GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

GlobalFood
www.uni-goettingen.de/globalfood

RTG 1666 GlobalFood

Transformation of Global Agri-Food Systems:
Trends, Driving Forces, and Implications for Developing Countries

University of Goettingen

GlobalFood Discussion Papers

No. 119

A Latent Class Analysis of the Demand for Food Diversity in India

Christoph Steffen
Xiaohua Yu

May 2018

RTG 1666 GlobalFood · Heinrich Düker Weg 12 · 37073 Göttingen · Germany
www.uni-goettingen.de/globalfood

ISSN (2192-3248)

Suggested Citation:

Steffen, C., X. Yu (2018). A Latent Class Analysis of the Demand for Food Diversity in India. GlobalFood Discussion Paper 119, University of Goettingen. http://www.uni-goettingen.de/de/213486.html.

A Latent Class Analysis of the Demand for Food Diversity in India

Christoph Steffen and Xiaohua Yu
Georg-August-University of Goettingen
37073 Goettingen, Germany

Abstract

In this paper we study the latent demand structure for food diversity in India using data from the 2012 Consumer Expenditure Survey. We assume that consumers who have not yet attained calorie sufficiency favor calories over food diversity and once passing a threshold of subsistence substitute away from staples towards a more varied diet.This implies a latent demand pattern as calorie sufficiency depends on unobservable individual characteristics. Latent classes and consumption patterns are identified by means of finite mixture models. Therefore we examine the link between food diversity indices and socioeconomic indicators, explain component memberships in order to characterize latent classes and evaluate nutritional implications. Two clearly distinct demand patterns for diversity could be identified, consistent with the initial assumptions. The identified classes differ substantially in income, household composition and nutritional adequacy.

Key Words: food diversity, India, finite mixture models, consumer demand
JEL: D12 I32

Acknowledgements:
This research was financially supported by the German Research Foundation (DFG) through the GlobalFood Program (grant number RTG 1666).

Introduction

Recent events such as the Agenda 2030 show that development goals are shifting beyond the mere elimination of hunger towards the improvement of food security and nutrition as an essential part of sustainable development. The Millennium Development Report 2015 accounts for the successful reduction in extreme poverty (UN, 2015) in India. However despite high growth rates in the last decade undernutrition levels in India remain higher than for most countries of sub-Saharan Africa, even though those countries are currently much poorer than India (Deaton and Drèze (2009)).

A study that grew famous for unraveling the severity of this situation with respect to the nutritional status of children is the National Family Health Survey 2005-06 that showed that 48% of all under five year old children are chronically malnourished (stunted) and 54% of all death of under five year old children are related to malnutrition. While an adequate intake of calories is essential for survival it does not suffice to maintain health. In fact Deaten and Drèze (2009) find that in some regions of India malnutrition is worse although calorie intakes are higher.

Nutritionists seem to agree that the consumption of a variety of foods is a key feature in achieving a healthy diet that can translate into the sufficiency of a variety of nutrients including vitamins and minerals that are required to maintain health. In the light of prevailing nutrient deficiencies a better understanding of the demand structure for food diversity could serve as a powerful tool in fighting malnutrition or health issues related to dietary patterns like chronic diseases, the double burden of malnutrition and other non communicable diseases.

The current literature on the demand for food diversity mostly focuses on the inherent relation between food diversity and income within linear demand models. However especially for developing countries, where poverty and hunger can be decisive when it comes to food choices, more complex demand structures should be taken into consideration. Bennett's Law (Bennett, 1941) states that consumers substitute away from starchy staples as their income increases. This implies that consumers add food items from food groups other than starchy staples to their diet, which usually leads to a higher level of food diversity and improved nutrition. Jensen and Miller (2011) extend this theory and assume that a consumer follows a different demand pattern depending on whether he faces subsistence concerns. Since a consumers status has a major influence on his marginal utility of calories and other food attributes it can be reasoned that it will also affect his choice of food diversity. However the minimum calorie requirements for subsistence vary individually and are unobservable.

Consequently the remaining challenge for demand analysis is that there are multiple observable demand behaviors based a consumer's unobservable status. If members of subgroups cannot be distinguished based on their observable characteristics it is common to talk about latent classes. While the literature on demand systems has already recognized the importance of identifying latent classes in demand system (Zhou and Yu, 1014), mixed behaviors were not yet considered in the analysis of the demand for food variety. In traditional demand analysis it is usually assumed that all households belong to the same population and hence exhibit identical consumption patterns. However there may be several distinct patterns that consumers follow with mixed probabilities.

In this paper we use the basic intuitions from Jensen and Miller (2011) to reason observed demand for food variety is the outcome of two distinct demand patterns that consumers follow with mixed probabilities. As a measure for food diversity we use food item counts. We estimate the demand model by means of a finite mixture model using data from the 68th round of the CES consumer survey, analyze posterior component probabilities and evaluate nutritional implications of class membership. Results suggest that two latent demand patterns exist that are consistent with the theoretical assumptions. The identified latent classes differ substantially in income, household composition and nutrient adequacy.

1 Literature Review

To date more attention has been paid to the supply of variety than the demand (Gronau and Hamermesh (2008)). Beyond income and prices, diversity has been recognized as an important determinant of the variation in consumer demand. An early study to consider the consequences of consumers' desire for variety in economic modeling is Dixit and Stigiltz (1977). They evaluate the implications of Chamberlin's monopolistic competition model at the social optimum regarding the relation of market and resource allocation. However instead of modeling diversity directly they consider desire for variety as an inherent property of a weakly separable utility functions with convex indifference surfaces. Consequently if a consumer is indifferent between the quantities (0,1) and (1,0) of two different products he will prefer the set (1,1) to either (0,2) and (2,0). Their results contributed to the "excessive diversity" vs. "excessive capacity" debate by showing that neither extreme is an outcome in the social optimum. Benassy (1996) uses the tool set proposed by Dixit and Stiglitz (1977) and concentrates on the derivation of a "taste of variety" parameter that expresses the utility gain from consuming a variety of goods instead of concentrating on a single one.

A study that focuses on the link between income and the demand for variety is presented by Jackson (1984). Jackson suggests a concept to model a hierarchy of purchases and proposes a modified utility function to adapt to such a hierarchy. If consumer demand follows a hierarchical structure consumers with a low income purchase a very limited set of items and will add more items as income increases leading to continuing growth in diversity at all income levels. Jackson emphasizes the gravity of non-negativity constraints in a demand system. His intuition was based on Houthakker (1954) who shows by means of quadratic utility functions that commodities can enter and leave the budget. By adding additional constraints to a general utility maximization problem Jackson shows that in his framework, at a given price, the number of items in the purchased set of commodities is an increasing function of income. For his analysis Jackson used the US Consumer Expenditure Survey which was conducted in 1972-1973. He confirms the strong link between income and variety using a count measure to represent diversity.

Another theoretical framework that aims to explain differences in consumers demand for variety comes from Gronau and Hamermesh (2001) who use a household production model in which households produce activities. They explain differences in demand through differences in the opportunity costs of activities in terms of time costs and pay especially attention to the positive correlation between the educational level and the demand for variety.

