Children's diets, nutrition knowledge and access to markets

Kalle Hirvonen [a], John Hoddinott [b], Bart Minten [a], David Stifel [c]

[a] International Food Policy Research Institute, Addis Ababa, Ethiopia

[b] Cornell University, Ithaca, USA

[c] Lafayette College, Easton, PA, 18040, USA

30 January 2016

Abstract

Chronic undernutrition in Ethiopia is widespread and many children consume highly monotonous diets. To improve feeding practices in Ethiopia, local policymakers place a strong focus on improving caregivers' nutrition knowledge. In this paper, we study the impact of improving households' nutrition knowledge and its complementary with market access. To test whether the effectiveness of the nutrition knowledge on children's dietary diversity depends on market access, we use survey data from an area with a large variation in transportation costs over a relatively short distance. This allows us to carefully assess the impact of household nutrition knowledge with varying access to markets but still within similar agro-climatic conditions. We find that nutrition knowledge leads to considerable improvements in children's diets but only in areas with relatively good market access.

Keywords: dietary diversity, food markets, remoteness, Ethiopia

JEL codes: O12, Q12, Q18, I15

Acknowledgements: Funding for this work was received through the Feed-the-Future project funded by the United States Agency for International Development (USAID). The survey was funded by International Growth Centre (IGC).

1. Introduction

Within development study and practice, the last ten years has seen a significant increase in interest in policies and interventions that improve the nutritional status – heights, weights and micronutrient status – of pre-school children. This interest, exemplified by the Sustainable Development Goal of eliminating stunting and wasting by 2030 (United Nations 2015) is based on two considerations. First, improvements in nutritional status are an intrinsically valuable development outcome. Second, the preponderance of evidence shows that the harm caused by undernutrition in early life – both lost physical growth and neurological damage – is not fully recovered leading to lower levels of height, schooling, cognitive skills and ultimately lower income in adulthood (Black et al. 2013; Hoddinott et al. 2013). Children's nutritional status is determined by the quantity and quality of the child's diet and their health status. While at the household level, raising incomes improves household food security (Hidrobo et al. 2015), there is now considerable evidence that the link between income and children's nutritional status is weak or non-existent (Manley, Gitter, and Slavchevska 2013).

In light of this, attention has shifted to other factors that may affect children's diets. One is caregiver knowledge regarding correct child feeding practices both during the first six months of life, when children should be exclusively breastfed, and in the introduction and increased use of complementary foods between six and 24 months. If caregivers do not understand the importance of providing children with certain foods, or if they perceive healthy foods to be harmful, they will not provide these to their children even when they are available in the household. For example, in Ethiopia – the focus of our work – a study found that mothers do not feed young children vegetables because these are perceived to be difficult to digest leading to stomach illnesses (USAID 2011). A second study, based on focus group discussions and observation found that

Ethiopian mothers did not feed pre-school children meat or other animal source foods because they believed that children cannot digest these (Alive & Thrive 2010). Examples such as these point to an absence of caregiver understanding of the importance of diet quality. In response, Behavioural Change Communication (BCC) interventions that seek to improve caregivers' nutrition knowledge have gained popularity among policymakers in low income countries (WHO and Unicef 2003; USAID 2014; African Union 2015). BCC has been found to be effective in improving child feeding practices in a number of randomized control trials in different settings but many of these have taken place in urban localities or areas characterized by high-population density implying good access to food markets. (Santos et al. 2001; Bhandari et al. 2004; Penny et al. 2005; Zaman, Ashraf, and Martines 2008).

But poor access to foods is likely to be a limiting factor on the effectiveness of BCC (Penny et al. 2005). This points to a second issue, the availability of a diverse set of foods for adults and children to consume. Sibhatu, Krishna, and Qaim (2015) report on a four country study that shows that households with higher levels of food production diversity exhibit higher levels of food consumption diversity. Evidence from Ethiopia suggests that households with better access to markets consume more diverse diets (Stifel and Minten 2015) and their food consumption is less dependent on their own agricultural production (Hirvonen and Hoddinott 2014; Hoddinott, Headey, and Dereje 2015).

To this point, however, explorations of the role of caregiver knowledge and the role of market access as determinants of diets, particularly child diets, have proceeded in parallel. In this paper, we bring these ideas together. We do so drawing on data from Ethiopia. Ethiopia is a good site to study this question. Its rugged terrain and poor, though improving, infrastructure make transportation difficult and expensive. Chronic undernutrition is widespread and many children consume monotonous, undiversified diets (Central Statistical Agency and ICF International 2012; Headey 2014). We use a novel data set from an area with a large variation in transportation costs over a relatively short distance but similar agro-climatic conditions. Our survey data contain detailed information on the diets of pre-school children, their mothers' knowledge of good feeding and nutrition practices and market access. Using instrumental variable techniques to address the endogeneity of household's nutrition knowledge, we find that nutrition knowledge leads to considerable improvements in children's diets – but only in areas with relatively good market access. Strikingly, improving nutrition knowledge in the most remote localities has no impact on children's dietary diversity.

2. Data

The data used in this analysis come from the second round of a household panel survey conducted in Alefa *woreda* (commune) in the rugged terrain of northwestern Ethiopia. This area was chosen because the large variation in transportation costs over a relatively short distance allows us to carefully assess the impact of these varying costs in a situation of similar physical and climatic conditions. The study site is an isolated area with little to no electricity and mobile phone access, and without any development or humanitarian assistance programs provided by non-governmental organizations. The starting point for the study area is the market town of Atsedemariam, which is connected to a major metropolitan area (Gonder) to the northeast by a gravel road that is passable year round.

Trucks regularly ply the road between Atsedemariam and the product markets in Gonder and beyond with goods originating from and destined for Atsedemariam. To the west of Atsedemariam there exist communities whose access to the outside markets is available for the most part only through Atsedemariam because of the difficult terrain. Further, access to Atsedemariam (and onward to Gonder) is limited to paths along the route that are accessible mainly to foot traffic only, though motorcycles can pass along some portions. To transport agricultural produce to Atsedemariam, and to transport agricultural inputs and consumer goods back from Atsedemariam, community members rely on donkeys. Farmers in the survey area rely on the cooperative office of Atsedemariam as their source of modern inputs.

Households were surveyed along a series of seven sub-districts (or *sub-kebeles*) along the route emanating from Atsedemariam. For sampling purposes, an equal number of households was interviewed in five different distance brackets (measured in travel time by donkey) from the market of Atsedemariam. 170 households were interviewed in each bracket, yielding 850 households in total. Households were sampled evenly from sub-districts within each category to assure a relatively homogenous spread of households over the space between Atsedemariam and the most remote households in Fantaye. The sampling objective was to obtain a representation of households in the districts along the route from the market at Atsedemariam to Fantaye, not to be representative of the population in the woreda. The first round of the survey took place over the course of five weeks in November and December 2011, which followed shortly after the main season (Meher) harvest. The second round, which is used in this analysis, was conducted over a similar period in 2015, and 775 of the original 850 households were interviewed.

Transport costs were measured based on information collected in the household portion of the survey. Using information that households provided on the cost of renting a donkey for a round-trip to Atsedemariam and on how many kilograms a donkey can carry for such a trip, we calculated the cost of transporting one quintal (100 kilograms) on a donkey to Atsedemariam. However, because farming households almost always take their own products to the market by donkey, rather than hiring porters, a more complete measure of transport costs is one that includes the opportunity cost of the farmers' time. Thus our measure of transport cost is based on augmenting the cost of renting a donkey with the imputed value of farmers' travel time. To determine the value of time, we use the median harvest-period wage in the village to value the amount of time that households report that it takes to walk to Atsedemariam and back.[1] This is the measure of transport costs that we use throughout the analysis as a measure of remoteness.[2] On average, it takes households 4.6 (5.2) hours to travel one-way during the dry (wet) season to the market in Atsedemariam. Households incur an average cost of 75.3 Birr to transport a kilogram to Atsedemariam. This varies from 25.6 Birr/kg for the least remote household, to 109.2 Birr/kg for the most remote.

