ajfand SCHOLARLY, PEER REVIEWED
AFRICAN JOURNAL OF FOOD, AGRICULTURE, NUTRITION AND DEVELOPMENT
Volume 17 No. 1
March 2017
PUBLISHED BY AFRICAN SCHOLARLY SCIENCE COMMUNICATIONS TRUST
ISSN 1684 5374

Afr. J. Food Agric. Nutr. Dev. 2017; **17(1):** 11477-11496 **DOI: 10.18697/ajfand.77.16680**

AN ADAPTIVE HOUSEHOLD SAMPLING METHOD FOR RURAL AFRICAN COMMUNITIES

Awuah R[1*], Douglass R[2], Agyepong S[3] and E Kuwornu[4]

Rebecca Awuah

[*]Corresponding author email: rawuah@ashesi.edu.gh

[1]Lecturer, Department of Arts & Sciences, Ashesi University, PMB CT3 Cantonments, Accra, Ghana

[2]Emeritus Professor of Health Administration, College of Health and Human Services, Eastern Michigan University, Ypsilanti, Michigan, USA

[3]Lecturer, Department of Business Administration, Ashesi University, Berekuso, Ghana

[4]Research Assistant, Department of Arts & Sciences, Ashesi University, Berekuso, Ghana

ABSTRACT

Investigators working in rural communities and small towns in Africa face many obstacles to obtaining a random and representative sample of households for their research. The civic infrastructure used as the building blocks of survey sampling in developed countries are mostly absent in rural Africa. The purpose of the study described in this paper was to pilot an innovative and cost-effective approach to household sampling designed to generate probability samples representative of the socio-economic diversity of the small town of Berekuso, in the Eastern Region of Ghana, without relying on existing census data, household registers, or a regular layout of roads and dwellings. Utilizing Google Earth images and a Graphical Information System (GIS) map of Berekuso, sampling units were defined as 15-degree wedge-shaped sectors radiating from the center of the original township. All households within randomly selected sectors were surveyed, and based on a household classification scheme, each household type was identified. Additional sectors were randomly selected and surveyed in sequence until no new household types were identified – a notion recognized by laboratory scientists as an ‘end point’. The adaptive sampling strategy was cost and time effective: freely available versions of Google Earth and QGIS software were employed along with inexpensive handheld Global Positioning System (GPS) devices; a total of 57 households were surveyed by teams of two enumerators over three consecutive Sundays. The survey method yielded a probability sample that is representative of the socioeconomic diversity of Berekuso, and produced generalizable results for median household size, median age of residents, sources of potable water and toilet types, among others. For example, based on the results of the survey, a 95% confidence interval estimate of the proportion of residents of Berekuso under the age of 20 years is between 0.49 and 0.58. These figures are consistent with results of Ghana’s 2010 census which pegged the proportion of the population of the Eastern Region under the age of the 20 years at 0.49. The authors believe that the methodology described in this paper may be applicable to household research in many rural African villages and small towns where little civic infrastructure exists to create more traditional sampling frames.

Key words: Google Earth, area sampling, household survey, impact evaluation, wealth ranking

ajfand SCHOLARLY, PEER REVIEWED
AFRICAN JOURNAL OF FOOD, AGRICULTURE, NUTRITION AND DEVELOPMENT
Volume 17 No. 1
March 2017
PUBLISHED BY AFRICAN SCHOLARLY SCIENCE COMMUNICATIONS TRUST
ISSN 1684 5374

INTRODUCTION

Whether for planning purposes, impact evaluation or longitudinal studies, many forms of research require data at the household level. If the household data are to be analyzed using inferential statistical methods, then obtaining a random and representative sample of households is essential. Investigators working in rural communities and small towns in Africa face many obstacles to obtaining such a sample: census data are aggregated to the district level rather than the town level; sections of the population are undocumented and without birth certificates, voter registration, or social security identities; dwellings do not have identifiable codes, such as house numbers; and there is no grid or formal structure of roads and dwellings. Nearly all of the building blocks of survey sampling, as dictated by well-defined and rigorous survey research in developed nations, are absent in rural Africa.

Yet the need for representative population data is more acute in rural Africa than in most other places. Growth and development in Africa depend on understanding migration patterns, changing age distributions, and shifting trends in health, education, and employment. Blind to the dynamics of populations when data are lacking, development planning often takes place anyway. This dilemma exists at the regional and national levels, but especially at the local level. Bringing contemporary development to a rural community requires documentation of need, accurate planning, and valid protocols for evaluation; the entire process depends upon current and reliable data. For these reasons, survey sampling methodology that is both scientifically adequate and practical is more than a theoretical consideration; such methodology is a necessary tool to identify, direct and evaluate planned development in rural African communities.

The global community and international development organizations offer little to this predicament. For instance, the excellent *WHO Immunization Coverage Cluster Survey Reference Manual* (2005) [1], subsequent WHO technical support manuals for survey sampling, or comparable publications from the United Nations, typically rely on elements of a member state's civic infrastructure from which to create sampling frames. These methods, developed in academic centers of industrialized nations, usually depend on registries for voting, vital statistics, licensure of drivers and vehicles, enrollments in schools, labor unions and other social organizations, and other resource lists that are commonplace in developed nations. Such methods have evolved over the last half century to produce precise and efficient samples of large and complex populations [1-4]. Unfortunately, such civic infrastructures are not commonplace in Africa, and for rural African populations with local development needs, such sophisticated tools are nearly irrelevant as noted by Zarkovich [5] and subsequently by many observers.

This paper describes an innovative survey sampling method that was developed for a longitudinal household survey in the village of Berekuso, in Ghana's Eastern Region. The purpose of the longitudinal survey, called The Berekuso Impact Study, is to generate a time-series dataset for multidisciplinary research on the demographic and socio-economic factors of change as Berekuso transitions from a rural agrarian settlement to an urban suburb of Greater Accra. The sampling method, joining a chorus of "adaptive" research methods world-wide, is drawn from the concept of "species-area graphs" and is

ajfand SCHOLARLY, PEER REVIEWED AFRICAN JOURNAL OF FOOD, AGRICULTURE, NUTRITION AND DEVELOPMENT Volume 17 No. 1 March 2017 PUBLISHED BY AFRICAN SCHOLARLY SCIENCE COMMUNICATIONS TRUST ISSN 1684 5374

an adaptation of quadrat or area sampling common in the field of ecology. The method is designed to generate independent probability samples representative of the socio-economic diversity of the town. Sampling units are 15-degree wedge-shaped sectors that radiate from the center of the original township. All households within randomly selected sectors were surveyed, and based on a household classification system described in detail below, each household type was identified. Additional sectors were randomly selected and surveyed in sequence until no new household types were identified – a notion recognized by laboratory scientists as an 'end point'. The authors believe that this method may be applicable to household level research in many small African communities and rural locations with little civic infrastructure from which to create more traditional sampling frames. The application of Google Earth, GPS devices and QGIS software to obtain a random sample of households in Berekuso is described here to offer a "cookbook" methodological tool for others involved with African development at the community level.

METHODS

Study Background

The leadership of Ashesi University, the newest "resident" of Berekuso in the Eastern Region of Ghana, recognized that it was incumbent upon the institution to measure the impact of imposing a modern university onto a traditional, rural Ghanaian community. The need to monitor the university's impact on the place and people of Berekuso evolved into The Berekuso Impact Study and is part of the stewardship and mission of the university. The objective of the pilot study was to conceptualize and test a method of repeated independent random selection and surveying of households in Berekuso, a rural community undergoing rapid changes. In Berekuso, traditional sampling frames, such as registers, house numbers, and formal structures for roads and dwellings, do not exist. A secondary purpose was to field test a survey methodology that could be applied to similar situations elsewhere in Africa.

