Three Essays on Changing Food Consumption Patterns in Indonesia

by

Mohammad Rafiqul Islam

THESIS

Submitted to the University of Adelaide

In partial fulfilment of the

requirement for the degree of

Doctor of Philosophy

in Economics

June 2018

Table of Contents

List of Tables and Sections

List of Figures

Abstract

Food consumption differs significantly across households over time. Economists tend to explain the differences in food consumption in terms of traditional economic variables such as prices and income. While these factors have been observed to explain the differences in consumption in a greater extent, other factors such as migration and education are assumed to substantially alter patterns of food consumption. In this thesis, I study the differences in food consumption, mainly focusing on non-traditional economic factors that have been perceived to be important drivers of changing food consumption patterns.

Previous research suggests that households' resources (i.e. total expenditure), demographics, migration, and education, are the key determining factors influencing household welfare. Because households have different levels expenditure and expenditure is related to household's welfare, factors that may affect expenditure such as household size, natural disasters, levels of education, out-migration and so on may also affect welfare. In this dissertation, I explore the impact that the above factors have on the patterns of food consumption, and hence welfare, of households in Indonesia using rich and comprehensive longitudinal Indonesian Family Life Survey (IFLS) data.

An Engel curve depicts the mean budget share of a particular food group at each level of household expenditure while prices of goods are held constant. Employing the correct specification of an Engel curve in food demand analysis plays a key role in estimating precise food consumption parameters. The first essay analyses the food consumption patterns by applying Lewbel and Pendakur's (2009) Exact Affine Stone Index (EASI) to the IFLS dataset. The EASI demand system is a powerful framework for analysing consumer food choices and policy evaluation, as it can be applied to any higher order polynomials of per capita food expenditure as a main explanatory variable when estimating the Engel curve. To my knowledge,

my first essay titled “Changing food consumption patterns: An application of the Exact Affine Stone Index demand system in Indonesia” is the first to apply the EASI method to investigate consumer food demand functions in the context of a developing economy, namely Indonesia. I find that the estimated food Engel curves have a variety of shapes with sufficient curvatures, and that the rank of the food demand functions (i.e. Engel curves) can be approximated by up to 3rd order polynomial functions of real household expenditure. Furthermore, poorer and richer households have statistically significantly different food consumption choices. The most striking and somewhat surprising finding is that the wealthier households do not appear to diversify their food consumption further when their income rises, whereas poorer households tend to diversify their food consumption significantly when their wealth increases.

The second essay addresses the impact of internal migration on food consumption patterns. This issue is pertinent to Indonesia as it has a large number of internal and interprovincial migrations throughout its history. I use distance to the migrant’s destination and propensity score-matching to generate plausibly exogenous variations in migration to identify the effect of migration on food consumption. Overall, I find that on average, the migrant-sending household’s per capita food consumption is larger (13.4%) than that of non-migrant sending households living in the same neighbourhood (10.7%). To claim that this finding is not driven by other unobserved variables, this study has employed both fixed effects (FE) and instrumental variable (IV) regressions and these estimates consistently support OLS findings. Moreover, migrant households appear to make a substantial shift from the consumption of rice, corn, and wheat towards the consumption of vegetables and fruits, dairy products, and ‘meat and animal’ foods. The results have a suggestive evidence about the value of internal migration for improving welfare in terms of changing food consumption patterns of migrant-sending households.

Since the 1970s, Indonesia has had an impressive record of educational extension, including six years of compulsory schooling (effective from 1984) and nine years of compulsory schooling (effective from 1994). Enrolment rates in primary schools are close to universal and about 75% for secondary education. There is an ongoing effort to expand secondary school attainment at the universal level. Whereas an overwhelming portion of the literature has focused on the labour market (monetary) returns to education in both the developed and developing economies, to date, only a few studies have investigated the impact of education on food consumption. The third essay attempts to fill in this gap by exploring the relationship between the household head's educational attainments and household consumption patterns in Indonesia. To obtain consistent and causal estimates, I employ a quasi-parametric selection model and instrumental variable approach to address the endogeneity of education (i.e. schooling). I use distance from the household to the institutions in which the household has attained education in Indonesia as a source of plausibly exogenous variation in schooling. The first-stage result shows that distance to the school is a strong predictor to education. The ordinary least squares (OLS) estimates suggest that households with graduates from the higher secondary schools tend to consume 31.5% more healthy foods and 22.8% less unhealthy foods than households with graduates from lower secondary schools. These findings have been confirmed by IV estimation and imply that the OLS estimates are not driven by other unobserved characteristics in the households. The results also demonstrate that households have heterogeneous food consumption returns due to different educational attainments.

Taken together, these three essays is an attempt to provide an empirical investigation into how household welfare, measured in terms of food consumption, could be influenced by important socio-economic variables such as household's resources, migration, and education. The findings from the three essays of the dissertation suggest that non-traditional economic variables (as opposed to tractional economic variable such as prices and income) market access,

natural disasters, migration and education such as also influence food consumption significantly across households. The findings may have policy implications that the government may perhaps undertake in relation to migration and education so as to enhance welfare within a household.

Key words: food consumption, IFLS, expenditure, migration, education, Engel curve, household welfare, policy issues

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library catalogue and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Signed

Date: 04 June 2018

Acknowledgments

First, I'd like to express my profound gratitude to Almighty Allah who is the creator of the universe. Indeed all praise is due to Him, as without Him my success throughout this journey would not have been possible.

Next, I am highly grateful to my principal supervisor Professor Christopher Findlay for his ideas to work on Indonesian food consumption patterns, every guidance, strong support and allocating his valuable time throughout all stages of the dissertation. I am highly thankful to co-supervisor Nicholas Sim for his helpful insights, research guidance, and editing. Without his strong intellectual leadership and mentorship, this dissertation wouldn't have been completed in a timely manner. He has provided invaluable comments and suggestions on all my drafts.

My PhD program was financially supported by Adelaide Scholarship International (ASI) at the University of Adelaide. Without this support the journey to accomplish dissertation would be in a very difficult state. I have my very best wishes for the University of Adelaide's endeavour to support for the PhD studies.

I am indebted to Miranda Roccisano for her excellent editorial assistance to make the thesis reader friendly to others. This help made my life a lot easier, and most importantly saved my time to put more attention on study.

My graduate school of economics experience was greatly enriched by my interactions with fellow students and friends, in particular Wenshou, Kaixing, Heng, Jacky, Samira, Jesmin, Manob, Ann, Dessie, and Habtamu. Their camaraderie made my graduate student life more enjoyable. I am also very grateful for the support from the administration, faculty and staff of the University of Adelaide's graduate school of economics.

During my Indonesian research visit, I have been greatly benefitted from a number of people including Nunung Nuryartuno, Shamsul, Dennis, Silvi, Pristi, Sigit, and all the staffs of the Inter-café at the Bogor Agricultural University in Indonesia.

I am especially indebted to Sharmina Ahmed, post-doctoral fellow at the Global food Studies in the University of Adelaide for recommending me Christopher Findlay as the supervisor and subsequent support to accomplish the program. I am also grateful to Professor Wahid Murad at the University of South Australia for his continuous support and assistance extended to me throughout the program.

I am especially grateful to my wife Kamrun Ahmed who been highly supportive during my PhD program. She has taken a lot of family responsibilities and has greatly motivated me to finish the program in a timely manner. She has also taken much care of my kids Raisa and Raiyan. I have every love and thanks to my kids.

Finally, I want to thank my parents for their unfailing and unconditional love.

Dedication

This dissertation is dedicated to Professor Shaheen Akhter, former Chair of Economics at Shahjalal University of Science and Technology Sylhet Bangladesh, who has reshaped my vision to start a career in Economics.

Acronyms

3SLS	three-stage least squares (3SLS)
AARES	Agricultural and Resource Economics Society
AIDS	almost ideal demand system
ATE	average treatment effect
ATT	average treatment on the treated
CPI	consumer price index
CSASE	Centre for the Study of African Economies
DCE	discrete-choice-experiment
DID	difference-in-difference
EA	enumeration area
EASI	Exact Affine Stone Index
FE	fixed effect
GMM	generalised methods of moments
HH	household
IFLS	Indonesian Family Life Survey
LOWESS	locally weighted scatterplot smoothing
MPRTE	marginal policy relevant treatment effect
MTE	marginal treatment effect
OLS	ordinary least squares
PRTE	policy relevant treatment effect
PSM	propensity score matching
QUAIDS	quadratic almost ideal demand system
SDG	sustainable development goals

UNDP	United Nations Development Program

Chapter 1 Introduction

1.1 Background and motivation

For both developing and developed nations, food consumption is an important determinant of welfare. Sen (1989) noted that enjoying a decent life, including freedom from both avoidable morbidity and premature mortality, is centrally connected to the provision of food. From the point of view of development practitioners and nutritionists, one relevant question is: How are patterns of food consumption affected when certain household characteristics such as demographics undergo a change, or when households are faced with natural disasters or sudden price hikes for essential commodities? These factors are assumed to change a household food consumption patterns are substantially. However, magnitude of the changes in food consumption may have different for developing countries than for developed countries. Moreover, internal migration and levels of education of a household head may affect food consumption patterns significantly.

With a population of more than 255 million (World Bank, 2015), Indonesia has become the fourth most populous country in the world, behind the US (321 million), India (1,251 million) and China (1,367 million). Indonesia is also the largest archipelago in the world with more than 17,000 islands (of which 6,000 are inhabited). A map of Indonesia has shown in the Figure 1 to show its diversity in Geography. Most importantly, Indonesia also has a diverse geography with numerous ethnicities (more than 300 ethnic groups), has enormous cultural differences, and more than 719 spoken languages (*Indonesia – the road ahead*, 2014). Furthermore, since its independence in 1945, Indonesia has undertaken various policy changes (e.g. compulsory schooling policies in 1974 and in 1984) and implemented a transmigration program during 1905–2015. The enormous diversity and drastic economic policy changes in response to a crisis or to boost development may have direct welfare implications for the households. Although quite a lot of studies devoted to analyse Indonesia's food consumption patterns, scanty of literature found to identify causally the impact of important variables such as migration and education on food consumption.

Figure 1: Map of Indonesia

Source: http://www.worldatlas.com/webimage/countrys/asia/id.htm

The prevailing food consumption studies for Indonesia are inadequate for wide-ranging food policy analysis (see for instance Alderman and Timmer, 1980; Deaton, 1987 Dixon, 1982; Jensen and Manrique, 1998; Muzayyanah and Maharjan, 2011; Teklu and Johnson 1986, 1987, 1988; Timmer et al. 1986). First, most of the food demand elasticities are estimated from ad hoc demand models and these models do not have enough flexibility to account for a variety of shapes of the Engel curves. Pendakur (2009) argued that typical consumer demand models cannot explain the variety of shapes (i.e. curvatures) of the Engel curve and do not incorporate unobserved preference heterogeneity. Second, the demand parameters are restricted to estimates of expenditure elasticities for some essential commodities and are explained by traditional economic variables such as income and prices. Third, existing food demand estimates in Indonesia have scarcely recognised the impact on food consumption of some critical food consumption drivers, such as migration and education. The existing literature on these issues in the Indonesian context is sparse and documentation is not neither systematic nor coherent.

This thesis attempts to add to our understanding of food consumption patterns in the context of developing countries. Using ongoing and longitudinal Indonesian Family Life Survey (IFLS) data, this study investigates three issues in the context of Indonesia: identification of important drivers of changing food consumption patterns by applying the Exact Affine Stone Index (EASI) demand system developed by Lewbel and Pendakur (2009), the causal relationship between internal migration and food consumption patterns, and the impact of education on food consumption.

The main dependent variable in this study is annualised food consumption. It has been well documented that (food) consumption is one of the most important measures of household-level welfare. Consumption is a core variable in measuring welfare as it provides a more accurate measure of material wellbeing than income, which varies seasonally throughout the year (consumption, poverty and welfare, 2015). Other reasons for why researchers are keen to use consumption as a measure of welfare are that consumption is more log-normally distributed than income (Battistin, Blundell, and Lewbel, 2009); recent household-level surveys record consumption in greater detail; and data on consumption are less noisy than data on income.[1] Adam Smith (1776) noted that "… [c]onsumption is the sole end and purpose of all production and the welfare of the producer ought to be attended to, only so far as it may be necessary for promoting that of the consumer" (p. 49).

1.2 Brief literature review

1.2.1 On food consumption patterns

Precise estimates of food consumption parameters are essential for designing food policies. This necessitates the application of an appropriate food demand function. However, researchers face significant challenges because popular and widely used demand models (including the translog model of Christensen, Jorgenson, and Lau (1975), the almost ideal demand system (AIDS) of Deaton and Muellbauer (1980), and the quadratic almost ideal demand system (QUAIDS)) cannot explain the variety of shapes of the Engel function. In other words, they are constrained by rank restrictions.[2] Deaton (1986) explained the standard econometric issues associated with demand system estimation. More recently, demand system literature has

[1] Some important implications of assuming log normal distribution is as follows: log normality implies that social welfare function can be parsimoniously specified, log normality of consumption can handle possible measurement errors in nonlinear models, and budget shares of some food groups (e.g. food) are close to log-linear in total consumption (for more detailed explanation see Lades (2013)).

[2] Gorman (1981) demonstrated that in the case of exactly aggregable demand systems, the matrix of Engle curve coefficients is to be represented by at most, rank 3. Full rank demand systems are those with rank equal to the columns of the coefficient matrix in the Engel specification (Lewbel 1990). Barnett and Serletis (2008) provided a good discussion of demand system rank, with examples.

focused on the higher rank of demand systems and functional specifications (see LaFrance and Pope, 2008, 2009).

Most parametric demand systems, including AIDS and QUAIDS, can explain up to quadratic Engel curves. However, recent empirical work finds that Engel curves have significant curvatures and variety of shapes across goods (Blundell, Chen, and Kristensen, 2007). Moreover, existing demand models cannot take into account of unobserved preference heterogeneity. To capture both the variety of shapes and unobserved preference heterogeneity, Lewbel and Pendakur (2009) developed a new methodology for modelling demand systems known as the EASI demand system. This new demand system has several features that other existing models cannot explain. First, the EASI functional form exhibits Engel curves with a high degree of flexibility. Second, the demand system has budget shares with linear parameters. Third, it can accommodate heterogeneity in preferences. Fourth, Engel curves can have polynomials of any order in expenditure. Finally, it can have any rank up to n-1, where n is the number of goods.