The survey instrument included a module on children's diets. Mothers were asked a series of yes/no questions about foods consumed by all children younger than 60 months who currently resided in the household.[3] Following the recommendations found in WHO (2008) for assessing infant and young child feeding (IYCF) practices, these were grouped into the following categories: Grains, roots and tubers (e.g. barley, maize, teff, and wheat); Legumes and nuts; Dairy products (milk, yogurt, cheese); Flesh foods (meat, poultry and fish products); Eggs; Vitamin A rich fruits and vegetables; and Other fruits and vegetables. This yields an index ranging in value from zero to seven. A more diverse set of foods is necessary if children are to meet both energy and micronutrient needs. Moreover, children who consume from at least four groups during the

[1] Wage data was collected at the household level for three periods: (a) planting and preparation, (b) general cultivation, and (c) harvesting. The mean of the "average daily wage" reported in the sample was the same for all three periods (Birr 43).

[2] To minimize measurement errors in estimating travel times and costs, each household's transport cost is calculated as the average cost of the household's reported cost and the costs reported by its five nearest neighbors. The nearest neighbors are determined using the GPS coordinates for each household.

[3] Note that these questions were not asked at the individual level. Rather they were asked at the household level; about all children less than 60 months old currently residing in the household. As a result, we cannot study differences in diets of children who reside in the same household. This could be problematic if there were gender differences in infant and young child feeding practices. However, recent econometric studies using data from Ethiopia do not find evidence that supports this (Headey 2014; Hirvonen and Hoddinott 2014).

previous day have a high likelihood of consuming animal-source foods and at least one fruit or vegetable as well as a staple food such as a grain, root or tuber (WHO 2008).

This relatively simple indicator is a good proxy of diet quality in the literature and is found to be highly correlated with more detailed and complex measures of food intake such as quantitative recall data of the quantity of all foods consumed by children. For example, Moursi et al. (2008) show that in Madagascar, this index is correlated with mean micronutrient density adequacy (MMDA). Using receiver-operator-curves, they show that children with a score of two or less have predicted low MMDA (less than 50 percent of daily requirements) with 64 percent specificity, 82 percent sensitivity and only 22 percent misclassification. Other studies that have shown that this index is correlated with micronutrient intake and density in both developing countries include Daniels et al. (2009), Kennedy et al. (2007) and Steyn et al. (2006) and the United States, Kant (1996, 2004). Building on earlier work by Arimond and Ruel (2004), this index has been shown to be correlated with longer term measures of children's nutritional status such as height in a variety of developing countries including Bangladesh, Ethiopia, India and Zambia (Jones et al. 2014). Disha et al. (2012) also find evidence that this index is correlated with child height in Ethiopia.

3. Descriptive analysis

Figure 1 shows the distribution of our dietary diversity indicator for the 448 children less than 60 months of age in our sample. Children residing in an average household in our sample eat from 3 food groups. Only in 7.5 percent of the households, the children meet the WHO recommendation of eating from 4 or more food group. There are a small number of children whose mothers reported that they consumed none of these foods during the previous day either because they were ill or because they only consumed breastmilk.

Figure 2 shows the content of this monotonous diet. Nearly all children consume staple crops (grains, roots and tubers) and legumes and nuts. About one third consume dairy products but the consumption of other animal source foods (meat and eggs) is uncommon. Similarly, only less than 10 percent of the children consume vegetables or fruits that are rich in Vitamin A but the consumption of other fruits and vegetables is relatively high. Results such as those shown in Figure 2 are not uncommon in Ethiopia; a similar pattern is found for example in Nguyen et al. (2013).

Households' nutrition knowledge is captured through seven statements about appropriate infant and young child feeding practises (Alive & Thrive 2014). The respondents were asked whether they agreed or disagreed with these statements. Table A1 of the Appendix A provides an overview of the statements and the distribution of the responses to each statement. We reduce the household responses to these statements into one index using principal components analysis. The seven statement variables are highly correlated (average correlation coefficient is 0.330) and the principal components analysis attempts to find components that account for most of the variation among these variables. The end product is a single variable that we take to represent household's nutrition knowledge. Appendix A provides a more detailed description of the principal components analysis and the results. Moreover, to facilitate interpretation, household's nutrition knowledge is expressed in units of z-scores.[4] Figure 3 shows how this nutrition knowledge score is distributed among the households. A higher number indicates better knowledge. We see that most observations lie within two standard deviations from zero.[5]

[4] The z-scores are computed by subtracting the initial knowledge value from the sample mean and then dividing this with the standard deviation of the sample.
[5] We considered 5 households as outliers for which the nutrition knowledge score was less than -2.5. However, our results are not sensitive to these extreme values; including these 5 households yields identical coefficients in all regression models.

As education, especially of mothers, has been shown to matter significantly for nutritional outcomes, we look in more detail at this variable. Formal education levels are extremely low in this context. Figure 4 shows the distribution of education in our sample. The average level of education in these data is 2 years. More than 50 percent of the households do not have a single adult member who has attended school. Figure 5 estimates a locally weighted regression of the association between level of education and household's nutrition knowledge. The relationship is flat throughout the education distribution implying that education does not explain differences in nutrition knowledge.

In Figure 6 we test whether formal education is correlated with children's dietary diversity. We do not find strong evidence that children originating from households with more education enjoy better diets. There is some evidence of this at the top end of the schooling distribution but the small number of households with higher level of education renders the confidence intervals extremely wide. [6]

Figure 7 shows the relationship between children's diets and household nutrition knowledge. At the first half of the knowledge distribution we see that dietary diversity is positively associated with nutrition knowledge. After a certain threshold, dietary diversity no longer increases with nutrition knowledge. The figure further shows that children's dietary diversity is still relatively low, even with good nutritional knowledge scores.

We then categorized households into low (bottom half in the nutrition knowledge distribution) and high (top half of the distribution) nutrition knowledge households. Figure 8 shows how children's

[6] This is consistent with results found in Alderman and Headey (2014) who study the role of (maternal and paternal) education on children's nutritional status. They find that while child nutrition status is positively correlated with parental education, the correlation gets stronger with secondary education.

dietary diversity varies across the transport cost gradient for both household types. We see that children in households with better nutrition knowledge enjoy more diverse diets across the remoteness gradient – except in the most remote localities.[7] However, the confidence intervals (not reported) are large and overlap across the two regression lines. The evidence is therefore at best suggestive. This motivates the use of econometric techniques to further study this relationship.

4. Econometric approach

We model the number of food groups consumed by children in household h located in community v (d_{hv}) as a function of household nutrition knowledge (k_{hv}):

$$d_{hv} = \beta k_{hv} + \gamma r_{hv} + c_{hv}{}' \delta + x_{hv}{}' \varphi + \omega + \varepsilon_{hv}, \quad (1)$$

where r_{hv} captures household's remoteness. More specifically, r obtains a value 1 if the household's transportation costs to the main market are among the top third of the transportation cost distribution (zero otherwise).[8] The term x_{hv}' is a vector of household level characteristics including household demographics (including representation of different age groups) and household assets. Table 1 provides the summary statistics for all variables used in the analysis. The observed and unobserved community (sub-kebele) level characteristics are captured through community dummies (ω).[9] The last term in the equation, ε_{hv}, represents the error term.

A problem in estimating Equation (1) arises from the fact that household's nutrition knowledge cannot be directly observed. Here we proxy for household knowledge through the seven nutrition

[7] Using data for the whole sample we find that children's dietary diversity decreases with transportation costs. This is in line with earlier research from this context, see Stifel and Minten (2015).
[8] All results are also robust to defining household remoteness as the top 4th and 5th quintile of the transportation cost distribution.
[9] The 448 households in the final sample group into 7 communities (sub-kebeles). All results are robust to estimating the model without the sub-kebele dummies.

statements (see Section 3). However, this – as any other attempt – to measure the 'true level nutrition knowledge' will result in a variable that is measured with some degree of error (Variyam et al. 1999). Such measurement error in the independent variable, if randomly distributed with zero mean, typically leads to a lower-bound estimate (e.g. Deaton 1997). To address this issue, we opt for an IV-strategy based on household's access to the health information. The IV strategy also addresses concerns that the estimated impact is driven by some unobserved household characteristics correlated with the knowledge variable.

