

Article

Linking the Sustainable Development Goals through an Investigation of Urban Household Food Security in Southern Africa

James Sgro ¹, Bruce Frayne ², Cameron McCordic ^{2,*}

¹ Balsillie School of International Affairs, University of Waterloo, 67 Erb Street West, Waterloo, Ontario, N2L 6C2, Canada

² School of Environment, Enterprise and Development, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, N2L 3G1, Canada

* Correspondence: Cameron McCordic, Email: c2mccordic@uwaterloo.ca.

ABSTRACT

Background: It has been theorised that there are a network of relationships linking the Sustainable Development Goals (SDGs), whereby achieving one SDG may have spillover effects for other SDGs. This discussion is relevant to the multidimensional stressors experienced by poor urban households in Sub-Saharan Africa.

Methods: We evaluate whether variations in the gender of a household head (SDG 5), education level of a household head (SDG 4) or household wages (SDG 8) are predictive of household food security (SDG 2) among over 6000 poor urban households surveyed in eleven cities in Southern Africa. These comparisons are made using regression analysis and machine learning techniques while controlling for, and comparing against, the contribution of household size, the age of the household head, and the number of household dependents to household food security prediction.

Results: Of the variables investigated, our study finds that household wages and the education level of the household head are important predictors of food security among the surveyed households. This investigation also identifies a potentially indirect relationship between the gender of the household head and household food security when other variables are controlled.

Conclusions: These findings suggest a predictive relationship between SDG 4 (Quality Education), SDG 8 (Decent Work and Economic Growth) and SDG 2 (Zero Hunger) while highlighting a curious indirect relationship between SDG 5 (Gender Equality) and SDG 2 among poor urban households in Southern Africa. By understanding these relationships, it may be possible to chart efficient policy pathways towards SDG achievement in Southern African cities.

KEYWORDS: Sustainable Development Goals; food security; nutrition security; poverty; urban; gender; Southern Africa

Open Access

Received: 22 April 2019

Accepted: 15 May 2019

Published: 21 May 2019

Copyright © 2019 by the author(s). Licensee Hapres, London, United Kingdom. This is an open access article distributed under the terms and conditions of [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

ABBREVIATIONS

AFSUN, African Food Security Urban Network; FAO, Food and Agriculture Organization; HDDS, Household Dietary Diversity Score; HFIAP, Household Food Insecurity Access Prevalence; HFIAS, Household Food Insecurity Access Scale; LOO, Leave-One-Out; PPP, Purchasing Power Parity; SDG, Sustainable Development Goal; USD, United States Dollar; VIF, Variance Inflation Factor; WHO, World Health Organization

INTRODUCTION

Since their inception, the Sustainable Development Goals (SDGs) have been an ambitious vision to transform global development in the face of pressing sustainability challenges [1]. The second of the United Nations' 17 SDGs is the achievement of Zero Hunger by 2030. With the global urban transition well underway, cities will play a pivotal role in achieving this goal [2–7]. Yet much of the food security literature and many policy responses emphasize food production as the solution to food and nutrition insecurity [8–10]. However, we argue that because food and nutrition security is also a social phenomenon which weaves economic, social, environmental, and physical factors together, investment in food production alone cannot meet the objectives set out in SDG 2—Zero Hunger [2,11]. Instead, by applying a broader development lens to the question of persistent food insecurity, it may be possible to expose the key factors that influence food security outcomes at the household and individual level [12,13]. Furthermore, it has been theorised that there are a network of relationships linking the SDGs [14], where, the achievement of one SDG may influence the achievement of other SDGs in the network. By understanding the nature of these relationships, it may be feasible to develop efficient policy pathways towards achieving the SDGs.

This paper aims to better understand these SDG relationships by analyzing the findings of the African Food Security Urban Network's (AFSUN) food security baseline survey which was undertaken in eleven cities in nine Southern Africa countries. More specifically, we evaluate whether variations in the gender of a household head (SDG 5—Gender Equality), education level of a household head (SDG 4—Quality Education) or household wages (SDG 8—Decent Work and Economic Growth) are predictive of household food security (SDG 2). Through our analysis, we also argue that food availability is not the principle determinant of food security outcomes in the urban context, but rather that factors associated with broader development objectives are important predictors of levels of household food security. Our results indicate that the household wages and education level of the household head are significant predictors of household food security. Importantly and controversially, the gender of the household head is found to be a statistically insignificant predictor of household food security when other factors are controlled.

The following subsection “Literature Review” provides a background on the state of food security globally before investigating the Southern African urban context. This precedes a review of the literature as it relates to the proposed predictors of food security. The “Materials and Methods” section outlines our research question, method for data collection and analysis, and outlines the significance of the machine learning model proposed. The “Results” section first investigates the impact of each predictor towards food security on a univariate basis, before delving into the “Multivariate Logistic Regression” and “Machine Learning” subsections. These latter two subsections bring together every variable previously discussed into a final model, and investigates each variable’s statistical significance, predictive capacity, and relative importance. A “Discussion” section investigates the broader implications of these statistical findings focusing on how these results translate to realizing the Sustainable Development Goals. The policy implications are discussed in further detail in the “Conclusions” section, addressing food security policy from the perspective of social protections, maternal health and parental autonomy, and universal basic education.

Literature Review

The World Summit on Food Security held in Rome in mid-November 1996 offered renewed impetus towards halving the number of undernourished people by 2015 [15]. This objective became the first Millennium Development Goal in 2000 with specific goals around raising the poorest incomes above \$1.25 (Purchasing Power Parity), and halving the number of underweight children [16].

Ten years later, the former Director General of the FAO—Jacques Diouf—stated that the number of undernourished was increased by 4 million people per annum [17]. However, by 2015, the official tally of those undernourished had declined to 795 million, which was 216 million fewer than the baseline year in the early 1990s [18]. 72 countries had succeeded in halving their undernourished population, but the report highlighted the impacts conflict had in protracting food crises for 24 countries in Africa. In Southern Africa, the financial target was never met, and today half the population earns less than \$1.25 per day [19]. In 2015, 23.2 percent of those living in Sub-Saharan Africa were undernourished, though investments notably in Western Africa have lessened this figure [20].

By 2050, the population is expected to exceed nine billion people [21,22]. As the world’s total population increases, urban population levels are projected to double to 6.4 billion people by 2050 [23]. Yearly, 60 million people become new urban residents which strains the capacity of cities, resulting in an ever-increasing informal sprawl as seen by the estimated 828 million individuals living in informal settlements today [23]. The vast majority of these settlements are located in cities in the global south where access to consistent nutritious food is a challenge for many residents. An

estimated 700 million Sub-Saharan Africans currently live in informal settlements, but this figure is expected to rise as the population in Africa rises from 1.1 billion to 2.3 billion by 2052 [24,25]. Rural to urban migration is also expected to rise from 40 percent to 60 percent [24].

While much of the work on food security focuses on the rural context, rural-urban migration, coupled with natural population growth, calls for research that aims to better understand the urban dimensions of food security [25]. More specifically, the complexity around the urban dimension of food security implores policy makers to ensure all members of society have access to adequate and nutritious food [26]. Vogel & Smith [27] reasserted that contemporary issues in food security had little to do with production, but rather the system of food distribution. These findings were echoed by Sidhu *et al.* [28] as their research showed that the depth of food insecurity was greater in urban settings, likely because rural areas had greater access to available food. Specifically, urban households were 47 percent more likely to experience food insecurity than their rural counterpart. This is particularly troubling in urban areas, given that food is generally widely available, but it is the means of access that can be limited by factors that include income and infrastructure [13].

The following subsections review some of the key predictors of household food security that have been identified in empirical studies on the topic. This literature review demonstrates the importance of the household wages, education of household heads and gender of household heads (as proxies for SDG 8, 4, and 5) in the prediction of household food security, and highlights the importance of age, household size, and number of dependents as control variables in this assessment.