Most empirical studies on the demand for variety focus on developed countries. Thiele and Weiss (2002) analyze the demand for food diversity in Germany. As a measure of diversity they use the logit transformed Berry Index and the Entropy index. As the main positive drivers of the demand for food variety they identify income and the number of children between 7 and 17 years. Single male households demand significantly less variety. In Addition they find strong regional effects and food variety is higher in large East German cities. A significant influence of the educational level could not be confirmed.

Consumers in developing countries may however exhibit different pattern than those in developed countries. For instance due to differences in income, infrastructure or food supply. So far only a few studies that perform an analysis on a micro level are available for developing and transition countries. Moon et al. (2002) presents a study on Bulgaria and find strong regional effects. High age, low incomes and low levels of education are associated with a less varied diet. Cupak et al.

(2014) present a 2-stage least squares approach to estimate demand for diversity in Slovakia and find significant increases of income elasticities over time resulting in a convergence towards the levels of Western European countries. Thorne-Lymann et al (2014) examine the link between dietary diversity scores and socioeconomic status in Bangladesh. They establish a strong link between diversity and income although a varied diet was low throughout all income groups. They also confirm that a low level of diversity increases the risk of malnutrition.

A diversified diet has important implication concerning the consumer's health and is hence of economic significance. The proposition that food diversity allows an inference on health implies a link between food diversity and nutrition or dietary quality. Hatloy et al (1998) show that such an inference is possible. They compare Diet Diversity Scores, which are defined as the number of food groups consumed, with nutrient adequacy ratios that compare actual nutrient intakes with intake recommendations. The authors conclude that it is possible to infer on nutritional adequacy from dietary diversity.

2 Theory

The basic characterization of latent classes that is used here to model observed demand for food diversity as the outcome of at least two distinct behaviors is based on the revealed preference framework from Jensen and Miller (2010). They propose a new measure of nutritional sufficiency based on observed consumption. The proposition was motivated by the problem that whether an individual has achieved calorie sufficiency or overcome hunger cannot be adequately determined by a calorie threshold. However individuals reveal their nutritional status in their consumption pattern. In order to determine an individual's nutritional status they assume that a consumer exhibits different food consumption patterns when facing subsistence concerns and when he is outside of subsistence concerns. Under subsistence concerns a consumer suffers a significant disutility from calorie deprivation. The disutility can appear in the form of hunger, headaches and other side effects arising from calorie deprivation. In this state the marginal utility of calories is very high and the cheapest sources of calories will be demanded by a utility maximizing consumer. Consequently the proposed measure of nutritional sufficiency is the staple calorie share (SCS) that is defined as the ratio between calories from staples and total calorie consumption. This ratio remains high as long as subsistence needs are not met and starts to decline once calorie sufficiency has been achieved.

The cheapest sources of calories in India are typically rice and wheat that are also available via the public distribution system to which a consumer has access when he possess a ration card. Once the subsistence threshold is passed marginal utilities of calories decline the consumer starts substituting away from staples towards more expensive calories that he favors over other attributes such as taste or diversity. This model implies a latent demand structure for food diversity. Diversity will have a low marginal utility in contrast to calories as long as an individual faces the primal instinct of survival or seeks to avoid hunger. In the absence of hunger marginal utility of additional calories declines due to repletion and consumer may favor attributes like diversity over calories to realize further utility gains from food consumption.

The decisive connection between the present study that analyzes the latent demand structure for food diversity and the revealed preference framework is the assumption on how consumer preferences change with calorie sufficiency. In terms of calories there is no clear cut off that would separate these two behavioral patterns. Jensen and Miller emphasize that there is no clear consensus on a minimum calorie threshold. Consequently the subsistence threshold is individually varying and is hence unobservable since it depends on the absorption efficiency of the individual and thus represents a source of unobservable heterogeneity. Deaton and Drèze (2009) also doubt the usefulness of calorie norms since there are too many sources of variation such as health and activity levels. Further support for the relevance of the underlying demand structure comes from Zhou and Yu (2014) who are concerned the resulting implications for calorie elasticities. They examine the relation between income and calorie consumption and verify the existence of a latent demand structure, which is consistent with the behavioral assumptions from Jensen and Miller (2010)

Similar to Drescher et al. (2009) who performed a demand analysis for Germany based on the healthy food diversity index, we assume a demand function of the form:

$$d = d(P, Y, K) \tag{1}$$

Here the demand for food diversity depends on prices, income and consumer characteristics. If we assume that observed demand for diversity is the result of C different demand behaviors that consumers follow with mixing probability π_j equation (1) can be rewritten as the weighted sum of

C distinct demand functions.

$$d = \sum_{j=1}^{C} \pi_j d_j(P, Y, K); \qquad \sum_{j=1}^{C} \pi_j = 1 \tag{2}$$

The revealed preference framework suggests that consumers behave differently depending on their calorie sufficiency. Consequently two different classes are to be distinguished to model the demand for food diversity.

$$d = \pi_1 d_s(P, Y, K) + \pi_2 d_{ns}(P, Y, K), \qquad \pi_1 + \pi_2 = 1 \tag{3}$$

While in concern about subsistence (d_s) calories are more important than diversity in order to avoid the disutility. A consumer with no subsistence concerns ($d_n s$) favors food diversity over calories as his basic needs are already satisfied. Disutility penalizes the process of substituting away from staples towards higher levels of food diversity if subsistence calories have not been attained yet. Consumers substitute away from staples if they do not experience a disutility in the form of hunger or other symptoms of deprivation. A consumer who is sufficiently wealthy to afford any food consumption bundle he desires will not have to face any disutility. A consumer's mixture probability is hence determined by the penalty he faces due to his individual distance to his calorie subsistence threshold. Ultimately each observed realization of this demand function is the outcome of both demand patterns with mixing probabilities π_1 and π_2.

Figure 1 shows the assumed relationship between food diversity and income. The line below the subsistence threshold corresponds to the demand behavior d_s and above to $d_n s$. The income elasticity of food diversity increases after passing the unobservable subsistence threshold. However food diversity has a natural limit and will converge to certain limit for high income classes. On the one hand diversity is limited by supply on the other it is reasonable to assume that increasing efforts are required to further diversify an already well balanced diet. This circumstance would be sufficient to consider the existence of a potential third class. While this insight deserves testing it is beyond the here considered theoretical framework that concentrates on behaviors around a subsistence threshold.

Table 1 gives an overview on the expected differences between the latent demand patterns and consumer characteristics. Under d_s in contrast to $d_n s$ the consumption basket is expected to have a lower level of food diversity and consumers are expected to earn more on average. The income elasticity of d_s is lower than of d_ns since consumers favor calories over diversity. This implies that the marginal utility of calories is higher than the marginal utility of diversity with d_s and the other way round under d_ns. Consequently the marginal utility of calories is higher under d_s than d_ns. Zhou and Yu (2014) test this implication of the revealed preference framework and find that calorie elasticities decrease in income.

Consumers reveal their preferences and with it their nutritional status in their consumption patterns. If they substitute away from staples it is a signal that the subsistence threshold has been surpassed and calorie elasticities are decreasing. Jensen and Miller thus suggest using the share of calories from staples in total calorie consumption as an indicator of nutritional sufficiency. In the empirical model this insight can be used to model mixture probabilities. Figure 2 shows the relation between the SCS and log income by means of a local polynomial regression. For India the two variables have a clear negative relation similar to the findings of Jensen and Miller (2010) for China. In the case of India however changes of calorie preferences with income appear to be smoother which renders the position of a subsistence threshold less obvious.

