Impact of Malawi's School Meals Program on Primary Education

Author: Jaime Fellers

Boston College Electronic Thesis or Dissertation, 2015

Impact of Malawi's School Meals Program on Primary Education

Jaime Fellers
May 2015

Advisor: Associate Professor Paul Cichello

Thesis submitted to the faculty of Boston College in partial fulfillment of requirements for Bachelor of Arts in Economics with Honors from the School of Arts and Sciences and for the Boston College Honors Program

Table of Contents

List of Abbreviations

Community Informant (CI)
Dedza, Ntcheu, and Salima (DNS)
Government of Malawi (GoM)
Malawian Kwacha (MWK)
Malawi's School Meal Program (SMP)
Ministry of Education, Science, and Technology (MoEST)
National Statistics Office (NSO)
Non-Governmental Organization (NGO)
Ordinary Least Squares (OLS)
Purchasing Power Parity (PPP)
Primary School Leaving Certificate (PSLC)
Propensity Score Matching (PSM)
Randomized Controlled Trial (RCT)
School Meals Program (SMP)
Take-Home Ration (THR)
World Food Program (WFP)

Abstract

Initially launched as a pilot program in 1996 by the World Food Program (WFP) at the request of the Government of Malawi (GoM), the School Meals Program (SMP) reached approximately 642,000 primary school children by 2011. According to the WFP, the objectives of the SMP are: 1) reduce drop out rates; 2) promote regular attendance; 3) increase enrollment; and 4) improve children's ability to concentrate and learn, through food provision (WFP, 2010). Given these aims, this paper aims to determine if Malawi's SMP affects the primary enrollment rate or attendance as measured as an impact on temporary withdraws. By applying a propensity score matching (PSM) model to the Third Integrated Household Survey data from 2010-2011, the estimation of the impacts will aim to mitigate selection bias using historic enrollment and other covariates, which include WFP selection criteria and theory-based community and political characteristics. The findings of the paper are contrary to the majority of studies that explore the impacts of SMPs on education outcomes. Using three different matching techniques, the model predicts that the SMP has no impact on primary enrollment and a statistically insignificant, but positive impact on attendance, here measured as a decrease in temporary withdraws. Explanations for these atypical results include the presence of exclusion errors, which were found in the pilot evaluation, model misspecification, and the lack of social desirability bias in my measures. Further research is needed to determine the extent to which previous results have been biased by Hawthorne effects or social desirability bias. Given the potential of the temporary withdraws for highlighting a positive impact of the program, further studies should include this measure as a potential outcome of any SMP program, especially in agrarian economies.

Introduction

According to the World Bank Development Indicators, Malawi's adjusted net enrollment rate[1] was 97% in 2009; yet, performance indicators reveal that the education system still has plenty of room for improvement. The literacy rate among adults (ages 15+) was only 61% in 2010, with the literacy rate for adult females at only 51%. While enrollment is nearly perfect, many children do not make it to the final year of schooling: the primary completion rate was 68% and the secondary rate was a staggering 18%.[2] The median number of years children attend primary school has remained constant from the 1990s through 2014 at six (compared to completion which ideally occurs after 8 years). Note that this does not necessarily mean people are reaching Standard 6, as grade repetition is extremely common. National poverty lines indicate approximately 51% of the population lived in poverty as of 2010, but according to the international standard of $2 (PPP) a day, 88% of the population lived in poverty. Not surprisingly, approximately 23% of the population was undernourished in 2010 (World Bank 2009, 2010). These combined statistics show a strong need for an intervention to bolster educational achievement and improve nutrition.

The WFP has adopted SMPs as one of their main platforms for increasing enrollment and attendance, providing a safety net during times of crisis, supporting local agriculture, and to a lesser extent, improving nutrition. Each year, WFP facilitates the provision of school meals to between 20 and 25 million children across 63 countries. In

[1] Measured as the percent of children in the official primary school age window (6-13) who are currently enrolled in primary or secondary school.

[2] Calculated by dividing the number of new entrants, not including repeaters, in the last grade of primary (secondary) education, by the population at the entrance age for the last year of primary education.

1996, the GoM requested the WFP's support in launching an SMP. Beginning as a pilot program in one district, the SMP was eventually taken over by the Ministry of Education, Science, and Technology (MoEST) and converted into the four-year School Feeding Development Program in 2008. With technical assistance from the WFP, the SMP reached an estimated 642,000 primary school children, approximately 30% of primary-aged children, in 13 districts by 2011. The WFP's targeting is done first on a district-level according to food insecurity, enrollment levels, and dropout rates, with specific schools chosen according to local conditions such as accessibility by road, availability of storage facilities and potable water, and evidence of community commitment to participate. All pupils in the selected schools receive a mid-morning hot serving of corn soya blend porridge each school day (WFP, 2010).

An initial evaluation of the pilot program showed a 5% increase in enrollment and a 36% increase in attendance following the introduction of school meals. However, further investigation found that not all of this increase in enrollment was new students, but rather students migrating from other schools that did not receive the feeding program, leading to disorder and disruption of classes (WFP, 1996). From 2010-2011, Nkhoma et. al. evaluated the nutritional impact of the SMP and found improvements in catch-up growth in lean muscle mass and certain types of learning. However, given their small sample size of one control school and one treatment school, these results should not be extrapolated to the entire program. Additionally, although the researchers conducted a wide array of cognitive testing, including measures of learning, set-shifting, memory, and attention, only one measure was significantly impacted (Nkhoma, et. al., 2013).

International literature strongly supports the educational benefits of SMPs, and to a lesser extent, the nutritional value provided to children. This paper will build on the existing literature by conducting an evaluation of an SMP that does not use data directly collected for the purpose of evaluating the SMP. This ensures that the results are less prone to the social desirability bias that causes teachers or village leaders to report commendable outcomes. This is particularly applicable here, as the WFP provides a plethora of funding for social programs and perceived successes may encourage increased investment. Additionally, due to the size of the survey, many community-level variables can be estimated that would not be included in a typical evaluation, such as a lagged education indicator, historic attendance pre-program implementation. As most previous studies are based on randomized controlled trials (RCTs), this paper may capture a more realistic effect since there are no researchers ensuring the food is properly delivered. This paper will also directly focus on the proclaimed objectives of the SMP, including enrollment and attendance. Unfortunately, the survey does not include attendance data in the typical sense, and therefore will have to be approximated through the percent of students who have temporarily withdrawn from school in the past 12 months, causing them to miss two or more consecutive weeks of class. The measure will therefore be an estimation of extended absences instead of daily fluctuations in children attending school. This may, however, be positive, as Malawi, given their agrarian economy and therefore cyclical labor patterns, may want to focus on reducing the number of children being pulled out of school for work for long periods of time. This would not only hold the overall achievement of the class behind, but would also contribute to grade repetitions and worse learning outcomes.

Literature Review

Aside from the unpublished WFP reviews of Malawi's SMP, both during the pilot phase and at a few intervals since, the only study conducted by academics on the program occurred from 2010 to 2011. The study was an impact evaluation on cognitive and anthropometric outcomes of 226 school children aged six to eight for one school year. The team found a statistically significant difference in the increase of the middle-upper arm circumference between the SMP school and the non-SMP school. However, no significant difference was found in height or weight, which is not particularly surprising given the short time period over which the study was conducted. Children at the SMP school also performed significantly better in reversal learning exercises, although they performed equally well in tests of memory and attention (Nkhoma, et. al., 2013). Unfortunately, these children were only located at two schools, calling into question the overall validity of the findings. Typically, international literature that reviews nutritional impacts of SMPs are RCTs conducted by the researchers; some find significant improvements (Ash, et. al., 2003; Grillenberger, et. al., 2003; Powell, et. al., 1998.) while others do not find any impact whatsoever (Abrams, et. al., 2003; Simeon, 1998; Chandler, et. al., 1995; Van Stuijvenberg, et. al., 1999). Given the mixed results and difficultly of isolating the nutritional impact without conducting a randomized trial, this paper will shift focus from nutritional benefits to education impacts.

Contrary to the findings of international literature on SMPs with regard to nutrition, education-based evaluations typically reveal strong positive impacts of SMPs. These outcomes are loosely divided into three categories, with most studies evaluating the SMPs impact on all three: enrollment, attendance, and performance indicators.

Results are mixed for performance-based indicators, and extremely difficult to compare across countries and different types of assessments. Studies from Kenya, Bangladesh, and Peru among others, have found significant, both statistically and economically, improvements to attendance (Omwami, et. al., 2011; IFPR, 2004; Jacoby, et. al., 1998). In Bangladesh, an SMP RCT was found to raise enrollment by 14.2%, reduce the probability of dropping out of school by 7.5%, and increase attendance by about 1.3 days a month (IFPR, 2004). The Peru study was another RCT, which was conducted in the first year of the program's rollout. The breakfast, distributed in school, improved attendance and lowered dropout rates (Jacoby, Cueto, & Pollitt, 1998). While RCTs have the potential to reduce confounding effects and can directly and accurately gather important information, the execution of the program during the trial and in reality may be extremely different and may display better outcomes than would be typically observed.