1.2.2 On migration and food consumption patterns

Migrants play a key driving force for the socio-economic development of Indonesia (International Organisation for Migration, 2010). Labour migration in the form of large-scale migration across provinces and to other countries plays a vital role in improving livelihood and social mobility (Anggraeni, 2006; Ford, 2001). In Indonesia, labour migration has been observed to be influential to newly acquired money, goods, ideas, and behaviour, which otherwise would not be the case (Hugo, 1995). The intensity of family network has found to be positively and significantly affects finding employment in the formal sector (Knerr, 2012).

Most studies demonstrate that short-term dominates over long-term and permanent migration in Indonesia, Viet Nam, China, and Cambodia (Hugo, 2003; International Institute for Environment and Development, 2004; Ping, 2003; Sheng, 1986; Zhao, 2003). This implies that migration duration could play an important role to affect welfare in the families of migrant's origin.

Several studies have investigated the impact of migration on food consumption. For example, Karamba et al. (2011) examined the impact of migration on food consumption patterns in Ghana. One potential concern is that migration is endogenous, as migrant-sending households that have unobservable characteristics, not controlled for otherwise, which could influence per capita food consumption. This could create an omitted variable bias, leading to inconsistent estimates when OLS is applied. To address this issue, they used the migration networks as a strategy to identify the effect of migration on food consumption. They found that internal migration does not markedly affect per capita food expenditure, and food consumption increased only in areas of high migration density. They also observed that consumption shifts towards sugar, beverages and eating out. All of which may be less nutritious. In terms of policy implications, these findings raise the question of whether migration does improve the overall wellbeing of households in Ghana. Employing panel data, Nguyen and Winters (2011) investigated the impact of migration on food consumption patterns in Vietnam. To address the endogeneity issue of migration, the authors employed an instrumental variable regression, where an interactions of household size and migration network has been used to instrument migration. Their findings suggest that short-term migration positively and significantly affects food consumption and long-term migration has an insignificant impact on food consumption. They proposed that the government of Vietnam should adopt the policy of stimulating short-term migration to improve food consumption patterns.

A large number of literatures have looked at the impact of migration on labor market outcomes, remittances, and health issues. However, there is little research on the impact of internal migration on food consumptions. This paper attempts to shed light on this issue by focusing on Indonesia. In terms of methodology, this study has endeavoured to bridge these gaps by employing both a semi-experimental method and an instrumental variable approach.

1.2.3 On education and food consumption patterns

Given the extensive epidemiological and experimental research that links diseases to the choices of food consumption, health and development practitioners need to know the main drivers (e.g. education) of the consumption patterns of the population (Chait et al., 1993; Denke, 1995; Fraser et al., 2000; WCRF/AICR, 2007).

Despite the economic progress made by developing countries in recent years, Pritchett's (2001) cross-national study found significant and negative association between educational capital growth and total factor productivity. Therefore, it is essential from the social policy perspective to derive good estimates of human capital.

The existing research literature of Indonesia has focused mainly on the monetary returns from education (Behrman and Deolalikar, 1993; Carneiro et al., 2016; Comola and Mello, 2010; Duflo, 2001; Dumauli et al., 2015; Maulana, 2012; Newhouse and Suryadarma, 2011; Patrinos, Ridao-Cano and Sakellarjou, 2006; Purnastuti et al., 2013; Purnastuti et al., 2015). While studies on monetary returns to education are important because of household allocation of (limited) resources to the priority areas that provide the highest returns, investigation of non-monetary returns to education (e.g. food consumption) are equally important because monetary returns do not always imply that individuals are eating well and consuming healthy food bundles. The increased consumption of healthy foods as a result of education has important policy implications for building the productive populace of the nation. This study is an attempt to meet this gap in the literature in the context of Indonesia.

1.3 Data

This thesis uses ongoing and longitudinal public available IFLS datasets to examine the food consumption dimension of welfare in Indonesian households. To date, the IFLS have conducted five waves of surveys from 1993 at irregular periods: IFLS 1993, IFLS 1997, IFLS 2000, IFLS 2007, and IFLS 2014. I drop IFLS 1993 and employ the rest four IFLS surveys, because the consumption module from the first survey is not comparable across the surveys.

The IFLS collects detailed information systematically at the individual, household and community levels. It contains a comprehensive consumption module of food and non-food consumption as well as information on several other topics that are essential in the assessment of household-level welfare changes. There is a particularly rich set of data for migration, education, receipt of central government sponsored (JPS) programs and other social safety net programs. Moreover, the IFLS contain an extremely rich set of data at the community level. Because it is a panel survey it is possible to analyse changes for specific circumstances for the individuals in specific households and their communities.

As well, by including IFLS 1997 and IFLS 2000 there is a unique opportunity to explore both the short-term and long-term impact on household welfare pre- and post–Asian Financial Crisis.

1.4.1 IFLS1, 1993

The first of the surveys, IFLS 1993, was fielded in the second half of 1993 and ended in January 1994. More than 30,000 individuals from 7,224 households were sampled. The first wave was stratified on province and then rural (and urban) areas were further stratified within provinces. Within these strata, enumeration areas (EAs) and households within EAs were randomly sampled. The sampling framework is based on the Central Bureau of Statistics sampling for the National Socioeconomic Survey 1993 in Indonesia (SUSENAS 1993, http://www.rand.org/labor/bps/susenas/1993.html). Provinces were selected to maximise representation of the size of the population, to capture the cultural and socioeconomic diversity of Indonesia, and to minimise the cost of the survey. The resulting sample covered 13 provinces on the islands of Java, Sumatra, Bali, Kalimantan, Sulawesi and Nusa Tenggara.

Some 321 EAs from 13 provinces were randomly sampled. More EAs were selected from urban areas and a smaller number of EAs were selected from the smaller provinces to make the sampling comparable between rural and urban areas and between Java and Non-Java. Twenty households were selected randomly from each urban EA, and 30 from each rural EA. This approach reduced intra-cluster correlation across urban households.[3]

In the IFLS 1993, a total of 7,730 households were selected as the original target sample. Of these households, 7,224 (93%) were interviewed.

[3] A household was defined as a group of people whose members reside in the same dwelling and share food from the same cooking pot (the standard Badan Pusak Statistikn (BPS) or Central Bureau of Statistics definition).

1.4.2 IFLS2, 1997

The main fieldwork for IFLS2 took place between June and November 1997. This was just before the worst hit of the Asian Financial Crisis on the economy of Indonesia.

The total number of households contacted in IFLS 1997 was 7,629, of which 6,752 were panel households. This represents a completion rate of 94.3% for the IFLS 1993 households that remained.

1.4.3 IFLS3, 2000

The main fieldwork for IFLS 2000 continued from June through to November 2000. The sampling approach for this survey was to recontact all of the original IFLS 1993 households. Over 10,500 households (the original plus split-off households) were contacted, containing more than 43,600 individuals. A 94.8% recontact rate was achieved for all "target" households.

1.4.4 IFLS4, 2007

The main fieldwork for IFLS 2007 continued from late November 2008 to May 2009.The target households for IFLS 2007 were the original IFLS1 households, minus those of which all members had died by 2000, plus all of the split-off households from 1997, 1998 and 2000. In total, 13,995 households were contacted. Of these, 13,535 households were interviewed. Of the 10,994 target households, the recontact rate was 90.6%: 6,596 original IFLS1 households and 3,366 old split-off households.

1.4.5 IFLS5, 2014

Household fieldwork took place between September 2014 and March 2015. The recontact rate (including deaths) in IFLS 2014 for IFLS1 individuals was thus 76%. The recontact rate for the main respondents from IFLS1, 1993, was higher, 82%.

Over the course of IFLS, 17,295 individual respondents were contacted in all 5 waves (52.3% of IFLS1 household members), and of these, 11,889 (54% of IFLS1 "main respondents") were interviewed in all five waves.

1.4 Research objectives

The main objective of this study is to identify the critical food consumption parameters in Indonesia. The specific objectives are to be explored in the form of three essays. This dissertation deals with the following objectives.

In Essay 1, the objectives are:

- to investigate the shape of the Engel curve including estimation of food consumption parameters
- to explore whether poor people diversify food consumption as they become affluent.

In Essay 2, the objectives are:

- to explore the extent to which migrant-sending household's welfare in terms of food consumption differs to that of non-migrant-sending households living in the same neighbourhood

- to examine whether migrant-sending households shift consumption from carbohydrate-rich foods towards nutrient-dense foods.

In Essay 3, the objective is:

- to inspect whether higher educated individuals make a healthy food choice or unhealthy food choice.

1.5 Contributions and key findings of the thesis

This dissertation contributes to the recent literature on food consumption patterns in several ways. First, to my knowledge, this study is the first to apply and estimate Lewbel and Pendakur's (2009) Exact Affine Stone Index (EASI) demand system to investigate food consumption patterns of Indonesian households. The model estimated and presented in this study differs from well-established models, such as AIDS and QUAIDS, and departs slightly from Lewbel and Pendakur (2009) in that I explicitly include the structural variables as explanatory variables while the existing literature typically uses traditional socioeconomic variables. Second, using distance-based measure of migration and propensity score matching (PSM) methods, this study attempts to find causal linkage between internal migration and food consumption. A key departure of the second essay from the literature is that it looks at the effect of migration on changing food consumption diversification, while the other literature considers mainly the effect of migration on earnings or overall consumption. Third, applying quasi-experimental methods, the third essay examines whether more education of the head of a household lead to choose healthy food bundles. A central difference of the third essay from the prevailing literatures is that it examines the non-monetary returns to education, while other studies explores mainly the monetary returns to education.

From the empirical investigation of three essays on food consumption patterns in Indonesia, the key findings, which may have implications for designing food policies in the developing country context, are summarised as follows.

The first essay employed the EASI method to estimate food demand semi-elasticicities and their determinants in Indonesia households. This study uses IFLS 1997, IFLS 2000, and IFLS 2007 data to capture consumption patterns before and after the Asian Financial Crisis and unobserved preference heterogeneities in consumption. The estimated EASI Engel curves are found to be a variety of shapes across food groups and empirical evidence supports the dimensions of the Engel curves up to rank 3. Moreover, Indonesian consumers are found to be highly sensitive even to small changes in the magnitudes of the own price elasticities.

The second essay studies the impact of internal (out) migration on food consumption patterns of the migrant-sending households living in the same locality of the non-migrants. I use IFLS 2007 and IFLS 2014 to estimate whether households having a migrant have different consumption patterns (i.e. welfare) compared to households not having migrants. The main empirical challenge of this investigation is the endogeneity of migration. I use a difference-in-difference (DID) approach to remove all potential sources of time-invariant unobserved heterogeneity. Nevertheless, if there are unobserved idiosyncratic household characteristics that affect both the migration process and the household welfare, DID estimates may still be biased and inconsistent. Therefore, I employ the instrumental variable (IV) method to address the confounding influence that these variables might have. The result shows that internal migration exerts a statistically significant impact on the overall wellbeing of the migrant-sending households compared to non-migrant-sending households. In particular, on average, per capita food consumption (i.e. welfare) of migrant households increases by about 13.4% compared to non-migrant households living in the same locality. To claim that this finding is not driven by other unobserved variables, this study has employed both fixed effects (FE) and instrumental variable (IV) regressions and these estimates consistently support OLS findings. Moreover, in the case of migrant households, there appears to be a substantial shift in consumption from rice, corn, and wheat towards vegetables and fruits, dairy products, and meat and animal food group consumption. This finding has direct nutrition and health implications for the population in Indonesia.

The final essay investigates the marginal food consumption returns to education. This study employs a quasi-parametric selection model and location-based measure of instrument to schooling to identify the causal parameter of interest. I find statistically significant evidence that households with graduates from the higher secondary schools tend to consume 31.5% more healthy foods and 22.8% less unhealthy foods than households with graduates from lower secondary schools. The results also demonstrate that households have heterogeneous food consumption returns owing to different educational attainments.

1.6 Structure of the thesis

The rest of the thesis is structured as follows. Chapter 2 (Essay 1) identifies the main drivers and parameters of food consumption. Chapter 3 (Essay 2) establishes an empirical causal link between internal migration and food consumption patterns. Chapter 4 (Essay 3) identifies the relationship between education and food consumption. Finally, Chapter 5 concludes with policy prescriptions.

References

Alderman, H., and C. P. Timmer (1980): "Food policy and food demand in Indonesia," *Bulletin of Indonesian Economic Studies*, 16(3), 83–93.

Anggraeni, D. (2006): *Dreamseekers: Indonesian women as domestic workers in Asia*. Equinox Publishing and ILO, Jakarta

Angus Deaton: Consumption, poverty and welfare (2015) Scientific Background on the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2015. Available at https://www.kva.se/globalassets/priser/ekonomi/2015/sciback_ek_en_15.pdf

Barnett, W. A., and A. Serletis (2008): "Consumer preferences and demand systems," *Journal of Econometrics*, 147 (2), 210–24.

Battistin, E., R. Blundell, and A. Lewbel (2009): "Why is consumption more log normal than income? Gibrat's Law revisited," *Journal of Political Economy*, 117 (6), 1140–1154.

Behrman, J. R., and A. B. Deolalikar (1993): "Unobserved household and community heterogeneity and the labor market impact of schooling: A case study for Indonesia," *Economic Development and Cultural Change*, 41 (3), 461–488.

Bianchi, C., and G. Mortimer (2015): "Drivers of local food consumption: a comparative study," *British Food Journal*, 117 (9), 2282-2299.

Blundell, R., X. Chen, and D. Kristensen (2007), "Semi-nonparametric IV estimation of Shape-Invariant Engel Curves," *Econometrica*, 75 (6), 1613–1669.

Carneiro, P., M. Lokshin, and N. Umapathi (2016): "Average and marginal returns to upper secondary schooling in Indonesia," *Journal of Applied Econometrics*, 31(4).

Chait, A., J. D. Brunzell, M. Derke, D. Eisenberg, N. D. Ernst, F. A. Franklin, H. Ginsberg, T. A. Kotchen, L. Kuller, R. M. Mullis, M. Z. Nichaman, R. J. Nocolosi, E. J. Schaefer, N. J. Store, and W. H. Weidman (1993): "Rationale of the diet–heart statement of the American Heart Association," *Circulation*, 88, 3008–3027.

Christensen, L. R., D. W. Jorgenson, and L. J. Lau (1975): "Transcendental logarithmic utility functions," *American Economic Review,* 65 (3), 367–83.

Comola, M., and L. de Mello (2010): "Educational attainment and selection into the labour market: The determinants of employment and earnings in Indonesia," Paris-Jourdan Sciences Economiques Working Paper 2010–06, Paris-Jourdan Sciences Economiques, Paris.