Study Site

Berekuso is a traditional agricultural town of about 2,000 residents located in the coastal highlands 20 kilometers north of Accra, Ghana's capital city. The residential area of Berekuso is approximately 0.16 square kilometers, though farms, stool lands (meaning land ownership by extended traditional stewardship), and family land extend far beyond. Most families in Berekuso engage in small-scale farming of pineapple, maize, plantain, teak, cassava or vegetables. Other than agriculture, trade, small-scale manufacturing (cement blocks, iron works), and crafts (such as Kente weaving) are the traditional livelihoods of residents.

Ghana's 2000 census recorded the population of Berekuso at 1,391 residents [6]. However, much has changed in the past 16 years. A paved road was constructed through the town in the early 2000s in order to reroute traffic between Accra and the Eastern Region hub of Aburi during several years of re-construction on the Accra-Aburi highway. In 2011, Ashesi University, a private liberal arts college, opened a modern residential campus on the hill above the center of town, bringing with it over 500 student residents, daily commuters of faculty and staff, and new tenants. Although the road through

DOI: 10.18697/ajfand.77.16680

ajfand SCHOLARLY, PEER REVIEWED AFRICAN JOURNAL OF FOOD, AGRICULTURE, NUTRITION AND DEVELOPMENT Volume 17 No. 1 March 2017 PUBLISHED BY AFRICAN SCHOLARLY SCIENCE COMMUNICATIONS TRUST ISSN 1684 5374

Berekuso no longer carries much diverted Accra-to-Aburi traffic (in fact, most of the pavement has washed away), daily volume is likely to increase as development progresses. Real estate development in the area has accelerated as low housing prices attract workers from the capital, transforming the vicinity just south of Berekuso into a new suburb of Accra. In 1996, Berekuso was connected to the national electricity grid and in 2014 its small community health center upgraded to a sub-district level facility.

Berekuso inhabitants of Akan descent speak a dialect of Twi, but there are also significant numbers of Ga speakers, descendants of families who migrated from coastal areas during conflicts in the Ga Kingdom. Today, residents with Akan, Ga, and Ewe roots practice a combination of Christianity and traditional religion. Present-day Berekuso also has a mosque and a small number of Muslim inhabitants who are laborers with roots in Northern Ghana.

The broader Akwapem Ridge, the mountainous ridge around Berekuso, is sprinkled with many villages and small towns sharing similar demographic and agricultural roots. A central goal of the longitudinal study for which the methodology will be applied, is to examine over time an example of changing rural landscapes and rapid modernization. Recognizable patterns in the expansion and urbanization of Greater Accra gives urgency to the need for information and data to fully understand rapid changes in the "peri-urban zones" – areas undergoing rapid rural to urban transitions – of Ghana's capital city [7]. By collecting time-series data and comparing changes in Berekuso with one or more comparison towns in the Akwapem Ridge, it may be possible to identify patterns in positive and negative long-term impacts on people, families, livelihoods, and quality of life due to urbanization, infrastructure development, new or expanded social services such as schools and clinics, and in the case of Berekuso, a new tertiary institution relocating to the area.

The goals of The Berekuso Impact Study are: (1) to understand how the demographic and socio-economic level of residents of Berekuso are changing over time; (2) to see how the changes in Berekuso compare with one or more similar towns in the Akwapem South District; (3) to understand what changes, positive or negative, can be associated with Ashesi University's presence in Berekuso; and (4) to generate a rich time-series dataset for multidisciplinary research on factors of change for African communities transitioning from rural agrarian societies to urban suburbs of rapidly growing cities.

Development of an Adaptive Household Sampling Methodology

Two approaches to sampling often recommended for household surveys in developed and developing countries are stratified and systematic sampling. Possible sub-populations of interest for the Berekuso Impact Study are households at different socio-economic levels and new versus long-term residents of the town. However, the civic infrastructure needed to create such strata is not readily available. Systematic sampling was similarly ruled out because there is no reliable register of residents and no grid or regular structure to the layout of dwellings, roads, and paths in Berekuso necessary to select households in a systematic fashion. As noted by Rozelle [8], data collection in rural areas within developing countries needs to deal with the problems of sample frames. According to Rozelle, such frames, if they exist at all, typically are inaccurate, incomplete, often have duplicate measures, are rarely current and are too expensive to maintain. In the absence of information necessary for stratified and systematic sampling based on traditional sampling frames, it was necessary to explore other methods.

As noted by Casale *et al*. [9] regarding survey research methods in Africa, the research team needed to be innovative and adapt its methodology to realities in the field. The dilemma is not particularly new, however, and was clearly stated by Carrothers and Chambers [10]:

> "In rural development and rural research, there is a tension between two approaches to understanding: that of the academic community, interested more in detail, precise observation and measurement and rigorous and respectable methodology and with a generally rather unhurried concern with knowledge for its own sake in the longer term (for example these papers, the original versions of which were written in 1979, emerged in published form only in 1981, in contrast to the need for valid information for practical purposes)." [10].

This was elaborated at the operational level by Casale *et al*. [9] in 2014:

> "However, conducting health survey research in remote rural areas comes with its own set of unique hurdles. In effect, the on-the-ground reality of conducting field research can be challenging in all settings, particularly in resource-scarce areas of the developing world. However, working in isolated rural areas may be especially daunting, as a result of factors such as relative geographic isolation, limited services and distrust of outsiders. Unless addressed effectively, challenges that arise during the course of field research may threaten to delay and even jeopardize the overall success of data collection."

Transect sampling, a traditional field sampling technique in ecology, field biology and agriculture, has proved useful in some household surveys where address information is either not available or too expensive and time-consuming to obtain [11, 12]. The transect method, as applied to household surveys, randomly selects starting positions from a map and then all households are surveyed, or households are surveyed systematically, along a defined transect line. One drawback to this approach is that transect lines can overlap. Getis *et al*. [13], echoed by Morrison *et al*. [14], noted that the transect method could be subject to geographical bias when similar types of households are in proximity to one another.

An alternative approach used by researchers conducting a malaria indicator survey in Malawi employed Google Earth and GIS software to identify and enumerate all structures at the community level and randomly select a sample of households. This methodology was thoroughly described by Escamilla *et al.* [19] and by Haenssgen [20]. Although the method is cost effective because it takes advantage of freely available satellite image technology to generate a sampling frame, the approach depends on the timeliness of images on Google Earth. In a place that is experiencing rapid economic and

ajfand SCHOLARLY, PEER REVIEWED AFRICAN JOURNAL OF FOOD, AGRICULTURE, NUTRITION AND DEVELOPMENT Volume 17 No. 1 March 2017 PUBLISHED BY AFRICAN SCHOLARLY SCIENCE COMMUNICATIONS TRUST ISSN 1684 5374

demographic change, these images may be obsolete. A further challenge noted by Pearson *et al*. [21] in 2015 is the fact that business and other structures, such as sheds and latrines, cannot be distinguished from similar sized household structures visible in the satellite image. At the time of the Berekuso household survey, Google Earth images of the town were two years old, and many new structures were easily identified as missing.