A rigorous evaluation of what is often considered one of the oldest and most efficiently managed SMPs was conducted using the Chilean education department's administrative data from 2001-2005. The program began in 1980 and by 2000, covered nearly a third of all primary-aged children. Since the evaluation was conducted 20 years after the program's implementation, any effects should be interpreted as the long-run benefits of an SMP in the Chilean context. The author used a regression-discontinuity model to estimate the program's impact on enrollment, grade repetition, attendance, and test scores on the national fourth-grade assessment and found no statistically significant difference in any variables over the various groupings of students. According to the author, these results are not particularly shocking given Chile's high enrollment rates and relative lack of malnutrition (McEwan, 2013). Although the conditions are obviously different in Malawi currently, these results are more promising than they appear, and may show that an SMP can be so effective, that it is eventually no longer necessary.

Data

a. General Description of Data

This study will utilize data from the Third Integrated Household Survey, which was collected from March 2010 to March 2011 by Malawi's NSO. The team collected detailed information on a sample of 12,288 households from 768 communities. The sample was statistically designed to be representative at the national, district, urban, and rural levels. In addition to a household questionnaire, community informants (CIs) were interviewed, averaging 8.44 per community. These CIs were predominantly headmen (20.3%), counselors to the headmen (17.8%), religious leaders (10.4%), and businessmen/women (9.7%). Approximately 70% of the CIs were male and ranged in age from 8 to 99, with 50% between the ages of 33 and 55. The CIs were better educated than the majority of the population, with 47.6% having attained at least a Primary School Leaving Certificate (PSLC).

This study will utilize community-level data, some provided directly by CIs and some aggregated from household data, to determine the difference in enrollment rates and temporary withdraw rates between communities with and without an SMP. Exactly 16 households were randomly chosen in each of the 768 communities, giving a randomized sample with at most 768 observations. Of the 31 observations that drop out due to missing variables, 27 of them are due to unreported distance to the nearest primary school. This could potentially skew our results if CIs are not reporting the distance because it is extremely far. Additionally, these communities are almost exclusively from the two northern-most districts in Malawi, Chitipa and Karonga, and have statistically significantly higher rates of poverty and enrollment. Of these, only one community had both CIs and households report an SMP. The final four observations drop out for varied reasons and should not cause any bias.

b. Dependent Variables—Enrollment

The first dependent variable is primary school enrollment, which this study defines as the percent of primary-aged children currently enrolled in school. Children are technically supposed to start school at age six and should therefore complete primary school at age 13, thus the official primary-aged window is 6-13. Children are classified as enrolled if they, or their parents, reported they are currently attending school or, if school was not in session at the time of the interview, attended school in the session just completed and plan to enroll next session.[3] Enrollment is estimated per community, and is a low estimate of the enrollment rate given the frequency of delayed starts.

Figure 1: Primary Enrollment Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
enrollment	737	.8810949	.1405121	.0588235	1

With these criteria, the average primary school enrollment rate is 88.1% in the sample, which is far different from the various reported enrollment rates touted by the GoM. Urban communities have an average primary enrollment rate of 94.6% while rural communities show much lower rates with an average of 86.6%. The correlation between primary enrollment and average poverty gap and ultra poverty gap of the community, as defined by national standards, is -0.40 and -0.34 respectively. This is unsurprising as

[3] For Standard 8, students were considered enrolled if they attended school in the last completed session.

poorer families typically send their children to school less out of economic necessity, because they either cannot afford school fees or need their children to supplement the family's income. However, it is shocking that the ultra poverty indicator does not have a larger impact on enrollment than simple poverty.

c. Dependent Variables—Attendance

To test the effectiveness of the program in improving attendance, this paper will estimate attendance through the number of students who have temporarily withdrawn from school, missing at least two consecutive weeks, in the past 12 months. To control for the differences in numbers of students, this is converted to a percent of primary-aged children that are currently attending school. The mean percent is 5.3 and the maximum percent is 100, with the majority (455) having no students temporarily withdraw in the past year. The majority of withdraws were due to students' illnesses (40.3%), followed by having no money for necessary expenses (38.9%) and distantly by having to help at home (5.1%). Withdraw rates are slightly higher in urban areas, 5.8% compared to 5.2%, although the sample is considerably smaller.

Figure 2: Percent Temporary Withdraws Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
p_withdraw	737	.0532735	.0976008	0	1

d. Independent Variables—Treatment

CIs reported SMPs across 25 districts in 219 communities, representing 28.5% of all communities.[4] However, the WFP only operated in 13 districts at this time, meaning 67 of the communities (30.6%) who reported having an SMP did not get them through

[4] For a detailed breakdown of the number of reported SMP schools per district, refer to Appendix A.

the WFP. Information is limited on the GoM's role in implementing these programs. It is uncertain if a non-sponsored SMP would have the same impact, and therefore these observations should be distinguished from other communities. Unfortunately, we are not confident if communities have actually received WFP assistance in the districts in which they work. Measurement error is certainly a confounding factor, as if only one CI in the sample reported an SMP, schools were identified as having the program. This is a major problem for the data and certainly impacts the quality of the results. However, there is another method of identifying SMP communities through household questionnaires that should help to clarify the origin of the program.

As part of the extensive interview process, the head of household was asked about their receipt of benefits from an extensive list of government-sponsored programs, including the SMP. Ideally, due to the placement of the question in this section, parents will only report government-sponsored, WFP-backed programs. In the communities where CIs reported SMPs, the percentage of households with children currently enrolled in primary school who reported an SMP is 54.4%. Just observing these communities, the percent of households reporting SMPs is 46.7% in non-WFP districts and 57.8% in WFP districts, a statistically significant difference. These figures appear to show that the placement of the question on the household questionnaire was beneficial in identifying government SMPs, however the aforementioned measurement error could play a large role in misidentification by CIs. The only caveat that could potentially refute this is differences in SMP effectiveness. The typical parent is more likely to recall or have knowledge of an SMP if the program is reliable, well defined, and possibly in conjunction with other development initiatives in the community. This type of program is also more likely to have an impact on behavior, since if the system is unreliable, it will not be enough incentive for families to enroll their children in school, knowing they will be more valuable if they are able to earn an income or help with household duties.

The household questionnaire does not ask about the presence of an SMP, but rather if households have actually received aid from the program in the past 12 months. The wording of the question is also ambiguous, as parents may not include in-school feeding as aid received since the household does not seemingly directly benefit from it. Putting this aside, there may be a way to determine if household or CI responses are impacted by the effectiveness of the program. The next question asks parents how many months in the past year they have received this aid, a measure that should give a good indication of reliability. To properly determine if the percent of months children receive food impacts CI or household reporting, OLS regressions were run on both indicators using all the independent variables as controls. The reliability measure is obviously only available for communities where at least one household reported an SMP and additionally answered the second question, reducing the sample size available. The percent of months children received food unsurprisingly significantly impacts both indicators of SMPs, meaning people are more likely to report a program if it is effective.[5] Interestingly, households' responses are less swayed by effectiveness; yet again the significance of these results is blurred by the fact that CIs responses are binary and could be swayed by one uninformed CI. Putting this aside, these results could be explained by CIs, typically headsmen or local leaders, desire to appear effective if he or she believe delays in distribution will be attributed to her or him. In this case, it would be more prudent to

[5] For detailed OLS output, view Appendix B.

deny the presence of an SMP so surveyors do not investigate further. This again leads us to question the quality of CIs' responses, especially given the potential measurement error, and encourages a more rigorous protocol for assigning communities to either the treatment or control group.

In summary, the percent of households with children enrolled in primary school that reported receiving aid from an SMP might be subject to bias caused by knowledge gaps, although is seemingly less influenced by reliability. Community leaders are more likely to be informed about all NGO and government activity, but may be more subject to purposeful misleading as a result of the effectiveness of the program. Finally, it is difficult to distinguish WFP/GoM programs from other programs in the areas in which the WFP does operate. Communities will therefore be placed in the treatment or control group according to a rigorous algorithm detailed in Appendix C that takes into account all the above information.