The identification strategy is based on insights drawn from Ethiopia's strategy to combat under-nutrition in the country. Since the start of the 2008 National Nutrition Programme, Ethiopia's nutrition strategy has followed a community based approach where community (kebele*)* serves as a delivery platform for various health services (Lemma and Matji 2013). The community based approach is part of the national Health Extension Programme (HEP), initiated back in 2003. The community based nutrition program is widespread, covering nearly all districts (woredas) of the country. Since its initiation, more than 30,000 Health Extension Workers have been trained and placed into communities (White and Mason 2012). One of the key tasks of the health extension workers is the provision of health education. This is done together with volunteers recruited from the communities (Wakabi 2008). The HEP typically deploys two health extension workers per health post and together these government employees are supposed to reach approximately 5,000 individuals.

The national nutrition programme also targets media through capacity building and by providing nutrition information materials to the media outlets (GFDRE 2013). Radio and TV broadcasts contain nutrition related messages that promote dietary diversity and discuss about the importance of micro-nutrients.

We use these insights to develop our identification strategy. Our first instrument is a dummy that obtains a value 1 if the household owns a radio (through which they can access the radio broadcasted nutrition messages). Moreover, the survey instrument contains a question whether the household was visited by the Health Extension Worker or a health volunteer in the past 12 months. This variable is our second instrument: we hypothesize that these visits lead to improvements in households' nutrition knowledge. Given the central role of the health workers and the radio stations in delivering nutrition information, these two instruments should perform as good predictors of household nutrition knowledge.

There are two obvious concerns related to the exclusion restriction. The first is that these variables are correlated with household's remoteness. Using a simple local polynomial regressions suggests that this is concern is valid for the health extension visits (more remote households are less likely to be visited) but not for the radio ownership. However, we address this issue by including the remoteness variable into our specification in addition to the community dummies. The second concern is that radio ownership captures some type of wealth effect. We address this issue by including controls for household wealth (livestock ownership, value of productive assets and land size). Furthermore, it is worth noting that the cost of owning a radio is rather low. [10]

Table 2 shows the first stage results for the two excluded instruments based on two-step linear IV-GMM approach.[11] The excluded instruments appear with expected signs: both radio ownership

---

[10] A recent paper by Pylypchuk and Norton (2014) also instrument knowledge using radio ownership to study the role of maternal knowledge on malaria prevention measures in Zambia. They make the same point about the relatively low cost of radios nowadays. They also point out that listening to the radio does not require literacy or schooling attainment.

[11] The linear two-step GMM model implemented here is more efficient than the conventional two-stage least squares model when standard errors are heteroskedastic and the equation is over-identified (see e.g. Cameron and Trivedi 2005, p. 187-8). The two conditions are satisfied in our application. First, the number of instruments (2) exceeds the number of endogenous regressors (1). Second, the null of homoscedasticity is rejected in our application: the White (1980) test, less sensitive to departures from normality, yields 411.6 exceeding all the conventional critical values.

and visits by the Health Extension Worker are associated with better nutrition knowledge. Both coefficients appear positive and significant at least at the 5 percent level. The IV-diagnostics further show that the instruments are relevant; the Cragg-Donald test yields a value of 9.82 and Angrist and Pischke (2009) test for weak identification is passed ($p<0.001$). Finally, according to the Hansen-Sargan test, we cannot reject the null of zero correlation between the instruments and the error term.

5. Results

Table 3 shows the regression results based on the estimation of Equation (1). Column 1 provides the OLS result that treats nutrition knowledge as exogenous and column 2 shows the IV-regression results. The coefficients on the control variables are *a priori* correct. Livestock ownership and other wealth are associated with better children's dietary diversity. The presence of infants lowers the dietary diversity for the simple reason that these children are likely to be exclusively breastfed.

The coefficient on the transportation variable appears negative in both columns suggesting that there is a remoteness penalty on children's diets. However, the coefficient is not statistically significant in the OLS model (column 1). The IV-model produces a larger coefficient (in absolute terms) but the coefficient is statistically significant only at the 10 percent level. Children residing in the most remote areas consume 0.35 food groups less than other children, on average and after controlling for household wealth, demographics and level of education.

Household nutrition knowledge appears with a positive coefficient in both columns. However, according to the OLS model, nutrition knowledge has a modest and statistically imprecisely estimated effect on children's diet. In contrast, the instrumented coefficient in column 2 appears highly significant ($p<0.01$) and large. This difference between the OLS and IV-estimates is likely due to the measurement error in the nutrition knowledge variable. [12] According to the IV-estimate, a one standard deviation increase in household's nutrition knowledge score leads to 0.8 food group increase in children's diets, on average and *ceteris paribus*. An alternative way of interpreting this effect is the following. The average household in the sample has a (standardized) nutrition knowledge score of 0.00 and their children consume from 3.08 food groups. Improving this household's nutrition knowledge to the level of the most knowledgeable household in the sample (knowledge score = 1.42) would result in a 1.15 food group increase in children's diets. As a result, children in this average household would now eat from 4.2 food groups thus satisfying the WHO (2008) guideline of having minimum of four food groups per day.

We assessed the robustness of this finding in several ways. First, the nutrition statements displayed in Table A1 of Appendix A measure households' general knowledge about feeding practises. One of the statements is directly linked to our outcome variable (dietary diversity): "Give a variety of foods to very young children (6-24 months)". Households that agree with this statement should feed their children with a more diverse diet than other households. In order to conduct this 'sanity check', we re-constructed the knowledge variable so that households that strongly agree with the statement receive 4 points while those who strongly disagree receive 0 points. The table in Appendix B re-runs Table 2 using this knowledge variable that runs from 0 to 4. The coefficient on the OLS model does not appear statistically significant (p=0.142) but the coefficient in the IV-model is statistically significant at the 5 percent level (p=0.013). As before, higher scores in this alternative knowledge variables lead to improvements in children's dietary diversity.

[12] The IV-strategy can be thought to correct this measurement error by 'forcing' the nutrition knowledge variable to respond to the BCC treatments received by the households. In other words, the identification in the IV-regression comes from households who respond to the changes triggered by the instruments. Therefore, our estimate – as any IV-estimate – should be interpreted in terms of a Local Average Treatment Effect (LATE) (see Angrist, Imbens, and Rubin 1996).

Second, while our outcome variable is essentially a count it does not seem to follow a Poisson distribution (see Figure 1). Still, we assessed the robustness of our findings using the Poisson model. Appendix C shows that the estimated coefficients (marginal effects) are very similar to the ones obtained using the linear models in Table 3. The results do not seem then to be driven by the non-linear nature of our left-hand side variable.

Third, it is reported that health workers may refer under-nourished children to the health post for 'therapeutic feeding' (White and Mason 2012). While we believe such practice to be rare, this would obviously violate the exclusion restriction. To check that our results are not driven by this possibility, we re-ran Column 2 in Table 3 without the health worker visit instrument (Appendix D). The coefficient on the nutrition knowledge variable is nearly identical to the one observed in Column 2 in Table 3.