Further, as a strategy for repeated sampling over time, which is an essential element of community planning and evaluation, depending on current Google Earth images is problematic. It is unlikely that Google Earth, or an equivalent source, would provide uniformly spaced snapshots of a specific community over a five-to-ten-year time span. Time-series analyses rely on uniformly spaced observations for internally-valid studies such as the interrupted times series quasi-experimental designs that would be ideal for applications to specific communities in rural Africa.

The final approach – a modified form of area sampling and an adaptation from field studies in ecology – was chosen to meet the needs of the repeated cross-sectional, time-series design of The Berekuso Impact Study and to overcome the challenges of sampling without dependence on the elements of civic infrastructure. Area sampling uses maps rather than lists to define the sampling frame but does not require individual dwellings to be identified prior to going into the field. Geographical units, called clusters, are delineated using a map, a random sample of clusters is obtained and every household in the cluster is surveyed by enumerators in the field.

Study Method

The target populations for the Berekuso Impact study are households and residents of Berekuso and households and residents of one or more comparison towns located in the South Akwapem Ridge area over a time span of five to ten years. Possible comparison towns for the longitudinal study are of similar or smaller population and area. However, for purposes of this methodology pilot, only households in Berekuso were included.

The goal was to adopt a household survey method that was valid, sufficiently rigorous to be reliable when repeated at regular intervals over a time span of five to ten years, and practical in terms of cost, including labor and time. A variant of the area sampling approach, in which clusters are operationally defined as 24 wedge-shaped radial sectors layered on top of a Google Earth image of Berekuso, was adapted for the study [Figure 1]. A small number of sectors were randomly selected and the household survey administered to each dwelling in the sector. Enumerators in the field used GPS devices to stay within the sector boundaries. It is important to note that this method is not equivalent to area or cluster sampling of large, complex, national or continental populations; its application is to a specific community in a limited space.

The concept of "species–area graphs" from the field of ecology was employed to determine how far to continue sampling sectors in sequence until an overall diversity in types of households was captured in the sample – an end point. The principal objective of any sampling strategy is to draw a representative sample that is as efficient as possible. In this project, it was realized that building materials and household types – sizes and

DOI: 10.18697/ajfand.77.16680

ajfand SCHOLARLY, PEER REVIEWED
AFRICAN JOURNAL OF FOOD, AGRICULTURE, NUTRITION AND DEVELOPMENT
Volume 17 No. 1
March 2017
PUBLISHED BY AFRICAN SCHOLARLY SCIENCE COMMUNICATIONS TRUST
ISSN 1684 5374

family living arrangements – offered a structural variable that not only defined Berekuso, but also reflected change and evidence of development. Identifying and classifying residential structures created an analogy to a biological species variable, used in ecology studies to determine abundance and diversity. The end point for sampling is defined as that point when measured diversity of the sample (cumulative number of unique “species”) is not changed by surveying additional samples. Applying transects in the context of survey sampling was developed by Ghirotti [22] in 1985, based on previous work by Carothers and Chambers [10]. These methods were used in agricultural inventory studies in the central highlands, Sidamo and Shoa regions in Ethiopia, the Kafue flats of Zambia, the Quiché area of Guatemala, Owamboland of Namibia, the Boké district of Guinea and the Nouhao Valley of Burkina Faso.

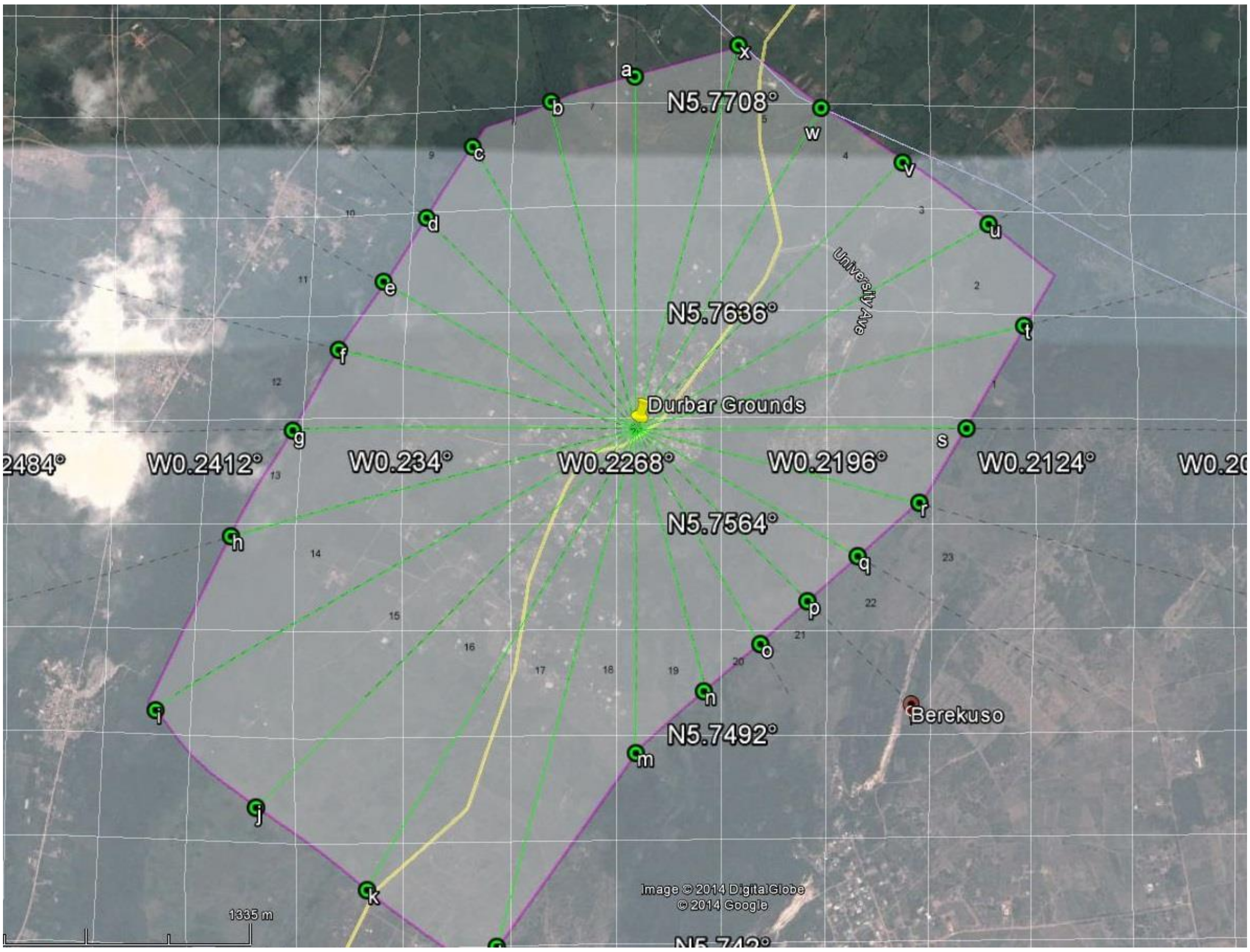

Figure 1: Sampling frame consisting of 24 wedge-shaped radial sectors layered on Google Earth image of Berekuso. Berekuso’s main road and the relative density of structures are visible

The sampling frame used for the pilot study consists of an aerial photo of Berekuso obtained from Google Earth with a Graphical Information System (GIS) map layered on top, showing the one main road through the town and the relative densities of structures in different parts of the town [Figure 1]. The Durbar Grounds, a small park in the center of Berekuso containing no dwellings, was identified as a convenient central point.