3 Data

Our analysis is based on the Consumer Expenditure Survey (CES) which has been conducted by the National Sample Survey Office (NSSO). The NSSO has been founded in 1950 by the government of India and belongs to the Ministry of Statistics and Programme Implementation. It regularly conducts consumer expenditure surveys nationwide throughout India. For our analysis we use the 68th round of the survey which was conducted in 2012. Over 4 sub rounds 101626 household were interviewed. The survey contains information on consumption expenditure over the last 30 days and provides a high level of detail on food expenditures. In total it provides quantity and expenditure information of 142 food items. After dropping tobacco and liquors 127 food items remain for the analysis.

Using a 30 day recall period for a dietary analysis is sometimes criticized as it is prone to memory errors. Ruel (2006) suggests that a 7 day recall might be optimal to minimize recall error. The NSSO (2002) found in a pilot study that frequently purchased items were recalled more precisely on a monthly than a weekly base while it was the other way round for infrequently purchased goods. While memory errors might represent a weakness in the assessment a longer period of observation may reveal dietary patterns more clearly. A very interesting study that allows inference on the development of food diversity over time is presented by Drewnowski et al. (1997) who collected detailed food consumption data from individuals on 15 consecutive one day recalls. They find that food diversity is increasing over the whole period of 15 days although converging towards a steady level during the last days. Hence observations over a longer run can be helpful in dietary pattern analysis, especially since some consumers may repeat certain consumption patterns more frequently than others which is not necessarily observable in shorter periods.

A well accepted measure to evaluate of the nutritional status of the households within the latent classes would is the nutrient adequacy ratio. The calculation of nutrient adequacy ratios requires the construction of a conversion table to calculate nutrients from food item quantities and the recommended dietary allowances. Conversion tables were extracted from the database Nutrisurvey and dietary allowances are provided by the national dietary guidelines from the National Institute of Nutrition (2010). The guidelines contain nutrient intake recommendations for 12 groups that differ by age and gender. Since the CES has details on age and gender of all household members the information from the guidelines can be used for accurate calculations of dietary allowances on a household level.

Outliers were removed that were in terms of calories too high or low to be realistic. Following the example of Wiesmann et al (2009), Foote et al (2004) and Lovon and Mathiassen (2014) all observation below 500 calories and above 5000 calories per capita per day were removed. As a result 2151 observations were dropped.

4 Empirical Model

If there are classes to be distinguished by their demand patterns, then these would follow different probability distributions. In order to identify latent behaviors that consumers follow with mixed probabilities finite mixture models (fmm) are an adequate instrument of analysis. In a finite mixture model the population is assumed to consist of an additive mixture of C subpopulations that are mixed in proportions $\pi_1, ...\pi_C$. Each of the distinct probability distributions is assumed to arise from the same parametric family $(d_i|\theta)$, where the diversity index represents the latent variable and θ the set of parameters. The joint density function of the population is then given by:

$$f(d_i|\theta_j) = \sum_{j=1}^{C} \pi_j(z) f_j(d_i|\theta_j) \tag{4}$$

The mixture probabilities π_j satisfy the restrictions $\sum_{j=1}^{C} \pi_j = 1$ and $\pi_j \geqslant 0$. The component probability π_j can be modeled as a function of z. This can be especially useful in the presence of large overlaps within the mixture distribution. In this case we use the staple calorie share to model the component probabilities and the number of meals consumed outside the household as an additional control variable. In a two component model the prior component probabilities are then estimated as follows (Ayyagari et al 2013):

$$\pi_{ij} = logit(Z_i|\delta), \qquad 0 \leqslant \pi_{ij} \leqslant 1, \qquad \sum_{j=1}^{C} \pi_{ij} = 1 \tag{5}$$

As a measure of food diversity we choose food item counts. In the light of the present analysis this measure has several advantages. Count indices can be analyzed with standard econometric methods without the necessity of any transformation to meet distributional assumptions. This allows a straightforward interpretation of the results. Beyond that counts of food groups or items are the standard unit used in dietary recommendations. Another popular choice for a diversity index is for instance the berry index. However since this variable is bounded between zero and one it requires a log transformation prior to its inclusion in a model with normal distribution assumption or a model that can be fitted to these properties such as a beta model. While the berry index usually has a better correlation with nutritional adequacy count indices allow for a straight forward interpretation and can be analyzed with standard econometric methods within a finite mixture approach.

Count data follows a discrete distribution. It is usually assumed to follow a Poisson or negative binomial distribution. However the Poisson distribution, as a one-parameter distribution, tends to underestimate the variance while negative binomial models allow for over dispersion. For the analysis of the count measure for food variety we follow Moon et al (2002) who use a negative binomial model. The probability to observe the consumption of n food items is given by the density function:

$$f(d_i|\theta) = \sum_{j=1}^{C} \pi_j \frac{\Gamma(d_i + \psi_{j,i})}{\Gamma(\psi_{j,i})\Gamma(d_i + 1)} \left(\frac{\psi_{j,i}}{\lambda_{j,i} + \psi_{j,i}}\right)^{\psi_{j,i}} \left(\frac{\lambda_{j,i}}{\lambda_{j,i} + \psi_{j,i}}\right)^{d_i} \tag{6}$$

In order to fit a model for the final analysis the Poisson, NB1 and NB2 models are estimated with two and three components. Ex post to the fmm estimation the posterior component probabilities can be estimated by means of Bayes rule:

$$Pr(d_i \in j|\theta, d_i) = \frac{\pi_j f_j(d_i|\theta)}{\sum_{j=1}^{C} \pi_j f_j(d_i|\theta)} \tag{7}$$

The estimated posterior probabilities allow the analysis of determinants of class memberships. This could be achieved by regressing a vector of socioeconomic variables on the probability of being member in component j. The posterior probabilities are however continuous on the open interval (0,1). An analysis by means of ordinary least squares would be flawed as the fact that the variates are bounded violates OLS assumptions (Kiesschnieck and McCullough (2003)). The error distribution of regression models using a bounded dependent variable is heteroskedastic as the conditional variance approaches zero when the mean approaches the boundaries. Kiesschnieck and McCullough (2003) analyze the fit of 7 different regression models for variables on the open interval (0,1) and recommend the use of a beta regression model for the analysis of this type of data. Smithson and Verkuilen (2006) agree with this conclusion. In order to explore the influence of household characteristics on posterior probabilities we hence use a beta regression model as proposed by Smithson and Verkuilen (2006) and Ferrari and Cribari-Neto (2004). With the component membership from the estimates it is possible to explore further nutritional implications of this class. An observation is attributed to a component if the respective component probability is greater or equal 0.5. The comparison of mean nutritional intakes and adequacy ratios across components yields information about distinctiveness of component members in terms of nutrients.

$$NAR_{li} = \frac{N_{li}}{\sum_{j=1}^{12} RDA_{lj} M_j} \tag{8}$$

The nutrient adequacy ratios are calculated by dividing all nutrients of nutrient l consumed by household i by the sum of dietary allowances over all household members. The sum of dietary allowances for a nutrient is here calculated by summing up the products of each of the 12 RDA groups with the number of household members in this group. The consideration of gender and age groups in the calculation of nutrient adequacy ratios gives a very precise picture of the households nutritional status.