Income

Previous research has indicated the importance of household income as a predictor of food access. Sidhu *et al.*'s [28] study of food security in India claimed that every incremental increase in monthly income of Rs 100 increased the probability of food security by 3 percent. Bashir *et al.* [29] found that households with two financial providers were 10.183 times more likely to be food secure. In this same study, households which broke even with their expenses (net zero income) were 8.146 times more likely to be food secure contrasted against indebted families, and those with a net positive income were 14.775 times more likely to be food secure. Sidhu *et al.* [28] found that families with higher incomes were 4.26 times more likely to be food secure. In a separate study, Babatunde, Omotesho, & Sholotan [30] suggested that credit (and governmental assistance) played a vital role in improving family flexibility against unexpected bouts of food inaccessibility. Gebre [12] affirmed this assertion by suggesting that social support and lending systems helped cushion periods of food inaccessibility.

Education

Education also appears to be a significant predictor of food security. Household food security was found to be 6.687 times more likely when the household head has a primary education, contrasted to heads with no formal education [29]. Babatunde *et al.* [30] found that education had a positive, statistically significant relationship to food security. In a study conducted in Nigeria by Omonona and Agoi [31], 67 percent of surveyed families experiencing food insecurity had household heads with no formal education. In pre- and post-natal situations, a mother's level of education also positively related to the child's nutrition levels [32].

Gender

Omonona and Agoi [31] found in their Nigerian study that food insecurity was 11 percent higher in female-headed households. By contrast, Amaza, Umeh, Helsen, & Adejobi [33] showed that female-headed households had a higher probability of being food secure. While both these studies took place in Nigeria the role gender plays in ensuring food security remains a topic of contestation. One possible explanation is rooted in the terminology around headship, where conjugal families are often considered "male-headed". Because of this, a single income household would also be categorized as a female-headed household. From this understanding it is entirely possible that the gender of the head better explains the number of earners rather than directly correlating to food security.

Household size

Amaza *et al.* [33] indicated that household size was a significant determinant of food insecurity, arguing that the number of dependents (elderly, infirm, and young residents) hampered a household's ability to be food secure. A meta-analysis conducted by Bashir *et al.* [29] suggested that families with more than 3 individuals were almost twice as likely to be food insecure. They suggested this impact was partially mitigated if newly introduced members were also income providers. Sidhu *et al.* [28] found that every additional family member decreased the food security likelihood by 44 percent holding all else constant. Mitiku, Fufa, & Tadese [34] suggested an alternative finding that every incremental increase in household size related to an 80 percent increase in the likelihood of food insecurity. Omonona and Agoi [31] showed in their study that food insecurity was lowest (27 percent) for households with less than four residents.

Age

Bashir *et al.*'s [29] study found that older family heads were less likely to be food secure, compared to the reference group in their 20s. In this study, the percent likelihood of food security was lowest with a family

head between the ages of 61–70. Food security was highest in the 21–30 age group category and the food security likelihood decreases by 83 percent when the household head is between 36 and 45 years old (contrasted to the younger age-group). Omonona and Agoi [31] shared similar findings, where the incidence of food security was lowest (58 percent) for heads aged between 61 and 70 and highest (30 percent) for household heads in their 20s. Babatunde *et al.* [30] indicated that older household heads had a moderately negative impact on food security at the 10 percent significance level.

Dependency percentage

Measured as the number of non-working age residents compared to the whole household, Iram & Butt [35] found that a high dependency had a significant effect on food security as it related to caloric intake. In a study of subsistence agriculturalists in Nigeria, Asogwa, Okwoche, & Umeh [36] found that the number of dependents was a statistically significant ($p < 0.05$) predictor of poverty and nutrition. In this study, a 1 percent increase in the dependency ratio increased household poverty intensity by 0.42 percent. That said, Gebre's [12] study of food security in urban centers in Ethiopia found the dependency ratio was not a statistically significant predictor of household food security. Mitiku *et al.* [34] also found the number of dependents to be statistically insignificant ($p = 0.207$) in their logistic regression model, suggesting that there was little consensus across the literature with regards to the significance of this variable. For the purposes of this paper, dependency will be measured as a percent of dependents since households with absolute dependency would have ratios divided by zero.

Synthesizing the literature

The urban food security literature broadly indicated that the following six variables displayed a potentially significant relationship to food security: gender, education, wage, age, household size, and dependency percentage. The former four variables were measured as they related to the household head. However, each of these factors were identified in different geographical contexts. As such, the first research question investigates whether these influences carry out similarly in the urban Southern African context as well.

The gap in the literature is two-fold. First, much of the work that has been done to outline the determinants of food security have been conducted in Pakistan, Ethiopia, Nigeria, and so on. Few current studies have specifically applied this approach to the Southern African context. Even then, too often the concept of food security was reserved for the rural setting, despite research indicating that urban areas were more prone to price-induced issues of food inaccessibility [28,29,31]. While countless studies have concerned themselves with specific countries in the Southern African region, there were fewer cross-sectional studies which regarded

the larger Southern African region (e.g., Frayne *et al.* [2]). Second, those studies which have conducted similar research in the Southern African context had yet to apply supervised machine learning techniques to further elucidate the lived experience of food insecurity internationally. There were studies which had applied machine learning models in relation to poverty as a whole [37–39], though these models stopped short of the topic of food security specifically. Third, there is an urgent need for, and a relatively limited, amount of research on the interlinked network of relationships between the SDGs among poor urban households in Southern Africa (and the efficiencies that these relationships could lend to urban sustainable development). This gap presents an opportunity to further explore the potential contributions of multiple SDGs that may explain the achievement of SDG 2 (Zero Hunger) in Southern African cities using supervised machine learning and regression analysis.

MATERIALS AND METHODS

Research Questions and Hypotheses

The aim of this paper is to assess whether variations in the gender of a household head (SDG 5), education level of a household head (SDG 4) or household wages (SDG 8) are predictive of household food security (SDG 2) among over 6000 poor urban households surveyed in Windhoek, Gaborone, Maseru, Manzini, Maputo, Blantyre, Lusaka, Harare, Cape Town, Pietermaritzburg, and Johannesburg. This paper will then apply supervised machine learning modeling to discern whether household wages, education of household heads or gender of household heads are more effective than others in identifying instances of food insecurity (as a proxy of SDG 2).

- *Question 1:* Do predictive variables such as household size, income, structure, and education display statistical significance in predicting food security in poor areas in the eleven sampled cities in Southern Africa?
- *Question 2:* Are household wages (as a proxy for SDG 8), education of household heads (as a proxy for SDG 4), and gender of household heads (as a proxy for SDG 5) more effective (relative to the other identified indicators) at categorizing food insecure households (SDG 2) in the poor areas in the eleven sampled cities in Southern Africa?

The first research question relies on regression analysis to better understand how efforts towards realizing the Sustainable Development Goals should be organized. By recognizing the Goals as an interrelated network, we hypothesize that efforts towards one goal may have spillover effects towards several other goals. The second research question helps ground policy initiatives by comparing between various SDG proxies to better identify leverage points where targeted intervention would have the greatest impact towards food security.

Data Collection

The African Food Security Urban Network (AFSUN) survey was conducted in the last quarter of 2008 in eleven cities across nine countries. The data constitutes a mix of seven capital cities and four secondary cities. Table 1 illustrates the distribution of the surveys issued across the selected cities. The cities were selected on the basis of local expertise, expressed interest and engagement from policy makers, and the fact that they collectively offer a wide platform from which to address the issues of urban food security more generally. One or more poorer urban neighbourhoods were identified for study in each city. In the larger cities, such as Cape Town and Johannesburg, a mixture of formal and informal urban neighbourhoods were chosen. The survey itself was conceptualized in June 2008 in Botswana and received ethical clearance from the Queen's University General Research Ethics Board before training local undergraduates in each city. All surveyors were supervised by faculty members and trained to ensure inter-rater reliability.

Table 1. Number of household surveys per city.