Deaton, A. S. (1986): "Demand analysis," in Z. Griliches and M. D. Intriligator (Eds) *Handbook of Econometrics,* vol. 3. North-Holland: New York.

Deaton, A. (1987): "Estimation of own and price elasticities from household survey data," *Journal of Econometrics,* 36 (1-2), 7–30.

Deaton, A. S., and J. Muellbauer (1980): "An almost ideal demand system," *American Economic Review,* 70 (3), 312–326.

Denke, M. A. (1995): "Cholesterol-lowering diets. A review of the evidence," *Arch Intern Med*, 155 (1), 17–26.

Dixon, J. A. (1982): "Food Consumption Patterns and Related Demand Parameters in Indonesia: A Review of Available Evidence," Mimeo, World Bank.

Duflo, E. (2001):"Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment," *The American Economic Review*, 91 (4), 795–813.

Dumauli, M. T. (2015): "Estimate of the private return on education in Indonesia: Evidence from sibling data," *International Journal of Educational Development*, 42, 14–24.

Ford, M. (2001): "Indonesian women as export commodity: Notes from Tanjung Pinang," *Labour Management in Development Journal*, 2 (5), 2–9.

Frankenberg, E. and L. Karoly. (1995): *The 1993 Indonesian Family Life Survey: Overview and field report.* RAND, Santa Monica, CA.

Frankenberg, E. and D. Thomas (2000): *The Indonesia Family Life Survey (IFLS): Study design and results from waves 1 and 2*," RAND, Santa Monica, CA. DRU-2238/1-NIA/NICHD.

Fraser, G. E., A. Welch, R. Luben, S. A. Bingham, N. E. Day (2000): "The Effect of Age, Sex, and Education on Food Consumption of a Middle-Aged English Cohort—EPIC in East Anglia," *Preventive Medicine* 30, 26–34.

Gorman, W. M. (1981): "Some Engle curves," in *Essays in the theory and measurement of consumer behavior: in honour of Sir Richard Stone*, ed. A.S. Deaton. Cambridge University Press, New York.

Hugo, G. (1995): "International labor migration and the family: Some observations from Indonesia," *Asian and Pacific Migration Journal*, 4, 273–301.

Hugo, G. (2003). *Urbanisation in Asia: an overview*. Paper prepared for Conference on African Migration in Comparative Perspective, Johannesburg, South Africa. Available from http://webcache.googleusercontent.com/search?q=cache:ppFQRoZnpmQJ:www.china-up.com:8080/international/case/case/1141.pdf+&cd=1&hl=en&ct=clnk&gl=au

International Institute for Environment and Development (2004): "Stay on the farm, weave in the village, leave the home: Livelihood diversification and rural-urban linkages in Viet Nam's Red River Delta and their policy implications," Briefing Paper 9, Rural–Urban Interactions and Livelihood Strategies, International Institute for Environment and Development: London.

Indonesia: The road ahead – Decoding a nation of 13,466 islands, 360 ethnic groups and 719 languages (2014). [Book review]. *The Economist*, 21 June. Available at http://www.economist.com/news/books-and-arts/21604528-decoding-nation-13466-islands-360-ethnic-groups-and-719-languages-road-ahead

International Organisation for Migration (2010): *International migration and migrant workers' remittances in Indonesia: Findings of baseline surveys of migrant remitters and remittance beneficiary households*. International Organisation for Migration: Makati City.

Jensen, H. H., and J. Manrique (1998): "Demand for food commodities by income groups in Indonesia. *Applied Economics*," 30 (4), 491–501.

Karamba, W. R., E. J. Quiñones, and P. Winters (2011): "Migration and food consumption patterns in Ghana," *Food Policy*, 36 (1), 41–53.

Knerr, B. (2012): "Transfers from international migration: A strategy of economic and social stabilisation at national and household level," in *International labour migration,* ed. B. Knerr. Kassel University Press, Kassel.

Kearney, J. (2010): "Food consumption trends and drivers, Philosophical Transactions of the Royal Society B, 365 (1554), 2793-2807.

Lades, L. K. (2013): "Explaining shapes of Engel curves: the impact of differential satiation dynamics on consumer behaviour," *Journal of Evolutionary Economics*, 23 (5), 1023–1045.

LaFrance, J. T. and R. D. Pope (2008): "Full Rank Rational Demand Systems," Working Paper, Washington, DC, Washington State University.

LaFrance, J. T. and R. D. Pope (2009): "The generalized quadratic expenditure system," in *Quantifying consumer preferences*, ed. Slottje, 83–117.

Lewbel, A. (1990): "Full rank demand systems," *International Economic Review*, 31: 289–300.

Lewbel, A., and K. Pendakur (2009): “Tricks with Hicks: The EASI Demand System,” *American Economic Review*, 99 (3), 827–863.

Maulana, A. (2012): *Returns to junior secondary education and the role of education in moving people out of poverty: The case of Indonesia*, USAID.

Muzayyanah, M. A. U., and K. L. Maharjan (2011): “Socioeconomic determinant of livestock products consumption in urban and rural Java, Indonesia,” *Journal of International Development and Cooperation*, 17 (1), 89–102.

Newhouse, D., and D. Suryadarma (2011): “The value of vocational education: high school type and labor market outcomes in Indonesia,” *The World Bank Economic Review*, 25 (2), 296–322.

Nguyen, M. C., and P. Winters (2011): “The impact of migration on food consumption patterns: The case of Vietnam,” *Food Policy*, 36 (1), 71–87.

Patrinos, H. A., C. Ridao-Cano, and C. Sakellariou (2006): “Estimating the returns to education: Accounting for heterogeneity in ability,” World Bank Policy Research Working Paper 4040, World Bank, Washington DC.

Pendakur, K. (2009): “EASI made easier,” in *Quantifying consumer preferences* (pages 179–206), ed. D. Slottje. Contributions to Economic Analysis series. Emerald Group Publishing, London:

Ping, H. (2003): “China migration country study,” Paper presented at the Regional Conference on Migration, Development and Pro-Poor Policy Choices in Asia organized by the Bangladesh Refugee and Migratory Movements Research Unit, Bangladesh / DFID UK, Dhaka, 22–24 June.

Potter, J. D. (Ed.) (1997): *Food, nutrition, and the prevention of cancer: A global perspective*. World Cancer Research Fund and the American Institute for Cancer Research: Washington, DC.

Pritchett, L. (2001): "Where has all the education gone?" *World Bank Economic Review*, 15 (3), 367–391.

Purnastuti, L., P. W. Miller, and R. Salim (2013): "The evolution of labour market returns to education in Indonesia: Evidence from the Family Life Survey," *Bulletin of Indonesian Economic Studies*, 49 (2), 213–236.

Purnastuti, L., R. Salim and M. A. Joarder (2015): "The returns to education in Indonesia: Post reform estimates," *Journal of Developing Areas*, 49 (3), 183–204.

Sen, A. (1989): "Food and freedom," *World Development*, 17 (6), 769–781.

Sheng, Y. K. (1986): *Urbanization and internal migration*, Poverty and Development Division, ESCAP.

Smith, Adam (1776): *An Inquiry into the Nature and Causes of the Wealth of Nations*, London, Methuen & Co (book IV, chapter 8, 49)

Strauss, J., Beegle, K., Sikoki, B., Dwiyanto, A., Y. Herawati and F. Witoelar (2004): *The third wave of the Indonesia Family Life Survey (IFLS): Overview and field report*. WR-144/1-NIA/NICHD.

Strauss, J., F. Witoelar, and B. Sikoki. (2016): *The fifth wave of the Indonesia Family Life Survey (IFLS5): Overview and, field report*, WR-1143/1-NIA/NICHD. RAND Labor & Population: Santa Monica, CA. Available at http://www.rand.org/content/dam/rand/pubs/working_papers/WR1100/WR1143z1/RAND_WR1143z1.pdf

Strauss, J., F. Witoelar, B. Sikoki and A. M. Wattie (2009): *The fourth wave of the Indonesian Family Life Survey (IFLS4): Overview and field report*. WR-675/1-NIA/NICHD. RAND Labor & Population: Santa Monica, CA.

Sukamdi and Mujahid, G. (2015): "Internal migration in Indonesia," *United Nations Population Fund publication*.

Teklu, T. and S. R. Johnson (1986): *A review of consumer demand theory and food demand studies on Indonesia.* Staff Report 86-SR7, Centre for Agricultural and Rural Development, Iowa State University, Ames, IA.

Teklu, T., and S. R. Johnson (1987): Demand system from cross section data: An experiment for Indonesia, Working Paper 87-WP. Centre for Agricultural and Rural Development, Iowa State University, Ames, IA.

Teklu, T., and S. R. Johnson (1988): Demand system from cross-section data: An application to Indonesia. *Canadian Journal of Agricultural Economics*, 36, 83–101.

World Bank (2015): Population 2015, Available at http://databank.worldbank.org/data/download/POP.pdf

World Cancer Research Fund / American Institute for Cancer Research (2007): "Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective," Washington DC: AICR. Available at http://www.aicr.org/assets/docs/pdf/reports/Second_Expert_Report.pdf

Zhao, Z. (2003). *Rural–urban migration in China – What do we know and what do we need to know?* China Centre for Economic Research, Peking University.

Chapter 2 Essay 1: Changing Food Consumption Patterns: An Application of the Exact Affine Stone Index Demand System in Indonesia

2.1 Introduction

Expenditure on food consumption is an important indicator of household welfare. There is a growing body of literature that looks at the welfare of households in terms of the amount of expenditure spent on food consumption.[4] Because food consumption relates to one's welfare, credible estimates of food consumption patterns, especially the relationship between household expenditure and food demand, are useful from the perspective of policymaking.[5] To obtain these estimates in precise and flexible way, it is necessary to devise an Engel curve, which formalizes the relationship between demand for food and household total income that has the ability to take advantage of large consumer expenditure data sets for modelling any nonlinearities that might be present in food consumption patterns.

One commonly encountered issue when estimating the food Engel curve lies in the fact that the Engel curve is constrained by the so-called rank restriction. The rank restriction essentially means the maximum number of linearly independent columns in the matrix of Engel equation coefficients.[6] To satisfy the rank restriction when estimating food consumption patterns,

---

[4] See, for example, Banks, Blundell, and Lewbel, 1997; Attanasio and Lechene, 2014; Attanasio et al., 2013; Blundell, Duncan, and Pendakur, 1988; Deaton, 2016; Bludell, Pashardes, and Weber, 1993; Jorgenson, Lau, and Stoker, 1980, Deaton and Irish, 1984; Deaton and Subramanian, 1996; Deaton, Ruiz-Castillo, and Thomas, 1989).

[5] Several studies document the relationship between food consumption and welfare implications within a household: Ravallion and Lokshin, 2007; Skoufias, Suryahadi, and Sumarto, 2000; Frankenberg, Smith, and Thomas, 2003; Vu and Glewwe, 2011; Juarez-Torres, 2015.

[6] The rank M of any demand system can be defined as the maximum dimension of the function space spanned by the Engel curves of the demand system (Lewbel, 1991). To put it differently, the rank of a demand system is defined as the rank of the matrix comprised of the coefficients of the Engel functions (Kebede, 2003). For instance, if we arrange the coefficients of Engel equations in the form of a demand system as a 3 (number of goods) by 3 matrix, then the maximum possible rank would be 3. Gorman (1981) illustrated that the maximum possible rank of an exactly aggregable demand system is 3. This means that no matter how many columns are in the Engel specifications, adding into additional columns in the matrix of linear expenditure system of Engel curve will be linearly dependent on the others.

researchers often do not explicitly incorporate unobserved preference heterogeneity.[7] However, it is important to do so, because as Pendakur (2009) has pointed out, the observables in typical consumer demand model explain no more than 50% of the variation in budget shares. The rest are attributed to customary uncertainties, including measurement error and consumers' unobserved preference heterogeneity. If unobserved preference heterogeneity is not accounted for, it may result in underestimating the impact of income changes across consumers.

In this chapter, I mainly study the shape of the Engel curve including estimation of food consumption parameters in Indonesia. To do so, I will estimate the modified Engel function proposed by Lewbel and Pendakur (2009), known as the EASI implicit Marshallian demand system. The EASI demand model possesses several advantages over the existing consumer demand models, such as the AIDS and the QUAIDS, and is suitable for capturing consumption patterns in a tractable way. First, the EASI model accounts for unobserved preference heterogeneity through the EASI budget-share functions where the utility (unobservable) can be represented by observables. Second, the EASI demand system can modelled in a flexible way by any number of polynomials of real expenditures. Third, the model error terms can be interpreted as random utility parameters representing unobserved heterogeneity. Fourth, apart from real expenditure, the model is linear in parameters and its error terms are additive, which makes it straightforward to estimate. Fifth, in the EASI demand model, the welfare of a consumer can be calculated from simple closed form expressions for welfare function. The EASI model can be estimated by nonlinear three-stage least squares (3SLS) regression and

[7] See, for example, (Deaton and Muelbauer, 1980a; Banks *et al.*, 1997; Lewbel, 1991; McFadden and Richter, 1991),

generalised methods of moments (GMM) and its linear approximation can be estimated by OLS regression.

The other objective is to compare welfare between the poor and non-poor groups in food consumption. A key departure of this study from the demand system of Lewbel and Pendakur (2009), known as EASI, is that this chapter assesses food consumption patterns, incorporating both traditional and structural variables in an Engel specification. Moreover, the EASI demand system does not assess welfare distinguishing poor and well-off groups across households.

On the issue of food consumption patterns, Indonesia is interesting for the following reasons. Firstly, the current population of Indonesia is about 260 million (United Nations, 2016), and it has 300 distinct ethnic and linguistic groups, cultural diversity, and heterogeneous distribution of the population among different islands. Hence, food consumption patterns in this diverse nation are expected to be different among different classes and groups.

Secondly, there has been a dramatic shift in the economic and political landscape in Indonesia after the Indonesian rupiah came under pressure in the second half of 1997 (Ahuja *et al.*, 1997; Cameron, 1999). Prices of many essential commodities surged during the first three quarters of 1998 and subsidies were withdrawn on several goods (e.g. rice, oil, and fuel). The net food consumers were severely affected by the crisis as food prices particularly that of staples, increased to a point where they are 20% more than the general price index (Thomas and Frankenberg, 2007).

Thirdly, national income accounts (BPS, 2013) show that household consumption accounts for the largest share of Indonesia's GDP (55.8%). From the year 2000 onwards, the average annual household consumption growth has been around 5%. This is faster than the 4.5 per cent average recorded in the rest of East Asia (excluding China, Japan and Taiwan). Per capita consumption growth has been at around 3% per year since 2000, lower than Indonesia's pre-crisis rate of 4% (Elias and Noone, 2011).