Finally, the nutrition literature recommends administering the dietary questions at the individual level (Ruel 2003). Another caveat in our data is that it covers a relatively small area raising concerns about external validity. We addressed both of these concerns by replicating the analysis using the Feed-the-Future midline survey implemented in June and July in 2015.[13] While not nationally representative, the survey is widespread, being administered in 252 villages in 84 of the 670 rural districts (*woredas*) in five regions of the country (Amhara, Oromia, Tigray, Somale and SNNP).[14] The sample consists of 4,107 children who are between 6 and 59 months of age. We added the same nutrition knowledge questions to this survey but the dietary diversity questions were asked at the individual level (child) level based on a 24 hour recall. Using the same IV-

[13] The main purpose of the survey was to obtain post-intervention (midline) information in localities that were to receive investments to improve agricultural production and nutrition under the Feed the Future (FtF) program funded by the United States Agency for International Development (USAID), or in localities that were to act as comparison sites for the evaluation of FtF.
[14] For further information about this survey (including sampling), see Bachewe et al. (2014).

strategy we get coefficients that are similar in magnitude (see Appendix E). According to the IV-specification, increasing household nutrition knowledge by one standard deviation leads to a 0.75 food group increase in children's dietary diversity. This implies that the results presented in Table 4 do not appear specific to this one woreda in the Amhara region. Furthermore, the use of household level dietary data does not seem to affect our findings. Next we return to the transportation survey data that permits a precise measure of households' market access (or remoteness).

6. Results by market access

The foregoing results show how improved nutrition knowledge leads to better diets. In this section, we study whether this relationship depends on households' access to markets. Specifically, we interact household nutrition knowledge (k_{hv}) in Equation (1) with the remoteness variable (r_{hv}). The analysis is somewhat complicated by the fact that by interacting the endogenous variable with an exogenous one results in two endogenous variables. To account for the possibility that the instruments work differently for remote and less-remote villages we also interacted the two instruments with the remoteness variable. As a result, we now have four excluded instruments. Table 4 reports results. Focusing on the IV-results in column 2, we see that for both variables, the Angrist and Pischke (2009, p. 217-218) F-statistic rejects the null that the endogenous regressor is weakly identified. The coefficient on the nutrition knowledge variable appears significant at the 1 percent level. For less remote households, improving household nutrition knowledge by one standard deviation increases children's dietary diversity by 1.8 food groups. The coefficient on the interaction term is negative and almost of the same magnitude in absolute terms as the coefficient on the non-interacted knowledge variable. This coefficient is statistically significant at the 5 percent level. Furthermore, the joint significance test implies that the knowledge variable coefficient for the children residing far from the markets is not statistically different from zero at conventional levels (p=0.746). This means that for remote households, improvements in nutrition knowledge do not lead to increases in children's dietary diversity.

7. Conclusions

This paper studied the impact of improving households' nutrition knowledge and its complementary with market access. Using novel survey data that permits a careful measure of market access we find that nutrition knowledge leads to considerable improvements in children's diets but only in areas with relatively good market access. Improving nutrition knowledge in the most remote localities has no impact on children's dietary diversity.

Bibliography

African Union. 2015. African Regional Nutrition Strategy 2015-2025. In *First meeting of the specialised Technical Committee on Health, Population and Drug Control*. Addis Ababa, Ethiopia: African Union.

Alderman, H., and D. D. Headey. 2014. The nutritional returns to parental education. *IFPRI Discussion Paper* 01379.

Alive & Thrive. 2010. IYCF practices, beliefs and influences in Tigray Region, Ethiopia. edited by Alive and Thrive. Addis Ababa, Ethiopia.

———. 2014. Overview of the approach for mobilizing families and communities in Ethiopia to adopt seven feeding actions. Addis Ababa, Ethiopia: Alive & Thrive.

Angrist, J. D., G. W. Imbens, and D. B. Rubin. 1996. "Identification of causal effects using instrumental variables." *Journal of the American statistical Association* no. 91 (434):444-455.

Angrist, J. D., and J.-S. Pischke. 2009. *Mostly harmless econometrics: an empiricist's companion*. Princeton: Princeton University Press.

Arimond, M., and M. T. Ruel. 2004. "Dietary diversity is associated with child nutritional status: evidence from 11 demographic and health surveys." *The Journal of Nutrition* no. 134 (10):2579-2585.

Bachewe, F., G. Berhane, K. Hirvonen, J. Hoddinott, J. Hoel, F. Tadesse, A. S. Taffesse, T. Woldehanna, I. Worku, F. Yimer, Y. Yohannes, K. Abay, B. Yigezu, H. Beyene, Y. Abebe, and N. Meshesha. 2014. Feed the Future (FtF) of Ethiopia – Baseline report 2013. Ethiopia Strategy Support Program II (ESSP II) and International Food Policy Research Institute.

Bhandari, N., S. Mazumder, R. Bahl, J. Martines, R. E. Black, and M. K. Bhan. 2004. "An educational intervention to promote appropriate complementary feeding practices and physical growth in infants and young children in rural Haryana, India." *The Journal of nutrition* no. 134 (9):2342-2348.

Black, R. E., C. G. Victora, S. P. Walker, Z. A. Bhutta, P. Christian, M. De Onis, M. Ezzati, S. Grantham-McGregor, J. Katz, and R. Martorell. 2013. "Maternal and child undernutrition and overweight in low-income and middle-income countries." *The Lancet* no. 382 (9890):427-451.

Cameron, A. C., and P. K. Trivedi. 2005. *Microeconometrics: methods and applications*. New York, NY: Cambridge University Press.

Central Statistical Agency, and ICF International. 2012. Ethiopia Demographic and Health Survey 2011. Addis Ababa, Ethiopia and Calverton, Maryland, USA: Central Statistical Agency and ICF International.

Daniels, M., L. Adair, B. Popkin, and Y. Truong. 2009. "Dietary diversity scores can be improved through the use of portion requirements: an analysis in young Filipino children." *European journal of clinical nutrition* no. 63 (2):199-208.

Deaton, A. 1997. *The analysis of household surveys: a microeconometric approach to development policy*. Baltimore, MD: Published for the World Bank by Johns Hopkins University Press.

Disha, A., R. Rawat, A. Subandoro, and P. Menon. 2012. "Infant and Young Child Feeding (IYCF) practices in Ethiopia and Zambia and their association with child nutrition: analysis of Demographic and Health Survey data." *African Journal of Food, Agriculture, Nutrition and Development* no. 12 (2):5895-5914.

GFDRE. 2013. National Nutrition Programme 2008-2015. Government of the Federal Democratic Republic of Ethiopia (GFDRE).

Headey, D. 2014. An Analysis of Trends and Determinants of Child Undernutrition in Ethiopia, 2000-2011. *Ethiopia Strategy Support Program (ESSP) Working Paper* 70.

Hidrobo, M., J. Hoddinott, N. Kumar, and M. Olivier. 2015. Social protection, food security and assets. *mimeo*.

Hirvonen, K., and J. Hoddinott. 2014. Agricultural production and children's diets: Evidence from rural Ethiopia. *IFPRI-ESSP Working Paper* (69).

Hoddinott, J., J. R. Behrman, J. A. Maluccio, P. Melgar, A. R. Quisumbing, M. Ramirez-Zea, A. D. Stein, K. M. Yount, and R. Martorell. 2013. "Adult consequences of growth failure in early childhood." *The American journal of clinical nutrition*:ajcn. 064584.

Hoddinott, J., D. Headey, and M. Dereje. 2015. "Cows, missing milk markets and nutrition in rural Ethiopia." *Journal of Development Studies* no. 51 (8):958–975.

Jones, A. D., S. B. Ickes, L. E. Smith, M. N. Mbuya, B. Chasekwa, R. A. Heidkamp, P. Menon, A. A. Zongrone, and R. J. Stoltzfus. 2014. "World Health Organization infant and young child feeding indicators and their associations with child anthropometry: a synthesis of recent findings." *Maternal & child nutrition* no. 10 (1):1-17.

Kant, A. K. 1996. "Indexes of overall diet quality: a review." *Journal of the American Dietetic Association* no. 96 (8):785-791.

———. 2004. "Dietary patterns and health outcomes." *Journal of the American Dietetic Association* no. 104 (4):615-635.

Kennedy, G. L., M. R. Pedro, C. Seghieri, G. Nantel, and I. Brouwer. 2007. "Dietary diversity score is a useful indicator of micronutrient intake in non-breast-feeding Filipino children." *The Journal of nutrition* no. 137 (2):472-477.

Lemma, F., and J. Matji. 2013. "Delivery platforms for sustained nutrition in Ethiopia. Comment." *Lancet*:2.