City	Number of Households	Percentage of Sample
Windhoek	448	6.9%
Gaborone	400	6.2%
Maseru	802	12.4%
Manzini	500	7.7%
Maputo	397	6.2%
Blantyre	432	6.7%
Lusaka	400	6.2%
Harare	462	7.2%
Cape Town	1060	16.4%
Pietermaritzburg	556	8.6%
Johannesburg	996	15.4%

Sample selection was conducted using systematic random sampling in specifically identified poorer urban neighborhoods. These poor areas were identified using a combination of census data and geoinformatics. As well, maps of the areas to be surveyed were prepared in advance and used in the field for household selection. When selected residents were not present to be surveyed, a substitution was made. Each survey required informed consent and was conducted by a responsible adult in the residence. Field supervisors and/or city partners checked completed questionnaires. To minimize data entry errors data computation and cleaning was centralized to the University of Namibia resulting in ~6500 household surveys and ~29,000 individual surveys. Specific questions spanned 7 sections, relating to food aid, consumption, acquirement, and of course, food insecurity.

Study Design

This paper aims to employ predictive modeling techniques to best understand the influence of predictor variables on the response variable—food security. However as Jarosz [40] mentioned in her paper on policy discourse, food security as a definition must go beyond the idea of simply being hungry. For the purpose of this paper, a working definition will follow the same one as coined by the FAO in 1996: that “all people, at all times, have physical, social and economic access to sufficient, safe and nutritious food to meet their dietary needs and food preferences for an active and healthy life” ([15], p. 28). The definition provided by the FAO stresses the parallel importance of access and nutrition. Great strides have been made to develop measurement scales like the Household Dietary Diversity Score (HDDS) and the Household Food Insecurity Access Scale (HFIAS) [41,42]. Employing these metrics refines the definition of food security to account for both access and nutrition. The FAO also acknowledged the triad of physical, social, and economic components to food security. With this in mind, any statistical method must account for the multidimensional nature of food insecurity.

Until now, the response variable for this study has only been termed ambiguously and overarchingly as “food security”. The AFSUN survey used the Household Food Insecurity Access Scale (HFIAS) as a measure which encompasses both the physical and cultural dimensions of food security. However, because the HFIAS is a scale from 1 to 27, it is challenging to classify security and find a cutoff point. For this reason, Coates *et al.* [42] derived an additional scale—Household Food Insecurity Access Prevalence (HFIAP)—which reduced the 27 different responses from the HFIAS into 4 measurements. Depending on the nature of the predictor variable, the HFIAS metric will be used to provide a higher resolution of the change a given variable has on food security. However most interpretations of the regressions will be done with HFIAP as the response variable to easily convert numbers to a real-world understanding of food security.

Analysis Plan

The first approach this paper will apply is logistic regression, used to understand the change in likelihood for food insecurity given incremental changes in each predictor variable. The formula for logistic regression is as follows:

$$\ln\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

The main difference between a linear probability model and logistic regression is that the latter applies a sigmoidal curve to the response variable and is most appropriate for normally distributed data where the response variable is binary in nature. This model uses a standardized recombination of each predictor variable (β_k) which then culminates in a

log-likelihood figure acting as a percent estimator of food security for any given observation. To measure the change in odds for any incremental increase of the predictor variables, the response variable $\left[\text{Ln} \left(\frac{p}{1-p} \right) \right]$ is then exponentiated by $(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)$. This model allows observers to interpret the influence each variable has in increasing the log-likelihood of food insecurity being observed while holding all other variables constant. To statistically confirm this model performs better than mean-comparison, the Pearson's Chi-Square Test of Independence will be applied.

This analysis will be conducted through the open source programming language R. To ensure that the variables are independent and homoscedastic, the Variance Inflation Factor (VIF) will also be incorporated, where observations with a value of one have little multicollinearity, and values above five are considered strongly correlated [43]. Lastly, to answer the second research question, this paper will use the Random Forest algorithm to rank the importance of each variable in categorizing different levels of food security, as follows:

$$VI^{(t)}(x_j) = \frac{\sum_{i \in \bar{\beta}^{(t)}} I(y_i = \hat{y}_i^{(t)})}{|\bar{\beta}^{(t)}|} - \frac{\sum_{i \in \bar{\beta}^{(t)}} I(y_i = \hat{y}_{i,\pi_j}^{(t)})}{|\bar{\beta}^{(t)}|}$$

Where, $\hat{y}_i^{(t)} = f^{(t)}(x_i)$ is the predicted class prior to permutation and, $\hat{y}_{i,\pi_j}^{(t)} = f^{(t)}(x_{i,\pi_j})$ is the predicted class after permuting X_j [44].

As a central method to answer the second research question, the role of the random forest algorithm is to produce a decision matrix model (or tree) which best categorizes observations of food security based on the proposed predictor variables. The underlying notion behind a single tree is to run data through binary inquiries progressing through the array of inquiries until it has been properly categorized. Unfortunately, any infinitely long tree with an infinite horizon will always perfectly categorize the training data; therefore it runs the risk of overfitting the model to the point that it becomes inoperable on new observations. As such, random forests manifest thousands of decision trees with simulated noise through bootstrapped aggregation. Because so many trees are being evaluated and cross-validated with different combinations of data, the final model has a predictive capacity more robust than the single decision tree mentioned earlier.

The main benefit of this algorithm—and the reason for its inclusion in this paper—is that it naturally measures variable importance as every split in the decision matrix provides an out-of-bag error rate. By aggregating these error rates, an overall misclassification rate is discerned which ranks the individual variable importance in the creation of the model [45]. As well, the leave-one-out (LOO) cross-validation method allows the model to pick up multicollinearity and variables which are dependent on one another. Because LOO drops out every variable with identical frequency as a process in model evaluation, any codependent variables will not artificially lower the error rate [45]. In short, two strongly collinear

variables with equal predictive capacities will not compound upon one another to produce an even greater overall predictive accuracy, even when both variables are present during cross validation.

From an applied perspective, the decision to include random forests as an analytical tool draws from research from McBride & Nicols [37], Xie *et al.* [38], and Jean *et al.* [39] where machine learning techniques are employed to measure poverty more generally. Such methods remedy the drawbacks of generalized linear models by allowing for researchers to directly compare between variables, rather than hold all other variables constant. Previous studies have applied machine learning models in relation to poverty more generally, but few specifically investigate food security specifically [46,47], presenting a critical opportunity for further analysis relating to sustainable development more broadly. Random forests and variable importance metrics have been used in bioinformatic research and phylogenetic analysis. In this field it has proven a useful tool to rank highly dimensional data with strong predictive capacity and informative pathway analysis. This paper applies the same method to identify which of the several variables are most relevant in categorizing varying levels of food security.

It is proposed that variables which hold statistical significance will indicate varying levels of importance for categorizing food security once processed through a supervised machine learning model. Algorithms including random forests act as helpful tools for ranking the role these variables play in categorizing hypothetical new data. For example, it might be found that a variable is strongly correlated to food insecurity through logistic regression. However, once incorporated into the random forest algorithm, it could be found that the income variable poses little use in separating between secure and insecure households. While this is a hypothetical example, it illustrates the role an unsupervised machine learning algorithm might have in further testing the predictive capacity of these proposed variables.

RESULTS

The following section delves into each of the proposed variables as follows:

$$\widehat{HFIAP} = \beta_0 + \beta_1 age + \beta_2 wage + \beta_3 educ + \beta_4 size + \beta_5 gender + \beta_6 dependency + u$$

Age, education, and gender are all assumed to relate to the household head prior to any analysis and wage, size, and dependency relate to the household level. In some cases, HFIAS will be used to measure food security and while HFIAP is derived from HFIAS, the specific findings from either are not considered synonymous to one another. Before modeling the above equation, this section first analyzes each variable on a univariate basis. Each subsection evaluates how exactly the metric should be measured and whether transformations with respect to x_i would better the final model. The methods used to evaluate the univariate

relationship between HFIAP and the variable will be limited to the Ordinary Least Squares regression where $\hat{\beta}_1 = \frac{cov(x,y)}{var(x)}$. Once each variable has been vetted, this section will move to multivariate logistic regression and finally, supervised machine learning.