DOI: 10.18697/ajfand.77.16680

ajfand SCHOLARLY, PEER REVIEWED AFRICAN JOURNAL OF FOOD, AGRICULTURE, NUTRITION AND DEVELOPMENT Volume 17 No. 1 March 2017 PUBLISHED BY AFRICAN SCHOLARLY SCIENCE COMMUNICATIONS TRUST ISSN 1684 5374

Twenty-four fifteen-degree wedge-shaped sectors radiating from the central point were constructed using QGIS software, layered on top of the map, and numbered. The radial sectors capture household diversity (household income, long-term versus new residents) by radiating from the center of the original village along the natural transition from old family compounds to new homes. Individual sectors are too geographically narrow to encompass all of an extended family cluster. There is some risk that rare characteristics may not appear in a sample on a given year. Methodologically, the dilemma is that, within a small community, the intentional selection of small minorities could threaten the objectivity presumptions of the sample. It might be necessary to supplement survey data with qualitative studies, such as focus groups, to ensure that small, unique, subsets of the population are represented, but not included in statistical analyses for purposes of generalization.

A series of random numbers from 1 to 24 was generated in Excel and every dwelling with its front entrance in the first three randomly selected sectors was surveyed. Handheld GPS devices loaded with the sector map were used by enumerators to stay within the boundaries of the selected sector. Each household was surveyed and its household type classified using a two-dimensional classification scheme [Figure 2]. The two-dimensional household classification scheme, by material type (mud on stick frame, mud with rendering, block without rendering, block with rendering, fully finished) and family type (single family, extended family, rental units, family + rental, extended family + rental) was created and tested for Berekuso, resulting in up to 30 possible household types. Prior to going into the field, the household classification system was tested and refined by a research team member with expertise in the field of building technology in Ghana. Research team members and enumerators were trained to accurately and objectively classify dwellings. Repeated surveys over time could add to this classification if new building materials or family type considerations emerge.

Berekuso Household Type Classification System

	(1) Single Family	(2) Multi/Extended Family	(3) Multiple Rental Units	(4) Single Family + Rental	(5) Multi/Extended Family + Rental	**Total**
(A) Mud on Stick Frame	1	1	0	0	0	**2**
(B) Mud with Rendering	4	7	1	0	2	**14**
(C) Block w/out Rendering	8	2	0	2	0	**12**
(D) Block with Rendering	9	6	1	7	2	**25**
(E) Fully Finished	1	1	0	2	0	**4**
Total	**23**	**17**	**2**	**11**	**4**	**57**

Figure 2: Number of households surveyed for each household type

Dwelling typologies have been used successfully in many studies to rank the relative wealth of households [23]. According to Ghirotti [22], for households in sub-Saharan Africa, cash availability may influence the type and quality of the building materials, the number of rooms in the house, the possibility to subdivide space and the provision of accessories. In the context of rural Ghana and much of rural Africa, nuclear family dwellings are traditionally part of extended family compounds and this is still very much evident in the present layout of Berekuso. However, one of the visible ways current residents are seeking to benefit from recent growth and development in the area is by

ajfand SCHOLARLY, PEER REVIEWED
AFRICAN JOURNAL OF FOOD, AGRICULTURE, NUTRITION AND DEVELOPMENT
Volume 17 No. 1
March 2017
PUBLISHED BY AFRICAN SCHOLARLY SCIENCE COMMUNICATIONS TRUST
ISSN 1684 5374

subdividing space or adding rooms and amenities to existing structures, such as plumbing and generators, for the purpose of generating rental income. Newcomers with access to capital have leased parcels of land in the outlying area previously used for farming, and have built dwellings both large and small.

After administering the household survey to all dwellings in a sector, a graph of the number of distinct household types as a function of the number of sectors surveyed was generated. This is analogous to using a species area graph in the field of ecology, often used to determine the minimum number of samples that should be taken as truly representative of a particular habitat [24-26]. In this sampling process, one cluster at a time is surveyed in the field. From data gathered from the first two or three clusters, the cumulative frequency curve of species type rises sharply, as each newly sampled cluster reveals one or more species not previously sampled. However, such a curve will eventually level off to an end point when sampling additional clusters reveals no new species. In the case of The Berekuso Impact Study, when the slope of the cumulative household type versus sector graph flattened out after the 4^{th} and 5^{th} sectors were surveyed (in other words no new household types were recorded) the sample was considered representative of household diversity in Berekuso [Figure 3]. This was determined to be the cluster selection end point when five of the wedge-shaped sectors were sampled, giving each household a probability of 5/24 or 20.8% chance of being included in the survey. A total 17 types of households were observed among 57 total households surveyed.

ajfand SCHOLARLY, PEER REVIEWED AFRICAN JOURNAL OF FOOD, AGRICULTURE, NUTRITION AND DEVELOPMENT Volume 17 No. 1 March 2017 PUBLISHED BY AFRICAN SCHOLARLY SCIENCE COMMUNICATIONS TRUST ISSN 1684 5374