5 Variables

As a measure of food diversity we choose simple food item counts. The index is calculated by summing up the number of distinct food items that were consumed by the household over the 30 day recall period. The log of monthly per capita expenditure and the amount of land possessed were included as measures of income and capital. Due to the existence of zeros and ones we added one ha before taking logs of the amount of land. The 2015 Millennium Development Goals country report suggests that there are still large differences in the nutritional status between rural and urban regions. Hence a dummy was added to indicate whether the household is located in a rural area. Another dummy indicates whether the household is in possession of a ration card. A ration card gives access to very cheap calories via the public distribution system. Jensen and Miller (2011) find that subsidizing staples can result in consumers substituting away from these. Consequently ration cards are expected to have a positive effect on food diversity. Especially in rural areas of India we find many semi subsistence farmers. The survey contains the additional information if a consumed food item was purchased or produced at home. In order to capture the effect of home produced food consumption the ratio of the quantity of consumed home produced food and total consumed food was added as a variable. Characteristics on the household head include years of education. This variable was here derived from the stated general educational level. The relevance of the educational level is regularly discussed in the context of demand for diversity and is usually expected to have a positive impact. The food consumption data does not contain detailed information on consumption outside the household. However we can calculate the number of meals per household member outside the household that were consumed over the observed 30 day recall period. Additionally the average number of meals per household member was added for the same time period. Similar to Thiele and Weiss (2001) we count household members of different age groups. We thus count the number of children from 1-6 years, 7-13 and 14-17. In addition we add a count of elderly members with age 60 and above, the number of women and add the size of the household. Eating habits in India are often shaped by beliefs. For instance Hindus do not eat beef and many are vegetarians. Muslims avoid pork and Sikhs beef. We hence include religion dummies for Hinduism, Islam, Christianity, Sikhism, Jainism and Buddhism to control for possible influences on dietary diversity. In order to control for price differences across regions 35 state dummies were added. The smallest seven of these which make up for 2.3 % of the observations were assigned to the base group. The assignment of too many state members of these small states in the latent class analysis resulted in the failure to calculate the respective coefficients and hence evaluate the model fit. As the survey was conducted in 4 sub rounds 3 sub round dummies were included to control for possible seasonal effects. Two variables were chosen to model the component probabilities. The first on is the number of meals outside the home. This part of nutrition is not captured by the data and hence needs to be controlled for. Households that eat out a lot may show lower levels of food consumption and diversity according to the data. Due to a supposedly large overlap in the distribution of dietary diversity such households may be attributed to a class with a lower mean of diversity then the one they rather belong to in terms of their actual level of diversity. The second variable to model the component probabilities is the staple calorie share. This indicator is here calculated as the ratio of calories obtained from rice and wheat in the total calorie consumption of the household. One of the basic insights from Jensen and Miller (2011) is that a consumer's subsistence threshold is unobservable as it depends on unobservable individual characteristics. However he reveals his status in his consumption pattern. In this case we can use the calorie staple share to model the mixture probabilities as a function of the consumer's revealed preferences.

6 Results

6.1 Model Selection

Among the single equation models the Akaike information Criteria the Baysian Information Criteria and the log likelihood all favor unanimously the negative binomial 2 model over the Poisson and the negative binomial 1 model.The χ^2 based goodness of fit test for the Poisson model rejects the Null that the count index follows a Poisson distribution. The χ^2 test statistics that were estimated along with the NB 1 and NB 2 both reject the hypothesis that the data follows a Poisson distribution. This is consistent with the summary statistics of the count index from table 7 that clearly shows that the variance exceeds the mean greatly. Hence there is evidence of overdispersion which is more adequately modeled by means of a negative binomial model than a Poisson model which assumes equidispersion where the mean is identical to the variance.

In total six finite mixture models were estimated with two and three components for Poisson, NB1 and NB2 models. None of the three component models converged. Among the two component models only the Poisson model converged. This circumstance could not be altered by changing model specifications or maximization algorithms. However all calculated information criteria show that the two component Poisson model fits the data better than the best fit among the single component models. This indicates that a mixture model might be better suited than a single equation models to analyze the demand for food diversity.

6.2 Characterization of Latent Classes

Before analyzing the demand model estimations differences between latent classes are compared in order to evaluate whether the class characteristics match the theoretical assumptions. Table 6 of Appendix B contains the summary statistics of the nutritional status and income for each component. Component 1 covers 28583 observations and 28.77 % of the data sample. Component 2 contains 70765 observations which make up for 71.23 % of the sample. Differences in income are striking. The mean per capita income in the first component is 1065 Rupees and the second one is with 2032 Rupees almost twice as large. On average the households in component 1 consume 30.44 distinct food items and the ones in component 2 consume 39.88 items. In terms of calories the classes are almost identical. This is also true for calorie adequacy and protein adequacy although fat adequacy of component 2 is with 96.81 % is clearly higher as compared to component 1 with 78.05 %. In component 2 adequacy ratios are higher for all tested micronutrients without any exceptions. The mean nutrient adequacy component 2 is with 79.48 % 12.74 percentage points higher than that of component 1. Figures that stick out are especially the adequacy of vitamin A that is critically low in component 1 with 50.12 %. The vitamin A adequacy of component 2 is significantly higher with 74.4 %.

So far the findings are consistent with theory. One class could be identified with lower income and a lower level of food diversity that also accounts for higher incidence of malnourishment. Additionally the level of calorie consumption is comparable in both classes indicating that members of component 2 have shifted their consumption towards more expensive calories while achieving higher levels of food diversity as compared to members of component 1. Therefore the probability to be a member of component 1 corresponds to π_1 of the theoretical model. As these characteristics are in line with the initial assumptions it is possible to match the empirical outcomes to the theoretical model. Thus component 1 refers to the class with subsistence concerns and component

2 to the class without subsistence concerns. Justified alternative labels in line with the identified characteristics for these classes could also be deprived/well-nourished, poor/rich, hungry/sated or one-sided and diversifying food consumers.

Table 5 contains the results of the beta regression and the corresponding marginal effects. OLS results were included as well to be available for model comparison. According to the F and χ^2 statistic both models show overall a very good fit although the coefficients in the models differ considerably in terms of magnitude and significance. However the Null hypothesis that the dependent variable follows a normal distribution is strongly rejected for tests on skewness and kurtosis. Also the AIC and BIC indicate that the beta model has a better fir than the OLS model. The average marginal effects of the beta model show the probability of being a member of the class with subsistence concerns decreases in income and the amount of land. The signs of the coefficients are as expected since these are determinants of the household's purchasing power that enables the household to consume a higher level of food diversity. Further variables that have a significant negative effect are the share of home produced food, the number of meals outside the hh, years of education, the number of children between 1 to 6 years and 7 to 13 years as well as the number of women and elders. The probability is positively affected by a location in rural areas, the possession of a ration card, the household size and the number of older children between 14 and 17 years.

Figure 3 shows the local polynomial regression of the posterior probability of being a member of the class with subsistence concerns on mean nutrient adequacy. There is a clear negative relation between the two variables. So a high probability to be a member of the class with subsistence concerns is associated with a lower level of nutrient adequacy. Consequently, since there are only two components, it is the other way round for the probability of being a member of the class without subsistence concerns.