Fourthly, the percentage of people living in urban areas had risen to 48% in 2005 and is expected to reach 60% by 2025 (Central Bureau of Statistics, 2011; Widarjono, 2012). Along with this, 5 million more are entering the urban consumer class every year, bringing the number of urban consumers to 70 million, and consumer spending is expected to increase by 7.7% a year. Given the current scenario of urbanisation and expected rapid growth, food demand is expected to change from high carbohydrate dependency to high protein consumption in the near future, if it has not already started.

Finally, there are several important demographic changes happened over the past two decades. For example, the population density in per square kilometres has increased from 94 to 126 between 1990 and 2010 and is projected to reach 154 by 2030. Over the same period, the percentage of the population over 65 years of age has increased from 3.8% to 5% and is predicted to reach 9.2% by 2030, which implies that Indonesia is facing an ageing population. In contrast, the dependency ratio has decreased from 67% to 53% and is expected to decrease further to 46% by 2030.[8] Along with this, the urban population has increased from 30.6% to 49.9% and is projected to increase to 63.1% by 2030 (Guilmoto and Jones, 2015). The

[8] The dependency ratio is a measure showing the number of dependents, aged 0 to 14 and over the age of 65, to the total population, aged 15 to 64.

combined effect of demographic change and rapid urbanisation is likely to influence food demand patterns.

Why is the study of food consumption patterns important? For low income countries, it concerns consumers who typically spend more than 50% of their budget on food consumption, such that food consumption may significantly affect their welfare. Moreover, because household demographics have been changing dramatically, it is important to understand to what extent they may change consumption patterns by altering the consumption of staple food items by poor people. In addition, having a good understanding of how consumption responds to both traditional variables (e.g. prices and income) and structural variables (e.g. natural disaster and urbanisation) is useful for informing the way micro- and macro-economic policies, such as future food policies, are designed. Finally, as consumption diversification (in the form of 'nutrition transition') takes place in many countries over time, consumption patterns could have health implications and may be capturing consumption heterogeneity which is important to take into account of the fact that it may influence price heterogeneity across households.

This chapter makes the following contributions. First, in terms of estimation approach, this study is the first to model the food consumption patterns of Indonesian households based on the EASI demand system (Lewbel and Pendakur, 2009). The model presented here differs from well-established models, such as Deaton and Muellbauer's (1980) AIDS and Banks *et al.*'s (1997) QUAIDS, and departs slightly from Lewbel and Pendakur (2009) in that I explicitly include the structural variables as explanatory variables while the literature typically uses traditional socioeconomic variables. Second, whereas policy makers and the researcher focus on studying the welfare of the people who are living below the poverty line, I also examine the welfare of the Indonesian people who are living just above the poverty line. The people who are living just above the poverty line (say, Group 1) are at least as vulnerable as those living below the poverty line (say, Group 2). The reason is that any shock to the household (e.g. flood) could affect people in Group 1 so severely that they could end up not just below the poverty line, but also poorer than those who were already living in poverty.

Using Indonesian Family Life Survey (IFLS) data and applying the most recent demand functions, known as EASI demand system, I find three important empirical results. Firstly, consumers are highly sensitive of purchasing a good to an increase of its price. This implies that if there is a scope to substitute across a wide range of food items, consumers may have well protected against changes of relative prices. Secondly, an estimated Engel curve of food may capture shape up to third order polynomial of household's real expenditures that may have important implications for welfare of the households. Finally, I observe that physical distance from the locality of the households to the market and natural disasters alter food consumption patterns significantly.

The rest of the chapter is organized as follows. Section 2 contains the brief relevant review of literature. Section 3 describes the data and variables. Section 4 discusses the methodology and identification issues. Section 5 presents and discusses the main result. Section 6 concludes with policy implications.

2.2 Literature review

Many empirical studies in the food consumption literature are based on the application of linear expenditure system (LES), AIDS or QUAIDS models, using traditional explanatory variables such as prices and income. The landmark contribution of Deaton and Muellbauer's (1980a) AIDS takes the theoretical demand function of food to micro-data sets. The AIDS model concludes that for a given level of price, there exists a linear relationship between budget shares and the log of total expenditures. The AIDS framework is popular mainly because of the convenience of its approximation of the Engel function in a linear way. A great number of empirical studies have relied on the AIDS or QUAIDS frameworks to analyse food consumption patterns in developing and developed economies. However, neither the AIDS nor the QUAIDS framework incorporates unobserved heterogeneity and these models are also constrained by rank restrictions, which must therefore be explained through traditional explanatory variables, such as prices and income. Most of the food consumption studies in Indonesia have used either the National Socioeconomic Survey (SUSENAS) repeated cross-section surveys of households or the IFLS - panel household surveys - with large applications of SUSENAS data sets. To study food demand function using SUSENAS in Indonesia, researchers typically apply linearised approximation of the AIDS (LA/AIDS) (see for example Tell and Johnson, 1987; Tabor et al., 1989; Jensen and Manrique, 1998; Hutasuhut et al., 2002; and Moeis, 2003). For cross-country food demand studies with the application of QUAIDS, detailed explanations can be found in Pangaribowo and Tsegai (2001) for Indonesia; in Attanasio and Angelucci (2012) and Attanasio and Lechene (2014) for Mexico; in Abdulai (2002) for Switzerland; and in Kebede (2003) for Ethiopia.

The earlier work on assessing food consumption patterns in Indonesia dates back to Kakwani (1977), using the SUSENAS 1969 survey. The study estimated expenditure elasticities in several linear and nonlinear functional forms of Engel curves for eight food groups and found that expenditure elasticities diverged across different forms of Engel curves. The magnitude of the bias for the elasticity estimates depends on the various forms of Engel curves, and expenditure elasticities for cereals, cassava and vegetables were found to be inelastic. Timmer and Alderman (1979) examined the demand for rice and cassava across income groups and by rural–urban areas. They found that demand for rice was elastic, whereas demand for cassava was inelastic. Skoufias, Tiwari and Zaman (2011) examined economic crises, food prices and income elasticity of micronutrients using SUSENAS 1996 (pre-crises) and SUSENAS 1999 (post-crisis) in Indonesia. They found that income elasticities of key micronutrients were significantly higher in times of crisis than in normal times. Their identification strategy depends upon nonparametric regression and IV approach.

Other studies have focused on the Indonesian crisis from a microeconomic perspective. For instance, Fallon and Lucas (2002) documented the effects of economic shocks on household welfare; Strauss et al. (2004) examined the long-term effects of crisis and Frankenberg, et al. (1999) documented of the immediate impact of the crisis on welfare; Levinson, Berry, and Friedman (2003) investigated the likely effect of the crisis before the actual crisis.

The EASI demand system has yet to be applied widely as it was developed only recently. Olivieri (2014) applied the EASI model to Italian consumption data to estimate the Engel curve with a policy simulation exercise and found that food and fuel had a diminishing Engel curve, whereas clothing and transport had increasing Engel curve. Wood et al. (2012) estimated the EASI food demand system with the GMM technique, using Mexican household data, and showed how households substitute goods in response to price changes. Using the same data set, Magana-Lemus et al. (2013) estimated the impact of rising food prices on poverty and welfare by applying a linearised approximation of the EASI framework. Song et al. (2013) investigated the profile of Chinese household consumption to find the rank of demand systems.

This study differs from the previous studies in two particular ways. First, this study adds to the functions of EASI important control variables like disaster and distance, which are referred to as structural variables. Although none of the preceding works have considered such variables explicitly in the functions, there is a possibility that both types of variables may affect food consumption differently. Second, it is predicted that the EASI nonlinear food demand function would capture more variety of curvatures of the Indonesian household's unobserved heterogeneity in food consumption than the linear EASI model. In addition, I have considered a relatively longer term effect of crisis than other studies focusing on the impact of the financial crisis.

2.3 Data and descriptive evidence

2.3.1 Data

This study employs the IFLS household level panel data sets comprising three waves carried out during 1997, 2000, and 2007 (known as IFLS2, IFLS3, and IFLS4, respectively). The IFLS sample is representative of about 83% of the Indonesian population, covering over 30,000 individuals living in 13 of the 27 provinces in the country (RAND, 2010). The IFLS is one of the few data sets from developing countries that collect vast amounts of information at the individual and household levels including detailed modules on household consumption, health, schooling, farm and non-farm assets, health, marriage, education, migration, labour-market characteristics, household decision-making and other socioeconomic indicators. (See Strauss *et al.* 2004 for more detail on sample selection.) The IFLS also collects important information from community surveys that are linked to the household surveys of the same time period. The community-level survey conducts interviews with the community leaders and heads of the village women groups. The community surveys module includes important information from the locality, such as transportation and infrastructure, availability of electricity, water source and sanitation, agriculture and industry, history and climate, health facilities, saving and borrowing, market prices, and poverty alleviation programs.

In the IFLS, provinces were selected to maximise population representation, capture socioeconomic and cultural diversities, and minimize the cost of the surveys. Of the 13 provinces, 321 enumeration areas (EAs) were randomly selected based on the SUSENAS (National Socioeconomic Survey in Indonesia) framework and within each EA, households were chosen randomly (Strauss *et al.* 2009). Thirty households were chosen from each rural EA and 20 households from each urban EA. This study utilises three waves of IFLS: IFLS2, IFLS3, and IFLS4. The first wave, IFLS1, was excluded, as consumption expenditure modules are not suitably comparable across the waves.

In the IFLS first wave, conducted in 1993-1994, a total of 7224 households were interviewed. The second wave of IFLS (IFLS2) was conducted four years later between August 1997 and January 1998. About 93.3% of the households contacted in IFLS1 were again contacted and interviewed in IFLS2. The final sample size of IFLS2 is about 7000 households. The corresponding households sampled in 2000 and 2007 are 10,000 and 12,000, respectively.

The IFLS data set, in addition to the demographic and other crucial household level variables, contains very detailed information on consumption, which conveniently allows us to estimate the Engel curve and its modern extension. The data set contains information on the value of commodities purchased as denominated in the local currency (the Indonesian Rupiah). This structure allows us to estimate food consumption more precisely, with no imputation of values and unit prices for the goods under consideration. The IFLS2 data has been collected before the Asian Financial Crisis and IFLS3 and IFLS4 after the crises. This study also enabled me to investigate the impact of the crisis on household consumption patterns and therefore on the welfare of the people.

2.3.2 Construction of food group expenditure and price indices

2.3.2.1 Food expenditure

The IFLS collects expenditure data for 38 individual food items. The data were collected by asking the respondents in the survey whether the households had purchased any specific food items during the week preceding the interview, with the expenditure expressed in rupiah. Hence the terms food consumption, food expenditure, and food purchased are used interchangeably in our analysis. However, self-produced household food items or foods received from similar sources for consumption are excluded from the food expenditure share calculation because of the difficulty of valuing home-produced goods. Even though the survey question explicitly asked how much consumption was during the past week and the expenditure on purchasing different food items, it is possible that purchases may have taken place at a time that did not coincide with the past week. As a result, for some households the recorded expenditure might be higher than consumption and other households might record a zero purchase, even though they consumed a positive amount of food. To estimate the responsiveness to prices and other demographics of budgetary shares of food consumption, it is necessary to include both types of household in the analysis (Attanasio, Di Maro, Lechene, and Phillips, 2013). Also note that households with zero purchases and hence zero budget shares were included. This is to find the total demand response of both consumers and non-consumers to a change in an important variable such as price (see, Nimi, 2005, for illustration).

As the IFLS does not provide information on the quantity of food consumed by each household, the weekly (i.e. over the recall period) food purchase expenditure was used instead. For practical reasons, it was not possible to model food demand separately for all 38 commodities. Firstly, it is computationally difficult to track and record all the commodities in a methodical way. Secondly, the IFLS does not provide information on individual prices in the household surveys, which is critical for modelling food demand. To overcome these problems and to simplify analysis, the 38 food items were aggregated into eight food groups: staple foods, starchy foods, vegetable, meat, fish, dried foods, condiments, and other foods. The food items were aggregated based on how closely substitutable they are with each other. Aggregation into food groups facilitates observation of changing household consumption patterns of basic food and nutrient-rich food in response to change in affluence (Pangaribowo and Tsegai, 2011). Table 2.1 shows the food group's name and its composition.

Table 2.1: Composition of food groups

Group No.	Food group	Food items
1	Main staple	Rice, corn, sago/flour
2	Starchy staple	Cassava, tapioca, dried cassava, and other staples like potatoes and yams
3	Vegetable and fruit	Green vegetables (e.g. kankung [water spinach], spinach) and fruits (e.g. papaya, banana)
4	Meat/Animal products	Meat (e.g. beef, mutton, chicken), fresh milk, canned milk, powdered milk etc.
5	Fish	Fresh fish, salted fish, etc.
6	Dried foods	Tofu, noodles, other chips etc.
7	Condiments	Cooking oil, (e.g. coconut oil, peanut oil, corn oil, palm oil), salt etc.
8	Other	Bottled drinking water, granulated sugar, cigarettes etc.

Notes: (a) The separation of starchy staples from main staples is made on the basis of micronutrient contents. From the nutritionist point of view, the micronutrients from the above-ground crops (main staple) are very different from underground crops (starchy staples). (b) Not all foods items surveyed (the 38 individual food items) are included in the classification of foods because of the unavailability of corresponding market prices.

Aggregation of expenditure for each individual commodity is required to make total household expenditure. This study utilised the expenditure the households reported on food that they consume. It is assumed that utility between food and non-food items is separate and model explicitly for food items. This means that total expenditure in the utility function in an empirical model is total expenditure on food only and expenditure shares refer to particular food items in total food consumption (Attanasio, Di Maro, Lechene, and Phillips, 2013).

2.3.2.2 Price indices

The absence of price data is common in household surveys of developing countries. Researchers often use regional price indices collected either by the government statistical agencies or by the purposive design of a questionnaire. Although the IFLS does not provide individual prices on food items, the corresponding community surveys of the areas in which IFLS households are located do contain individual food prices information. These community-level food prices correspond with the food items in the household surveys during the same period and the information is contained in each EA. When quantity data in the household's consumption module is not available, price information from the community questionnaire is preferable (Deaton and Zaidi, 2002). In this study, the IFLS price data were collected from markets near the village office. Out of the 38 food items in the household surveys, 31 price items were collected in the market. These items correspond to the commodities listed in the food expenditure module. Hence, 31 out of 38 food items were used to construct eight food groups and the corresponding food group price indices.