Manley, J., S. Gitter, and V. Slavchevska. 2013. "How Effective are Cash Transfers at Improving Nutritional Status?" *World development* no. 48:133-155.

Moursi, M. M., M. Arimond, K. G. Dewey, S. Trèche, M. T. Ruel, and F. Delpeuch. 2008. "Dietary diversity is a good predictor of the micronutrient density of the diet of 6-to 23-month-old children in Madagascar." *The Journal of nutrition* no. 138 (12):2448-2453.

Nguyen, P. H., R. Avula, M. T. Ruel, K. K. Saha, D. Ali, L. M. Tran, E. A. Frongillo, P. Menon, and R. Rawat. 2013. "Maternal and child dietary diversity are associated in Bangladesh, Vietnam, and Ethiopia." *The Journal of nutrition* no. 143 (7):1176-1183.

Penny, M. E., H. M. Creed-Kanashiro, R. C. Robert, M. R. Narro, L. E. Caulfield, and R. E. Black. 2005. "Effectiveness of an educational intervention delivered through the health services to improve nutrition in young children: a cluster-randomised controlled trial." *The Lancet* no. 365 (9474):1863-1872.

Pylypchuk, Y., and S. W. Norton. 2014. "Preventing malaria among children in Zambia: the role of mother's knowledge." *Health Economics* no. 24 (11):1389–1402. doi: 10.1002/hec.3093.

Ruel, M. T. 2003. "Operationalizing dietary diversity: a review of measurement issues and research priorities." *The Journal of nutrition* no. 133 (11):3911S-3926S.

Santos, I., C. G. Victora, J. Martines, H. Gonçalves, D. P. Gigante, N. J. Valle, and G. Pelto. 2001. "Nutrition counseling increases weight gain among Brazilian children." *The Journal of nutrition* no. 131 (11):2866-2873.

Sibhatu, K. T., V. V. Krishna, and M. Qaim. 2015. "Production diversity and dietary diversity in smallholder farm households." *Proceedings of the National Academy of Sciences* no. 112 (34):10657-10662.

Steyn, N., J. Nel, G. Nantel, G. Kennedy, and D. Labadarios. 2006. "Food variety and dietary diversity scores in children: are they good indicators of dietary adequacy?" *Public health nutrition* no. 9 (05):644-650.

Stifel, D., and B. Minten. 2015. Market Access, Welfare, and Nutrition: Evidence from Ethiopia. *ESSP Working paper* 77.

United Nations. 2015. Transforming our world: the 2030 Agenda for Sustainable Development. In *United Nations Resolution A/RES/70/1*. New York.

USAID. 2011. Integration of nutrition education into the Ethiopia Urban Gardens Program: Results of recipe trials and focus group discussions. edited by USAID: Infant & Young Child Nutrition Project. Washington DC.

———. 2014. USAID’s 2014-2025 Multi-Sectoral Nutrition Strategy

Variyam, J. N., J. Blaylock, B.-H. Lin, K. Ralston, and D. Smallwood. 1999. "Mother's nutrition knowledge and children's dietary intakes." *American Journal of Agricultural Economics* no. 81 (2):373-384.

Wakabi, W. 2008. "Extension workers drive Ethiopia's primary health care." *The Lancet* no. 372 (9642):880.

White, H. 1980. "A Heteroskedasticity-Consistent Covariance-Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica* no. 48 (4):817-838.

White, J., and J. Mason. 2012. "Assessing the impact on child nutrition of the Ethiopia Community-based Nutrition Program." *Report to UNICEF of an evaluation study. Addis Ababa: Tulane SPHTM/UNICEF*.

WHO. 2008. Indicators for assessing infant and young child feeding practices: part 1: definitions: conclusions of a consensus meeting held 6-8 November 2007 in Washington DC, USA. edited by World Health Organization.

WHO, and Unicef. 2003. Global strategy for infant and young child feeding.

Zaman, S., R. N. Ashraf, and J. Martines. 2008. "Training in complementary feeding counselling of healthcare workers and its influence on maternal behaviours and child growth: a cluster-randomized controlled trial in Lahore, Pakistan." *Journal of health, population, and nutrition* no. 26 (2):210.