6.3 Demand Model Estimations

Table 3 contains the parameter estimates of the NB2 model and two component finite mixtures Poisson model. The Wald chi statistic shows that models have an overall good fit. The results of the mixture model reveal that the staple calorie share is significantly related to the component probabilities while the number of meals consumed outside the house appears to be unrelated.

Table 4 shows the average marginal effects of the models. The effect of income on diversity is 21% higher for the class that has attained a subsistence level of nutrition. Below the subsistence threshold consumers are rather concerned with survival where the marginal utility of calories is still very high. Consequently we observe a higher income elasticity above subsistence threshold where the marginal utility of calories is expected to be lower as compared to the marginal utility of other food attributes such as diversity. Living in a rural area has a negative effect only on the class with no subsistence concerns while the poorer class remains unaffected. The infrastructure and supply in rural areas appears sufficient to meet basic needs but impedes diets beyond that. Consumption from home produced food has a negative effect for both classes. Since the diversity in home production usually has sensible limits increasing this share in food consumption decreases the number of food items consumed. The possession of a ration card has a positive effect in both classes. In the class that struggles with subsistence the possession of a ration card increases the number of food items by 2.3 on average. The effect is slightly lower with 1.9 food items on average in the class that has attained that has attained subsistence calories. However these effects are to be related to the mean diversity of the respective classes. This confirms the finding of Jensen and Miller (2011) that subsidizing staples can result in consumers substituting away from these. This again results in a higher level of food diversity. The years of education have a significant positive effect in both classes. Moon et al. (2002) suggest that more educated consumers are more concerned about nutritional balance and demand more diverse diets. The effect of education is however more than 2 times larger in the class with no subsistence concerns. A possible explanation for this is that while higher education promotes better diets consumers still require the resources to realize better diets. Religions have no effect on food diversity in the poorer class however all considered religions except for Janism have a positive effect on diversity within the class without subsistence concerns. Culturally specific eating habits appear to require a certain level of food diversity that is not achieved as long as calorie sufficiency is the primary focus of a diet. The household composition reveals some interesting differences across classes. The number of children between 1 and 6 years as well as the number of children between 7 and 13 has significantly negative effect on diversity in the class with subsistence concerns. In the class with no subsistence concerns the effect is significantly positive. In developed countries we usually find that the number of children positively affect the households food diversity as for instance in the study of Thiele and Weiss (2001) who analyze German consumption data. One possible explanation for this effect is that parents want to offer their children healthy diets. The negative effect in the class that struggles with subsistence indicates that consumers are unable to do so and budget constraints force households to substitute towards cheaper calories if they need to provide for more children. The number of elder household members of age 65 and older has a small negative effect in both classes. Moon et al. (2002) argue that older consumers are less likely to seek variety than younger generations. The coefficient on the number of women is insignificant in the poorer class but has significant positive effect at a 10 per cent level in the richer class. Lee and Brown (1989) also find that an additional female household member has a greater positive effect on the households food diversity than an additional male member and suspect that women have a greater interest in the preparation of various foods or nutrition. Most of the included state variables were highly significant suggesting strong regional effects. These may occur due to differences in supply or regional price differences.

7 Conclusions

In this paper some intuitions of Jensen and Miller (2011) are extended to a latent demand model for food diversity that can be estimated by means of a finite mixture model. The estimates from the demand model clearly show two distinct demand patterns that are consistent with our basic assumptions. Consumers that live near their individual subsistence threshold strive for survival resulting in lower income elasticity while the marginal utility of calories remains high. Once having passed that threshold the income elasticity of food diversity increases as consumers do not suffer disutility from deprivation anymore. The classes differ widely in terms of nutrition consistent with the theory. Under subsistence concerns more staples are demanded resulting in higher levels of malnutrition than the better off class that fares better with all micro nutrients while maintaining a comparable calorie intake. Beyond the confirmation of the underlying hypothesis the estimates reveal further interesting insights. It is often argued that better education could improve nutrition. While a significantly positive effect can be confirmed, the magnitude of the effect is much lower in the class with subsistence concerns. On the one hand this suggests that the effect of education might be overestimated when looking at the whole population on the other hand the effect might be constrained by income. The finding that children have a negative effect in the poorer class although they affect food diversity usually in a positive effect might be interpreted as a severe sign of neediness since households reduce their dietary quality to ensure the survival of all household members. This effect is easy to overlook since the one component NB2 model shows an overall positive effect. Beyond that the present analysis reveals some shortcomings of calorie based indicators as we find classed with vast differences in their nutritional status and identical calorie intake. The existence of latent behaviors and unobservable heterogeneity in the demand for food diversity poses challenges not only for demand analysis but also for assessment of nutrition and poverty. The mere insight that such patterns exist can however aid to better identify households that are vulnerable to malnutrition.

References

Ayyagari, P., P. Deb, J. Fletcher, W. Gallo, and S. L. Sindelar (2013) *"Understanding Heterogeneity in Price Elasticities in the Demand for Alcohol for older Individuals"* , Health Econ. 2013 January, 22(1) 98-105

Benassy, J. (1996) *"Taste for Variety and Optimum Production Patterns in Monopolistic Competition",* Economics Letters, 52, 41-47

Bennett, M. K. (1941) *"Wheat in National Diets",* Wheat Studies of the Food Research Institute, Stanford University, California

Deaton, A. and J. Drèze (2009) *"Food and Nutrition in India: Facts and Interpretations"* , Economic & Political Weekly, Vol XLIV No 7

Deb, P. and P. K. Trivedi (2002) *"The Structure of Demand for Health Care: Latent Class versus Two-Part Models",* Journal of Health Economics 21 601-625

Dixit, A. and J. E. Stiglitz (1977) *"Monopolistic Competition and Optimum Product Diversity",* American Economic Review 67, 292-308

Drescher, L., S. Thiele and G. Mensink (2007) *"A New Index to Measure Healthy Food Diversity Better Reflects a Healthy Diet Than Traditional Measures",* Journal of Nutrition 137:647-651

Drescher, L., S. Thiele, J. Rosen and G. Mensink (2009) *"Consumer Demand for Healthy Eating Considering Diversity- An Economic Approach for German Individuals",* International Journal of Consumer Studies ISSN 147-6423

Drewnowsky, A. , S. A. Henderson, A. Driscoll and B. J. Rolls (1997) *"The Dietary Variety Score: Assessing Diet Quality in Healthy Young and Older Adults"* , Journal of the American Dietetic Association Volume 97 Number 3

Ferrari, S.L.P. and F. Cribari-Neto (2004) *"Beta Regression for Modelling Rates and Proportions",* Journal of Applied Statistics 31(7): 799-815.

Foote, J. A., S. P. Murphy, L. R. Wilkens, P. Basiotis and A. Carlson (2004) *"Dietary Variety Increases the Probability of Nutrient Adequacy among Adults",* American Society for Nutritional Sciences.