To estimate the demand systems, I need to construct two types of price indices: a general price index to deflate the nominal expenditures in a household, and price indices of the eight food groups. I constructed Stone price indices using prices of individual food items to construct price indices for aggregate commodity groups, separately by locality (EA). The sub-group food items weights were built by summing expenditure on each good for a locality, and dividing by the total locality expenditure on that food group. These Stone price indices are used as general price index.

2.4.3 Descriptive evidence

Table 2.2 provides the summary statistics of the key variables in the estimation across three survey waves. In all survey waves, the main staple budget share is the largest among all food group consumption shares, followed by vegetables and fruits, meats and animal products, except ‘other’ food group. The main staple budget share fell from 29% in 1997 to 25% in 2000 and increased slightly to 26% in 2007. This may reflect the increase in prices of the main staples between 1997 and 2007 by 0.75 percentage point and for the same time period the increment is 1.30 percentage point.

Table 2.2: Descriptive statistics of the key variables

		1997		2000		2007	
Variable		Mean	SD	Mean	SD	Mean	SD
Budget Shares	Main staple	0.29	0.19	0.25	0.18	0.26	0.19
	Starchy staple	0.02	0.04	0.02	0.03	0.02	0.03
	Veg and fruit	0.15	0.12	0.15	0.11	0.13	0.12
	Meat/animal products	0.12	0.13	0.12	0.14	0.12	0.15
	Fish	0.09	0.09	0.1	0.1	0.09	0.09
	Dried foods	0.11	0.09	0.12	0.1	0.12	0.1
	Condiments	0.06	0.05	0.06	0.05	0.07	0.06
	Other	0.15	0.14	0.18	0.17	0.21	0.2
Log prices	Main staple	6.87	0.8	7.62	0.85	8.17	0.82
	Starchy staple	2.96	0.98	4.72	0.92	4.1	0.85
	Veg and fruit	5.85	0.85	4.98	0.75	6.05	0.92
	Meat/animal products	7.91	0.78	8.33	0.86	9.21	0.89
	Fish	6.44	1.08	6.63	1.16	7.75	1.1
	Dried foods	4.01	0.85	4.11	0.92	5.33	0.91
	Condiments	5.01	0.81	4.95	0.74	6.61	0.79
	Other	5.69	0.92	6.19	0.97	7.72	0.98
Demographics	Household size	5.22	2.44	5.27	2.73	5.4	2.98
	Age of head	48.44	31.49	46.28	35.82	45.22	33.32
	Education of head	3.15	4.35	11.11	21.46	13.94	23.74
	Head is male	0.82	0.38	0.82	0.38	0.82	0.39
	Head is employed	0.83	0.37	0.85	0.35	0.86	0.35
	Transfer	0.06	0.24	0.3	0.45	0.13	0.33
Log of real expenditure	x	12.8	0.74	13.46	0.75	14.18	0.8
Structural	Proximity to the nearest market	2.6	2.28	2.32	3.4	4.03	5.23
	Proximity to district	9.45	12.6	9.1	12.37	13.89	30.51
	Natural disaster	0.27	0.45	0.44	0.5	0.19	0.39

Notes: 1) In IFLS 1997 about 7,000 households were sampled; the corresponding sample household were about 10,000 in IFLS 2000 and were about 12,000 in IFLS 2007.

2.4 Relationship between log real food expenditure and budgetary share of staple food across years

To show the correlation between log of real food expenditure and budgetary shares of staple food, I consider the following estimating equation:

$$foodconshare_i = Constant + \mu_t + \beta_1 y_i + \beta_2 y_i^2 + \beta_3 y_i^3 + \beta_4 y_i^4 + \sum_{n=1}^{4} y_i^n * \mu_{2000} + \sum_{n=1}^{4} y_i^n * \mu_{2007} + \varepsilon_i \qquad (1)$$

The dependent variable, $foodconshare_i$, is the household level food consumption share of staple food. The regressors include the fourth power of log of real expenditure, y; μ_t is survey year dummies (dummy 2000 and dummy 2007, dummy 1997 serves as a reference category); and the last two terms in Equation 1 is the interaction of survey year dummies 2000 and 2007 with log of real expenditures.

In column (1) I estimate a model with a quartic polynomial that has a common intercept for all year. In column (2) I allow intercepts to differ 2000 and 2007. In column (3) I allow the coefficients on log real expenditure to differ in different periods. In column (4) I do the same for the 1997 and 2007 sub periods dummy.

In the pooled sample, OLS is applied (shown in Table 2.3) to show the relationship between higher order polynomials of log of nominal expenditures and food consumption shares. In all alternative specifications, support is obtained from the quintic relationship between staple food share and log of per capita food expenditures in the households. The results are also robust to alternative specifications. The results support the hypothesis of Eichengreen and Gupta (2009) in the estimation of service sector growth.

Table 2.3: Quartic relationship between log real expenditure and share of staple food (dependent variable: staple food share)

Variables	(1)	(2)	(3)	(4)
Log real expenditure	-4.112**	-5.221***	-90.102***	-20.108***
	(1.521)	(2.614)	(12.578)	(6.891)
Log real expenditure, squared	0.882**	2.433***	12.100***	11.001***
	(0.655)	(0.311)	(2.355)	(3.119)
Log real expenditure, cube	-0.092***	-0.231***	-2.355***	-2.008***
	(0.038)	(0.037)	(0.110)	(0.099)
Log real expenditure quartic	0.008***	0.008***	0.058***	0.048***
	(0.001)	(0.001)	(0.004)	(0.006)
Dummy 2000		-0.019***		
		(0.002)		
Dummy 2007		0.046***	-101.662***	-80.212**
		(0.008)	(35.101)	(30.525)
Log real expenditure *dummy2000			7.056***	
			(2.211)	
Log real expenditure squared *dummy 2000			-2.113***	
			(0.395)	
Log real expenditure cube *dummy 2000			0.190***	
			(0.054)	
Log real expenditure quartic *dummy 2000			-0.110***	
			(0.005)	
Log real expenditure*dummy 2007			48.212***	48.331***
			(10.331)	(15.301)
Log real expenditure squared *dummy 2007			-6.258***	-8.611***
			(2.215)	(2.364)
Log real expenditure cube *dummy 2007			0.674***	0.782***
			(0.130)	(0.110)
Log real expenditure *dummy 2007			-0.039***	-0.056***
			(0.008)	(0.005)
Constant	10.334	20.584**	124.218***	252.189***
	(7.888)	(6.311)	(25.311)	(28.347)
Observations	30,000	30,000	30,000	30,000

Notes: Standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1 indicates significant at 1%, 5%, and 10% levels respectively. Column (1) shows the common intercept for all the years. Column (2) shows intercepts to differ in 2000 and

2007. Column (3) allows the coefficients on log per capita income to differ in 1997, 2000, and 2007 sub-periods. Column (4) allows the coefficients to differ 1997 and 2007 sub-periods.

2.5 Separability of preferences

Following Varian (1992) and Kebede (2003), I test weak separability for eight commodity groups to the demand function $w_i = h_i(x_i, p_i)$. The consumption expenditures on the individual commodities constitute a specific food group are regressed on the log of total group expenditure, the log of the price index on the same group, and household size. In particular, the following function has been estimated to test for separability:

$$conexp_i = \alpha + \beta_j \ln(fgroup\ exp)_j + \gamma_j \ln(price\ index)_j + \delta_j hhsize + \mu_j \qquad (2)$$

where i indicates a particular food (say rice), which has the largest share of expenditure in food group j (say staple food group), $conexp_i$ indicates consumption expenditure on individual food item i, $fgroup\ exp$ is the sum of expenditure for all the individual food item constitute group, $price\ index$ is the constructed price index for food group j, and $hhsize$ indicates household size, and μ_j is the error.

If these regression equations are significant for the commodity which has the largest consumption share in a group, these food groups are said to have weakly separable from the other food groups.

It is plausible for food group expenditure and the error terms in the sub-group demand functions to be correlated. To address the concern that total expenditure could be endogenous, I perform two-stage least squares (2SLS) estimation techniques in the food group demand functions. In addition, Tobit estimates of the same sub-group demand functions were conducted, as zero observations were found in the data. Following Kebede's (2003) investigation of demand systems in rural Ethiopia, the thesis uses the total value of household assets and the size of the cultivated lands as instruments for total expenditure in the IV regression, and the predicted value of total group expenditure as an instrument for total expenditure in Tobit estimation. Table 2.4 shows the result for sub-group demand functions tests applying Equation 2. A conceptual explanation of the separability of preferences is further discussed in appendix A2.3.

Table 2.4: Tests of weak separability for sub-group food demand functions

Food groups	2007		2000		1997	
	Estimator	Estimator		Estimator		
	IV	Tobit	IV	Tobit	IV	Tobit
	(*F*-Statistics)	(Wald Chi-Square)	(*F*-Statistics)	(Chi-Square)	(*F*-Statistics)	(Chi-Square)
Main staple	40.95	200.05	25.31	250.05	29.22	242.12
Starchy staple	32.39	311.01	28.32	321.01	30.51	323.25
Veg and fruit	58.22	205.13	46.87	255.15	49.27	207.83
Meat/animal products	35.06	155.07	20.33	165.02	33.28	178.22
Fish	15.30	102.11	15.87	122.01	19.98	105.22
Dried foods	18.12	85.50	26.22	90.00	11.22	90.05
Condiments	15.54	125.31	20.80	136.21	17.11	28.33
Other	10.70	332.01	12.78	532.08	13.11	338.05

Notes: Both in IV and Tobit, the F-and Chi-square are for the significance of the food demand functions. All statistics are significant at the 1%vel.

All the *F*-statistics and Chi-square statistics for food demand regressions are significant at the 1% level. This indicates the existence of sub-group demand functions and weak separability among eight food groups. Therefore, the classification of eight food groups is empirically supported.

All the Wald chi-square statistics for the IV and IVTOBIT regressions are significant at the 1% level. This indicates the existence of sub-group demand functions with weak separability between in the eight food functions.

2.6 Methodology

2.6.1 Empirical analyses: EASI demand system

Following Lewbel and Pendakur (2009), the EASI budget shares, can be expressed in matrix notation as follows:

$$w = \sum_{r=0}^{4} b_r y^r + Cz + Dzy + \sum_{l=0}^{L} z_l A_l p + Bpy + \varepsilon \qquad (3)$$

where y is a measure of real expenditure and the coefficient of the real expenditure (y), b_r, governs the shapes of the Engel curve, which is assumed to be a fourth-order polynomial in y; p is the log-price (price index) of each food group; L expresses different demographic characteristics z. The terms C and D rule on how demographics (or taste shifters) enter into the budget-share equation through both intercept and slope terms on y. The compensated price effects are captured by both A_l and B, which allow for flexible price effects and interaction with expenditure and observable demographic characteristics. The unobserved preference heterogeneity is represented by the random utility parameter ε, which enters into the equation as a simple additive term. The derivation of EASI implicit Marshallian demand function, from the EASI budget shares have been derived, is shown in appendix A2.2.

For a well-behaved utility function, the budget-share equations must satisfy the adding up, homogeneity, and symmetry restrictions, and the Slutsky matrix of compensated price responses must be negative semi-definite.

2.6.2 Rationale for choosing EASI specification

The empirical problem is to characterise the household resources (budget) allocation to several food groups to portray the shape of the Engel curve. In this study, the household is the unit of analysis, hence the model relates the economic aggregates of average expenditure (economy wide, per household) on the different commodity groups. Accordingly, I focus on how average food group expenditure relates to total budgets per household across the economy after controlling for prices and other factors (see Blundell and Stoker, 2005, for more comprehensive treatment of heterogeneity and aggregates).[9]

We are concerned with heterogeneity in total household expenditure. By now, it is known from the Engel's law of food expenditure that food group spending varies nonlinearly with total budget size. Even some well-known early demand models have had budget shares in semi-log forms. See, for instance, the translog model of Jorgenson, Lau and Stoker (1980, 1982) and the AIDS model of Deaton and Meuellbauer (1980a, 1980b). More recent empirical studies show the importance of including nonlinear terms in budget-share functions. There is some evidence that quadratic logarithmic income terms are necessary (see, for example, Atkinson, et al, 1990; Blundell, et al., 1993; Lewbel, 1991). Empirical work on large consumer expenditure survey data sets also finds an Engel curve that is more S-shaped (See Blundell, et al. 2007). The nonlinearity here implies that the share of a particular food expenditure will be influenced by total budget size or total expenditure, as well as the degree of budget inequality across consumers. Nonlinear Engel relationships may also be obtained due to physical saturation of demand or non-homothetic consumer preferences. For instance, low-income households may have unfulfilled demand for meat or protein rich foods, so an extra income leads to larger meat

[9] As consumer preferences (i.e. heterogeneity) become more diverse, aggregation over economic variables becomes more complex and measured with error.

purchases. At higher income levels, the demand for meat may reach its maximum or consumers may choose to spend an additional food budget on a wider variety of meats.

Given that typical parametric demand models cannot incorporate this variety of Engel curve shape, the EASI demand system has a nice feature in that it allows for flexibility in the specification of budget shares through the use of polynomials of any order to capture any form of nonlinearity that might exist in total budgetary expenditures.[10]

2.6.3 Estimation and identification issues

2.6.3.1 Endogeneity of total food expenditure

When identifying the Engel curve for several food groups in Indonesia, we are interested in the estimation of a system of J food groups by treating each good separately. We can discard one food group from the specification. This group can be recovered from adding up the properties of the demand functions.[11] We place attention on the demand share function, which is up to fourth-order polynomials. In general, hypothetically, identification issues can be explained by the following equation:

$$w_j^* = \alpha_{i0} + \alpha_{i1} logx^* + \varepsilon_j \tag{4}$$

where $w_j^* = {\omega_j^*}/{x^*}$ is the budget share on the *j*th good, ω_j^* is the expenditure on the *i*th good and $x^* = \sum_{j=1}^{J} \omega_j^*$ is the total food expenditure. The parameters of interest are contained in the

[10] It is plausible that if there are 10 food groups, there could be 10 distinct shapes for the 10 food groups Engel curves. This implies that a demand system can assume any rank in the Engel specifications of demand systems.
[11] Adding up properties implies that demands must lie within the budget set: $p' f(y, p) \leq y$.

vector $\alpha_j = (\alpha_{i0}, \alpha_{i1})$. The star-indexed variables are assumed to be true unobserved variables or to be measured error-free.