Classifying communities according to the algorithm gives us 637 control and 131 treatment communities, or 17.1% treatment. However, as mentioned previously, 31 observations drop out, leaving 737, of which 608 are control and 129 are treatment, or 17.5% treatment. The average poverty gap to the national poverty line and ultra-poverty line in the control group is 13.9 and 4.5 respectively and 18.1 and 7.1 in the treatment. Although poorer communities are expected in treatment communities, the gap is far less than expected given that government selection is conditioned on poverty levels. This may be due to communities being classified as in the treatment group although they do not have a government-backed program. Communities in the non-WFP areas who would otherwise be classified as having an SMP have significantly lower levels of poverty.[6] Since these schools also have higher levels of enrollment, the estimator could be biased towards zero if classification is not accurate, as these communities would not be aiming to bolster overall enrollment and will have similar rates to non-SMP schools. Selection is additionally conditioned on historic enrollment rates; as defined above, the average historic attendance rate is 58.0% in the control group and 46.9% in the treatment. By our definition, primary enrollment is 88.8% in the control group and 85.0% in the treatment. Rates of temporary withdraws have much smaller variation, however, differences can be seen between treatment and control. The average withdraw rate is 5.5% in the control and 4.6% in the treatment. These percentages are contrary to what we would expect given the other statistics and the supposed targeting objectives. However, this is not conditioned on many other factors, and the difference is not significant enough to cause concern. The differences between non-treated communities and SMP communities is summarized in the following table:

[6] For detailed t-test output and discussion, view Appendix D.

Figure 3: Summary Statistics by Treatment Status

-> TR = 0

Variable	Obs	Mean	Std. Dev.	Min	Max
gap_poor	612	13.88308	10.79419	0	58.10249
gap_ultra	612	4.466805	5.4875	0	36.41922
historic_a~d	612	.5806637	.2345521	0	1
enrollment	612	.8870638	.13887	.0588235	1
p_withdraw	612	.0549242	.0992436	0	1

-> TR = 1

Variable	Obs	Mean	Std. Dev.	Min	Max
gap_poor	125	18.4903	13.67559	0	67.30995
gap_ultra	125	7.335303	8.65912	0	48.4246
historic_a~d	125	.4637373	.1990979	.047619	.9375
enrollment	125	.8518709	.1453497	.2142857	1
p_withdraw	125	.0451916	.089051	0	.4

e. Independent Variables—Covariates

The independent variables are broken down into five distinct categories: Expressed Targeting Objectives of the SMP, Connectedness, School Characteristics, Community Characteristics, and Government Attention. The WFP's targeting is done first on a district-level according to food insecurity, enrollment levels, and dropout rates, with specific schools chosen according to local conditions such as accessibility by road, availability of storage facilities and potable water, and evidence of community commitment to participate. Enrollment levels and dropout rates will be estimated using a historic attendance variable. Since food insecurity cannot be directly included as it is hopefully positively impacted by the presence of an SMP, poverty measures, specifically the poverty gap and the ultra poverty gap will provide covariates that will be correlated with food consumption. Finally, the feasibility of implementation indicator will be the percent of months the main road is accessible by a lorry, which is the main method with which food would be transported. In summary, the four variables in the Expressed Targeting Objectives category are historic attendance (historic_attend), the average poverty gap (gap_poor), the average ultra poverty gap (gap_ultra), and the percent of months the main road is accessible by a lorry (p_access). Historic attendance is an extremely important variable not only for predicting treatment, but is also an essential covariate for the equations predicting enrollment and temporary withdraw rates. This figure captures a great deal of information about a community's attitude towards education that cannot be gathered in a simple survey. This also serves as parents' education level, which is a very important variable for predicting individual educational achievement. The lagged variable is calculated as the percent of adults older than 28 who completed at least Standard 4. The age was calculated based on the minimum age a respondent could be given they were in the last year of primary school (age 13) when the program began in 1996, which is done to ensure complete independence. According to World Bank figures already mentioned, the average student completed six years of schooling at this time. Standard 4 was chosen to account for the frequency of grade repetitions. Given this definition, the average historic attendance rate is 56.4% overall, with a rate of 49.8% in rural areas and 82.7% in urban areas. The poverty and ultra poverty gap is calculated as the consumption shortfall relative to the poverty line and ultra poverty line respectively. The poverty lines are national standards, calculated according to current food prices, caloric requirements, and non-food expenses.[7] The

[7] The poverty line is 37,002 MWK and the ultra-poverty line is 22,956. Households who are ultra-poor do not have the necessary amount of consumption to meet a minimum of 2,400 kilocalories per person per day.

average poverty gap is 14.6 Malawian Kwachas (MWK), and the average ultra poverty gap is 5.0 MWK in the sample. Only 49 communities have an average poverty gap, and hence ultra poverty gap, of zero, and 112 communities have positive values for the poverty gap but zero for the ultra poverty gap. These figures are necessarily highly correlated, as the ultra poverty gap will typically increase with an increase in the poverty gap in a particular community. The gaps are significantly larger in rural areas, 17.2 verses 3.6 for the poverty gap and 5.9 verses 0.86 for the ultra poverty gap. The poverty gaps are highly negatively correlated with the historic attendance rates, -0.60 for the poverty gap and -0.47 for the ultra poverty gap. The final variable, percent of months the community is accessible by lorry, averages 82.8%, with rates significantly higher in urban areas.

The next grouping of variables aims to determine a community's level of Connectedness. The first variable is a dummy that takes the value of one if a member of the community is currently the Member of Parliament for Constituency, labeled resMP. The next variable is another dummy that determines the presence of a non-governmental organzation (NGO). If the CIs report an NGO that provides bed nets or cares for chronically ill, either Tuberculosis or HIV/AIDS patients, the dummy takes a value of one. The final variable is a continuous measure of the distance to the nearest doctor or clinical officer (dist_doc). To reduce the skewedness of the variable, the maximum distance is set at 100 km. With this restriction, the average distance is 23.3 km in rural areas and 5.7 km in urban areas.

The third group of variables is School Characteristics. Controlling for the size of the school is perhaps challenging if the number of students is influenced by the presence of an SMP. Therefore, the size of the school will be estimated through the logarithm of the number of teachers at the nearest government primary school (log_teachers), which is less likely to be influenced by the presence of an SMP. The next characteristic is whether or not the nearest primary or secondary school is electrified (electric). Nearly 90% of the communities are near an electrified primary or secondary school. The final variable is the distance of the nearest primary school to the community (dist_prim) as is similarly capped at 20 km to reduce skew. This variable will certainly influence enrollment and is therefore essential to the estimations.

The next set of variables aims to capture all the Community Characteristics that could impact enrollment or treatment status. The first variable is the logarithm of the number of polygamous marriages in a community (log_poly), which is frequently used as a gage for social progressiveness. The next variable, which clearly has a large impact, is an urban dummy (urban). The next few variables aim to capture the economic life in the communities. The first set is the percent of households who reported they were negatively impacted by 1) drought; 2) an economic shock, either unusually high costs of agricultural inputs or unusually high prices for food; and 3) conflict or violence (p_drought, p_econ_shock, p_conf_shock). The final variable is a dummy that takes on a value of one if the area typically has a strong agricultural sector (big_agri). This is determined by the main occupation of community residents, and if people typically move into the community during certain times of the year to look for agricultural work. There are 251 communities, or 34% of the sample that classify as agricultural hubs. Given these communities are all in rural areas, they have lower poverty gaps and ultra poverty gaps on average than the typical rural community.

The final set of variables, Government Attention, addresses the fact that the most worthy communities, unfortunately, are not necessarily the communities that receive the program. These decisions are often influenced by political will and village leader connections. The first variable, the percent of registered voters that voted in the previous election (voters), indicates the level of importance a community carries during elections. The values range from 1, nearly none, to 5, nearly all. This is perhaps an idealistic measure that assumes politicians are swayed predominantly by the needs of their constituents. The second variable therefore aims to capture prior government attention by controlling for whether or not the community has a government works program sponsored by the Malawi Third Social Action Fund (MASAF), which finances community projects and transfers cash to unemployed participants. The ability to secure funds certainly shows a level of clout that may influence whether or not a school has an SMP. The summary statistics, ordered as presented in the above section, follow:

Figure 4: Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
TR	737	.1750339	.3802542	0	1
gap_poor	737	14.6645	11.45551	0	67.30995
gap_ultra	737	4.953321	6.228276	0	48.4246
historic_a~d	737	.5608322	.2329759	0	1
p_access	737	.8288105	.2881162	0	1
resMP	737	.1289009	.3353178	0	1
NGO	737	.7367707	.4406849	0	1
dist_doc	737	19.93712	21.57681	0	100
log_teachers	737	2.381251	.7310767	0	4.584968
electric	737	.9050204	.2933859	0	1
dist_prim	737	1.961289	3.048867	0	20
log_poly	737	2.445236	1.606025	0	9.433484
urban	737	.1899593	.3925351	0	1
p_drought	737	.4092605	.3391063	0	1
p_econ_shock	737	.1774932	.1356602	0	.5833333
p_conf_shock	737	.0319708	.0519323	0	.3125
big_agri	737	.3405699	.4742227	0	1
voters	737	4.176391	1.117468	1	5
MASAF	737	.2279512	.4197958	0	1

