward i will have an equal effect on each of its neighbours. However, if the spill-over effects of ward i characteristics become more diluted the more neighbours it has, it is common to use row normalization. For example, assuming ward i has many health facilities, the ward's neighbours will experience spill-over effects in the form of increased access to services. However, those spill-over effects will be smaller the more neighbours are using the health facilities because the health facilities will be providing services to more people. Therefore, the spatial weighting matrix $\boldsymbol{W}$ will be row-normalized.

In order to establish the level of spatial clustering, multiple statistical tests have been created. One such test is Moran's test for spatial dependence. The null hypothesis for this test is that observations are spatially independent and identically distributed over space. In running the Moran test for spatial dependence, the spatial distribution of the error term in a linear OLS model is tested for clustering. In running this test on the most complete specification of the baseline OLS model, the result suggests that the null hypothesis can be rejected at the 0.01-per-cent level of significance (P=0.000). This implies that the spatial distribution of the high and low values of the variables included in this specification are more spatially clustered than would be the case if the underlying data-generating process were random. Specifically, this null hypothesis can be rejected in favour of the alternate hypothesis that the residuals are correlated with nearby residuals, where 'nearby' wards are defined by matrix W (StataCorp 2017). Therefore, SAR models are highly appropriate for modelling these data.