Figures

Figure 1: Distribution of the dietary diversity indicator

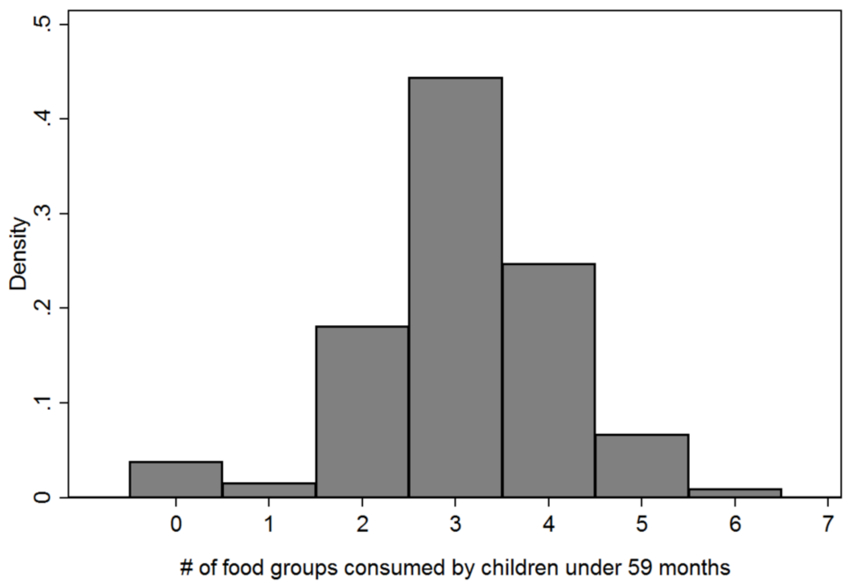


Figure 2: Diet content: share of children consuming from different food groups

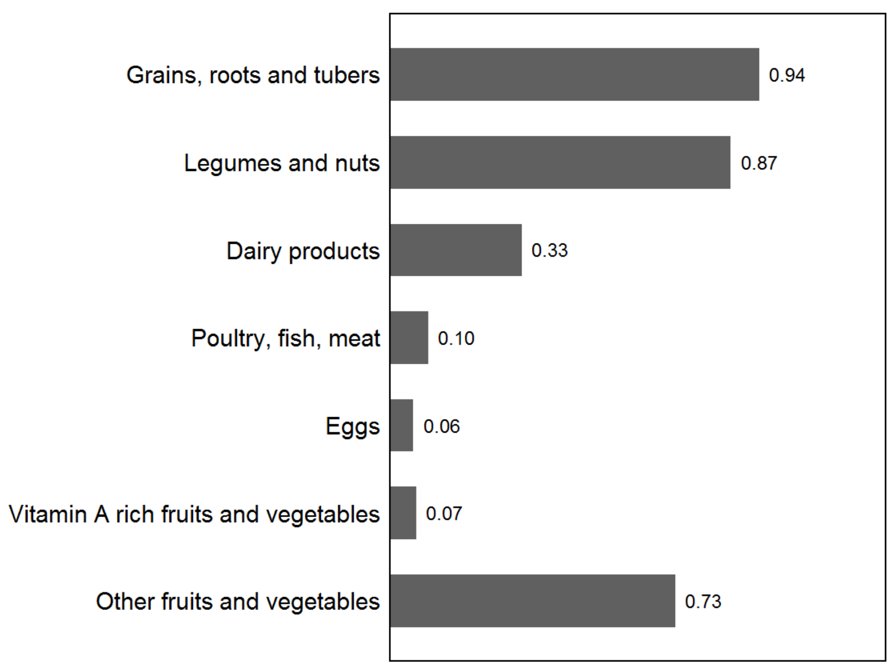


Figure 3: Distribution of household nutrition knowledge score

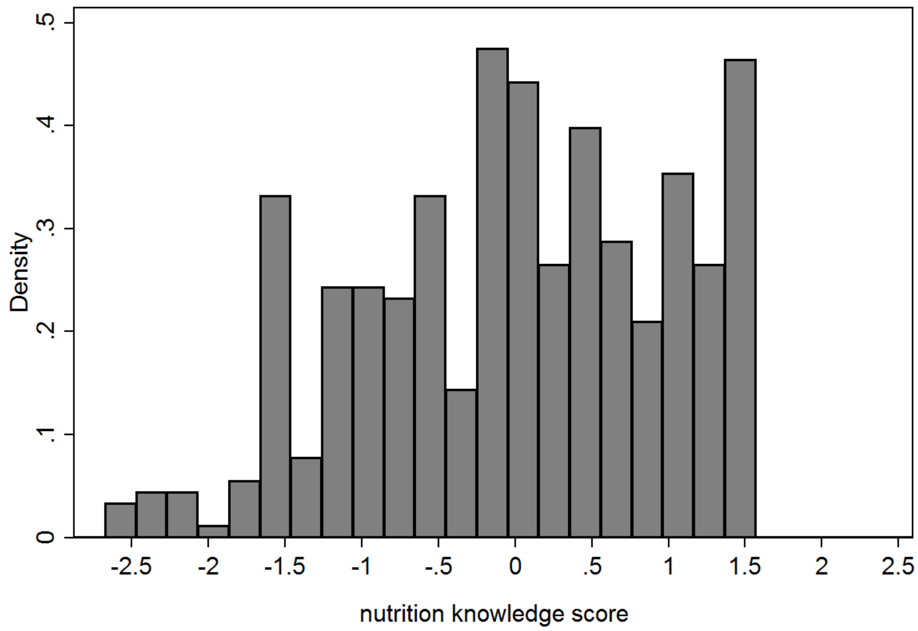
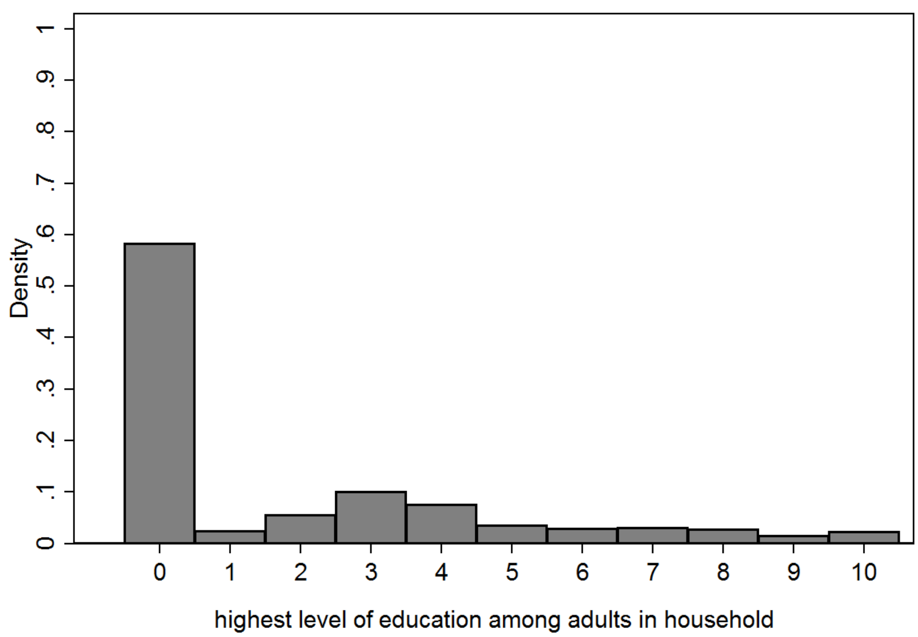
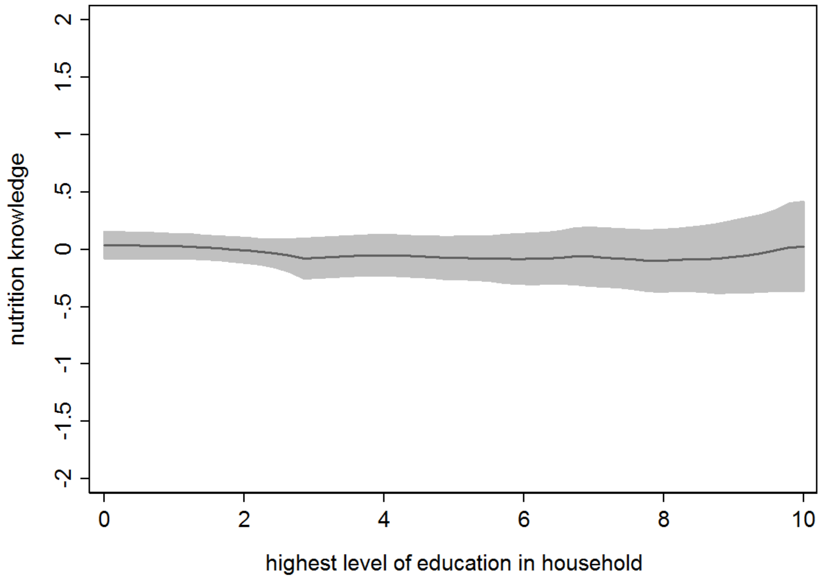


Figure 4: Distribution of education levels



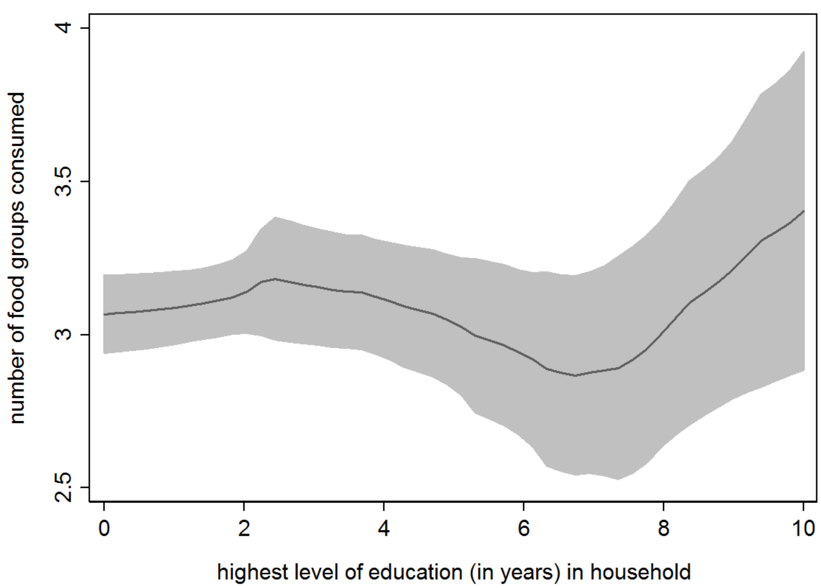
Note: horizontal axis measures the highest level of education (in years) in the household among adult members (aged 15 or more).

Figure 5: Formal education and nutrition knowledge



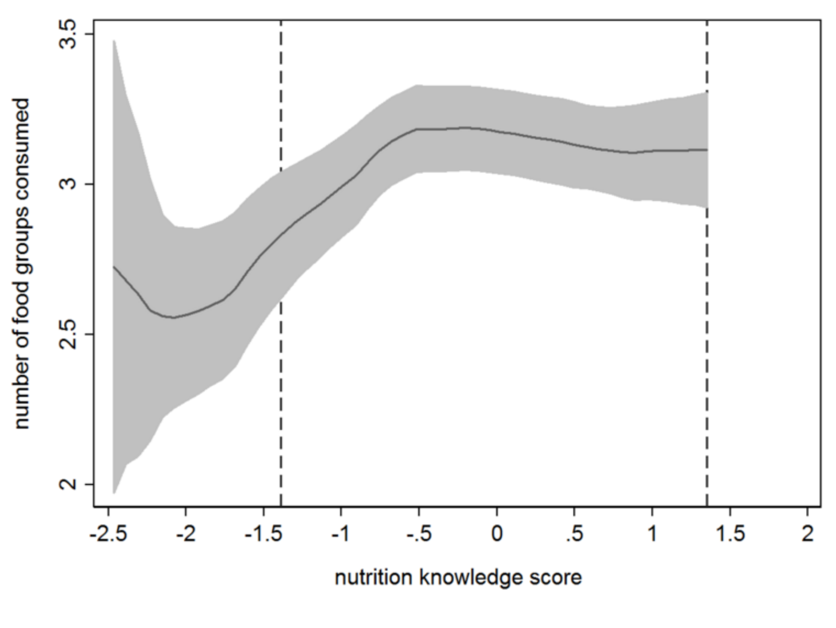
Note: Local polynomial regression. Shaded area refers to 95%-confidence interval. Horizontal axis measures the highest level of education (in years) in the household among adult members (aged 15 or more).