Methodology

The impact of the SMP on education will be estimated in three different ways. The first method will rely on individual data from household surveys to construct a difference-in-difference model. To serve as a comparison, a difference-in-difference model will first be run on the percent of adults, within a certain window, who attended school before and after the adoption of free universal education. The estimation of the SMP's impact on this rate will be estimated in a similar manner, and will utilize the fact that the WFP offered SMPs in different districts at different times. To control for overall differences in enrollment trends between the different districts, changes in non-SMP communities' rates will also be estimated. However, this model does not control for any of the aforementioned variables that will affect enrollment. To improve estimation, the next section will estimate the program's impact using community-level data and an Ordinary Least Squares (OLS) model. Yet this model can still be improved, as we know the treatment effect will be skewed due to selection bias; therefore the last section will determine the treatment effects we accept as the final results by using a PSM model with three different methods of matching to determine how robust the results are.

a. Difference-in-Difference Model

Free universal education was introduced in Malawi in 1994 and significantly impacted entrance to and continuation of school. To determine the extent of this, a difference-in-difference model will compare the percent of adults in a 10-year window before and after the policy's implementation who have ever enrolled in primary school, referred to as the ever-enroll rate for simplicity. The control group, assumedly not affected by the change in policy, is the group of adults who are identified as non-poor. While this is not perfect, it is reasonable to assume that non-poor individuals would not have based their enrollment decisions on monetary concerns, whereas the individuals identified as poor and ultra-poor would likely have been incredibly impacted by the policy change. There are therefore two treatment variables, and the difference-in-difference equation is:

$$ever\_enroll = ß_0 + ß_1 poor + ß_2 ultra + ß_3 young + ß_4 poor \times young + ß_5 ultra \times young + \mu$$

The poor and ultra variables capture if an individual is identified as poor or ultra-poor respectively. The young variable is a dummy that takes the value of one if the individual is in the age range where he/she could have taken advantage of the change in policy, age

24-33. The interaction variables will capture the treatment effects for the poor and ultra-poor populations.

Figure 5: Diff-in-Diff Estimation for Universal Primary School

	Non-Poor	Poor	Ultra-Poor	Differences of 1) NP to Poor 2) NP to UP
Age Group 24-33	92.8% (0.004)	83.3% (0.009)	73.7% (0.012)	1) 9.5% (0.008) 2) 19.1% (0.009)
Age Group 34-44	85.5% (0.007)	72.4% (0.011)	61.3% (0.013)	1) 13.1% (0.012) 2) 24.2% (0.012)
Difference	7.3% (0.009) [5.6%, 9.0%]	11.0% (0.014) [8.1%, 13.8%]	12.4% (0.018) [8.9%, 15.8%]	1) 3.6% (0.015) [0.6%, 6.7%] 2) 5.1% (0.016) [1.9%, 8.3%]

Just these rudimentary estimates give us a clear picture of the quick spike in entrance, especially for the ultra-poor population, after free universal primary school was instituted. The treatment effect for both populations is statistically significant, with the poor population experiencing a 3.6% boost in ever-enrolled rates and the ultra poor population receiving a 5.1% boost.

Thanks to the staggered rollout of the SMP, we are able to estimate the impact of the program on ever-enrolled rates using a difference-in-difference model as above, using a five-year window instead to ensure individuals in control communities had a very low likelihood of exposure to treatment. As of 2000, three districts had school feeding programs through WFP: Dedza (the pilot district), Ntcheu, and Salima (DNS). In order to create a good comparison, we will use data from communities in these districts that are identified as having an SMP by the algorithm, who presumably got the program before 2000, and treatment communities who eventually received an SMP outside these districts. This does mean that some included communities in DNS will not have had the SMP at that time, but hopefully this is a negligible percentage, as communities that needed the program should have been given it at that time. By only looking at communities who would eventually receive an SMP, we are hopefully controlling for certain characteristics that allowed both these areas to be chosen. Looking at ever-enrolled rates in the DNS communities as compared to the communities in the other districts for the age groups 24-28 and 28-32 should give us rough estimates of the impact of the program on entry to school. The difference in difference equation is:

$$ever\_enroll = ß_0 + ß_1 DNS + ß_2 young + ß_3 DNS \times young + \mu$$

The results of the regression are summarized in the following table:

Figure 6: Diff-in-Diff Estimation for SMP Communities

Age Group	DNS (SMP 2000)	Other Districts (SMP 2006+)	Difference
24-27	80.0% (0.054)	83.1% (0.016)	-3.1% (0.051) [-13.2%, 7.1%]
28-32	75.3% (0.049)	76.9% (0.017)	-1.6% (0.050) [-11.4%, 8.2%]
Difference	4.7% (0.073) [-19.1%, 9.8%]	6.1% (0.024) [1.4%, 10.9%]	-1.5% (0.074) [-16.0%, 13.1%]

This estimation, using 1205 observations, shows a 2.4% lower difference in ever-enrolled rates in districts with feeding programs as of 2000. Putting a 95% confidence interval around the difference given a standard error of 0.078, the difference is between -17.7% and 7.6%. These results are unexpected, however, this regression ignores that the trends in these areas could be different overall, which invalidates the common trend assumption. To correct for this problem, we will rerun the estimation using all the adults aged 24-32, even in non-SMP communities. The trend of the adults in the non-SMP communities will be the trend we assume we would see in the treatment groups had the areas not had SMPs. The difference-in-difference equation is:

$$ever\_enroll = ß_0 + ß_1 DNS + ß_2 SMP + ß_3 young + ß_4 DNS \times SMP + ß_5 DNS \times young + ß_6 SMP \times young + ß_7 DNS \times SMP \times young + \mu$$

The SMP variable is a dummy for if the community was determined to have an SMP as of 2011. The young variable and the DNS variable are as above. The first interaction term is the overall trend in the DNS area; the second interaction is the overall trend in SMP communities; and the third interaction is the additional trend in SMP communities in the DNS area, or the treatment effect we are aiming to identify. The trend, as identified by the non-SMP areas, is summarized in the following table:

Figure 7: Diff-in-Diff Estimation for non-SMP Communities

Age Group	DNS (Non-SMP)	Other Districts (Non-SMP)	Difference
24-27	87.4% (0.023)	91.0% (0.005)	-3.6% (0.019) [-7.3%, 0.1%]
28-32	79.5% (0.022)	87.8% (0.005)	-8.2% (0.019) [-12.1%, -4.4%]
Difference	7.9% (0.032) [1.6%, 14.1%]	3.3% (0.008) [1.7%, 4.4%]	4.6% (0.028) [-0.9%, 10.1%]

The trend for the non-SMP communities reveals that the DNS areas actually had a higher increase in entrance than non-DNS communities. However, since the first statistics revealed the trend was negative, our final results are even more negative for SMP communities than the original statistics lead us to believe. This estimation reveals a 6.1% lower entrance rate in the treatment communities, those that had an SMP introduced in 2000, than in the control communities, those that would eventually receive an SMP, holding the difference in trends between the two areas constant. However, this result is not statistically significant, likely due to the small sample size and low percent of treatment communities; a 95% confidence interval reveals the treatment effect is between -19.3% and 7.2%. This result and proves the need for further investigation and re-estimation using more timely data.

While this method does correct for common trends, it does not use any of the additional information on the communities that we know will impact enrollment rates. Additionally, this method looks at the impacts only on entrance, and not on the amount of students it encourages to continue attending school. The method also evaluates the programs retroactively, and ignores the possibility that adults could move after schooling. The next method will use this lagged information as a control, while adding a variety of other community level variables to begin to control for some of the other differences between communities that we know will have an impact on current enrollment.

b. Ordinary Least Squares Regression

Estimating a simple OLS regression will give us an initial approximation of the treatment effect on the two education measures we believe could be impacted by the presence of an SMP, primary school enrollment and temporary withdraw rates:

$$\widehat{prim\_enroll} = ß_0 + ß_1 TR + ß_2 X + \mu$$

$$\widehat{withdraw} = \partial_0 + \partial_1 TR + \partial\,_2 X + \mu$$

where prim_enroll is primary enrollment, withdraw is the percent of students who attended school in the last year who have temporarily withdrawn in the past year, TR is the dummy for treatment, and X is a vector of control variables. Just observing the above differences in sample statistics shows that communities with SMPs are systematically different than communities without one. A simple OLS regression will give us biased result due to selection criteria for treatment assignment. Since communities were chosen based on their poor education outcomes, treatment communities will have lower rates and therefore the treatment effect will be downward biased. Regressing treatment on enrollment will still be informative and is worth discussing.