Variable Analysis: Wage

To best approximate the monthly earnings for each household in the study, the wage variable sums household income from all sources in the last month. This includes both wage and casual work, but also goes a step further to incorporate remittances, informal business arrangements, agricultural income, aid, grants, and gifts. These figures span 16 different potential avenues for income and were identified in the country's currency. Outliers were defined as amounts exceeding 1.5 times the interquartile range (in USD), omitting a total of 466 households from the analysis.

On the note of financial restraints, currency devaluation remains a major limitation to codifying wage. Globally, the cost of staple foods rose 63 percent from 2007 to 2008 across 11 different Sub-Saharan countries [48,49]. Most notably, the time period in which the Zimbabwean dataset was collected was between Oct 16th, 2008, and Oct 25th, 2008. During this ten-day period, the value of the Zimbabwean dollar shifted from 227 ZWD per USD to 451 ZWD per USD [50]. When comparing between models which applied a per diem currency conversion and a single currency conversion based at the median date of October 11th, 2008, it was found that the single-date conversion performed better as a predictive model. A logistic regression is applied using a binomial distribution for the Generalized Linear Model. With a p -value <0.001 , the model returns an exponentiated value for β_1 of 0.998, or more specifically:

$$\text{logit}(\widehat{\text{foodInsecurity}}) = 2.228 - 0.001872\text{wage} + u$$

Using the binomial *foodInsecurity* variable which dichotomizes whether a household's HFIAP value is 1 or otherwise, it can be concluded that households with one additional dollar in monthly wage have, on average, a 0.2 percent decrease in the likelihood they are food insecure, holding all else constant. On the note of sampling bias, an evaluation was conducted to ensure that no single HFIAP value captured 80 percent of the total sample, and that less than 80 percent of the HFIAS values were below the value of five.

Variable Analysis: Education

A study conducted by Omonona & Agoi [31] found that as education increased, the food insecurity incidence decreased. However, a very large population (52.3 percent) of the food insecure group in the sample group has a tertiary education. While the percentage of food secure tertiary education households is still higher (69.0 percent), tertiary level education still constitutes the largest group of food insecure households. Omonona &

Agoi [31] hypothesized that this phenomenon may be a result of unemployment, though they also attributed the high level of food secure tertiary education households to being due to the new occupational opportunities that came with higher education. This in turn spills over as additional income in many cases, though the rationale for education as it relates to food security seems to be confounded. Overall, the largest group of respondents were the tertiary education group ($n = 103$), and of that large group, two thirds were food secure.

Babatunde *et al.* [30] made the case that a spouse's level of education could be more important than the household head as it related to food security. However, when evaluating possible models in the Southern African context, it was found that the model with the highest 'goodness of fit' measured education based on the household head. In the urban Southern African sample, a univariate linear regression analysis confirmed that average HFIAS continually decreased for groups where the household heads achieved higher education. Education was measured as a categorical variable indicating the highest level of education achieved (e.g., primary, secondary, tertiary). Even though the standard error grows as education increases, each level of education is statistically significant ($p < 0.001$). This model implies that as education increases, food insecurity decreases at a consistent rate. To test the hypothesis that wage is collinear to education, a variance inflation test is introduced. With a variance inflation factor of 1.156, this model satisfies the linear regression assumption that no two variables have perfect collinearity [43].

Variable Analysis: Gender

The gender of the household head has been largely documented to hold a statistically significant relationship to food security (e.g., [29,35,51]). However, the exact relationship gender holds in food security is up for debate. None of the literature reviewed found a statistically insignificant relationship between gender and food security, but where some found the presence of a female head to be detrimental, others found evidence that it was a boon to the family's food security outlook. In Omonona & Agoi's [31] research sample, 62.5 percent of male-headed households and 53.1 percent of female-headed households were food secure. The researchers hypothesized that the difference in the food security index was a result of differences in dependency ratios. They pointed towards household structure to say that female-headed households were more often widowed, and less likely to be a part of a dual income family.

Before delving into the role gender plays in food security, it is worth noting that a strong statistical correlation is evident in the research sample supporting a wage gap between genders heading the household. With a correlation coefficient of 0.67 percent and a p -value < 0.001 , the regression is as follows:

$$\widehat{wage} = 283.423 - 75.591female + u$$

This model suggests that female-headed houses in the research sample have, on average, a monthly wage of \$75.59 less than male-headed households holding all else constant. This in turn relates to Omonona & Agoi's [31] hypothesis that in patrilineal societies, dual income houses are more often termed 'male-headed' and only transfer to 'female-headed' when the male figure is absent. In this case, the gender used to define headship relates more so to whether or not it is possible that dual income might exist.

It is hypothesized that gender is strongly related to food security which has already shown statistical significance in the literature review. As well, 95 percent of all female-headed households in the sample did not have a partner present. Whether they are unmarried, widowed, or married to a partner who is a migrant worker, the potential for a single income household is greater for female-headed households than it is for men. A simple linear regression with wage as a function of cohabitation yields a beta-value of 73.41, indicating that cohabiting households, on average, have a monthly wage of \$73.41 more than non-cohabiting households ($R^2 = 6.7\%$, $p < 0.001$). When a multivariate model is introduced with wage as a function of both gender and cohabitation, the following function results, with $p < 0.05$ for both variables and an R^2 of 7.9 percent:

$$\widehat{wageUsd} = 249.22 + 43.80cohab - 44.27female + u$$

The maximum change in average income based on these two variables is $\pm\$88.07$ per month. The introduction of a cohabitation variable helps transfer some of the beta-weights from gender and offers a richer understanding of the interplay between wage, gender, and household income sources. Food insecurity (expressed as a Boolean) as a function of gender can be modelled with the following:

$$\text{logit}(\widehat{foodInsecurity}) = 1.5538 + 0.3012female + u$$

This logistic function results in an 86.5 percent likelihood that a family is food insecure when the household head is female and 82.5 percent food insecure when the head is male. The probability value is discerned by:

$$p = \frac{e^{(\beta_0 + \beta_1 x_1)}}{1 + e^{(\beta_0 + \beta_1 x_1)}}$$

Holding all else equal, the model for food security as a function of gender runs against some of the findings from the literature review. That being said, there is sufficient evidence that gender plays a large role in influencing wage. Holding all else equal, the percent likelihood that a family is insecure in the sample is higher for female-headed households than it is for male-headed households.

Variable Analysis: Household Size

A study of Nigerian urban food security conducted by Omonona & Agoi [31] found that food secure households most often had a household size of 4 or less, and that all of the households with 12 or more members in the

sample were food insecure. Insights from Garrett & Ruel's [52] comparison on urban and rural food security found that household size was one of the few variables which had significant influence on caloric intake in Mozambique. It was also one of the two variables to produce largely different results between urban and rural settings. This falls in line with Sidhu *et al.*'s [28] finding that the influence of household size plays out differently in rural and urban areas. However, the quadratic measurement of household size used in Garrett & Ruel's [52] study implied that food insecurity rose with household size, but at a decreasing rate. The influence household size poses on food security would be expected to plateau given large enough families. To borrow from the findings from Garrett & Ruel's work [52], an OLS model transformed as a quadratic relationship between HFIAS and household size as follows:

$$\widehat{HFIAS} = 9.6565 + 0.0241size^2 + u$$

When a quadratic transformation is applied, the correlation coefficient (R^2) increases from 0.69 percent to 0.99 percent. Because $\beta_1 > 0$, the quadratic function implies that food insecurity increases at an increasing rate for every additional household member. The quadratic function differs from Garrett & Ruel's [52] study in that the caloric intake model they proposed plateaus as it approached infinity, while the quadratic model of HFIAS in the Southern African sample is ever-increasing. Granted, both models measure different aspects of food security, but the exact form of a quadratic function of this type is a topic of further investigation for future studies.

Variable Analysis: Age

Previous studies conducted by Bashir *et al.* [29] and Omonona & Agoi [31] found that food security was highest for household heads in their 20s and lowest for those in their 60s. In both studies, food security decreases as the household head's age increases. The largest gap in the literature with respect to the headship age is for household heads under 20. It is too facile a conclusion to state that younger heads relate to greater food security, since the previous studies had the lowest age threshold set to 20 years.