Figure 6: Formal education and children's dietary diversity



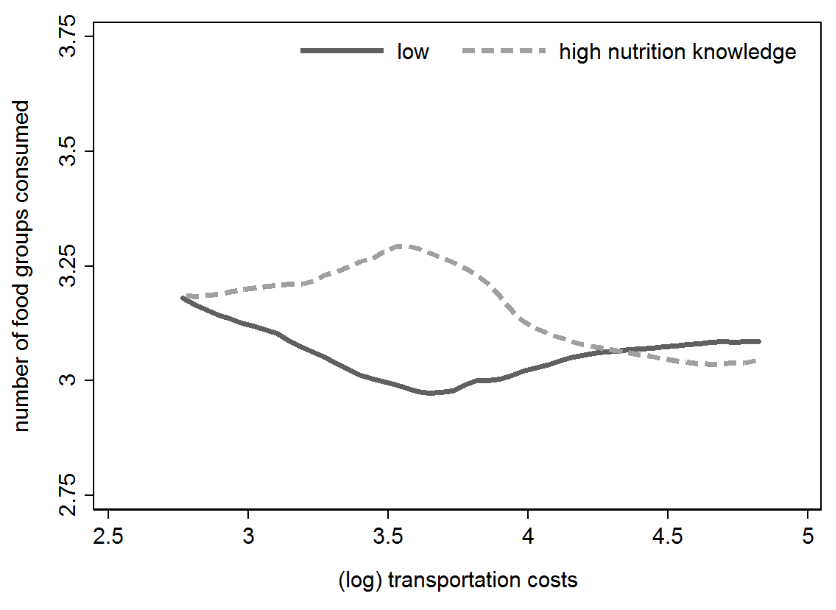
Note: Local polynomial regression. Shaded area refers to 95%-confidence interval. Horizontal axis measures the highest level of education (in years) in the household among adult members (aged 15 or more).

Figure 7: Nutrition knowledge and children's dietary diversity



Note: Local polynomial regression. Shaded area refers to 95%-confidence interval. Dashed lines represent the bottom and top 5% of the nutrition knowledge distribution.

Figure 8: Remoteness and children's dietary diversity, by level of nutrition knowledge



Tables

Table 1: Summary statistics

	mean (std. dev.)
# of food groups consumed by children under 60 months	3.089 (1.075)
household nutrition knowledge (z-score)	0.00 (1.000)
remote household (*)	0.324 (0.468)
# of HH members less than 1 years old	0.161 (0.380)
# of HH members 1 years old	0.192 (0.405)
# of HH members 2 years old	0.315 (0.474)
# of HH members 3 years old	0.306 (0.461)
# of HH members 4 years old	0.395 (0.489)
# males 5-9 years of age in HH	0.746 (0.764)
# females 5-9 years of age in HH	0.701 (0.730)
# males 10-15 years of age in HH	0.634 (0.803)
# females 10-15 years of age in HH	0.547 (0.725)
# males 16 or more years of age in HH	1.355 (0.737)
# females 16 or more years of age in HH	1.257 (0.582)
mother's (head's spouse's) education	0.308 (1.137)
age of the head	38.23 (10.08)
HH has a safe drinking water source (*)	0.221 (0.415)
HH owns chicken (*)	0.931 (0.254)
HH owns sheep or goats (*)	0.623 (0.485)
HH owns calves, cows or heifer (*)	0.846 (0.361)
(log) value of productive assets owned	6.592 (0.659)
(log) land size (in acres)	3.398 (0.881)

Included instruments:	
household owns a radio (*)	0.248
	(0.432)
HH was visited by a health worker (HEW/volunteer) (*)	0.795
	(0.404)

Note: (*) indicates a dummy variable. Household's location is considered remote if it belongs to top third in the transportation cost distribution.

Table 2: First-stage regression results

outcome variable: nutrition knowledge	**(1)**
household owns a radio	0.380***
	(0.103)
household was visited by a health worker	0.282**
	(0.128)
Included instruments? [a)]	Yes
Sub-Kebele dummies? [a)]	Yes
Weak Identification tests:	
Cragg-Donald Wald F statistic	9.82
Kleibergen-Paap rk Wald F statistic	10.00
Angrist-Pischke F-test – p-value	0.000***
Over-identification test:	
Hansen/Sargan test	0.921
- p-value	0.337
Number of observations:	448

Note: Remoteness dummy obtains a value 1 if the household belongs to top third in the remoteness distribution, zero otherwise. Robust standard errors in parentheses. Statistical significance denoted at *** p<0.01, ** p<0.05, * p<0.1.

[a)] Coefficients omitted to preserve space.

Table 3: Impact of household nutrition knowledge on children's dietary diversity

Dependent variable: number of food groups consumed	OLS (1)	IV (3)
nutrition knowledge score	0.089*	0.809***
	(0.049)	(0.292)
remoteness dummy	-0.177	-0.350*
	(0.151)	(0.197)
# of HH members less than 1 years old	-0.325**	-0.218
	(0.141)	(0.177)
# of HH members 1 years old	0.368***	0.179
	(0.136)	(0.194)
# of HH members 2 years old	0.681***	0.811***
	(0.127)	(0.167)
# of HH members 3 years old	0.616***	0.705***
	(0.141)	(0.177)
# of HH members 4 years old	0.542***	0.571***
	(0.131)	(0.154)
# males 5-9 years of age in HH	-0.089	-0.075
	(0.070)	(0.090)
# females 5-9 years of age in HH	0.002	-0.007
	(0.076)	(0.089)
# males 10-15 years of age in HH	-0.004	0.042
	(0.073)	(0.095)
# females 10-15 years of age in HH	0.018	0.140
	(0.076)	(0.107)
# males 16 or more years of age in HH	-0.089	-0.074
	(0.078)	(0.092)
# females 16 or more years of age in HH	0.088	0.088
	(0.105)	(0.118)
Education of the spouse (years)	0.037	0.085*
	(0.034)	(0.044)
age of the head	-0.053*	-0.067*
	(0.029)	(0.038)
age of the head squared	0.001**	0.001**
	(0.000)	(0.000)
HH has a safe drinking water source	0.171	0.123
	(0.135)	(0.167)
HH owns chicken	0.451**	0.488*
	(0.205)	(0.263)
HH owns sheep or goats	-0.020	-0.103
	(0.099)	(0.122)
HH owns calves, cows or heifer	0.287**	0.377**
	(0.141)	(0.166)
(log) value of productive assets owned	0.287***	0.327***

	(0.085)	(0.110)
(log) land size (in acres)	0.047	-0.006
	(0.073)	(0.087)
Sub-Kebele dummies? [a)]	Yes	Yes
R^2	0.204	-
Weak Identification tests:		
Cragg-Donald Wald F statistic	-	9.82
Angrist-Pischke F-test	-	10.00
p-value	-	0.000***
Over-identification test:		
Hansen/Sargan test	-	0.921
p-value	-	0.337
Number of observations	448	448

Note: Remoteness dummy obtains a value 1 if the household belongs to top third in the remoteness distribution, zero otherwise. Robust standard errors in parentheses. Statistical significance denoted at *** p<0.01, ** p<0.05, * p<0.1.