Gronau, R. and D. Hamermesh (2008) *"The Demand for Variety: A Household Production Perspective",* The of Economics and Statistics August 2008, 90(3): 562-572

Government of India (2015) *"Millenium Development Goals Report 2015"*

Hatloy, A., L. E. Torheim and A. Oshaug (1998) *"Food Variety - a Good Indicator of Nutritional Adequacy of the Diet: A Case study from an Urban Area in Mali, West Africa",* European Journal of Clinical Nutrition 52, 891-898

Jackson, L.F. (1984) *"Hierarchic Demand and the Engels Curve for Variety",* The Review of Economics and Statistics, Vol.66, No 1 (Feb.1984) pp.8-15

Jensen, R. T., and N.H. Miller (2010) *"A Revealed Preference Approach to Measuring Hunger and Undernutrition",* NBER working paper no. 16555, Cambridge, MA

Jensen, R. T. and N. H. Miller (2011) *"Do Consumer Prices Subsidies Really Improve Nutrition?",* The Review of Economics and Statistics, November 2011, 93(4) 1205-1223

Lancaster, K. (1990) *"The Economics of Product Variety: A Survey ",* Marketing Science 9, 189-206

Lee, J. and M. G. Brown (1989) *"Consumer Demand for Food Diversity ",* Southern Journal of Agricultural Economics Volume 21, Number 02, December 1989 pp47-52

Lovon, M. and A. Mathiassen (2014) *"Are the World Food Programme's Food Consumption Groups a Good Proxy for Energy Deficiency?",*Food Security (2014) 6:461-470

McLachlan, G. J. and D. Peel (2000) *"Finite Mixture Models ",* New York: Wiley

Ministry of Health and Familiy Welfare, Government of India (2006) *" National Family Health Survey (NFHS-3) India2005-06",*International Institute for Population Sciences Deonar, Mumbai

Moon, W., W. J. Florkowski, L. R. Beuchat, A. V. Resurreccion, P. Paraskova, J. Jordanov and M. S. Chinnan (2002) *" Demand for Food Variety in an Emerging Market Economy",* Applied Economics 34: 573-581

National Institute of Nutrition (2010) *"Dietary Guidelines for Indians: A Manual ",* Hyderabad - 500 007, India

National Institute of Nutrition (2009) *"Nutrient Requirement and Recommended Dietary Allowances for Indians ",*

National Sample Survey Organization Ministry of Statistics and Program Implementation Government of India (2002) *"Key Indicators of Household Consumer Expenditure in India ",*

National Sample Survey Organization Ministry of Statistics and Program Implementation Government of India (2002) *"Results of a Pilot Survey on Suitability of Different Reference Periods for Measuring Household Consumption ",*

International Institute for Population Sciences (IIPS) and Macro International (2007) *" National Family Health Survey (NFHS-3), 2005-06: India: Volume I",* Mumbai: IIPS

Ruel, M. (2006) *"Operationalizing Dietary Diversity: A Review of Measurement Issues and Research Priorities ",*The Journal of Nutrition

Shonkwiler, J. S., J. Lee and T. G. Taylor (1987) *" An Empirical Model of the Demand for Varied Diet",* Applied Economics, 19:10, 1403-1410

Smithson, M. and J. Verkuilen (2006) *"A Better Lemon Squeezer? Maximum Likelihood Regression with Beta-Distributed Dependent Variables ",* Psychological Methods 11(1): 54-71

Subramanian, S., and A. Deaton (1996) *" The Demand for Food and Calories",*Journal of Political Economy 104 (1): 133-162.

Thiele, S and C. Weiss (2002) *"Consumer Demand for Food Diversity: Evidence for Germany ",* Food Policy 28 (2003) 99-115

Theil, H. and R. Finke (1983) *"The Consumers Demand for Diversity",* European Economic Review 23 (1983) 395-400 North Holland

Thorne-Lyman, A.L., N. Valpiani, K. Sun, R. D. Semba, C. L. Klotz, K. Kraemer, N. Akhter, S. de Pee, R. Moench-Pfanner, M. Sari and M.W. Bloem (2010) *"Household Dietary Diversity and Food Expenditures are Closely Linked in Rural Bangladesh Increasing the Risk of Malnutrition due to the Financial Crisis ",*J Nutr 140(1): 182S-8S

United Nations (2015) *" Millenium Development Goals Report 2015",* Weiss, C. and W. Briglauer (2002) *"Determinants and Dynamics of Farm Diversification ",* Paper prepared for the Xth EAAE Congress 2002

Wiesmann, D., L. Bassett, T. Benson and J. Hoddinott (2009) *" Validation of the World Food Programme's Food Consumption Score and Alternative Indicators of Household Food Security"*,IFPRI Discussion Paper 00870

Zhou, D and X. Yu (2014) *" Calorie Elasticities with Income Dynamics: Evidence from the Literature"*,Global Food Discussion Paper N0.35

Table 1: Expected class differences

	With Subsistence Concerns (d_s)		Without Subsistence (d_{ns})
Mean level of food diversity	Mean food diversity (d_s)	$<$	Mean food diversity (d_{ns})
Income elasticity	Income elasticity (d_s)	$<$	Income elasticity (d_{ns})
Marginal utility of calories	Marginal utility of calories (d_s)	$>$	Marginal utility of calories (d_{ns})
Mean income	Mean Income (d_s)	$<$	Mean Income (d_{ns})
$\frac{Marginal\, utility\, of\, calories}{Marginal\, utility\, of\, diversity}$	> 1		< 1

Table 2: Model selection criteria

Model	**Df**	**ll**	**AIC**	**BIC**
Poisson				
1 component*	52	-789890401	1579780906	1579781400
2 components	107	-776290487	1552581189	1552582207
Negativ Binomial 1				
1 component**	53	-789501912	1579003931	1579004435
Negativ Binomial 2				
1 component***	53	-789291816	1578583738	1578584242

*Goodness of Fit Test:	$\chi^2 = 267000000$	$\bar{\chi}^2 = 0.000$
**Likelihood-ratio test of $\delta = 0$:	$\bar{\chi}^2 = 780000$	$\bar{\chi}^2 = 0.000$
***Likelihood-ratio test of $\alpha = 0$:	$\bar{\chi}^2 = 1200000$	$\bar{\chi}^2 = 0.000$

Table 3: Parameter estimates of the demand model

	NB2		**FMM-Poisson**			
Component	-		Component 1		Component 2	
	Coef.	$P > \|z\|$	Coef	$P > \|z\|$	Coef	$P > \|z\|$
Log income	0.1674	0.000	0.1391	0.000	0.1408	0.000
Rural	-0.0322	0.000	0.0024	0.746	-0.0276	0.000
Log Land	0.0044	0.000	0.0026	0.051	0.0041	0.000
Home Produced Share	-0.0729	0.000	-0.0388	0.002	-0.0964	0.000
Ration card	0.0556	0.000	0.0729	0.000	0.0509	0.000
P.c. Meals out	-0.0037	0.000	-0.0008	0.696	-0.0026	0.000
Pc meals at Home	0.0004	0.000	0.0020	0.000	0.00002	0.880
Education(years)	0.0064	0.000	0.0028	0.042	0.0067	0.000
Hinduism	0.0273	0.113	-0.0257	0.445	0.0748	0.001
Islam	0.0319	0.068	-0.0055	0.873	0.0729	0.001
Christianity	0.0142	0.426	-0.0356	0.290	0.0690	0.003
Sikhism	0.0349	0.061	-0.0218	0.622	0.0819	0.000
Jainism	-0.0104	0.633	-0.0539	0.294	0.0354	0.169
Buddhism	0.0275	0.162	0.0218	0.594	0.0681	0.004
Household Size	0.0342	0.000	0.0403	0.000	0.0347	0.000
Children 1-6y	0.0063	0.000	-0.0157	0.000	0.0086	0.000
Children 7-13y	0.0086	0.000	-0.0088	0.029	0.0121	0.000
Children 14-17y	-0.0041	0.030	-0.0030	0.452	-0.0027	0.214
Elders $\geq$60y	-0.0159	0.000	-0.0241	0.000	-0.0133	0.000
Women	0.0040	0.004	0.0040	0.164	0.0030	0.063
constant	1.3797	0.000	1.4980	0.000	1.7066	0.000
			π_1 :			
			Staple Calorie Share		28.1565	0
			P.c. Meals out		-0.0755	0.157
			constant		-20.0602	0
Wald χ^2	42571		Wald χ^2		31892	
Prob> χ^2	0		Prob> χ^2		0	
N	99508		28583		70765	