To test whether there is in fact a statistically significant difference in food security across age (binned at the 40-year threshold), this investigation relies on Welch's *t*-test since the variance between groups is unequal. With a test statistic of -5.42, we reject the null hypothesis that the average HFIAS for the younger sample is the same as the older sample. Even though average HFIAS is higher in the younger sample, the linear relationship between HFIAS and age seems to weaken once the age group 15-19 is categorized on its own. Average food security is highest for household heads in their 20s, just as previous studies [29,31] suggested. The major divergence, however, is that in including sample groups where the household head is less than 20 years old, food insecurity rises. The

following linear model for HFIAS as a function of the age of the household head is:

$$\widehat{HFIAS} = 9.34 + 0.04age + u$$

This model has a p -value <0.05 and an R^2 value of 0.0064. Such a small R^2 value is unsurprising since the model attempts to measure a real-world experience with a single metric. Since HFIAS increases at the lowest age-group, the data for age holds a slightly convex relationship with HFIAS. Therefore, it is proposed to apply a quadratic transformation to the regression model as follows:

$$\hat{y} = \beta_0 + \beta_1 x_1^2 + u$$

Given the various possible transformations, the model which provided the highest correlation coefficient was a quadratic transformation with respect to the predictor variable.

Variable Analysis: Dependency

Dependency ratio is calculated as the number of non-working residents for every working member of the household expressed as a percent of the household population. When modeling HFIAS as a function of dependency the residual and q-q plots appear to be relatively homoscedastic, however, since the predictor variable is expressed as a percentage of two small integers, the list of possible values jumps from percent to percent. The corresponding equation for the univariate model is as follows:

$$\widehat{HFIAS} = 6.8947 + 5.3325depend + u$$

With a p -value <0.001 and an R^2 value of 3.96 percent, the dependency percentage is statistically significant. The beta value for dependency percentage is unusually high because it only can take value between 0 and 1. When a cubic transformation is applied to the dependency variable, the correlation coefficient increases from 3.96 percent to 5.7 percent. This specific transformation is informed by the two distinct bends found in the variable's smoothed density plot, suggesting a standard linear model may be insufficient to explain variations in the response variable. When dependency percentages are modeled as a function of household size, a statistically significant ($p < 0.001$) relationship with a correlation coefficient of 11.79 percent results. Since the response variable is a continuous function, households with one additional member, on average, have an increased dependency percentage by 4 percent.

More interestingly, once a multivariate linear model for HFIAS is devised that includes both household size and dependency, household size significance drops off quite notably. A univariate model of HFIAS as a function of household size by itself has a t -value of 6.615. This t -value decreases to 0.931 once dependency is introduced. This helps to explain situations where previous researchers found dependency ratios to be statistically insignificant. However, in the logistic models other scholars

created, dependency lost significance rather than household size. Once household size is quadratically transformed, household size returns to a statistically significant level ($p < 0.001$) with a t -value of 3.8. Both variables are significant on their own, but household size only retains statistical significance when it's expressed as a quadratic. While each study's results will vary, this insight acts as a diagnostic tool to better understand food security as it relates to household structure. Both Gebre [12] and Mitiku *et al.* [35] measured household size and dependency ratio, and both dismissed dependency ratio due to its statistical insignificance. Gebre [12] found household size to be significant ($p < 0.01$) but did no transformation with regards to x_i . Second, Mitiku *et al.* [34] found household size to be statistically significant ($p < 0.05$) to food security and conducted a VIF test confirming that multicollinearity was not a potential concern. Similar to Gebre's [12] study, no transformations were applied with regards to x_i . While it cannot be confirmed whether a quadratic transformation on household size would change the significance dependency plays in their models, it poses a valuable question to help better understand the role dependency plays in the larger issue of food security.

Multivariate Logistic Regression

The previous sections outlined the variables to be included in the final model. After contemplating the ways to measure very similar factors of food security (e.g., education of a spouse *versus* education of the head) and applying the optimal linear transformation, the following logistic regression is proposed:

$$\text{logit}(\widehat{\text{foodSecure}}) = \beta_0 + \beta_1 \text{age}^2 + \beta_2 \text{wage} + \beta_3 \text{educ} + \beta_4 \text{size}^2 + \beta_5 \text{sex} + \beta_6 \text{depend}^3 + u$$

Food security (measured as a Boolean where 1 = food secure and 0 = food insecure) as a function of the previously proposed variables and their respective transformation is as follows:

Table 2. Linear model of food security and all transformed proposed predictors.

Response variables	Dependent variable:
	Food secure
β_0	-2.610*** (0.238)
Age ²	1.051e-04*** (0.00004)
Wage (USD)	1.502e-03*** (0.0001)
Educ_2	0.514** (0.227)
Educ_3	0.393 (0.241)
Educ_4	0.837*** (0.225)
Educ_5	1.324*** (0.227)
Educ_6	1.418*** (0.255)
Educ_7	1.251*** (0.344)
Educ_8	1.644*** (0.303)
Educ_9	1.936*** (0.466)
Household Size ²	-0.011*** (0.002)
Female	-0.127 (0.089)
Dependency (%) ³	-0.774*** (0.158)
Observations	4,991
Log Likelihood	-1952.664
Akaike Inf. Crit.	3933.328

Note: ** $p < 0.05$; *** $p < 0.01$.

Table 2 indicates that the age of the household head retains statistical significance at the 5 percent level with an exponentiated beta value of 1.0001. The model suggests that completing primary school (Educ_3) does not hold a statistically significant relationship to HFIAP at 5 percent in this model. With a variable inflation factor reaching no higher than 1.32 for all non-categorical variables, no excessive amount of collinearity is present within this model [43]. When dependency and household size are untransformed, the resulting correlation coefficient is 0.343, suggesting a weak positive linear relationship. However, when both variables are transformed as expressed in this model, the correlation coefficient decreases to 0.157.

As well, the gender of the household head also does not hold a statistically significant relationship to HFIAP in this model. That is not to say that gender does not affect food security, since several previous studies (e.g., [31,33,35]) have found statistically significant relationships both for and against the notion that female-headed households were more food (in)secure. The previous section on gender found that, in a univariate model with food security as a function of gender, there was a beta value of 0.3012. In the previous univariate model, families which are headed by a woman were 3.9 percent more likely to be food insecure. A wage gap of \$75.59 on average was identified in the study sample and as such, the gender of the household head may relate more strongly to the external conditions imposed around gender (e.g., wage and the number of earners) rather than gender in and of itself.

Another notable phenomenon is that dependency percentage holds a statistically significant relationship to HFIAP in this model, running contrary to Gebre [12] and Mitiku *et al.*'s [34] findings where this variable was statistically insignificant. As previously alluded to in the section on dependency, the determining factor of whether the number of individuals in the household is statistically significant depends on whether household size is expressed as a quadratic. Even though some multicollinearity exists between the number of dependents and the number household members, the statistical significance retained by both variables suggest they contribute positively to the model.

This model has a McFadden's R^2 value of 0.29 and a Cox and Snell R^2 of 0.28. When applying a chi-squared test on the logistic model with all variables included, the χ^2 value is 1643.9 with 14 degrees of freedom ($p < 0.01$), indicating this model performs better than randomly assigning variables. A variable-by-variable analysis of the chi-square test returns similar findings as previously described, building the case that gender and lower levels of education are statistically insignificant. Overall, while lower levels of education and gender as a whole are statistically insignificant, all other previously investigated variables contribute positively to better understanding food security in urban areas across Southern Africa.