[a)] Coefficients omitted to preserve space

Table 4: Impact of household nutrition knowledge on children's dietary diversity by remoteness

Dependent variable: number of food groups consumed	OLS (1)	IV (2)
nutrition knowledge (*A*)	0.101	1.789**
	(0.065)	(0.696)
nutrition knowledge X remoteness (*B*)	-0.033	-1.704**
	(0.097)	(0.788)
remoteness dummy	-0.176	-0.410*
	(0.152)	(0.243)
Other controls? [a)]	Yes	Yes
Sub-Kebele dummies? [a)]	Yes	Yes
χ^2-test: joint significance: (*A*)+(*B*) = 0	p = 0.346	p = 0.746
R^2	0.204	-
Weak Identification tests:		
Angrist-Pischke F-test: (*A*)	-	2.45
p-value	-	0.063*
Angrist-Pischke F-test: (*B*)	-	4.30
p-value	-	0.005***
Over-identification test:		
Hansen/Sargan test	-	1.017
p-value	-	0.601
Number of observations	448	448

Note: Remoteness dummy obtains a value 1 if the household belongs to top third in the remoteness distribution, zero otherwise. Robust standard errors in parentheses. Statistical significance denoted at *** p<0.01, ** p<0.05, * p<0.1.

[a)] Coefficients omitted to preserve space

Appendix A: Measuring nutrition knowledge

The survey instrument had 7 nutrition statements that attempted to test households' nutrition knowledge. The statements are based on the "7 Excellent Feeding Actions" by promoted by the Alive & Thrive in Ethiopia (Alive & Thrive 2014). Agreeing with the statement indicates that the respondent is knowledgeable about the best infant and young child feeding practices. Table A1 shows these statements and the distribution of the responses.

Table A1: Nutrition knowledge statements and the distribution of the responses (in %)

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Colostrum should be given to the baby	15.4	22.4	19.6	26.7	15.9
Give only breast milk for the first six months of life	2.2	4.1	13.3	46.3	34.1
Babies should eat thick porridge once they stop breastfeeding	2.7	6.2	18.7	43.1	29.3
Very young children (6-24 months) should eat eggs and meat	2.7	6.5	17.3	43.5	30.1
Porridge should be made by adding vegetables, eggs, milk	1.4	2.1	14.3	49.4	32.8
Give a variety of foods to very young children (6-24 months)	1.8	3.5	13.9	43.6	37.2
Give your sick children (6-24 months) more food than usual	2.8	9.6	18.8	43.2	25.6

Table A2 shows the eigenvalues for each of the seven components. We see that the first two component explains 66 percent of the variation in the data. We follow the Kaiser-rule that states that only components that obtain an eigenvalue larger than one should be retained.

Table A2: Principal components and eigenvalues

	Eigenvalue	Proportion
component 1	3.57	0.51
component 2	1.04	0.15
component 3	0.75	0.11
component 4	0.56	0.08
component 5	0.48	0.07
component 6	0.33	0.05
component 7	0.27	0.04

Columns 1 and 2 in Table A3 then provide the principal component loadings based on the first two components. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy is displayed in column 3. We see that all KMO values are close to 1 suggesting that the 7 statement variables are indeed measuring a common component.

Table A3: principal component loadings and sampling adequacy

	1	**2**	**3**
Statement	**Comp 1**	**Comp 2**	**Kaiser-Meyer-Olkin measure**
Colostrum should be given to the baby	0.14	0.84	0.79
Give only breast milk for the first six months of life	0.32	0.40	0.87
Babies should eat thick porridge once they stop breastfeeding	0.40	-0.04	0.89
Very young children (6-24 months) should eat eggs and meat	0.44	-0.14	0.86
Porridge should be made by adding vegetables, eggs, milk	0.44	-0.02	0.88
Give a variety of foods to very young children (6-24 months)	0.46	-0.11	0.83
Give your sick children (6-24 months) more food than usual	0.35	-0.31	0.88
Overall			0.86

Appendix B: Alternative measure of nutrition knowledge

Dependent variable: number of food groups consumed	**OLS (1)**	**IV (2)**
nutrition knowledge score	0.091	1.315**
	(0.062)	(0.527)
remoteness dummy	-0.167	-0.329
	(0.151)	(0.223)
Other control variables? [a)]	Yes	Yes
Sub-Kebele dummies? [a)]	Yes	Yes
R^2	0.205	-
Weak Identification tests:		
Cragg-Donald Wald F statistic	-	5.82
Angrist-Pischke F-test	-	5.58
p-value	-	0.007***
Over-identification test:		
Hansen/Sargan test	-	0.723
p-value	-	0.395
Number of observations	448	448

Note: Remoteness dummy obtains a value 1 if the household belongs to top third in the remoteness distribution, zero otherwise. Robust standard errors in parentheses. Statistical significance denoted at *** p<0.01, ** p<0.05, * p<0.1.

[a)] Coefficients omitted to preserve space

Appendix C: Poisson models

Dependent variable: number of food groups consumed	**Poisson (1)**	**IV-Poisson (2)**	**IV-Poisson (3)**
error structure:	**n/a**	**additive**	**multiplicative**
nutrition knowledge score	0.093*	0.747***	0.857**
	(0.048)	(0.265)	(0.351)
remoteness dummy	-0.176	-0.367*	-0.385*
	(0.150)	(0.193)	(0.229)
Other control variables? [a)]	Yes	Yes	Yes
Sub-Kebele dummies? [a)]	Yes	Yes	Yes
Pearson goodness-of-fit test	145.1	-	-
--- p-value	1.000	-	-
Over-identification test:			
Hansen/Sargan test	-	0.524	0.439
p-value	-	0.469	0.508
Number of observations	448	448	448

Note: Remoteness dummy obtains a value 1 if the household belongs to top third in the remoteness distribution, zero otherwise. Robust standard errors in parentheses. Statistical significance denoted at *** p<0.01, ** p<0.05, * p<0.1.

[a)] Coefficients omitted to preserve space.

Appendix D: Dropping the health-visit instrument

	IV (1)
nutrition knowledge	0.995***
	(0.376)
remoteness dummy	-0.394*
	(0.218)
Other control variables? [a)]	Yes
Sub-Kebele dummies? [a)]	Yes
Weak Identification tests:	
Cragg-Donald Wald F statistic	13.74
Angrist-Pischke F-test	14.59
p-value	0.000***
Number of observations	448

Remoteness dummy obtains a value 1 if the household belongs to top third in the remoteness distribution, zero otherwise. Robust standard errors in parentheses. Statistical significance denoted at *** p<0.01, ** p<0.05, * p<0.1. Hansen/Sargan test cannot be computed because the equation is exactly identified.

[a)] Coefficients omitted to preserve space.

Appendix E: Replicating Table 3 using the Feed-the-Future midline (2015) data

Dependent variable: number of food groups consumed	**OLS (1)**	**IV (2)**
nutrition knowledge score	0.049*	0.747**
	(0.029)	(0.376)
Control variables? [a)]	Yes	Yes
R^2	0.163	-
Weak Identification tests:		
Cragg-Donald Wald F statistic	-	11.37
Angrist-Pischke F-test	-	4.28
p-value	-	0.017**
Over-identification test:		
Hansen/Sargan test	-	0.858
p-value	-	0.354
Number of observations	4,107	4,107

Note: Standard errors (clustered at the woreda level) in parentheses. Statistical significance denoted at *** p<0.01, ** p<0.05, * p<0.1. Unit of observation is a child 6 to 59 months of age. Child level controls include: sex and age of child. Household level controls include: assets (logged value of durable and productive assets), livestock ownership (in tropical livestock units), access to electricity, household demographics (including household size), head's age and highest level of education in the household. The community level controls include: distance to the nearest market town, access to health care, radio signal and electricity and dummies for each region.

Data: Feed the Future midline survey (2015), Ethiopia.

[a)] Coefficients omitted to preserve space.