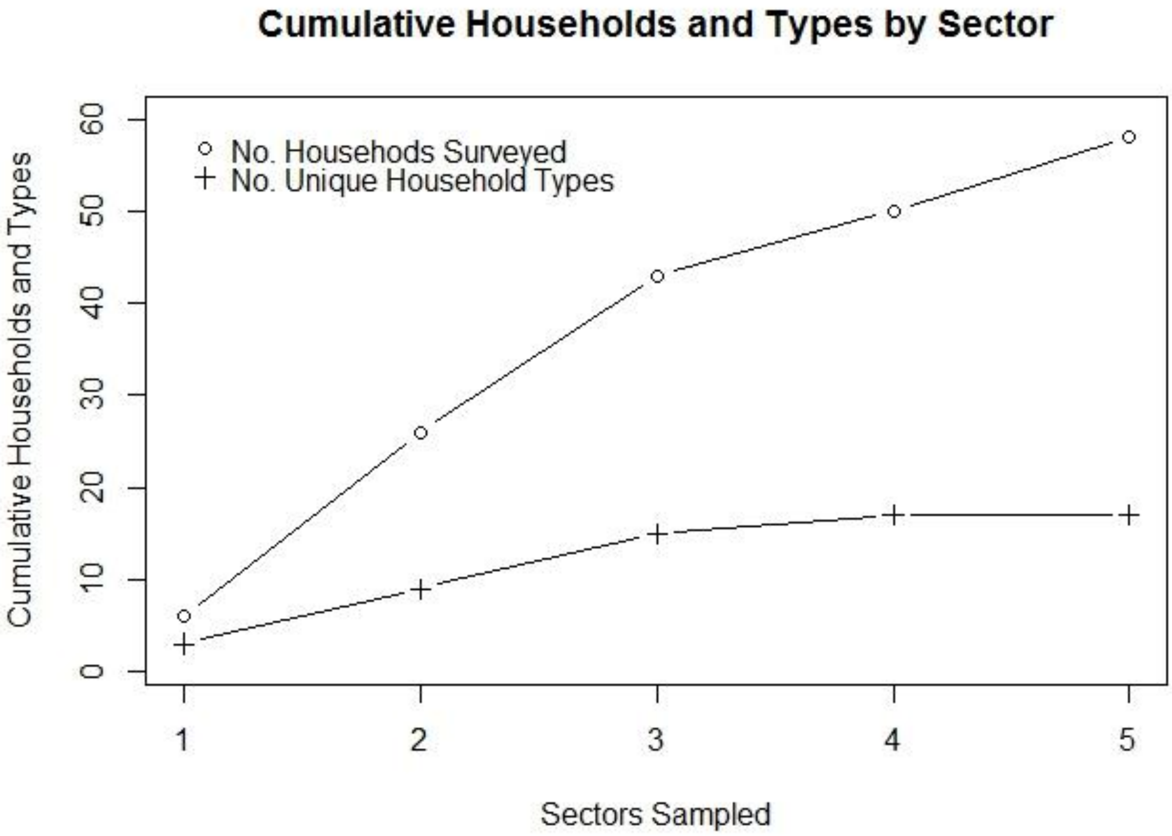

Figure 3: Cumulative households and cumulative household types by sector

Method Justification

The importance of the end point in the cluster selection is not just to avoid the costs associated with additional field work and data collection but also because the process is to be repeated over time. With repeated surveys over several years, the study could risk "response burden," as described by Creel *et al*. [27] , associated with individual respondents recalling their responses to previous survey interviews and conditioning their responses to what they might recall, or think they remember, from previous interviews. This threat is also called "testing/re-testing" by many social scientists [28]. In other words, The Berekuso Impact Study sought equally representative samples, over time, from a small community, that would avoid response burden threats as much as possible. Avoiding oversampling in small communities is one way to avoid response burden. Sampling efficiency was defined, for the purpose of this study, as the minimal number of selected clusters necessary to fully represent the household diversity of the population as represented by the classification of dwelling types.

It was also realized that insights about Berekuso's evolution based on dwelling conditions can be considered a proxy indicator for household income and economic and social status. Changes in family living arrangements, from embracing traditional extended family compounds to favoring newer single-family dwellings, are expected to be an important indicator of change in Berekuso. Measures of social diversity are needed to understand transformations taking place as a result of the rural-to-suburban transition and changes associated with new economic opportunities. The Berekuso Impact Study will need to monitor a shift from agriculture to other livelihoods, increased family income, wage-earning roles of multiple family members, and many other factors, in order to understand the evolutionary social dynamics of Berekuso.

Finally, the process of sampling and surveying proved efficient in terms of time, human resources, and cost. Generating the sampling frame using Google Earth and QGIS took about two weeks as the process and software were new to the investigators. Another couple of weeks were spent testing the Garmine Navigation Etrex 30 GPS devices in the field and training enumerators**.** Each sector took teams of two trained enumerators between two and four hours to survey, depending on the density of households. The entire data collection process took place over three Sundays, plus one weekday needed to return to the small number of dwellings where residents were originally not at home. The basic versions of Google Earth and QGIS software used in the study are freely available**.**

RESULTS

The following observations are based on household data collected during this pilot study. The purpose of the present project was to field test the sampling methodology with a modest data collection instrument. Future data collection will include substantially more variables and more complex constructions, such as knowledge and beliefs about health, nutrition, economic options and opportunities, social or political perspectives, personal finances, education and technical training interests, and more. The following analyses demonstrate that generalizable findings from this methodology can be achieved.

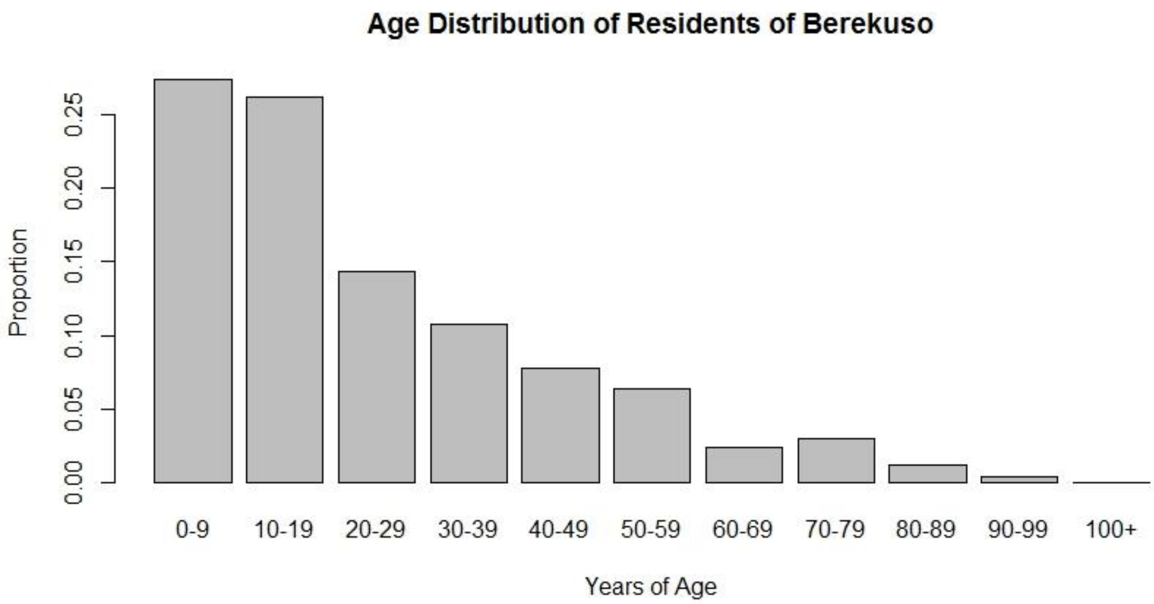

Figure 4: Household survey age distribution

The 95% confidence interval estimate of the proportion of residents of Berekuso who are less than 20 years of age was between 0.49 and 0.58 [Figure 4]. These figures are

ajfand SCHOLARLY, PEER REVIEWED
AFRICAN JOURNAL OF FOOD, AGRICULTURE, NUTRITION AND DEVELOPMENT
Volume 17 No. 1 March 2017
PUBLISHED BY AFRICAN SCHOLARLY SCIENCE COMMUNICATIONS TRUST
ISSN 1684 5374

consistent with results of Ghana's 2010 census which pegged the proportion of the population of the Eastern Region under the age of the 20 years at 0.49 [6].