Machine Learning: Random Forests

With the final set of variables ascertained, focus now turns to the role each variable plays in HFIAP prediction contrasted to one another. Within generalized linear models, researchers are unable to easily compare the effectiveness of each variable. Typical interpretation of these linear models must hold every other variable constant when interpreting the explanatory value of a single variable [53]. To evaluate the relative importance of each variable within this model, we propose applying the random forest supervised machine learning technique. This algorithm is a classifier which attempts to create several decision trees (or dendrograms) with which to categorize all observations. Since the dendrograms split the data through Boolean evaluations at given points in the model, it is possible to track which variables were relied on the most to properly categorize each observation. The variable importance chart for HFIAP as a function of the variables proposed in the previous section are in Figure 1, as follows:

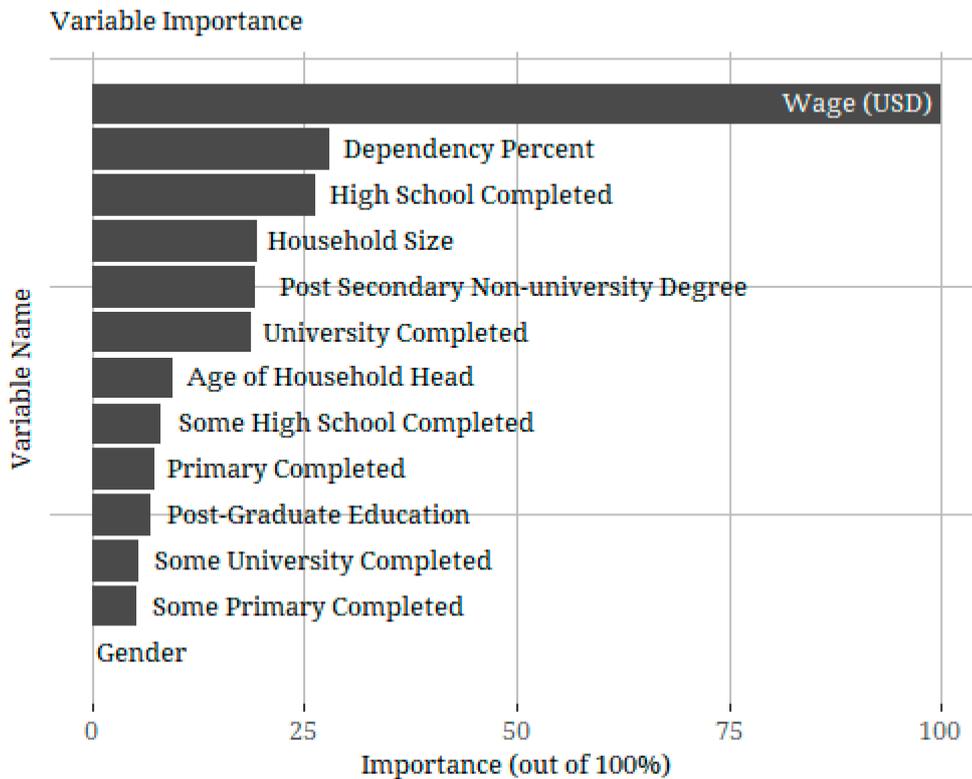


Figure 1. Random forest variable importance.

Some important aspects of Figure 1 to note are that the top variable is pinned to 100 percent. Each subsequent variable is expressed as a percent compared to the reference—in this case, wage. Second, each iteration of this algorithm will produce some level of variance despite creating several dendrograms. Rather than interpret the order of each value, it is helpful to instead interpret groupings of variables. For the model in Figure 1, wage in US dollars is always the top variable, followed most often by the dependency percentage and a high school education. Next, the third group is comprised of: a postsecondary non-university degree, the number of members in a household, and completing a university degree. Age and partial completion of high school are both in the fourth most important group, followed by the varying levels of education (most of which relate to both the very lowest and highest level of education obtainable). Invariably, gender holds no variable importance in predicting HFIAP through random forests.

Within this model, education has eight different levels—each acting as a dummy variable. Therefore, when interpreting values for their variable importance, it is important to note any given household head's highest level of education can only be one of the 8 possible levels of education (In the research sample education is measured at nine discrete levels, however the lowest level of education is used as the reference group. Therefore, all measures of education compare themselves to “no formal education”). The findings from the variable importance function in Figure 1 suggest that wage is a highly important factor in categorizing

groups into varying levels of food security, falling in line with Bashir & Schilizzi's [54] finding that income is the most important factor in influencing food security. As alluded to in the previous section, gender fails to hold statistical significance, and correspondingly provides no insight in the random forest algorithm. Lastly, education appears to matter at varying degrees. Based on the graph, it appears that either end of the education spectrum (primary and post-tertiary) matter significantly less than a level of education at about the halfway mark. Most importantly, households where the household head has a high school degree or tertiary diploma seem to be better categorized into the 'food secure' category. This could relate to household heads making more informed food-related decisions for the household as Babatunde *et al.* [30] suggested, or better employment opportunities for higher paying jobs, as suggested by Omonona & Agoi [31]. Either way, education has a positive effect on food security at varying levels.

DISCUSSION

This paper identified predictors of food security among poor households across primary and secondary cities in Southern Africa. This investigation found that age, wage, household size, gender of the household head, dependency percentage, and education were indeed significant predictors of urban household food security, but with qualifications. The univariate analysis of each variable indicated that age, gender and education were best measured according to the household head when predicting household food security. The inclusion of dependency percentages and the transformation of household size also represented a significant finding.

In previous studies [12,34], dependency ratio was statistically insignificant. However, borrowing from the work of Garrett & Ruel [52], a quadratic transformation was applied to household size. Given the sign of the quadratic value, the transformation suggested that groups with larger households were more food insecure, but the rate that household size impacted food security changed as larger families were introduced. Garrett & Ruel [52] hypothesized a possible economy of scale to food costs as households grew sufficiently large. When dependency percent and household size were both incorporated into the same model, household size only retained statistical significance in this sample when it was recognized that a quadratic relationship existed. In turn, this validated policies that attempted to increase food security by focusing on family planning initiatives. However, further investigation into the causality of this relationship is required to make a concrete policy recommendation. Additionally, policies that relate to food security may want to give particular attention to household groups with larger family sizes. This falls in line with the dependency percentage variable which suggests that more members of the household who do not earn a wage relate to higher levels of food insecurity.

The second notable finding from this investigation was that the gender variable was not statistically significant once additional measurements were introduced. While additional research must be conducted to better understand this relationship, it is hypothesized that once the real-world conditions that affect both gender and food security are held constant (e.g., differences in wage and number of earners), the gender of the household head is of no statistical significance in this research sample (As suggested by the variable importance chart that consistently reports a 0 percent importance in gender in the model). Further research should be conducted to better understand the role of gender as it relates to food security. Beyond this, the final set of important variables relating to food security in urban cities across Southern Africa in descending order are: wage, number of dependents, completion of high school or post-secondary diploma, household size, and age of the household head, and the remaining possible levels of education.

Throughout the analysis, education continued to be a statistically significant variable, and while the exact extent to which education related to food security varied, it can be suggested that the decision to attend school contributes positively to one's food security outlook. Not only for themselves, but for the family as a whole. Since wage was held constant for every level of education it also elucidated that having an education had spillover effects that go beyond simply earning a higher wage (in most cases, as suggested by [31]). Education may contribute to making more informed decisions around purchasing behavior that supports higher levels of food security. This supports policy initiatives which attempt to mitigate the effects of food insecurity by means of increasing school attendance rates in the urban Southern African context, for example.

Lastly, wage is consistently the most important variable for predicting varying levels of food security. While food security revolves around the financial capacity to command enough food, researchers should understand it as a metric driven by a variety of factors that can be influenced to ensure a more equitable world. This finding also continues to validate the use of wage as a means of inferring broader human insecurity in urban areas and the importance of food access through markets over food production. Together, these findings contribute to the broader conceptual model of food insecurity by highlighting, and validating, predictors of urban household food access. In the context of a rapidly urbanizing Global South, this study empirically grounds the necessary evolution of food security scholarship by elucidating the urban experience of food insecurity. Within the framework of the Sustainable Development Goals, the findings from this investigation also highlight the network of variables relating SDG 2 (Zero Hunger). This paper suggests that the achievement of these goals will require effective monitoring and evaluation practices that appreciate the urbanized context of food insecurity in the Global South.

CONCLUSIONS

The SDGs represent an ambitious set of targets and goals that will compel transformational changes in global development [55]. Within this context, cities (as engines of demographic, ecological and economic change) are set to play a potentially catalytic role in the global achievement of these SDGs [56]. That said, the adaptation of the SDGs to the urban context will require the participation of relevant stakeholders from multiple sectors [57]. The implementation of the SDGs in cities may highlight synergistic relationships among the goals that could provide efficient policy pathways towards the achievement of the SDGs in the urban context [58].