One complicating factor in the theoretical model is that the EASI budget shares, *w*, appear on both right and left hand sides of Equation 5. This suggests that *w* is an endogenous. By construction, as total expenditure is often determined by the individual expenditure shares of the commodities, total consumption expenditures are not immune from the issue of endogeneity. Moreover, there may be measurement errors in the calculation of consumption expenditure, as it is common for the consumer to make irregular purchases of some commodities (Meghir and Robin, 1992). These issues imply that to obtain the correct shape of any estimated Engel curvatures, either parametric or nonparametric, it is necessary to account for endogeneity due to measurement errors of total expenditure (Briggs and Chowdhury, 2014).

In this study, other important sources of statistical endogeneity may arise from unobserved heterogeneity, measurement error, and reverse causality (simultaneity). Each of these sources of endogeneity can lead to biased estimates of the model's coefficients. Bellemare, Christopher, and David (2016) have expressed a concern that the endogeneity problem, due to unobserved heterogeneity or reverse causality, is difficult to be eliminated completely when conventional econometric methods are applied. The next section discusses each of these issues in more detail.

a. Unobserved heterogeneity: If preferences are correlated with unobserved taste-shifters in the budget-share functions, then the residuals of the food group share would be correlated with log expenditure. For example, individuals who have a stronger preference for rice are also relatively impatient and so there are higher levels of current consumption of rice, with high budget shares of rice. Unobserved heterogeneity may arise because households that are correlated with the regressors in the main budget-share functions differ from each other. The residuals of the food demand functions can be interpreted as unobservable components of taste that affect budget shares. In this model, for any taste shocks in total food consumption that are correlated with unobserved shocks to food components, total food expenditure will be endogenous. IV regression can be used to reduce bias and inconsistency of the estimates. However, if unobserved preference heterogeneity ε is correlated with any observed taste-shifters (e.g. demographics), then those variables may be excluded from the instruments list. The measure of total consumption expenditure is household-specific. Its coefficient is identified by the within-household variation in consumption expenditure, as well as by the between-households-within-community variation in food consumption at a particular time.

b. *Measurement error:* One potential correlation between residuals and log total expenditure is the presence of measurement error in the consumption module. It is highly plausible that with IFLS data, as with many other survey data, consumption is measured with error. Baltagi (2005) has pointed out that measurement errors may arise because of faulty response arising from ambiguous questions, imperfect memory, intentional distortion of responses, inappropriate informants, misreporting of responses, and interviewer effects. As with other data sets, in IFLS, consumption is measured over a seven-day recall period, in which consumption is prone to recall-based error. Ahmed, Brzozowski, and Crossley (2006) stated that measurement error is pervasive in micro-data and a key challenge to empirical work. This is evident in the proportion of zero budgetary shares in the data (roughly 25% of observations). Measurement errors in total food expenditure would affect both the total expenditure as a regressor and the corresponding construction of the dependent variables and budget shares. In this context, a suitable source of variation may arise from a variable that is associated with the cross-sectional variation of log of total expenditure, but it is unlikely to be correlated with taste variables and/or with measurement error. In the literature, income is often used as an instrument for this purpose. However, income displays huge time variations that may be smoothed out in consumption decisions (Lewbel and Pendakur, 2009). Attanasio, Battistin and Mesnard (2009) have expressed another concern when total consumption is instrumented by total household income: that if labour supply enters the utility function in a non-separable manner, income may be correlated with one of the taste shifters in the main specification of budget shares in the same manner as total consumption is correlated.

c. *Reverse causality (simultaneity):* When estimating Engel curves, another source of endogeneity may arise in total expenditure being potentially jointly determined with budget shares. If total consumption expenditure is endogenous for individual food group demand, the conditional mean estimated by parametric regression will not identify the structural Engel curve relationship. In other words, the statistical Engel curve will not recover the essential income or expenditure expansion paths of consumer preferences. Nonetheless, under the assumption of two-step budgeting and separability of preferences, food shares and total expenditure form a triangular or recursive system that is amenable to simple estimation techniques.

2.6.3.2 Estimator

The EASI budget shares specified in Equation 3 are linear in parameters and nonlinear in y, which depends on the terms $\sum_{l=0}^{L} z_l p^{'} A_l p/2$ and $p^{'} Bp/2$, which are both nonlinear. We can estimate the model either by nonlinear estimation methods or by substituting y with an observable approximation. Let us define $\tilde{y}$ as

$$\tilde{y} = x - p^{'}\overline{w} \quad (5)$$

for some set of budget shares $\overline{w}$. Then, substituting y with the $\tilde{y}$ in Equation 5, we have

$$w = \sum_{r=0}^{4} b_r \tilde{y}^r + Cz + Dz\tilde{y} + \sum_{l=0}^{L} z_l A_l p + Bp\tilde{y} + \tilde{\varepsilon} \quad (6)$$

Equation 6 is defined as the **approximate EASI with interactions model**. The approximate model is unrestricted in the sense that it can nest both the AIDS, where demand is linear in real expenditure *y*, and the QUAIDS, where demand is quadratic in real expenditure y. We can think of $\overline{w}$ as an overall sample average of budget shares across consumers or community averages. Assuming that the $\tilde{\boldsymbol{\varepsilon}}$ is uncorrelated with the explanatory variables in Equation 8 and without imposing symmetry on the A_l and B matrices, we can consistently and separately estimate the approximate EASI model by OLS or equivalently by linear seemingly unrelated regressions.

To address for endogeneity in *y*, nonlinearity in parameters and possible unknown heteroscedasticity in $\tilde{\varepsilon}$, we can use IV estimator. Let q be an *N*-vector of observable variables that are not correlated with the model error term $\tilde{\varepsilon}$. If $E(\varepsilon|x, p, z) = 0_J$, then q can take any bounded functions of *p*, *z*, and *x*. Yet, if an unobserved heterogeneity is correlated with any of the observed covariates, such as *x* or any elements of *z*, then those elements must be excluded from the potential instruments list. Hence, $E\left[\varepsilon^{'} q_n\right] = 0_J$ implies

$$E[(w - \sum_{r=0}^{4} b_r (x - p^{'}\overline{w})^r - Cz - Dz\,(x - p^{'}\overline{w}) - \sum_{l=0}^{L} z_l A_l p - Bp(x - p^{'}\overline{w}))q_n] = 0_J$$

$$\text{for } n = 1, \ldots, N \quad (7)$$

As the parameter b_r^j controls the shape of the Engel curve with the restriction of $R < J$ to avoid an arbitrarily complex Engel curve (see Lewbel, 1991), the nonlinearity of the parameters would arise from the fact that b_r multiplies (a power of) A.

2.6.4 Diversity of household food consumption patterns

We use a standard measure of diversity index to analyse food consumption diversity. To measure the diversity of household spending patterns on foods, I use the popular Gini-Simpson index, which measures the probability that two individuals drawn randomly from a given population belong to two different groups. Formally, the Gini-Simpson diversity index is defined as

$$DIX_{\mathrm{GS}} = \sum_j^J s_j\left(1 - s_j\right) = 1 - \sum_{j=1}^J s_j^2 \qquad (8)$$

where DIX_{GS} stands for the Gini-Simpson diversity index and S_j is the share of total expenditure allocated to food group j and total number of food groups is J. This diversity index is complementary to the Herfindahl-Hirschman index (H), where $H = \sum_{j=1}^J s_j^2$. This function is a convex (quadratic) and additive function of food group shares. If all resources are spent in one product category, then H equals 1, and if all groups have the same weight, H becomes zero.

2.6.5 Food consumption patterns with traditional and structural variables[12]

To demonstrate food consumption patterns, let us define some demand elasticities (budget-share semi-elasticities) with respect to log prices, *ln p*, implicit utility, *y*, and demographic

[12] Following Ray's (1999) study of food consumption patterns in urban Java households, this paper defines traditional variables (e.g. prices and income) that have a short-run impact on household welfare and structural variables (e.g. natural disaster and distance from the household to the capital) that have long-run impact on household welfare.

characteristics, z. From the two-way interaction EASI expression (7), the budget-share semi-elasticities are defined as follows:

Compensated price semi-elasticities with respect to log prices are:

$$\nabla_{p'} w(p, y, z, \varepsilon) = \sum_{l=0}^{L} z_l A_l + By \tag{9}$$

Real expenditure semi-elasticities with respect to y are:

$$\nabla_y w(p, y, z, \varepsilon) = \sum_{r=1}^{5} b_r r y^{r-1} + Dz + Bp \tag{10}$$

Demographic budget-share semi-elasticities with respect to z are:

$$\nabla_{zl} w(p, y, z, \varepsilon) = c_l + d_l y + A_l p \tag{11}$$

The EASI with no interactions including structural variables can be expressed as:

$$w = \sum_{r=0}^{5} b_r y^r + Cz + Ap + \gamma s + \varepsilon \tag{12}$$

where s indicates structural variables.

From Equation 12, several elasticities can be defined that are stated as follows:

Compensated price semi-elasticities with respect to log prices are:

$$\nabla_{p'} w(p, y, z, \varepsilon) = \sum_{l=0}^{L} A_l \tag{13}$$

Real expenditure semi-elasticities with respect to y are:

$$\nabla_y w(p, y, z, \varepsilon) = \sum_{r=1}^{5} b_r r y^{r-1} \tag{14}$$

Demographic budget-share semi-elasticities with respect to z are:

$$\nabla_{zl} w(p, y, z, \varepsilon) = c_l \tag{15}$$

In this paper, I have used both parametric EASI model with interactions, specified by Equation 7, and EASI Model with no interactions, specified by Equation 13. I have estimated EASI with no interactions model to make estimation and interpretation simpler. The empirical analysis uses annualised food group shares of $J = 8$ categories. First, I applied the OLS method and then the conventional IV method, assuming that per capita food expenditure in the model is endogenous. Then I applied the 3SLS method to estimate the EASI equation specified in Equation 7. The Equation 13 has been estimated under the similar approximation assumption made in the specification for Equation 7 to analyse food consumption patterns across eight food groups. Lewbel and Pendakur (2009) noted that 3SLS is asymptotically efficient under the assumption of identically and independently distributed error terms. They also found that 3SLS estimation remains consistent, whether the error term is normal or homoscedastic.

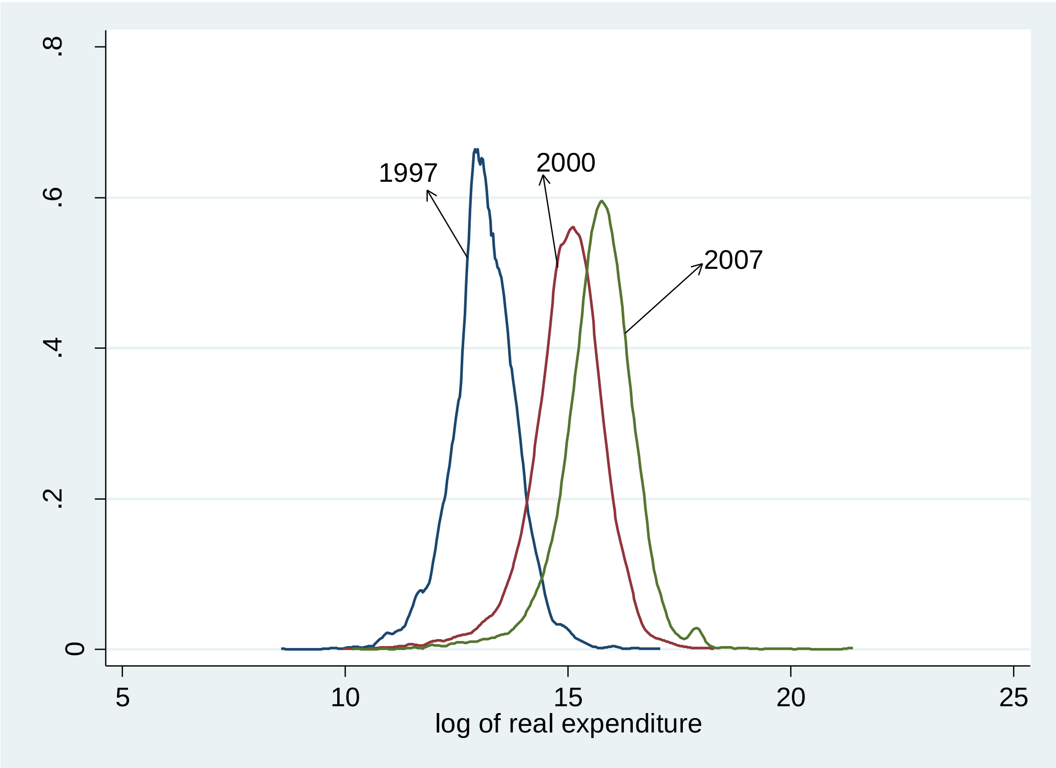
2.7 Results and discussion

2.7.1 Household food expenditure distribution across the survey years

It is possible to shed some light on the magnitude of the financial crisis in terms of household expenditure patterns. We begin with the expenditure patterns within the household before (1997) and after the crisis (2000 and 2007). In 1997 the average monthly expenditure was about Rp 1 million. Between 1997 and 2000, the mean total household expenditure declined by 10% and by 2007 had increased by 5%. A similar result is found in the changes in per capita monthly expenditures. The mean expenditure difference between 1997 and 2000 was very large and quite significant (declining 15% between 1997 and 2000).

Figure 2.1 illustrates the structure of the distribution of the log of per capita household expenditure.[13] Although the spread of the distribution is relatively stable across the survey years, the centre of the food expenditure distribution has shifted substantially. It indicates that both poor and better-offs households were affected by the crisis.

Figure 2.1. Distribution of log of real expenditure at the household level



2.7.2 Predicted shapes of the Engel curves

One of the aims of this thesis is to estimate Engel curve at different points of time. This would reflect the potential changes in budgetary allocations by the households over time. Three panels show the impact (upper panel a, middle panel b, and the lower panel c) for five main food

[13] This is a nonparametric estimate of the density of log of per capita food expenditure, based on Epanechikov kernel with an 8% bandwidth.

groups: staple food, vegetable and fruit, meat and/or animal products, fish, and condiments. These figures represent the predicted budget shares with 95 percent confidence intervals after the parametric 3SLS EASI Engel curve estimation, having up to third-order polynomial in real expenditures (rank-3 food demand system).