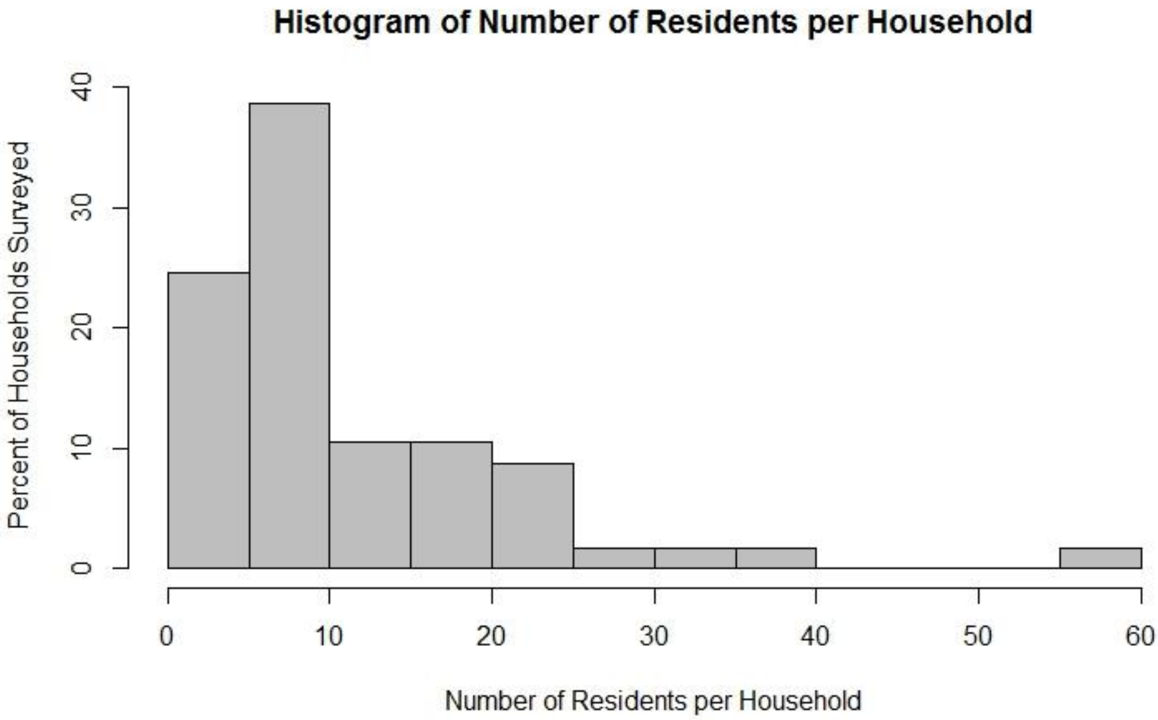

Figure 5: Household size

The sample median number of residents per household [Figure 5] was 8 people, with the middle 50% of households having between 6 and 16 residents. A 95% confidence interval gave an estimate for the median number of residents per household in Berekuso between 7 and 12 individuals. The average (mean) household population for the Eastern Region reported in the 2010 census is 6.1 [6].

ajfand SCHOLARLY, PEER REVIEWED
AFRICAN JOURNAL OF FOOD, AGRICULTURE, NUTRITION AND DEVELOPMENT
Volume 17 No. 1
March 2017
PUBLISHED BY AFRICAN SCHOLARLY SCIENCE COMMUNICATIONS TRUST
ISSN 1684 5374

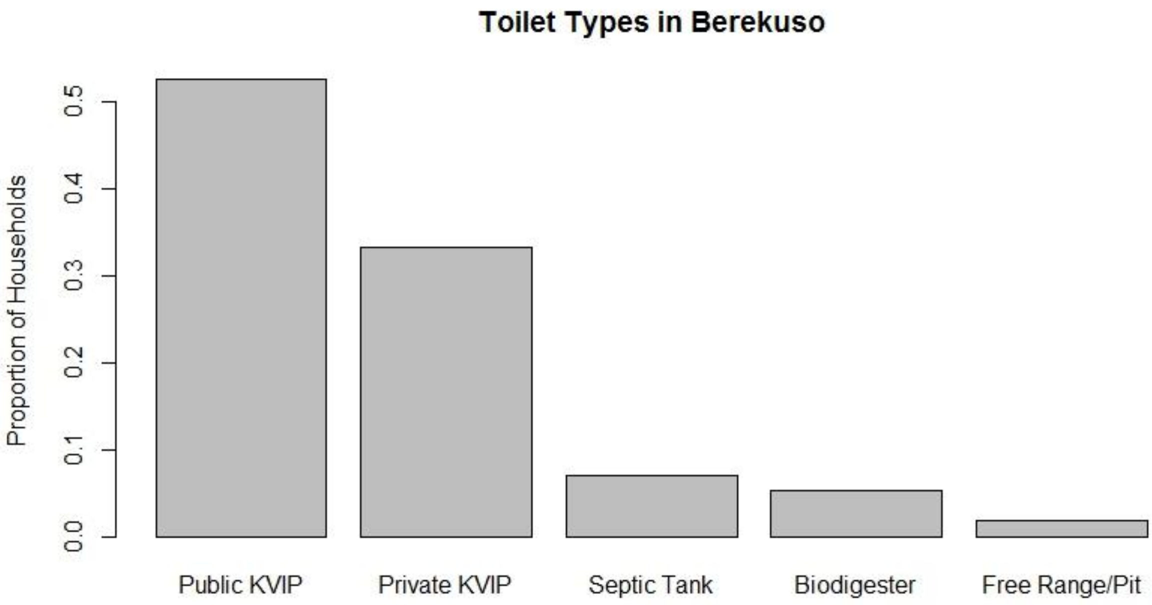

Figure 6: Toilet facilities used in Berekuso

The most common toilet facility used by households in Berekuso was the public KVIP (Kumasi Ventilated Improved Pit) [Figure 6]. The sample proportion of households in Berekuso who use the community KVIP as their primary toilet was 0.53. The 95% confidence interval estimate of the proportion of households using the community KVIP was between 0.39 and 0.66 households.

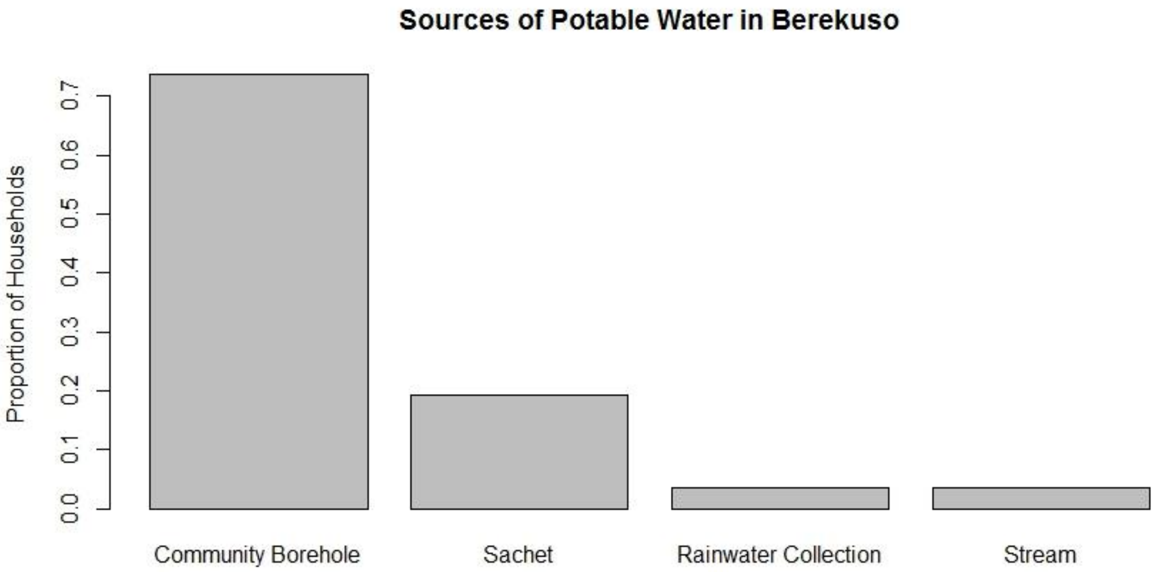

Figure 7: Household sources of potable water

ajfand SCHOLARLY, PEER REVIEWED AFRICAN JOURNAL OF FOOD, AGRICULTURE, NUTRITION AND DEVELOPMENT Volume 17 No. 1 March 2017 PUBLISHED BY AFRICAN SCHOLARLY SCIENCE COMMUNICATIONS TRUST ISSN 1684 5374

At the time the survey was conducted there were three community borehole sites in Berekuso and these were the primary source of potable water for residents. The sample proportion of households who use a community borehole as their primary source of potable water was 0.74 [Figure 7]. A 95% confidence interval estimate of the proportion of households using the community borehole as their primary source of potable water was between 0.60 and 0.84 households.