Figure 8: OLS Regression

```
Linear regression                                      Number of obs =     737
                                                       F( 19,    717) =   14.84
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.3407
                                                       Root MSE      =  .11559
```

enrollment	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
TR	-.0139634	.0141542	-0.99	0.324	-.041752	.0138252
gap_poor	-.0010586	.0013279	-0.80	0.426	-.0036656	.0015483
gap_ultra	-.0002038	.0025009	-0.08	0.935	-.0051137	.0047062
historic_attend	.2877178	.0269768	10.67	0.000	.2347549	.3406807
p_access	.0083916	.0168431	0.50	0.618	-.0246762	.0414594
resMP	-.0048811	.0121221	-0.40	0.687	-.0286802	.018918
NGO	-.002151	.0105444	-0.20	0.838	-.0228526	.0185507
dist_doc	-.0003527	.0002341	-1.51	0.132	-.0008123	.0001068
log_teachers	.0125926	.0072066	1.75	0.081	-.0015559	.0267411
electric	.0565298	.017629	3.21	0.001	.0219191	.0911405
dist_prim	-.0024579	.0016088	-1.53	0.127	-.0056163	.0007006
log_poly	-.0079788	.002944	-2.71	0.007	-.0137587	-.002199
urban	-.0600466	.0160914	-3.73	0.000	-.0916386	-.0284547
p_drought	-.0234922	.0158249	-1.48	0.138	-.0545608	.0075764
p_econ_shock	.055546	.0374765	1.48	0.139	-.0180309	.1291229
p_conf_shock	-.0650179	.0851852	-0.76	0.446	-.2322601	.1022243
big_agri	.0161073	.0088878	1.81	0.070	-.001342	.0335566
voters	.0040085	.0037067	1.08	0.280	-.0032689	.0112858
MASAF	-.0164185	.0099498	-1.65	0.099	-.0359528	.0031158
_cons	.6789429	.0420135	16.16	0.000	.5964587	.7614271

As expected, the coefficient on the dummy for the treatment variable is slightly negative, but not statistically different from zero. According to this model, whether a school has an SMP or not has no impact on their primary school enrollment rate. Historic enrollment is a large driver of primary enrollment rates in this model, along with whether the school is electrified or not. Interestingly, agricultural hubs have higher rates of enrollment, however, this is holding the percent of households who reported loss of income due to drought constant and it is not significant. Areas more prone to drought, holding all else constant, have lower rates of primary enrollment, which is unsurprising as households are more likely to rely on children's labor in areas with extreme weather conditions. The urban dummy negatively impacts enrollment, meaning urban areas have lower primary enrollment rates than urban areas, holding all else constant. Finally, the number of polygamous marriages impacts enrollment negatively. The other variables have no statistically significant impact on primary school enrollment, including the poverty indicators, the accessibility of the main road, the number of teachers at the nearest primary school and the distance to the school. Theoretically, these indicators should have an impact on enrollment and may imply that selection bias is having a confounding effect or this model uses the wrong functional form.

d. Propensity Score Matching (PSM) Model

A PSM model has the benefit of no functional form assumptions. For example, the poverty indicators, regardless of the type or form in which they are used in the model, will control for income and unobservables that are correlated with income and affect the outcome of interest. Although the correct functional form is not required, the conditional independence assumption must be satisfied. This means that selection is solely based on observable characteristics. Additionally, all the variables that affect treatment assignment or potential outcomes are observed and included in the model. This is a strong assumption, however, given the amount of data that was collected in the household survey and the CI survey, it is not unreasonable if the variables are selected strategically. Selecting too many variables, however, is not necessarily better, as the more variables, the more difficult it will be to find suitable pairs. Ideally, a PSM model aims to compare outcomes for two communities that are identical in every way except in their treatment status. Adding more variables makes it more difficult to find similar communities, reducing the area of common support. The common support condition ensures that two communities with the same characteristics have a positive probability of being in the treatment or control group. Balancing these two conflicting conditions requires including enough variables to control for a wide variety of possible covariates while excluding extraneous variables that would over-identify treatment communities.

The first grouping of variables, Expressed Targeting Objectives, is obviously essential as it includes the characteristics that the WFP seeks when choosing to expand. The most important of these is historic attendance, which will be correlated with a lot of potentially confounding attitudes towards education. Of the WFP's targeting objectives, a few conditions cannot be directly accounted for: food insecurity, availability of storage facilities, and evidence of community commitment to participate. The first, food insecurity, should be correlated with other measures such as the poverty indicators, the percent of community that reported an economic shock in the past year, including droughts, and the strength of the farming industry. The second variable is not concerning as it shouldn't be correlated with outcomes in any logical way, and may even be correlated with other good amenities in the community, such as an electrified primary school. The final criteria poses the biggest concern as it starts to underline the importance of political will power and connections.

Both school and community characteristics, which make up the second and third groups of variables, will begin to correct for or be correlated with important and potentially influential variables that affect treatment assignment.

The final set of variables, Government Attention, aims to address the role of government leaders in providing communities with SMPs. First, communities that have a MASAF Program are likely to be better connected or at least have capable local leadership to implement and organize the projects. The second variable aims to determine which communities a political leader would want to target, that is, the people who voted and will vote again to keep him or her in office. It is extremely difficult if not impossible to capture the unique and complex linkages that make some communities more favorable than others, thus we hope that the inclusion of measures of connectedness and community characteristics will be correlated significantly enough with political clout that it will not blur our impact analysis. NGO presence should be a good indication of efficient local leaders, as NGOs seemingly would not go into areas with poor leadership and organization. If a resident of the community is a Member of Parliament for Constituency, it is a clear signal that the leaders are well connected. By not including unnecessary variables, the model can retain common support while still meeting the conditional independence assumption.

Accepting these conditions, the next phase is estimating the propensity scores, or the predicted probability that a community has an SMP. The first stage of this is a probit regression, used because our dependent variable, if the community is in the treatment or control group, is binary. These coefficients should be interpreted as the average marginal effect for every community. This means for a one-unit increase in an independent

Figure 9: Results of Probit Estimation of Probability of Treatment using Marginal Effects

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
gap_poor	-.0081407	.0033631	-2.42	0.015	-.0147321	-.0015492
gap_ultra	.011553	.0050993	2.27	0.023	.0015586	.0215474
historic_attend	-.1780016	.0784714	-2.27	0.023	-.3318026	-.0242005
p_access	-.0212056	.0457127	-0.46	0.643	-.1108009	.0683897
resMP	-.0139165	.0387576	-0.36	0.720	-.08988	.062047
NGO	-.0067057	.0287087	-0.23	0.815	-.0629737	.0495623
dist_doc	-.0015491	.0006962	-2.23	0.026	-.0029137	-.0001845
log_teachers	.0942565	.0246736	3.82	0.000	.045897	.1426159
electric	.1054491	.0520192	2.03	0.043	.0034933	.2074048
dist_prim	-.0168	.0060727	-2.77	0.006	-.0287023	-.0048976
log_poly	-.0054909	.008517	-0.64	0.519	-.022184	.0112022
urban	-.1284285	.0513306	-2.50	0.012	-.2290347	-.0278223
p_drought	.2984868	.0461168	6.47	0.000	.2080994	.3888741
p_econ_shock	.1023488	.098726	1.04	0.300	-.0911506	.2958482
p_conf_shock	-.141522	.2456991	-0.58	0.565	-.6230834	.3400394
big_agri	-.0495811	.0300646	-1.65	0.099	-.1085067	.0093445
voters	.0287572	.0125518	2.29	0.022	.0041562	.0533582
MASAF	-.0325064	.0328876	-0.99	0.323	-.0969649	.0319522

variable, it will increase (or decrease) the predicted probability by the coefficient. For the dummy variables, this is the equivalent of the change in predicted probability for a community that moves from not included in a group to included. For example, looking at the coefficient on the urban dummy, there is a 0.13 percentage point decrease in the predicted probability of receiving treatment for a community in an urban area compared to a rural community that is alike in every other way. Rural areas could have higher propensity scores either because they are more vulnerable to food insecurity or because they are closer to food production and hence logistically easier to serve. The other dummy variable that has a significant impact on the probability of treatment is electric, which is one if a community is near an electrified school; for a community near an electrified school, their predicted probability increases by 0.10 percentage points compared to a community that is not, but is alike in every other way. This increase could be an indication of good political connections or political capability. The rest of the dummy variables are not statistically significant.

For non-dummy variables, the coefficients still represent the change in predicted probability for a one-unit increase, yet this does not necessarily represent a feasible change for some of the variables. The historic attendance variable is one of these variables, as a one-unit increase would represent a change from 0% enrollment to 100% enrollment. A more feasible change is an increase of 10%, which would decrease the likelihood of treatment by 0.018 percentage-points. This is surprisingly low and may reflect an inadequate tabulation of the historic enrollment rates. However, this percentage change is holding a lot of other variables constant, and may just reflect the inter-dependence of many of the independent variables.