For instance, the estimated staple food Engel curve shown in Figure 2.2 is approximately linear in 1997 and quadratic in 2000 and in 2007. This may be the result of household budget re-allocations in response to changes in financial conditions. One empirically contrasting scenario can be observed in Figure 2.2 (b): the share of staple food is a decreasing function of log of real expenditures, while both in Figures 2.2 (a) and 2.2 (c), the staple food share is an increasing function of real expenditure. One plausible explanation is that goods may be categorised into necessary, normal or luxury, depending on the level of expenditures that the households have at different time periods (Philips, 1983). An Engel curve's linearity in the lower level of expenditure implies that average budget shares of staple foods increase with the level of expenditure, whereas expenditure elasticities decrease and tend to 1. The vegetable and fruit Engel curve is a decreasing function of log of real expenditure and it is approximately linear in 1997 and 2000, and quadratic in 2007. Meat Engel curve is approximately linear in 1997, and quadratic in both 2000 and 2007 and the curve is an increasing function of log of real expenditure across the years. This indicates that households tend to consume more protein-rich foods as they become wealthier. The Fish Engel curves are approximately linear and a decreasing function of log of real for all of the survey years. Finally, the condiments Engel curves show irregular patterns and can be approximated by third-order polynomials for all the years.

Figure 2.2. Staple food Engel curves in 1997, 2000, and 2007

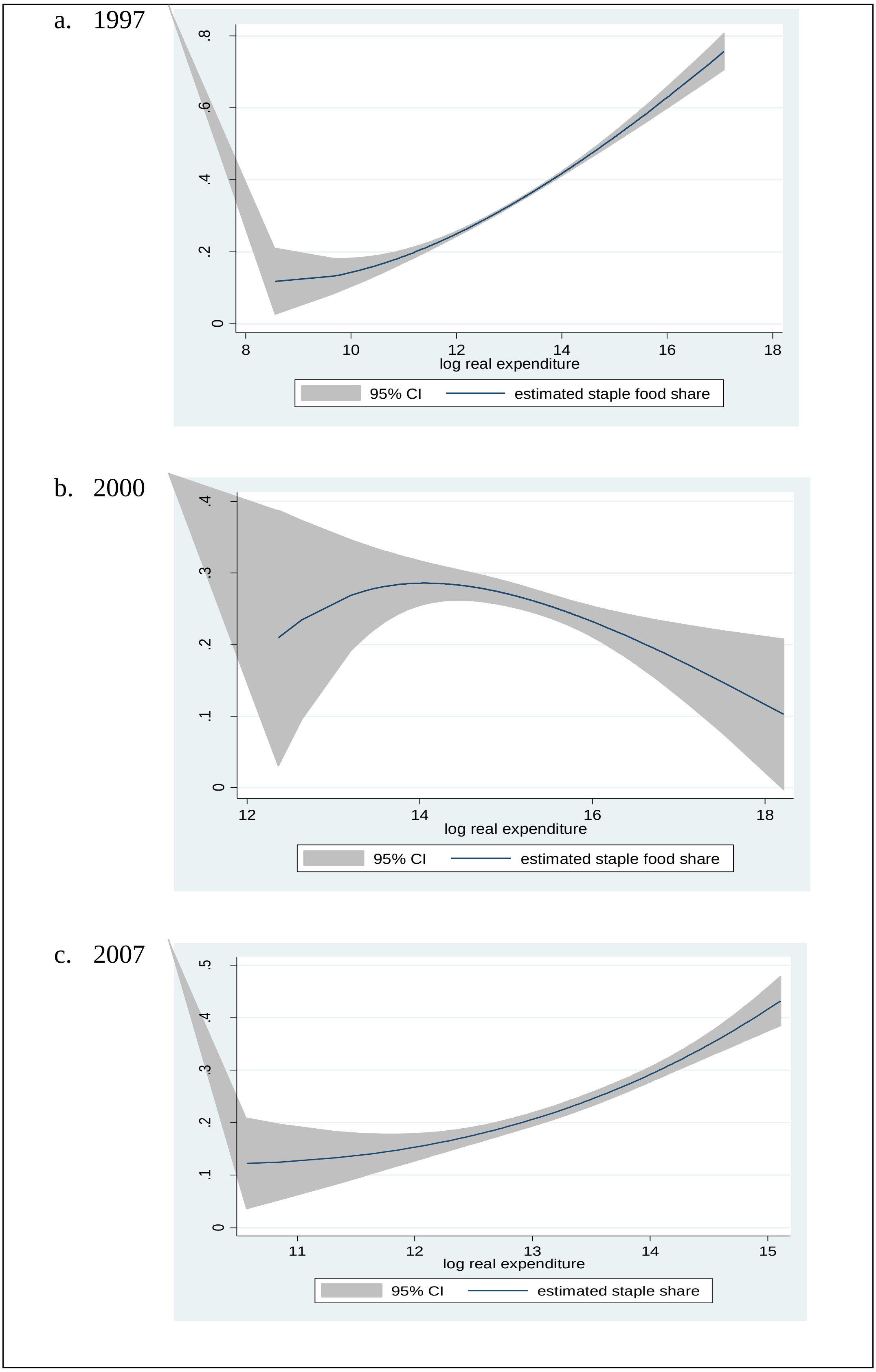


Figure 2.3. Vegetable and fruit Engel curves in 1997, 2000, and 2007

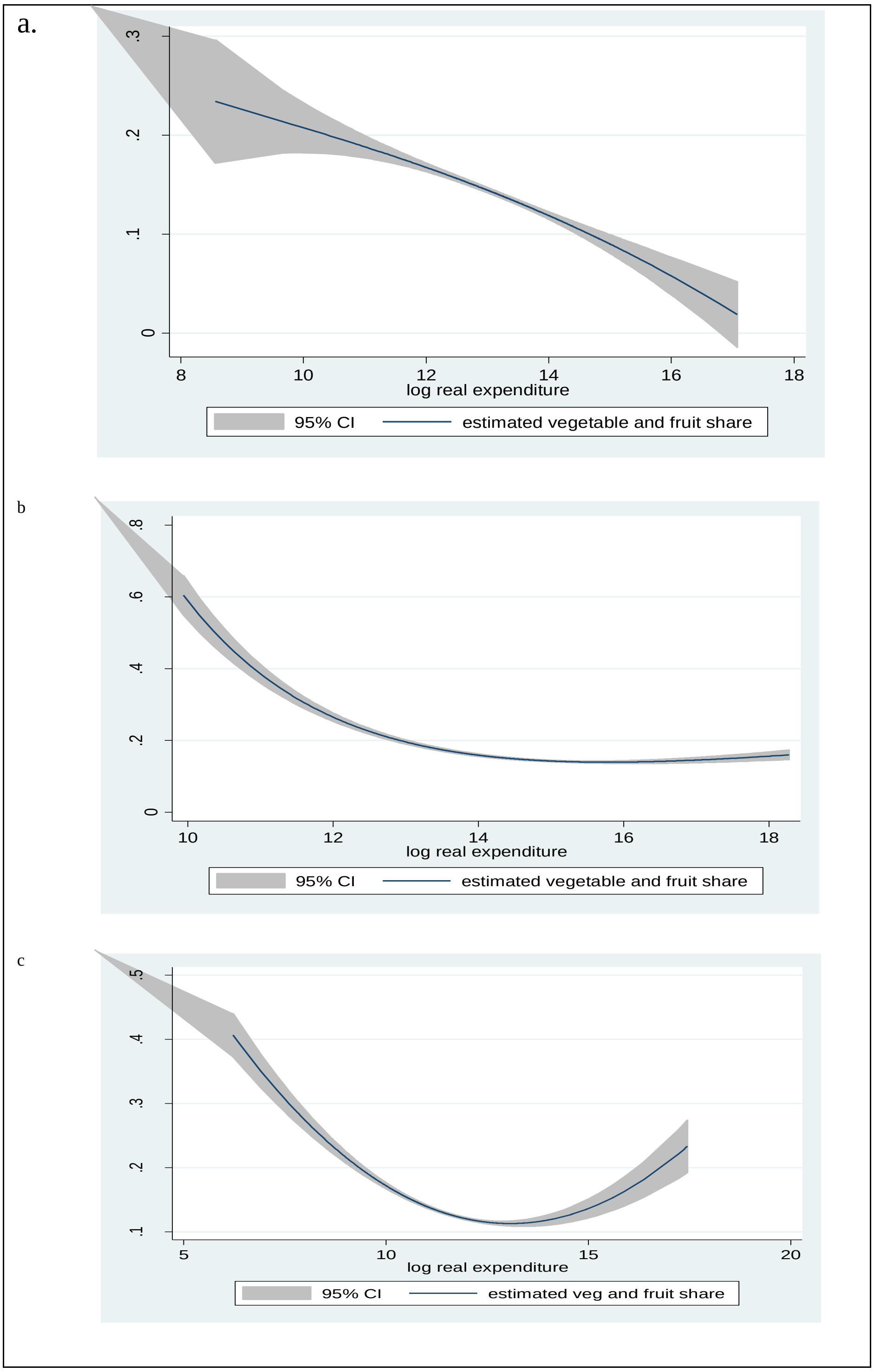


Figure 2.4: Meat Engel curves in 1997, 2000, and 2007

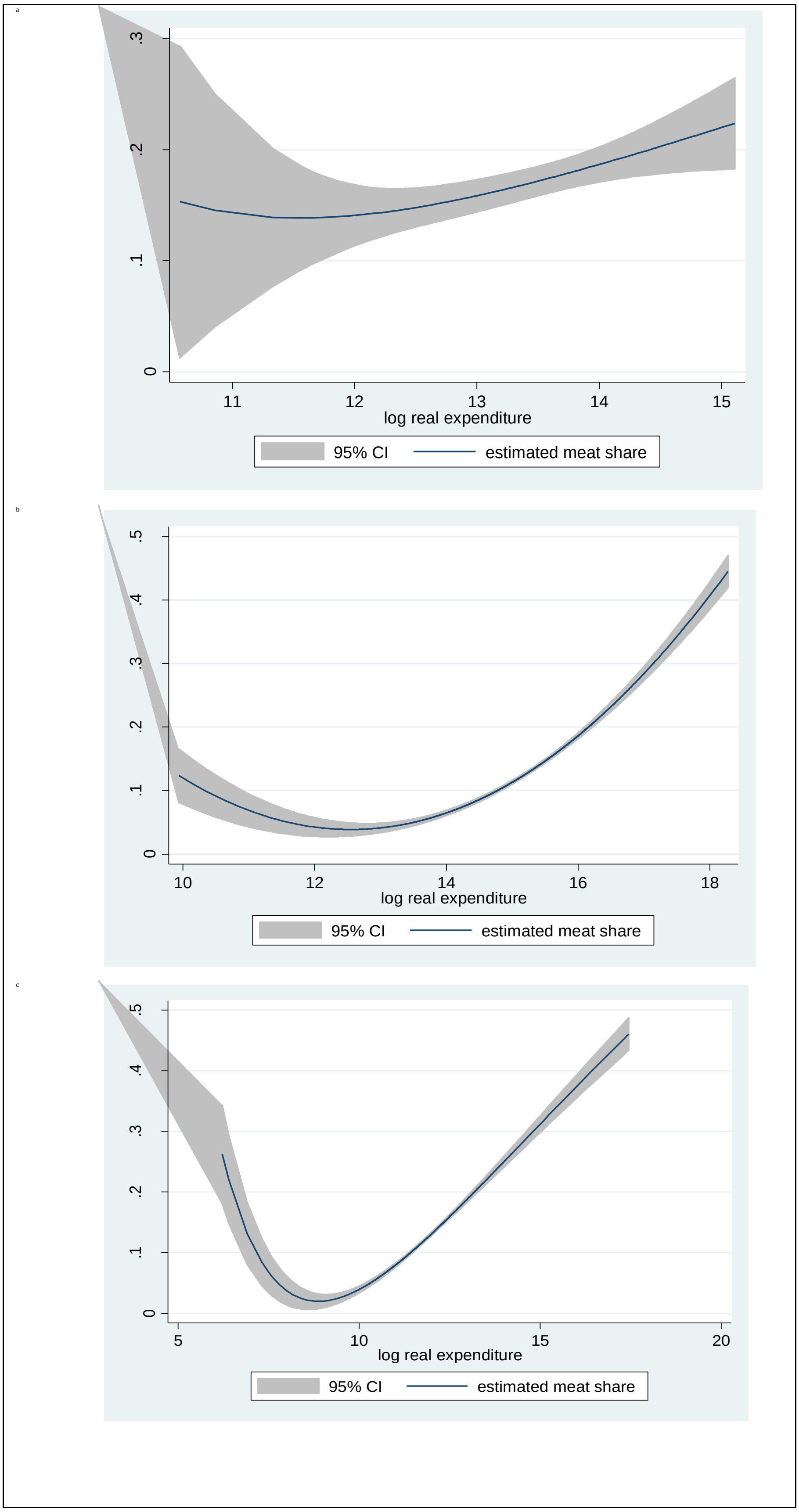


Figure 2.5: Fish Engel curves in 1997, 2000, and 2007

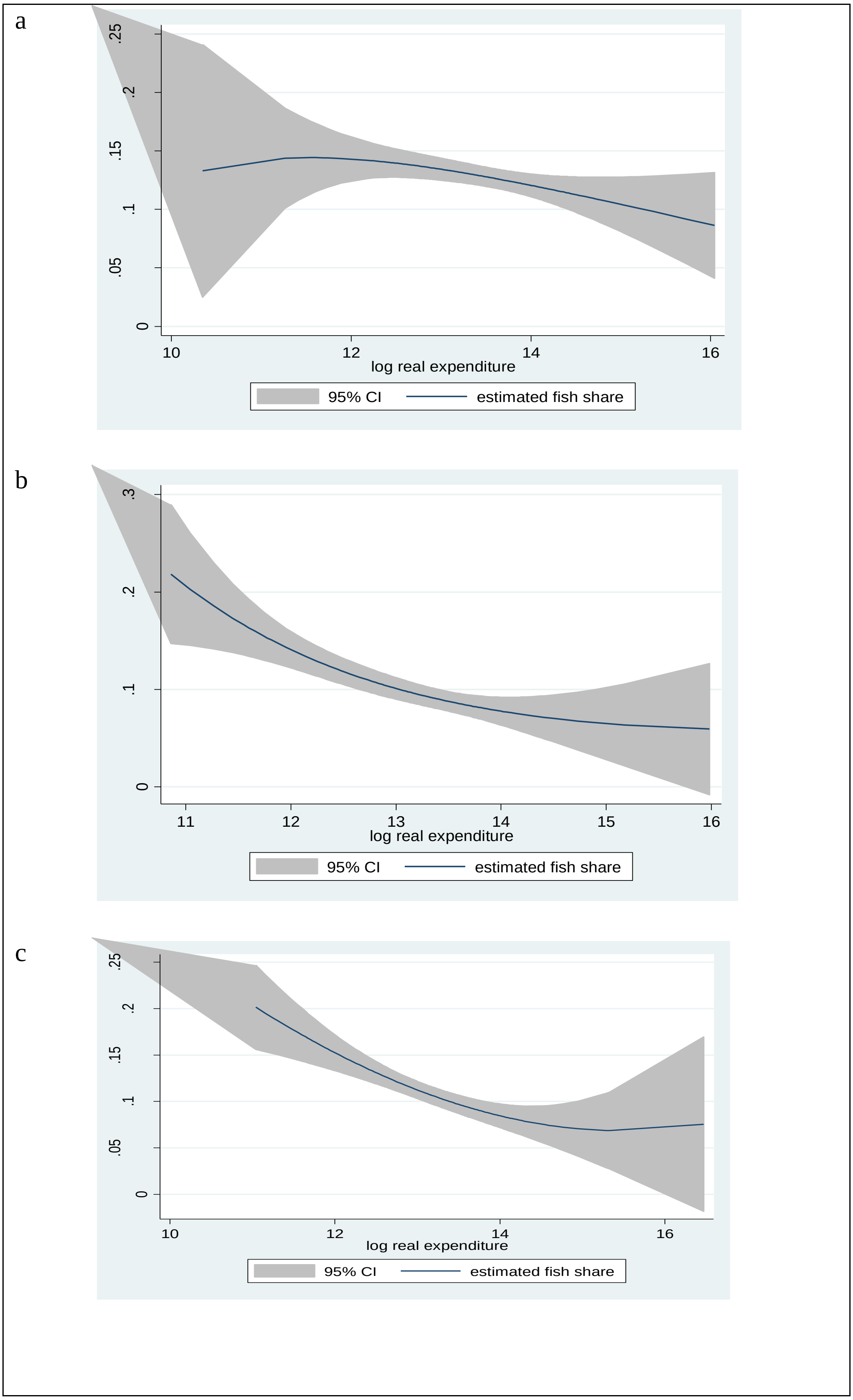
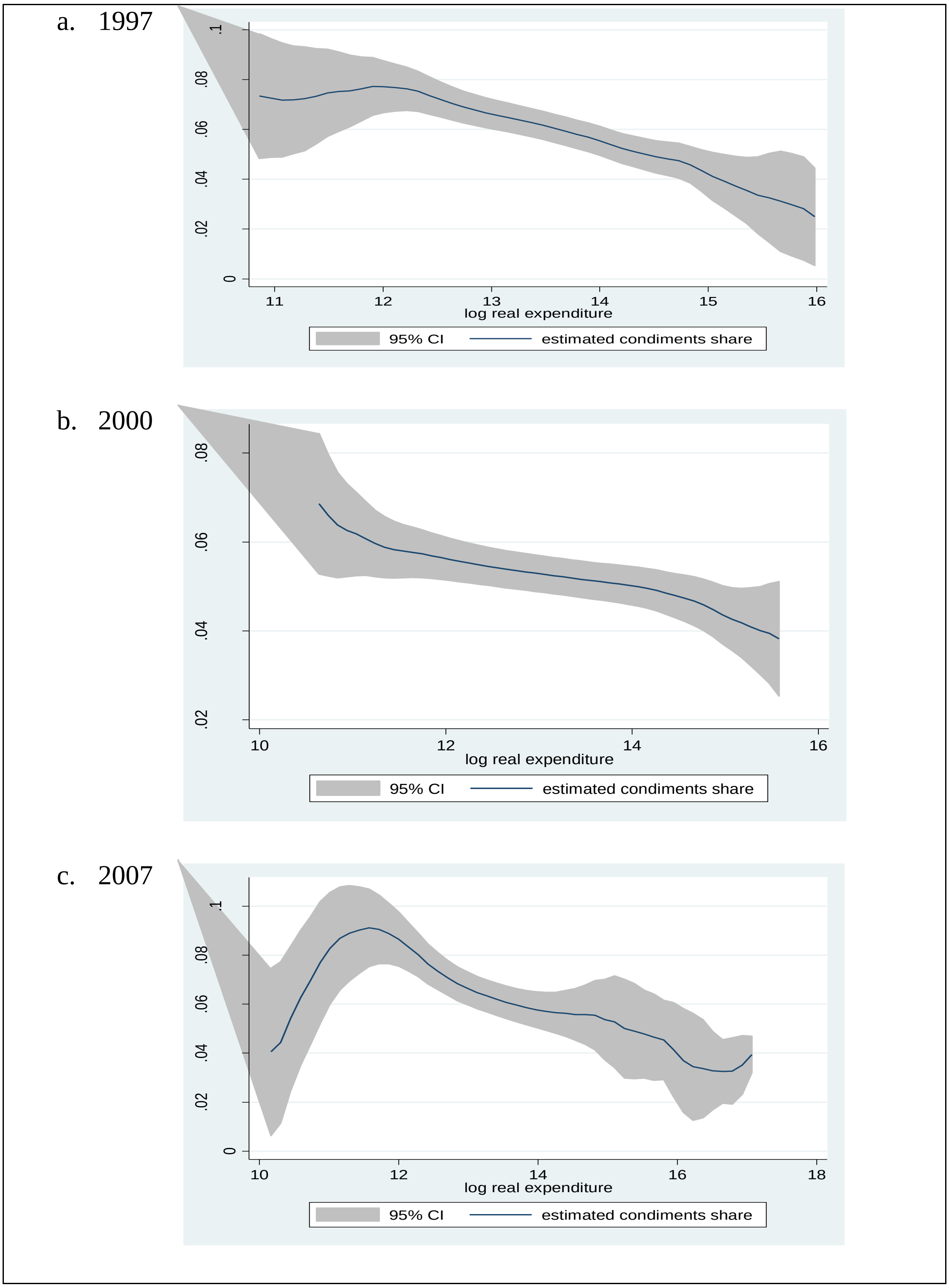


Figure 2.6. Condiments Engel curves in 1997, 2000, and 2007



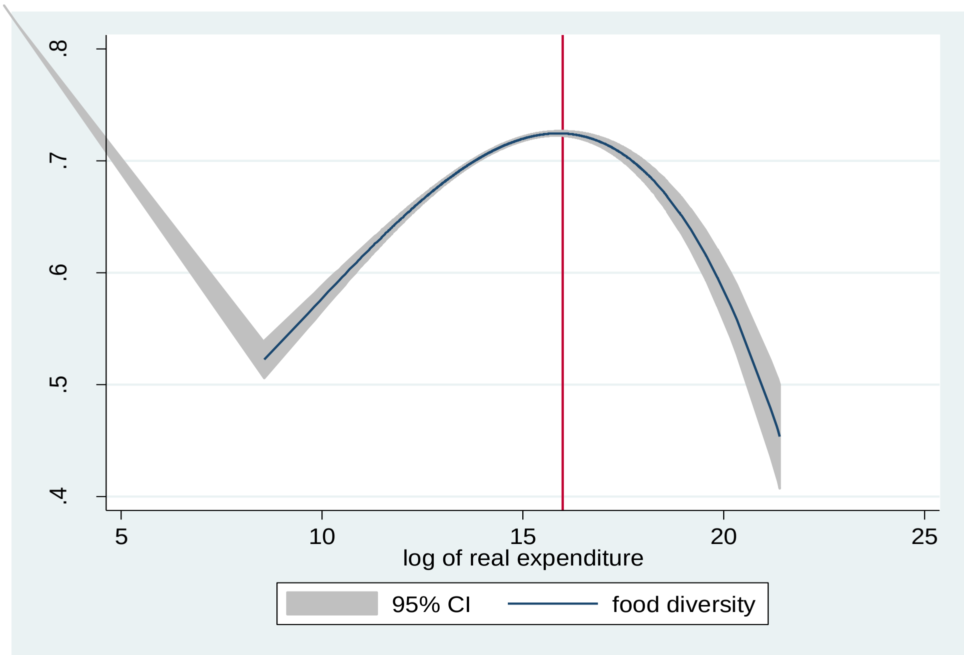
The diagrams in Figures 2.2 to 2.6 indicate that the Engel curve could explain up to third order polynomials. Inspection of the figures endorses the initial presumption that staple Engel curve and the vegetable and fruit Engel curve are best described by a quadratic curve, the meat Engel curve by a cubic curve, and the fish Engel curve by an approximately linear. Though not shown here, the estimated predicted Engel curves for dried foods and condiments appear quartic. ‘Vegetables and fruits consumption’ and ‘meat consumption’ are perceived to be substitutable goods over time.

2.7.3 Food consumption diversification

Figure 2.6 illustrates food consumption diversification across various households at different expenditure levels using the pooled sample. The sample covers all the three waves of the survey, approximating 28,000 nationally representative households.

The locally weighted scatterplot smoothing (LOWESS) regression shown in Figure 2.6 demonstrates that the diversity in food expenditure is very low for poor households and generally increases with total expenditure. A maximum diversity is achieved when household’s total expenditure reaches 15.8 log-point. It is difficult to estimate the exact location of the maxima; however, the stable plateau runs from about log of total expenditure 16.5 per year. An important observation here is that the curve tends to decline for high-expenditure households and this decline starts from about log of per capita expenditure of 15.8. The curve provides the evidence of nonlinearities of food Engel curves. The flattening of the curve (not shown here) implies that after a certain level of expenditure the households do not seem to diversify their spending across different levels.

Figure 2.7: Food consumption diversification: Gini-Simpson index in a pooled sample



2.7.4 Parameters estimates of food consumption

Equation 6, which refers to the EASI with interactions budget shares model and Equation 12, which refers to the EASI with no interactions model, is the next system of equations to be estimated. I divide the results section into five parts, in which I present:

a) Compensated price effects in EASI interactions model specified by Equation 6

b) OLS estimates of food consumption in EASI Equation 12

c) IV estimates to deal with potential endogeneity of food consumption expenditures in EASI Equation 12

d) EASI real expenditure semi-elasticities estimates in the pooled sample in EASI Equation 12