Table 4: Average marginal effects of the demand model

	NB2		FMM-Poisson			
Component	-		Component 1		Componant 2	
	Coef.	$P > \|z\|$	Coef	$P > \|z\|$	Coef	$P > \|z\|$
Log income	6.2307	0.000	4.5254	0.000	5.4580	0.000
Rural	-1.1993	0.000	0.0785	0.746	-1.0710	0.000
Log Land	0.1620	0.000	0.0845	0.052	0.1589	0.000
Home Produced Share	-2.7139	0.000	-1.2614	0.002	-3.7365	0.000
Ration card	2.0695	0.000	2.3704	0.000	1.9748	0.000
P.c. Meals out	-0.1374	0.000	-0.0265	0.695	-0.1000	0.000
Pc meals home	0.0148	0.000	0.0647	0.000	0.0007	0.880
Education(years)	0.2374	0.000	0.0914	0.046	0.2591	0.000
Hinduism	1.0144	0.113	-0.8369	0.444	2.8996	0.001
Islam	1.1860	0.068	-0.1780	0.873	2.8243	0.001
Christianity	0.5296	0.426	-1.1569	0.289	2.6760	0.003
Sikhism	1.2992	0.061	-0.7081	0.622	3.1754	0.000
Jainism	-0.3854	0.633	-1.7527	0.295	1.3743	0.170
Buddhism	1.0241	0.162	0.7091	0.594	2.6402	0.004
Hhsize	1.2718	0.000	1.3096	0.000	1.3435	0.000
Children 1-6y	0.2353	0.000	-0.5108	0.000	0.331	0.000
Children 7-13y	0.3193	0.000	-0.2871	0.029	0.4687	0.000
Children 14-17y	-0.1530	0.030	-0.0961	0.453	-0.1048	0.214
Elders $\geq$60y	-0.5932	0.000	-0.7853	0.000	-0.5141	0.000
Women	0.1477	0.004	0.1303	0.163	0.1150	0.063

Table 5: Regression on posterior probabilities of component 1

	OLS		Beta Regression Model			
	Coef.	$P > \|z\|$	Coef	$P > \|z\|$	dy/dx	$P > \|z\|$
Log income	-0.1794	0.000	-0.8058	0.000	-0.1427	0.000
Rural	0.0555	0.000	0.2122	0.000	0.0376	0.000
Log land	-0.0053	0.000	-0.0275	0.000	-0.0049	0.000
Homeratio	-0.0621	0.000	-0.2830	0.000	-0.0501	0.000
Ration Card	0.0041	0.387	0.0796	0.000	0.0141	0.000
P.c. Meals out	-0.0003	0.588	-0.0088	0.000	-0.0015	0.000
Pc meals home	0.0012	0.000	0.0049	0.000	0.0009	0.000
Education(years)	-0.0029	0.000	-0.0115	0.000	-0.0020	0.000
Hinduism	0.0083	0.827	-0.0350	0.702	-0.0062	0.702
Islam	-0.0142	0.711	-0.0825	0.371	-0.0146	0.371
Christianity	0.0557	0.151	0.1409	0.135	0.0249	0.135
Sikhism	0.0003	0.994	-0.0507	0.610	-0.0090	0.610
Jainism	0.0260	0.517	-0.0149	0.897	-0.0026	0.897
Buddhism	0.0031	0.939	0.1004	0.326	0.0178	0.326
Hhsize	0.0127	0.000	0.0521	0.000	0.0092	0.000
Children 1-6y	-0.0253	0.000	-0.1212	0.000	-0.0215	0.000
Children 7-13y	-0.0029	0.254	-0.0198	0.000	-0.0035	0.000
Children 14-17y	0.0112	0.000	0.0302	0.000	0.0053	0.000
Elders	0.0013	0.625	-0.0217	0.000	-0.0038	0.000
Women	-0.0060	0.010	-0.0253	0.000	-0.0045	0.000
constant	2.2210	0.000	7.9380	0.000		
N	99348		N	99348		
R^2	0.4223		Wald χ^2	56827		
Prob >F	0		Prob > χ^2	0		
F(51, 99296)	1423		Log likelihood	194621		

Table 6: Characterization of latent classes

Model	FMM	
Component	Component 1	Component 2
P.C.Income (Rs)	1064.99	2032.15
Count Index	30.4426	39.8841
Staple Calorie Share	0.7739	0.5564
Nutrients		
Calories	1905.809	1937.234
NAR		
Calories	0.8132	0.8159
Protein	0.8487	0.8944
Fat	0.7805	0.9681
Vit. A	0.5097	0.7440
Magnesium	0.9185	0.9509
Zinc	0.7567	0.8123
Iron	0.4932	0.6318
Vit B1	0.6794	0.7922
Vit B2	0.5012	0.6745
Vit B6	0.6688	0.7182
Folate Acid	0.8070	0.9246
Vit. C	0.7179	0.8408
Calcium	0.3277	0.5864
Mean Adequacy	0.6674	0.7948
Observations	28583	70765

Table 7: Variable description

Variable	Description	Mean	SD
Count Index	Number of distinct food items consumed by the hh	37.4230	9.7778
Log income	Log of total per capita expenditure in Rupees (0.00)	11.8343	0.6258
Rural	Dummy that takes the value 1 if hh lives in a rural area	0.7008	0.4579
Log land	Log of the amount of land possessed in ha +1	3.7467	2.5895
Home ratio	Ratio of consumed home produced food quantity and total consumed food quantity	0.1518	0.2446
Ration Card	=1 if household posses a ration card	0.8229	0.3817
Education	Years of education of household head	4.7187	3.7626
Pc meals out	Number of meals per capita that were on average consumed within the 30 day recall period outside the household	1.9076	4.3334
Pc meals	Average pc number of meals within 30 day recall period	71.0242	13.8866
Hh size	Number of household members	4.5427	2.1584
Hinduism	= 1 if religious affiliation is Hinduism	0.8312	0.3746
Islam	= 1 if religious affiliation is Islam	0.1189	0.3236
Buddhism	= 1 if religious affiliation is Buddhism	0.0061	0.0781
Christianity	= 1 if religious affiliation is Christianity	0.0241	0.1532
Sikhism	= 1 if religious affiliation is Sikhism	0.0152	0.1225
Jainism	= 1 if religious affiliation is Jainism	0.0027	0.0515
Children 1-6y	Number of children in household age 1-6	0.5167	0.8325
Children 7-13y	Number of children in household age 1-6	0.6740	0.9445
Children 14-17y	Number of children in household age 1-6	0.3728	0.6466
Elders	Number of household members $\geq$ 60 Years	0.3661	0.6402
women	Number of women in Household	2.2013	1.3131