The poverty coefficients are also challenging to interpret because of their inter-dependence on each other. While it is certainly possible to have a community where the poverty gap is higher and the ultra-poverty gap is lower than another community, this is would mean the community has a larger percent of poor people, but less ultra-poor people. Regardless of the differences, the first community would have a lower predicted probability of receiving an SMP. This could mean the government is targeting areas with large numbers of poor people, and not just areas with small concentrations of ultra-poor people. The reverse is also possible, where a community could have a smaller poverty gap, but a larger ultra-poverty gap than another community, if the first has a smaller percent of poor people, but more people in ultra-poverty. The impact on predicted probability here depends on the magnitude of these differences, but due to the greater coefficient on the ultra-poverty gap, it is likely the first community will have a higher probability of being treated, but to a smaller extent than in the first case. Yet, these results point to the fact that the government may be targeting areas where the program would have an impact on the greatest number of people. However, typically an increase (or decrease) in one of the indicators will correspond with a similar magnitude increase (or decrease) in the other indicator. A one-unit increase here would represent the average person in the community moving 1 MWK farther from the ultra poverty line, and hence the poverty line. This would increase the probability of treatment by .0035 percentage-points.

In the Connectedness category, the only significant variable is the distance to the nearest doctor. A feasible increase in this variable would be a 5 km increase, which would decrease the likelihood of treatment by 0.008 percentage-points. The direction of this variable is what we predicted: that the farther removed a community is, the less likely they will receive an SMP.

Unsurprisingly, all of the School Characteristics variables are statistically significant, the first being the electrified dummy that was already discussed. The number of teachers at the primary school is also significant and reveals that the government is targeting larger schools to increase their penetration with minimal logistical impact. The final school characteristic, distance to the nearest primary school, negatively impacts the likelihood of an SMP, meaning either the government is targeting schools in central locations that are close to many communities or both CIs and households are more likely to report an SMP if the school is close by. It could also reflect political clout to an extent, as good leaders are able to attract both schools and government programs to their communities.

The drought indicator is the only other Community Characteristic, aside from the urban dummy, which significantly impacts the probability of treatment. This confirms the theory that droughts serve as a proxy for food insecurity. The percent drought represents the amount of people in the sample that reported being negatively economically impacted by a drought. A feasible increase is therefore a .25 increase instead of a one-point increase. This would mean that the community went from not experiencing a drought to experiencing one where 25% of the people in the sample were negatively impacted, a conservative estimate given the reliance on agricultural work in Malawi. This would increase the probability of treatment by .075 percentage-points. This is the largest increase that has been observed, and while the chosen increases are not standardized across the board, the magnitude of the potential impact certainly shows the importance of food insecurity in choosing SMP communities. The economic shock variable also has a high coefficient, but lacks significance due to a lack of power that may stem from the variable's high correlation with drought. Big agricultural centers also have a lower probability of being selected for an SMP, although the result is not significant.

The final category of variables, Government Attention, has one significant variable, voters, which is the percent of the community that votes in elections. A reasonable increase in this variable is 1-point, which represents an increase of one-quarter of the community population voting in the election. This would increase the likelihood of receiving an SMP by 0.029 percentage-points, showing the power of voting even in countries that are thought to be particularly corrupt. The variable could also indicate that communities that are more politically involved or connected have SMPs more frequently. The correct signs on the important variables assure us that the model is properly defined and satisfies the conditions we assumed to be true. Additionally, there are variables in each category that are statistically significant, meaning we have captured all the areas we believed to be important in the determination of selection criteria.

The next stage of the PSM model is determining the region of common support, which is where there are both treatment and control communities with equal, or very close to equal, propensity scores. This region is between 0.005 and 0.745, with 99% of the observations within the region of common support having a score less than 0.67. As only a little more than 16% of the sample is in the treatment group, it is not unexpected that the majority of the communities have low probabilities of having an SMP. This region eliminates 25 control communities that have scores too low to be in the region of support. In order to properly estimate standard errors, the regressions will be repeated 200 times in order to bootstrap. In each iteration, N pairs (where N is the total number of matches) will be selected with replacement to estimate the treatment effect. The variation of the estimators will determine the standard error of the treatment effect.

The final stage of the model is selecting which type of matching to use to determine treatment effects. Nearest neighbor matching is the simplest, and compares outcomes for treatment communities against a control community with the closest score. Thus, the set of control communities, C(i) matched to the treated unit, i, with an estimated value of the propensity score, p_i can be written as:

$$C(i) = \min \|p_i - p_j\|$$

This can be done with replacement, which is allowing control communities to be selected as the nearest neighbor twice, or without. Replacement ensures quality matches, however, it reduces the amount of control communities that are utilized to determine the treatment effect and therefore reduces variation. A reduction in variance will increase standard errors and will make it more difficult to achieve statistical significance.

The next model is radius matching, which compares outcomes for not only the nearest neighbor, but also all the control communities whose propensity scores lie within a certain radius. Assuming a radius of r, the set can be written as:

$$C(i) = \{p_j \mid \|p_i - p_j\| < r\}$$

A benefit of this method is an increase in the amount of control communities that will be included, and hence a decrease in standard errors. Additionally, it allows as many comparisons as possible while avoiding the possibility of comparing to a community with far different characteristics. For these models, the treatment formula can be written the same, with Y_i^T representing the outcome for treated community i, Y_j^C the outcome for control community j, N_T the number of communities in the treated group, N_C the number in the control group, and weights, w_{ij} equal to ${}^{1}\!/_{N_C}$ if $j \in C(i)$ and $w_{ij} = 0$ otherwise:

$$\tau = \frac{1}{N_T}\left(\sum_{i \in T} Y_i^T - \sum_{i \in T} \sum_{j \in C(i)} w_{ij}\, Y_j^C\right)$$

The final model, kernel matching, expands on the radius model, but weights observations according to how close their propensity score is to the treatment community. Kernel matching uses all the variation within the sample, and hence minimizes the standard errors, giving the best chance of achieving statistical significance. The treatment effect is calculated according to the following equation:

$$\tau^K = \frac{1}{N_T}\sum_{i\in T}\left\{Y_i^T - \frac{\sum_{j\in C} Y_j^C G\left(\frac{p_j - p_i}{h_n}\right)}{\sum_{k\in C} G\left(\frac{p_k - p_i}{h_n}\right)}\right\}$$

where $G(p,h)$ is a kernel function and h_n is a bandwidth parameter. A kernel is used as a measure of similarity; a kernel function defines the distribution of similarities of points around a given propensity score. The Gaussian basis function is used, thus the bandwidth is optimally calculated according to the standard deviation of the sample and the number of communities. Thus the second half of the equation is simply a consistent estimator of the counterfactual outcome that cannot be observed. In order to ensure the treatment effect is robust to differing matching techniques, this paper will utilize all three methods.

The results of all three matching techniques for both dependent variables, are summarized in the following table:

Figure 10: Propensity Score Matching Estimations

Method	Number of Treated	Number of Control	ATT[8]	t-score
Primary School Enrollment				
Nearest-Neighbor	129	89	-0.041 (0.024)	-1.747
Radius (0.05)	125	583	-0.033 (0.015)	-2.121*
Kernel	129	583	-0.023 (0.018)	-1.319
Percent of Temporary Withdraws				
Nearest-Neighbor	129	89	-0.004 (0.013)	-0.305
Radius (0.05)	125	583	-0.009 (0.009)	-1.091
Kernel	129	583	-0.005 (0.009)	-0.539

*Statistically significant at the 5% confidence level.

[8] ATT is the Average Treatment Effect on the Treated, which assumes all communities in the treated group actually received the treatment.

The first matching technique utilizes all the variation in treatment communities, but a very small percent of the variation in control communities. As predicted, the standard errors are significantly higher for both estimations using nearest neighbor matching. This method predicts a statistically insignificant decrease in both enrollment and temporary withdraws. The next model is radius matching, using a 0.05 radius, which is half the standard radius to account for the smaller variation in propensity scores seen in the sample. The radius matching takes greater advantage of the area of common support, utilizing 125 treatment communities and 583 control communities, leaving four treatment communities not included. This method again shows a negative impact of the program on enrollment, this time with a large magnitude, 3.1 percentage points, that is just statistically significant at the 5% confidence level. This method also predicts a 0.9 percentage point decrease in temporary withdraws, which would be extremely significant economically if it was statistically significant. The final method, kernel matching, uses all the possible variation, but has slightly higher standard errors than the radius matching method. This estimation, the most complete, has a coefficient between the other two methods at a 2.3% decrease in enrollment for SMP communities that is again statistically insignificant. The kernel method finds a smaller impact on temporary withdraws than the radius matching method that is again not statistically different from no impact. As previously mentioned, the variation in the percent of temporary withdraws is very small, and is therefore be very difficult to find any significant impact.