DISCUSSION

The sampling and data collection methods employed in the pilot study appear to meet long-term goals of The Berekuso Impact Study. The adaptive sampling strategy was cost-and-time-effective. The probability sample of 57 households in Berekuso was surveyed over three Sundays plus one additional weekday. Enumerators working in pairs took approximately 16 hours to collect the survey data. The images and software required for generating the sampling frame are freely available online, though the Google Earth and QGIS software took some time to learn.

The innovation of applying a household classification system to the concept of "species–area graphs" from ecology yielded a probability sample that is representative of the overall diversity of types of households in Berekuso. Importantly, the authors believe that the household classification system, which takes into account materials used in the construction and maintenance of dwellings and the type of family compound, offers a reasonable proxy for diversity of household income and traditional versus contemporary living arrangements in Berekuso.

The research team intends to deploy the sampling methodology described in this paper at regular intervals in Berekuso and one or more comparison towns over a five-to-ten-year period to meet the goal of establishing a longitudinal and multidisciplinary database. A repeated cross-sectional design was selected over a panel design for two primary reasons. First, Berekuso is experiencing rapid migration: middle-class families are arriving due to low cost housing and new roads that connect the area to Accra; Ashesi University has brought jobs to Berekuso and construction workers, security officers, janitorial and catering staff have taken up residence in and around the town; and Ghanaian elites are beginning to develop parcels of land on the scenic hillsides. At the same time, many young people leave the town every year for further education or job opportunities elsewhere. A randomly selected panel of households for a longitudinal study would quickly become unrepresentative due to the constant flow of families, workers and young people moving in and out of Berekuso. Second, it was felt that the issue of panel conditioning (respondent burden or testing/re-testing threats), when respondents recall questions and responses from earlier waves of the study, could affect the objectivity, reliability and validity of data over several waves of interviews. Therefore, a cross-sectional quasi-experimental design that gives regular "snapshots" of the demographic and socio-economic level of residents of Berekuso, and one or more comparison towns, appears better suited than a panel design to the Berekuso Impact Study's long-term objectives.

ajfand SCHOLARLY, PEER REVIEWED AFRICAN JOURNAL OF FOOD, AGRICULTURE, NUTRITION AND DEVELOPMENT Volume 17 No. 1 March 2017 PUBLISHED BY AFRICAN SCHOLARLY SCIENCE COMMUNICATIONS TRUST ISSN 1684 5374

The field experiences that have been discussed here required a modest commitment of time and labor. By drawing from volunteer faculty members and students of Ashesi University's Research Club, there was no need for a personnel budget, but even if that were not the case, costs would have been very limited compared to a typically-financed, formal survey based on methodological guidelines that would reflect the state of the art in a highly developed nation with long traditions of civic infrastructure. But there were other issues, common to all field research, which also had to be addressed. As noted by Casale *et al*. [9]:

> "No matter the extent of scientific rigor, consultation and planning, unexpected obstacles will inevitably emerge during field research, and these will require adaptive strategies to be overcome. In particular, conducting survey research in remote rural areas of the developing world may pose unique challenges that require appropriate context-relevant responses. Unfortunately, text books and the current body of scientific papers do little to equip researchers for the experience that awaits them."

The "context-relevant" responses in this project included research design and technology, but also the softer side of research with human populations that is too often conducted with an arrogance defined by the researchers' or official agency's need for information. The goal of repeating the data collection process over a number of years would stumble and fail if the team did not proceed with concern for the privacy and traditions of Berekuso's population.

Before taking to the field, the team made certain to discuss the effort with Berekuso's traditional leaders. A qualitative focus group study of the perceived needs of Berekuso residents was conducted and distributed, and members of the enumeration team were oriented to the task of remaining rigorous and objective, yet respectful of the community. It was important to acknowledge that, just as imposing Ashesi University onto the Berekuso community, the solicitation of interviews within a sample of people's homes was invasive and foreign to the experiences of most respondents.

CONCLUSION

This pilot study demonstrates that it is both possible and advisable, when considering research methods and technologies in Africa, particularly in rural areas, to extend thinking beyond conventional social science, public policy, and methodological standards or habits in order to collect badly needed data for practical purposes with severe economic constraints. By shifting away from traditional fixed sampling frames, permanent random number assignments to individuals, and dependence on civic infrastructures that are routine in well-developed nations, to methods common to field ecology, biology and agriculture, useful research regarding human communities was accomplished. This should not be considered a radical approach; applied statistics and sampling methods all are pedigreed to academic agriculture and botany in the mid-20th Century. Others have called for innovation and adaptive, interdisciplinary perspectives, as seen in the work of Fortrell and Byass [29] and Shannon *et al*. [30]. There are many models from a wide range of fields that could, and should, inform researchers' methodologies with human populations, such as the work of Alhaji and Babalobi [31], Wampler *et al*. [32], Pearson *et al*. [21] or LaCon *et al*. [33]. In a sense, in order to "get the job done" without blind adherence to research methodological dogma of the global community's expectations, it is believed that this pilot study shows that good science, in the public interest, need not be avoided with the excuse of minimal resources. This project produced transparent methodology, replicability, and timely and efficient execution. Field training and supervision produced objective and verifiable data. Is this not the goal of innovative and adaptive, interdisciplinary attention to pressing needs throughout Africa and other developing regions of the world?

It is the intention of The Berekuso Impact Study to replicate this work, over time, to produce a useful and valid time-series database for planning and evaluation in Berekuso. By publishing the outputs of this research, it is hoped, that other researchers will try this methodology in the field, throughout Africa, and begin to produce data for planning and evaluation that will contribute to economic and social development without taxing local resources that are needed for immediate purposes such as food security, potable water availability, sanitation, public education, decent housing, and health care. Research methods should exhibit the same fervor for adaptability and innovation as all other disciplines, looking for what is right and adequate in particular places and communities today. This view is consistent with the ethos and mission of Ashesi University which motivated The Berekuso Impact Study: to educate Africans to develop innovative solutions to Africa's problems, and to lead by example.

Human Subjects Protection

The household survey instrument and sampling methods were reviewed by the Ashesi University's Human Subjects Review Committee before any data collection was initiated. The approval process included obtaining informed consent (verbal consent) from the head of household prior to the household's participation in the survey. The approved informed consent included the ability of any respondent to elect to participate in full, or in part, as an individual decision.

Acknowledgments

Authors would like to acknowledge the contributions and assistance with data collection of Ashesi University colleagues Charlie Jackson, Stephen Armah, Joseph Oduro-Frimpong, and Edward Opoku.

REFERENCES

1. **World Health Organization.** Immunization Coverage Cluster Survey: Reference Manual. 2005. WHO/IVB/04.23.

2. **Kish L** Sampling Methods for Agricultural Surveys. FAO Statistical Development Series 3, FAO, Rome. 1989.

3. **United Nations.** Household Sample Surveys in Developing and Transition Countries. Studies in Methods, Series E, No. 96. New York: United Nations, Statistics Division, Department of Economic and Social Affairs, 2005.

4. **Levy PS and S Lemeshow** Sampling of Populations: Methods and Applications. Hoboken, NJ, John Wiley & Sons, Inc., 2013.

5. **Zarkovich SS** Some Problems of Sampling Work in Underdeveloped Countries. **In**: Bulmer M and DP Warwick (Eds). Social Research in Developing Countries: Surveys and Censuses in the Third World. London: UCL Press, 1993; **3**: 103-110.

6. **Ghana Statistical Service.** 2010 Population and Housing Census: Summary Report of Final Results, 2012. http://www.statsghana.gov.gh/docfiles/2010phc/Census2010_Summary_report_of_final_results.pdf. Retrieved June 8, 2016.