This investigation provides novel insight into a potential network of relationships linking SDGs 4 (Quality Education), 5 (Gender Equality), and 8 (Decent Work and Economic Growth) to SDG 2 (Zero Hunger) among the urban poor in Southern Africa. While further research is needed, these results indicate the relative importance of each of the potential contributions made by these SDGs to the achievement of SDG 2 (and identified a potentially circuitous link between SDG 5 and SDG 2). In other words, these findings may identify influential SDGs that should be targeted when implementing these SDGs (and justifies the monitoring of multiple SDG outcomes on the basis of one implementation).

The relationship between household size and food security in this study suggests that food insecurity increases at an increasing rate for each additional household member. This finding runs contrary to previous findings that suggested economies of scale for feeding increasingly large households. Given this potentially exponential relationship, policy geared towards food security should reinforce social protection programs that support household dependents. Larger households are more likely to be food insecure, and fewer income earners in the household further spells financial challenge. Therefore, social safety nets like child grants and state pensions for those in old age can specifically target households whose higher percentage of dependents not only leads to financial challenge, but also lower levels of food security.

As well, policies geared towards increased access to health information and services as they relate to maternal health and reproductive rights not only reinforces informed and voluntary healthcare decisions, but could also improve a household's food security outlook. Municipal government should provide access to resources that aid women and their partners in making free, informed decisions for the timing and desired number of children for their family. Such programs should be domestically funded to avoid external management of internal affairs. A highlighted focus on parental autonomy helps preserve reproductive rights, while also making resources available that benefit child and maternal health and improve the likelihood of consistent adequate food consumption. These local efforts should fall in line with a broader coalition from higher-level political bodies to promote family wellbeing as it relates to food and nutrition security. Documents like the 2010 Maputo Plan of Action on Sexual and

Reproductive Health help lay the groundwork for an actionable strategy for such reproductive rights.

Secondary school completion is among the top predictors of food security in our analysis, suggesting that education not only increases one's own food security, but their household as well. This finding is in support for universal education, a priority that have been echoed to varying degrees through initiatives similar to Malawi's Free Primary Education policy or Zimbabwe's 1996 Education Act. Whether household heads earn more or make more informed decisions around consumption habits, the finding remains the same; investment in education benefits has spillover effects that reach the dinner table. With this in mind, policy involving education needs to be implemented in gradual, well-defined stages to ensure policy visions are transformative yet attainable.

An engendered dynamic to childhood school attendance rates suggest that gender must not be overlooked in the creation of educational policy [59]. Social protection programs geared to education including subsidized uniforms and afterschool feeding programs can specifically target disadvantaged groups, offering a small step towards maintaining school enrollment at the local level. More broadly, these findings provide initial guidance as cities in Southern Africa begin localizing and implementing the SDGs. Empirically, this area of research provides new application for broader machine learning techniques that could contribute to the identification of network relationships among the Sustainable Development Goals.

AUTHOR CONTRIBUTIONS

JS, BF and CM designed the study. BF managed the data collection. JS analyzed the data. JS, BF and CM wrote the paper with input from all authors.

CONFLICTS OF INTEREST

The authors declare that there is no conflict of interest.

FUNDING

The survey was funded by the Canadian International Development Agency (CIDA) through its University Partners in Cooperation and Development (UPCD) Tier One Program (CIDA Agreement No. S63441).

ACKNOWLEDGMENTS

The authors would like to thank the following for their contributions to AFSUN and the baseline survey: Caryn Abrahams, Ben Acquah, Jane Battersby-Lennard, Eugenio Bras, Mary Caesar, Abel Chikanda, Asiyati Chiweza, David Coetzee, Jonathan Crush, Belinda Dodson, Scott Drimie, Rob Fincham, Miriam Grant, Alice Hovorka, Florian Kroll, Clement Leduka, George Matovu, Sithole Mbanga, Chileshe Mulenga, Peter Mvula, Ndeyapo

Nickanor, Sue Parnell, Wade Pendleton, Akiser Pomuti, Ines Raimundo, Celia Rocha, Michael Rudolph, Shaun Ruysenaar, Christa Schier, Nomcebo Simelane, Joe Springer, Godfrey Tawodzera, Daniel Tevera, Percy Toriro, Maxton Tsoka, Daniel Warshawsky, Astrid Wood and Lazarus Zanamwe. The Food and Nutrition Technical Assistance Project (FANTA) is gratefully acknowledged for providing the methodology and questions used in this survey to collect food insecurity data.

REFERENCES

1. United Nations. Transforming our world: The 2030 agenda for sustainable development. New York (US): UN General Assembly; 2015 Sep 25. Resolution No.: A/RES/70/1.
2. Frayne B, Pendleton W, Crush J, Acquah B, Battersby-Lennard J, Bras E, et al. The state of urban food insecurity in southern Africa. Cape Town (South Africa): AFSUN; 2010. Urban Food Security Series Report No. 2.
3. Parnell S, Pieterse EA, editors. Africa's urban revolution. London (UK): Zed Books; 2014.
4. Palen JJ. The urban world. Oxford (UK): Oxford University Press; 2015.
5. Sonnino R. The new geography of food security: exploring the potential of urban food strategies. *Geogr J*. 2016;182(2):190-200.
6. Crush J, Battersby J, editors. Rapid urbanisation, urban food deserts and food security in Africa. London (UK): Springer; 2016.
7. Tacoli C. Food (in)security in rapidly urbanising, low-income contexts. *Int J Environ Res Public Health*. 2017;14(12):1554. doi: 10.3390/ijerph14121554
8. Tilman D, Clark M. Global diets link environmental sustainability and human health. *Nature*. 2014;515(7528):518-22. doi: 10.1038/nature13959
9. Food and Agriculture Organization of the United Nations. FAOSTAT statistics database. Available from: <http://www.fao.org/faostat/en/#home>. Accessed 2017 Nov 17.
10. Battersby J, Watson V. Addressing food security in African cities. *Nat Sustain*. 2018;1(4):153-5.
11. von Grebmer K, Bernstein J, Nabarro D, Prasai N, Amin S, Yohannes Y, et al. 2016 Global Hunger Index: Africa edition. Washington (US): International Food Policy Research Institute; 2017.
12. Gebre GG. Determinants of food insecurity among households in Addis Ababa city, Ethiopia. *Interdiscip Description Complex Syst: INDECS*. 2012;10(2): 159-73.
13. Frayne B, McCordic C. Planning for food secure cities: Measuring the influence of infrastructure and income on household food security in Southern African cities. *Geoforum*. 2015;65:1-11. doi: 10.1016/j.geoforum.2015.06.025
14. LeBlanc D. Towards integration at last? The sustainable development goals as a network of targets (DESA Working Paper No. 141). New York (US): United Nations Department of Economic and Social Affairs; 2015.
15. Food and Agriculture Organization of the United Nations. Rome Declaration on World Food Security and World Food Summit Plan of Action. World Food Summit; 1996 Nov 13-17; Rome, Italy. Rome (Italy): FAO; 1996.