Conclusion

The results of this paper are contrary to most previous studies, which show strong improvements to enrollment and attendance in SMP communities. The only study that did not find improvements to educational outcomes was the evaluation of the long-run impacts of Chile's SMP. However, Chile has achieved nearly universal enrollment and low levels of malnourishment, making it extremely different from Malawi. Thus while Chile's non-impact is likely due to an inability to improve further using this type of policy, the results of this paper, assuming the model is correctly specified, are due to different circumstances (McEwan, 2013). The pilot evaluation may shed light on the atypical results of this study. Although enrollment increased in the pilot communities, most of this was due to transfers from other areas, and thus overall enrollment was not improved. The report concluded that the main problem with the SMP was exclusion errors: the poorest families either do not send their children to school due to the direct and indirect costs of education or withdraw them during hard times when they are needed for work (WFP, 1996). While the negative coefficients are difficult to explain, the statistical insignificant results could be due to the inability of families to send their children to school regardless of the nutritional benefits. Another potential explanation is that more families are moving into areas with SMPs, and bringing children they do not intend to enroll along with them. This seems less likely and the statistics do not show any difference in community size between treatment and control communities.

Although these explanations are feasible, it is also possible that the model does not capture all the potential factors that influence either treatment assignment or education outcomes. This could bias the results and lead to an inaccurate estimation. Since it is such a large driver of the likelihood of selection, an error in the calculation of historic attendance could cause large bias. The unexpected coefficients on the poverty indicators could also be a sign that those measures are not adequate for identifying food insecure communities. Finally, it is extremely difficult to capture the governmental connections aspect of treatment assignment, and the variables included may not completely account for this. Another potential flaw lies in the identification of SMP communities, which relies on CI and parent reported data as opposed to government or WFP reports. Improperly identified communities would likely be wealthier areas that have internally funded programs that have the potential to dampen our treatment effect.

Assuming the model does in fact predict enrollment and withdraw impacts accurately, Malawi's SMP is not effectively drawing or keeping primary-aged children in school. Although these results conflict previous evidence, this could be due to the difference in methods. This evaluation uses national-level data that is comprehensive, but not targeted for estimating the impacts of the SMP. The results are therefore free from social-desirability bias that could sway community leaders to report favorable outcomes or even bribe families to temporarily enroll their children in school. It also estimates impacts according to a realistic operation of the program as opposed to what would happen if researchers were ensuring food was continually supplied. Families would also be better informed of the terms and benefits of the program during an RCT. This paper therefore provides a good avenue for further research into investigating how much these effects have influenced past studies by re-conducting the evaluations under less evident and more realistic circumstances. It also holds promise for an impact on temporary withdraws if more variation is included in the sample. Perhaps by focusing just on poorer communities, more variation could lead to a statistically significant negative decline in the number of students temporarily withdrawing. This measure of attendance is particularly important in agrarian economies like Malawi's and perhaps a benefit of the program that should be included in further evaluations.

Appendices

<u>Appendix A</u>: District Statistics and Number of CI Reported SMPs per District

Region	**District**	**Sample Size**	**# of SMP Communities**	**Percent Poor**	**Percent Ultra-Poor**
Central	Dedza	1500	2	56.0	23.2
Central	*Dowa*	1555	0	44.2	14.8
Central	Kasungu	1590	10	35.0	10.9
Central	Lilongwe	2337	4	53.9	29.3
Central	Lilongwe City	2206	1	15.6	4.3
Central	*Mchinji*	1519	*2*	54.7	31.0
Central	*Nkhota Kota*	1523	0	32.8	11.0
Central	Ntcheu	1442	5	46.0	13.7
Central	*Ntchisi*	1539	0	38.8	9.9
Central	Salima	1518	7	44.7	17.4
	10	**16729**	**31**	**41.6 (45.5)**	**16.6 (18.5)**
Northern	*Chitipa*	1553	0	73.3	42.2
Northern	Karonga	1571	3	62.4	26.6
Northern	*Mzimba*	1430	*4*	59.4	29.8
Northern	*Mzuzu City*	1521	*18*	14.4	1.7
Northern	*Nkhata Bay*	1617	0	44.0	17.6
Northern	*Rumphi*	1548	*2*	35.6	10.2
	6	**9240**	**27**	**48.1 (54.8)**	**21.3 (25.2)**
Southern	*Balaka*	1408	*4*	65.3	31.2
Southern	*Blantyre City*	1384	*24*	6.1	1.6
Southern	*Blanytyre*	1363	*6*	43.4	14.5
Southern	Chikwawa	1407	13	79.2	56.1
Southern	Chiradzulu	1353	14	43.1	13.1
Southern	*Machinga*	1440	0	74.9	40.2
Southern	Mangochi	1448	3	71.3	41.6
Southern	Mulanje	1356	15	62.4	31.6
Southern	*Mwanza*	1412	*5*	61.0	31.2
Southern	*Neno*	1495	*2*	65.6	31.9
Southern	Nsanje	1361	20	79.3	54.2
Southern	Phalombe	1386	11	64.5	41.3
Southern	Thyolo	1274	18	34.5	11.1
Southern	Zomba	1330	8	55.0	25.0
Southern	Zomba City	1415	18	13.3	2.8
	15		**136**	**54.8 (61.8)**	**28.7 (32.8)**

*Districts in italics do not have WFP sponsored SMPs.

*Numbers in parenthesis are averages excluding city regions.

Appendix B: Influence of Reliability on Reporting Treatment

OLS Regression: Percent of HHs Reporting SMPs (p_smp) on the Percent of Months HHs Report Receiving Aid from SMPs (p_months)

```
Linear regression                               Number of obs =     307
                                                F( 19,   287) =    7.05
                                                Prob > F      =  0.0000
                                                R-squared     =  0.2635
                                                Root MSE      =  .27996
```

p_smp	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
p_months	.3229444	.0792567	4.07	0.000	.1669462	.4789426
gap_poor	.0017003	.0045853	0.37	0.711	-.0073248	.0107254
gap_ultra	.004971	.0065918	0.75	0.451	-.0080034	.0179453
historic_attend	.1232049	.1078286	1.14	0.254	-.0890303	.3354401
p_access	-.1252152	.0611275	-2.05	0.041	-.2455303	-.0049001
resMP	-.0210092	.0482502	-0.44	0.664	-.1159784	.0739601
NGO	-.0054512	.0345203	-0.16	0.875	-.0733963	.0624939
dist_doc	-.0002188	.0010946	-0.20	0.842	-.0023731	.0019356
log_teachers	.0672059	.0301753	2.23	0.027	.0078129	.126599
dist_prim	-.0116367	.0058734	-1.98	0.049	-.0231972	-.0000763
electric	.1110509	.0658301	1.69	0.093	-.0185201	.2406219
log_poly	-.0008437	.0128565	-0.07	0.948	-.0261487	.0244614
big_agri	-.069476	.0408963	-1.70	0.090	-.1499706	.0110187
urban	-.0112327	.06023	-0.19	0.852	-.1297813	.1073159
p_drought	.2320684	.0687905	3.37	0.001	.0966705	.3674663
p_econ_shock	.2952037	.1528451	1.93	0.054	-.0056357	.5960432
p_conf_shock	.3258921	.3320194	0.98	0.327	-.3276097	.9793938
voters	.0252934	.0184193	1.37	0.171	-.0109607	.0615475
MASAF	-.024295	.0467509	-0.52	0.604	-.1163132	.0677231
_cons	-.2222535	.1686372	-1.32	0.189	-.5541761	.1096691