7. **Petra DP and CY Oduro** Patterns of Population Growth in Peri-Urban Accra, Ghana. *International Journal of Urban and Regional Research*. 2012; **36**(**6**): 1306-1325. doi: 10.1111/j.1468-2427.2011.01075.x.

8. **Rozelle S** Rural Household Data Collection in Developing Countries: Designing Instruments and Methods for Collecting Farm Production Data. Working Papers in Agricultural Economics, Cornell University, Department of Applied Economics and Management, 1991.

9. **Casale M, Lane T, Sello L, Kuo C and L Cluver** Conducting Health Survey Research in a Deep Rural South African Community: Challenges and Adaptive Strategies. *Health Research Policy and Systems.* 2013; **11**: 14. Retrieved June 9, 2016 doi: 10 1 186/1478-4505-11-14.

10. **Carothers I and R Chambers** Rapid Appraisal for Rural Development. *Agricultural Administration.* 1981; **8**: 407-422.

11. **Ma L, Stein ML, Wang M, Shelton AO, Pfister CA and KJ Wilderb** A method for unbiased estimation of population abundance along curvy margins. http://www.stat.uchicago.edu/~meiwang/research/pubpreview.pdf, doi: 10.1002/env.1053. Retrieved June 8, 2016.

ajfand SCHOLARLY, PEER REVIEWED AFRICAN JOURNAL OF FOOD, AGRICULTURE, NUTRITION AND DEVELOPMENT Volume 17 No. 1 March 2017 PUBLISHED BY AFRICAN SCHOLARLY SCIENCE COMMUNICATIONS TRUST ISSN 1684 5374

12. **Maxwell BD, Backus V, Hohmann MG, Irvine KM, Lawrence P, Lehnhoff EA and LJ Rew** Comparison of Transect-Based Standard and Adaptive Sampling Methods for Invasive Plant Species. 2012; **5(2):** 178-193. http://weedeco.msu.montana.edu/publications/pdfs/Maxwell_etal_adaptive_2012.pdf Retrieved June 8, 2016.

13. **Getis A, Morrison AC, Gray K and TW Scott** Characteristics of the Spatial Pattern of the Dengue Vector, *Aedes aegypti*, in Iquitos, Peru. *American Journal of Tropical Medicine and Hygiene.* 2003; **69**: 494–505. doi: 10.1007/9783642-019760_15.

14. **Morrison ML, Block WM, Strickland MD, Collier BA and MJ Peterson** Wildlife Study Design. 2nd ed. New York, Springer-Verlag, 2008.

15. **Morrison ML, Marcot BG and RW Mannan** Wildlife-habitat Relationships: Concepts and Applications. 3rd ed. Washington, DC, Island Press, 2006.

16. **Melville GJ and AH Welsh** Line Transect Sampling in Small Regions. *Biometrics.* 2001; **57**: 1130-1137.

17. **Cavada N, Barelli C, Ciolli M and F Rovero** Primates in Human-Modified and Fragmented Landscapes: The Conservation Relevance of Modelling Habitat and Disturbance Factors in Density Estimation *PLoS ONE.* 2016*;* **11**(**2**): e0148289. doi:10.1371/ journal.pone.0148289.

18. **Yoccoz NG, Nichols JD and T Boulinier** Monitoring of Biological Diversity in Space and Time. *Trends in Ecology and Evolution.* 2001; 446–453. doi: 10.1016/S0169-5347(01)02205-4.

19. **Escamilla V, Emch M, Dandalo L, Miller WC, Martinson F and I Hoffman** Sampling at Community Level Using Satellite Imagery and Geographical Analysis. *Bulletin of the World Health Organization*. 2014; **92**(**9**): 690–694. doi:10.2471/BLT.14.140756.

20. **Haenssgen MJ** Satellite-aided Survey Sampling and Implementation in Low- and Middle-income Contexts: a Low-cost/Low-tech Alternative. *Emerging Themes in Epidemiology*.2015; **12**: 20. doi**:** 10.1186/s12982-015-0041-8.

21. **Pearson AL, Zwickle A and AT Rzotkiewicz** Using Remote, Special Techniques to Select a Random Household Sample in a Dispersed, Semi-nomadic Pastoral Community: Utility for a Longitudinal Health and Demographic Surveillance System. *International Journal of Health Geographics*. 2015; **14**: 33. DOI 10.1186/s12942-015-0026-4.

22. **Ghirotti M** Rapid appraisal: benefiting from the experiences and perspectives of: livestock breeders. 1992: Retrieved June 10, 2016 at: http://www.fao.org/livestock/agap/frg/feedback/war/v1650b/v1650b0d.htm.

23. **Malleson R, Asaha S, Sunderland T, Burnham P and MA Egot Ekuri** Methodology for Assessing Rural Livelihood Strategies in West/Central Africa: Lessons from the Field. *Ecological and Environmental Anthropology*. 2008: 1, Retrieved June 10, 2016 http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1026&context=icwdmeea

24. **Krebs CJ** Ecological Methodology. Harper Collins Publishers, New York, 1989.

25. **Corker B** Ecological Sampling Methods: Lessons from the Offwell Ecological Study. Retrieved June 2014 from http://www.countrysideinfo.co.uk/biol_sampl_cont.htm.

26. **Buckland ST, Borchers DL, Johnston A, Henrys PA and TA Marques** Line Transect Methods for Plant Surveys. *Biometrics.* 2007; **63**:989-998.

27. **Creel D, Jang D, Kasprzak D and T Williams** Permanent Random Number Technique to Minimize Response Burden in Repeated Surveys. Joint Statistical Meetings - Section on Survey Research Methods. U.S. Department of Defense, Washington, DC, 2002.

28. **Campbell DT and JC Stanley** Experimental and Quasi-Experimental Designs for Research. Rand McNally, Chicago, 1966.

29. **Fortrell E and P Byass** Population survey sampling methods in rural African setting: measuring mortality. *Population Health Metrics*. 2008; **6**:2. Retrieved June 8, 2016 doi: 10.1186/1478-7954-6-2 2008.

30. **Shannon H, Hutson R, Kolbe A, Stringer B and T Haines** Choosing a Survey Sample When Data on the Population are Limited: a Method Using Global Positioning Systems and Aerial and Satellite Photographs. *Emerging Themes in Epidemiology*. 2012; **9**: 5. doi:10.1186/1742-7622-9-5. 4.

31. **Alhaji NB and OO Babalobi** Participatory Epidemiology of Ethno-veterinary Practices Fulani Pastoralists Used to Manage Contagious Bovine Pleuropneumonia and Other Cattle Ailments in Niger State, Nigeria. 2015 Mar 23. doi: 10.1155/2015/460408.

32. **Wampler P, Rediske R and A Molla** Using ArcMap, Google Earth, and Global Positioning Systems to Select and Locate Random Households in Rural Haiti. *International Journal of Health Geography*. 2013; **12**: 3. doi:10.1186/1476-072X-12-3. 5.

33. **LaCon G, Morrison AC, Astete H, Stoddard ST, Paz-Soldan VA, Elder JP, Halsey ES, Scott TW, Kitron U and GM Vasquez-Prokopec** Shifting Patterns of *Aedes aegypti* Fine Scale Spatial Clustering in Iquitos, Peru. *PLoS Neglected Tropical Diseases. 2014;* **8**(**8**): e3038. doi:10.1371/journal.pntd.0003038.