16. United Nations. The Millennium Development Goals Report 2015. New York (US): United Nations; 2015. Available from: [http://www.un.org/millenniumgoals/2015 MDG Report/pdf/MDG%202015%20rev%20\(July%201\).pdf](http://www.un.org/millenniumgoals/2015%20MDG%20Report/pdf/MDG%202015%20rev%20(July%201).pdf). Accessed 2019 Apr 19.
17. Food and Agriculture Organization of the United Nations. The State of Food Insecurity in the World 2006: Eradicating world hunger—taking stock ten years after the World Food Summit. Rome (Italy): FAO; 2006.
18. Food and Agriculture Organization of the United Nations. The State of Food Insecurity in the World: Meeting the 2015 international hunger targets: taking stock of uneven progress. Rome (Italy): FAO; 2015.
19. Frayne B, Crush J, McLachlan M. Urbanization, nutrition and development in Southern African cities. *Food Secur.* 2014;6(1):101-12. doi: 10.1007/s12571-013-0325-1
20. Porter G, Lyon F, Potts D. Market institutions and urban food supply in West and Southern Africa: a review. *Prog Dev Stud.* 2007;7(2):115-34. doi: 10.1177/146499340600700203
21. Godfray HC, Beddington JR, Crute IR, Haddad L, Lawrence D, Muir JF, et al. Food security: the challenge of feeding 9 billion people. *Science.* 2010;327(5967):812-8. doi: 10.1126/science.1185383
22. Nelson GC, Rosegrant MW, Palazzo A, Gray I, Ingersoll C, Robertson R, et al. Food security, farming, and climate change to 2050: scenarios, results, policy options. Washington (US): Intl Food Policy Res Inst; 2010.
23. World Health Organization. Centre for Health Development. Hidden Cities: unmasking and overcoming health inequities in urban settings. Geneva (Switzerland): World Health Organization; 2010.
24. Pieterse E. Rethinking African urbanism from the slum. *Power.* 2011;26(14.1):40-8.
25. Mougeot LJ, editor. *Agropolis: The Social, Political and Environmental Dimensions of Urban Agriculture.* London (UK): Routledge; 2005.
26. Crush JS, Frayne GB. Urban food insecurity and the new international food security agenda. *Dev South Afr.* 2011;28(4):527-44. doi: 10.1080/0376835X.2011.605571
27. Vogel C, Smith J. The politics of scarcity: conceptualising the current food security crisis in southern Africa: commentary. *S Afr J Sci.* 2002;98(7):315-7.
28. Sidhu RS, Kaur I, Vatta K. Food and nutritional insecurity and its determinants in food surplus areas: the case study of Punjab state. *Agric Econ Res Rev.* 2008;21:91-8.
29. Bashir MK, Naeem MK, Niazi SA. Rural and peri-urban food security: a case of district Faisalabad of Pakistan. *World Appl Sci J.* 2010;9(4):403-11.
30. Babatunde RO, Omotesho OA, Sholotan OS. Socio-economic characteristics and food security status of farming households in Kwara State, North-Central Nigeria. *Pak J Nutr.* 2007;6(1):49-58.
31. Omonona B, Agoi G. An analysis of food security situation among Nigerian urban households: Evidence from Lagos State, Nigeria. *J Cent Eur Agric.* 2007;8(3):397-406.

32. Kabubo-Mariara J, Ndenge GK, Mwabu DK. Determinants of children's nutritional status in Kenya: evidence from demographic and health surveys. *J Afr Econ*. 2008;18(3):363-87.
33. Amaza PS, Umeh JC, Helsen J, Adejobi AO. Determinants and measurement of food insecurity in Nigeria: some empirical policy guide. International association of agricultural economists annual meeting; 2006 Aug 12-18; Gold Coast, Australia.
34. Mitiku A, Fufa B, Tadese B. Empirical analysis of the determinants of rural households food security in Southern Ethiopia: The case of Shashemene District. *Basic Res J Agric Sci Rev*. 2012;1(6):132-8.
35. Iram U, Butt MS. Determinants of household food security: An empirical analysis for Pakistan. *Int J Soc Econ*. 2004;31(8):753-66.
36. Asogwa BC, Okwoche VA, Umeh JC. Analysing the determinants of poverty severity among rural farmers in Nigeria: A censored regression model approach. *Am Int J Contemp Res*. 2012;2(5):166-76.
37. McBride L, Nichols A. Improved poverty targeting through machine learning: An application to the USAID Poverty Assessment Tools. 2015. Available from: econthatmatters.com/wp-content/uploads/2015/01/improvedtargeting_21jan2015.pdf. Accessed 2019 Apr 19.
38. Xie M, Jean N, Burke M, Lobell D, Ermon S. Transfer learning from deep features for remote sensing and poverty mapping. Thirtieth AAAI Conference on Artificial Intelligence; 2016 Feb 12-17; Phoenix, USA.
39. Jean N, Burke M, Xie M, Davis WM, Lobell DB, Ermon S. Combining satellite imagery and machine learning to predict poverty. *Science*. 2016;353(6301):790-4. doi: 10.1126/science.aaf7894
40. Jarosz L. Defining world hunger: scale and neoliberal ideology in international food security policy discourse. *Food Cult Soc*. 2011;14(1):117-39. doi: 10.2752/175174411X12810842291308
41. Swindale A, Bilinsky P. Household dietary diversity score (HDDS) for measurement of household food access: indicator guide. Version 2. Washington (US): Food and Nutrition Technical Assistance; 2006.
42. Coates J, Swindale A, Bilinsky P. Household food insecurity access scale (HFIAS) for measurement of food access: Indicator guide. Version 3. Washington (US): Food and Nutrition Technical Assistance; 2007.
43. Fox J, Monette G. Generalized collinearity diagnostics. *J Am Stat Assoc*. 1992;87(417):178-83.
44. Strobl C, Boulesteix AL, Kneib T, Augustin T, Zeileis A. Conditional variable importance for random forests. *BMC Bioinformatics*. 2008;9(1):307. doi: 10.1186/1471-2105-9-307
45. Breiman L. Random forests. *Mach Learn*. 2001;45(1):5-32.
46. Biffis E, Chavez E. Satellite data and machine learning for weather risk management and food security. *Risk Anal*. 2017;37(8):1508-21.
47. Okori W, Obua J. Machine learning classification technique for famine prediction. In Proceedings of the world congress on engineering; 2011 Jul 6; London, UK. Hong Kong (China): International Association of Engineers, 2011. p. 991-6.

48. International Monetary Fund. IMF Primary Commodity Prices, Monthly Data (table), IMF (database). Last updated February 01, 2018. Available from: <http://www.imf.org/external/np/res/commod/index.aspx>. Accessed 2018 Feb 18.
49. Minot N. Transmission of world food price changes to markets in Sub-Saharan Africa. Washington (US): International Food Policy Research Institute; 2010.
50. Hanke SH, Kwok AK. On the measurement of Zimbabwe's hyperinflation. *Cato J.* 2009;29:353.
51. Quisumbing AR, Brown LR, Feldstein HS, Haddad L, Peña C. Women: The key to food security. *Food Nutr Bull.* 1996;17(1):1-2.
52. Garrett JL, Ruel MT. Are determinants of rural and urban food security and nutritional status different? Some insights from Mozambique. *World Dev.* 1999;27(11):1955-75.
53. Weisberg S. Applied linear regression. Hobokon (US): John Wiley & Sons; 2005.
54. Bashir MK, Schilizzi S. Determinants of rural household food security: a comparative analysis of African and Asian studies. *J Sci Food Agric.* 2013;93(6):1251-8. doi: 10.1002/jsfa.6038
55. Kroll C. Sustainable Development Goals: Are the rich countries ready? New York (US): Sustainable Development Solutions Network; 2015.
56. Klopp JM, Petretta DL. The urban sustainable development goal: Indicators, complexity and the politics of measuring cities. *Cities.* 2017;63:92-7. doi: 10.1016/j.cities.2016.12.019
57. Zinkernagel R, Evans J, Neij L. Applying the SDGs to Cities: Business as Usual or a New Dawn? *Sustainability.* 2018;10(9):3201. doi: 10.3390/su10093201
58. Valencia SC, Simon D, Croese S, Nordqvist J, Oloko M, Sharma T, et al. Adapting the Sustainable Development Goals and the New Urban Agenda to the city level: Initial reflections from a comparative research project. *Int J Urban Sustain Dev.* 2019;1-20. doi: 10.1080/19463138.2019.1573172
59. Wamba N, Mgomozulu V. The crisis in public education in Malawi. *Int J Adv Res.* 2014;2(4):323-32.

How to cite this article:

Sgro J, Frayne B, McCordic C. Linking the Sustainable Development Goals through an Investigation of Urban Household Food Security in Southern Africa. *J Sustain Res.* 2019;1:e190004. <https://doi.org/10.20900/jsr20190004>