OLS Regression: Likelihood CIs Report an SMP (com_SMP) on the Percent of Months HHs Report Receiving Aid from SMPs (p_months)

```
Linear regression                               Number of obs =     307
                                                F( 19,   287) =    7.74
                                                Prob > F      =  0.0000
                                                R-squared     =  0.2213
                                                Root MSE      =  .44017
```

com_SMP	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
p_months	.5460336	.1400603	3.90	0.000	.2703579	.8217093
gap_poor	-.0185316	.0071781	-2.58	0.010	-.03266	-.0044033
gap_ultra	.0304734	.0099103	3.07	0.002	.0109673	.0499794
historic_attend	-.2001564	.1648097	-1.21	0.226	-.5245455	.1242327
p_access	.0723249	.0994306	0.73	0.468	-.1233809	.2680306
resMP	.0017881	.070662	0.03	0.980	-.1372934	.1408696
NGO	-.0519658	.0525249	-0.99	0.323	-.1553486	.051417
dist_doc	.0003182	.0016651	0.19	0.849	-.0029591	.0035955
log_teachers	.1442478	.0493099	2.93	0.004	.0471928	.2413027
dist_prim	-.0092247	.0087835	-1.05	0.294	-.0265129	.0080635
electric	-.0565903	.1047966	-0.54	0.590	-.2628576	.1496771
log_poly	-.0282903	.0218187	-1.30	0.196	-.0712354	.0146547
big_agri	-.075473	.0664873	-1.14	0.257	-.2063376	.0553915
urban	.0919888	.0964038	0.95	0.341	-.0977593	.2817369
p_drought	.2506486	.1053445	2.38	0.018	.0433028	.4579944
p_econ_shock	.0104785	.2326309	0.05	0.964	-.4474006	.4683576
p_conf_shock	-.1781958	.5208754	-0.34	0.733	-1.203416	.8470246
voters	.0187375	.0277118	0.68	0.499	-.0358067	.0732817
MASAF	-.0477844	.0735959	-0.65	0.517	-.1926405	.0970717
_cons	-.0025238	.2706634	-0.01	0.993	-.535261	.5302133

Appendix C: Treatment Algorithm

1. If a community is in a district in which the WFP does not operate, they are in the control group.
2. If the number of households with children in primary school is less than 5 in the sample of a given community, the CIs' responses will determine treatment status.
3. If the percent of households with children in primary school who report a SMP is greater than or equal to 50%, the community is in the treatment group.
4. If the number of households with children in primary school is greater than 7 and the percent that report a SMP is less than 20%, the community is in the control group.
5. If the number of households with children in primary school is less than or equal to 7 and the percent that report a SMP is less than 20%, the CIs' responses will determine treatment status.
6. If the percent of households that report an SMP is between 20-50%:
    a. And the CIs report nearly every child receives food under the program, the community is in the control group.
    b. And the percent of months households reported receiving benefits is greater than 60%, the CIs responses will determine treatment.
    c. And the percent of months households reported benefits is less than or equal to 60%, the community is in the control group.

Criteria	# Placed in Control	# Placed in Treatment	# Conflicts with CIs
1	329	0	64
2	20	16	N/A
3	0	108	18
4	162	0	16
5	51	10	N/A
6a	21	0	21
6b	8	3	N/A
6c	9	0	0
Total	**600**	**137**	**119**

*Only communities with all variables not missing.

Appendix D: Poverty in non-World Food Program Areas

```
Two-sample t test with unequal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	45	4.327023	.7662871	5.14041	2.782673	5.871374
1	127	18.51073	1.205404	13.58421	16.12527	20.89619
combined	172	14.79988	1.028262	13.48553	12.77015	16.8296
diff		-14.1837	1.428353		-17.00339	-11.36402

```
    diff = mean(0) - mean(1)                                      t =  -9.9301
Ho: diff = 0                     Satterthwaite's degrees of freedom =  169.258

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 1.0000
```

This test aims to determine the difference in the poverty gap between Group 0, communities who would be classified as treatment, but are not in WFP areas, and Group 1, communities classified as treatment in WFP areas. The schools in Group 0 have significantly less poverty, indicating these SMPs are most likely not targeted at poor communities, but may be at schools with better resources that can afford their own programs. The same is true for ultra poverty levels:

```
Two-sample t test with unequal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	45	.7876198	.2314677	1.552732	.3211274	1.254112
1	127	7.287615	.7639155	8.608891	5.775848	8.799381
combined	172	5.587035	.6073171	7.96489	4.388231	6.785838
diff		-6.499995	.7982131		-8.077481	-4.922509

```
    diff = mean(0) - mean(1)                                      t =  -8.1432
Ho: diff = 0                     Satterthwaite's degrees of freedom =  146.658

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 1.0000
```

Additionally, enrollment is significantly higher in these areas:

Two-sample t test with unequal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	45	.9637627	.0097175	.065187	.9441784	.983347
1	127	.850126	.0130813	.1474186	.8242385	.8760135
combined	172	.8798565	.0106792	.1400564	.8587765	.9009366
diff		.1136367	.0162957		.0814575	.145816

```
    diff = mean(0) - mean(1)                                      t =   6.9734
Ho: diff = 0                     Satterthwaite's degrees of freedom =  162.086

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 0.0000
```

This means that an improperly identified school will not see any higher enrollment than a similar community based on the propensity score. The effect of this is a smaller coefficient than expected, or no impact on enrollment or withdraws.

Bibliography

Abrams, Steven A., Alex Mushi, David C. Hilmers, Ian J. Griffin, Penni Davila, and Lindsay Allen. "A Micronutrient-Fortified Beverage Enhances the Nutritional Status of Children in Botswana." The Journal of Nutrition 133.6 (2003): 1834-840. Web. 22 Apr. 2015.

Ash, Deborah M., Simon R. Tatala, Edward A. Frongillo, Jr., Godwin D. Ndossi, and Michael C. Latham. "Randomized Efficacy Trial of a Micronutrient-Fortified Beverage in Primary School Children in Tanzania." The American Journal of Clinical Nutrition 77.4 (2003): 891-98. Web. 22 Apr. 2015.

Chandler, Ann-Marie K., Susan P. Walker, Kevin Connolly, and Sally M. Grantham-McGregor. "School Breakfast Improves Verbal Fluency in Undernourished Jamaican Children." The Journal of Nutrition 125.4 (1995): 894-900. Web. 22 Apr. 2015.

Grillenberger, Monika, Charlotte G. Neumann, Suzanne P. Murphy, Nimrod O. Bwibo, Pieter Van't Veer, Joseph G. Hautvast, and Clive E. West. "Food Supplements Have a Positive Impact on Weight Gain and the Addition of Animal Source Foods Increases Lean Body Mass of Kenyan School Children." The Journal of Nutrition 133.11 (2003): 3957S-964S. Web. 22 Apr. 2015.

International Food Policy Research Institute (IFPR). Impact of Feeding Children in School: Evidence from Bangladesh. By Akhter U. Ahmed. Washington, D.C.: United Nations U, 2004. Web. 22 Apr. 2015.

Jacoby, Enrique R., Santiago Cueto, and Ernesto Pollitt. "When Science and Politics Listen to Each Other: Good Prospects from a New School Breakfast Program in

Peru." The American Journal of Clinical Nutrition 67 (1998): 795S-97S. Web. 22 Apr. 2015.

McEwan, Patrick J. "The Impact of Chile's School Feeding Program on Education Outcomes." Economics of Education Review 32 (2013): 122-39. Elsevier. Web. 22 Apr. 2015.

Nkhoma, O. W. W., M. E. Duffy, D. A. Cory-Slechta, P. W. Davidson, E. M. Mcsorley, J. J. Strain, and G. M. O'brien. "Early-Stage Primary School Children Attending a School in the Malawian School Feeding Program (SFP) Have Better Reversal Learning and Lean Muscle Mass Growth Than Those Attending a Non-SFP School." Journal of Nutrition 143.8 (2013): 1324-330. 26 June 2013. Web. 18 Mar. 2015.

Omwami, Edith Mukudi, Charlotte Neumann, and Nimrod O. Bwibo. "Effects of a School Feeding Intervention on School Attendance Rates Among Elementary School Children in Rural Kenya." Nutrition 27.2 (2011): 188-93. Elsevier. Web. 22 Apr. 2015.

Powell, Christine A., Susan P. Walker, Susan M. Chang, and Sally M. Grantham-McGregor. "Nutrition and Education: A Randomized Trial of the Effects of Breakfast in Rural Primary School Children." The American Journal of Clinical Nutrition 68.4 (1998): 873-79. Web. 22 Apr. 2015.

Simeon, D. T. "School Feeding in Jamaica: A Review of Its Evaluation." The American Journal of Clinical Nutrition 67.4 (1998): 790S-94S. Web. 22 Apr. 2015.

Van Stuijvenberg, Elizabeth, Jane D. Kvalsvig, Mieke Faber, Marita Kruger, Diane G. Kenoyer, and AJ S. Benade. "Effect of Iron, Iodine, and Beta-Carotene-Fortified

Biscuits on the Micronutrient Status of Primary School Children: A Randomized Controlled Trial." The American Journal of Clinical Nutrition 69.3 (1999): 497-503. 25 Aug. 1998. Web. 22 Apr. 2015.

World Development Indicators 2009. (April 2009). Washington, DC: World Bank.

World Development Indicators 2010. (April 2010). Washington, DC: World Bank.

World Food Program Malawi Country Office. (May 2010). *WFP Malawi: School Meals Programme.* Retrieved from http://documents.wfp.org/stellent/groups/public/documents/newsroom/wfp220250.pdf.

World Food Program. (1996). *Report on Pilot School Feeding Program Evaluation Report WFP/Malawi*. (Unpublished).