Figure 1: Expected relation between income and food diversity

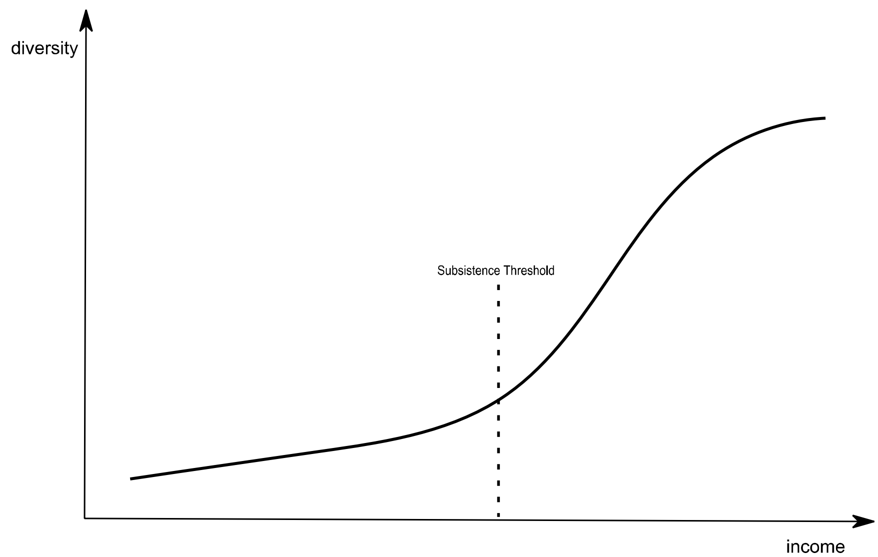

Figure 2: Local Polynomial Regression of the Staple Calorie Share on Log Income

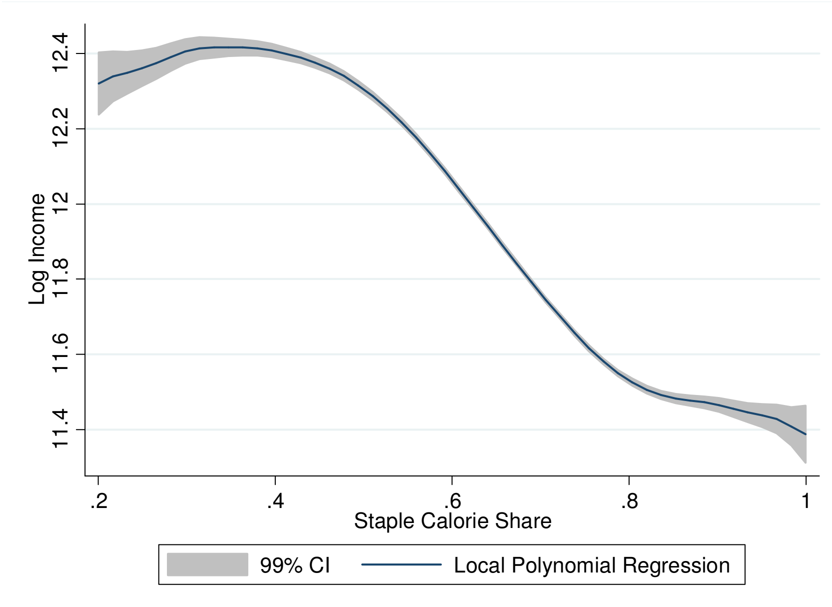

Figure 3: Local Polynomial Regression of Posterior Probabilities on Mean Nutrient Adequacy

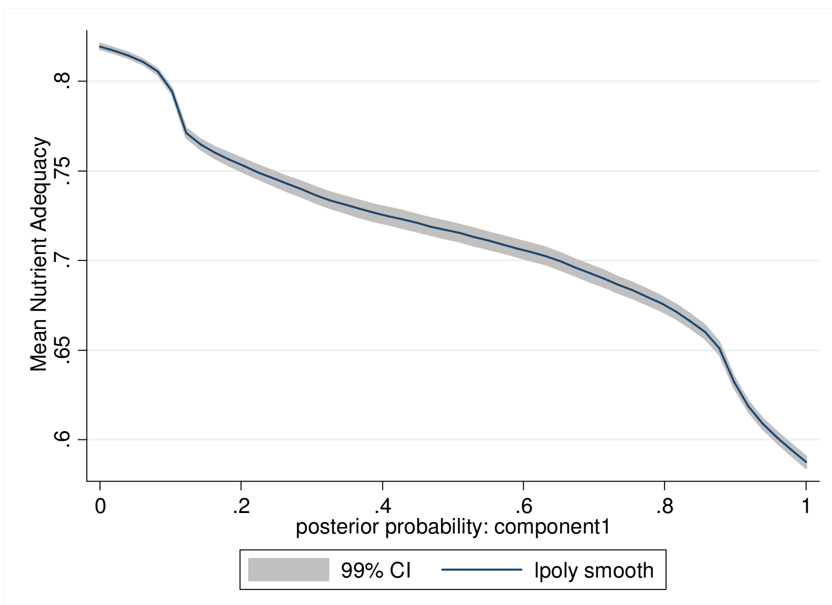

Figure 4: Histogram of the count index

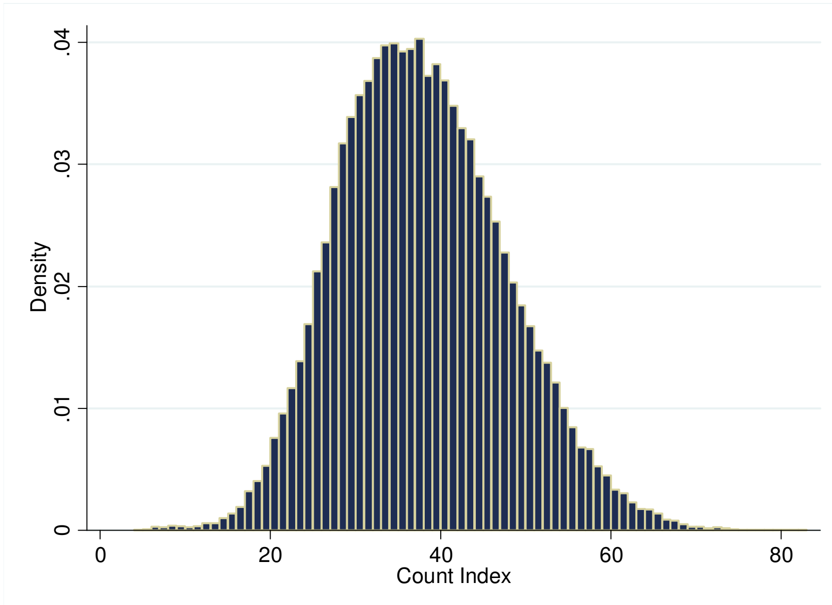

Figure 5: Histogram of the count index by component