e) EASI compensated price semi-elasticities for the pooled sample in EASI Equation 12

f) Welfare comparisons from poor and non-poor food consumptions.

A) Estimates of the compensated price effects in EASI interactions model

Table 2.5 presents various measures of the matrix of price effects estimated from the EASI budget share Equation 6. The first column of estimates in Table 2.7 comprises of the estimated own-price elements of **B**, which show the magnitudes of the interaction between own-prices and with log real total expenditures. These parameters allow us to evaluate whether or not compensated semi-elasticities are the same for rich and poor households. The second and third column show the **Slutsky** matrix and compensated **expenditure** elasticity, respectively. The fourth to twelfth columns of Table 2.5 exhibits the budget share semi-elasticities with respect to own-prices.

Table 2.5: Compensated price effects in EASI with interactions model

	Own-price	Own-price	Own-price	Budget share semi-elasticities							
	B element	**Slutsky** terms	**Expenditure** Elasticity	Main staple	Starchy staple	Veg and fruit	Meat	Fish	Dried	Condiments	Other
Main staple	0.052	-0.159	-0.084	-0.043							
	(0.012)	(0.031)	(0.217)	(0.026)							
Starchy Staple	-0.007	-0.130	-0.116	0.218	-0.018						
	(0.003)	(0.028)	(0.218)	(0.015)	(0.021)						
Veg and fruit	0.076	-0.032	0.367	0.115	0.015	-0.032					
	(0.028)	(0.010)	(0.431)	(0.011)	(0.011)	(0.005)					
Meat	0.080	-0.042	0.287	-0.241	0.042	-0.54	0.063				
	(0.023)	(0.021)	(0.378)	(0.012)	(0.002)	(0.003)	(0.028)				
Fish	0.005	-0.067	0.073	0.447	0.040	0.035	0.094	-0.021			
	(0.002)	(0.021)	(0.025)	(0.130)	(0.021)	(0.810)	(0.007)	(0.001)			
Dried	0.090	-0.115	0.087	0.267	-0.025	-0.052	-0.028	0.082	-0.036		
	(0.080)	(0.012)	(0.072)	(0.031)	(0.002)	(0.003)	(0.007)	(0.021)	(0.410)		
Condiments	0.003	-0.071	0.774	0.364	0.061	0.068	0.071	0.041	0.871	-0.031	
	(0.001)	(0.029)	(0.124)	(0.031)	(0.028)	(0.021)	(0.052)	(0.031)	(0.054)	(0.003)	
Other	-0.050	-0.038	-2.267	0.311	0.724	-0.030	0.061	0.064	0.154	0.044	-0.011
	(0.020)	(0.033)	(0.732)	(0.009)	(0.057)	(0.005)	(0.009)	(0.021)	(0.012)	(0.031)	(0.210)

Notes: Standard errors are shown in the parenthesis. The compensated budget share semi-elasticities with respect to prices are given by the matrix $\Gamma = \sum_{l=0}^{L} A_l z_l + By$. The own-price expenditure elasticity with respect to prices are given by $W^{-1}(\Gamma + ww')$, where $W = diag(w)$. The own-price Slutsky terms can be approximated from the compensated semi-elasticity matrix and given by $S = \Gamma + ww' - W$. The first column, own-price B element, is the derivative of Equation 7 with respect to interaction variable py.

The first column of Table 2.5 shows the magnitudes of the interaction between own-prices and with log of real expenditures. The estimated coefficient of the meat own-price semi-elasticity on y is 0.080, which is highly statistically significant. Consider the comparison between the meat own-price compensated semi-elasticity for a reference household at the fifth percentile of expenditure (=-0.85) versus that for a person at the ninety-fifth percentile of expenditure (=0.48). The median expenditure of a reference household is -0.062. At the fifth-percentile, its value is -0.062-(0.85*0.080) =-0.006, and is insignificantly different from zero. On the other hand, its value at the ninety-fifth percentile is -0.062+ (0.48*0.080) =0.025, and appears highly statistically significant. The corresponding own-price Slutsky terms are -0.097 at the fifth percentile and -0.014 at the ninety-fifth percentile, and both are negative and statistically significantly. These results suggest that poor households substitute much more than rich households in the face of an increase in the price of meat. The second column of Table 2.5 exhibits the values of the own-price Slutsky terms that are all negative and most of the Slutsky terms are statistically significant. The third column shows the own-price expenditure elasticities. For example, the compensated expenditure elasticity of fish with respect to price is 0.073, which is positive and statistically significant. Finally, the columns fourth through eleventh of Table 2.5 own-price effects. Several of the own-price effects are highly statistically significant. The own-price semi-elasticity for main staple budget share is 0.052, which implies that a main staple price increase of 10 percent is associated with a 0.52 points higher budget share when expenditure is increased to equate utility with the initial living condition.

B) Estimates of EASI no interactions model

2.7.4.1 OLS estimates of food consumption parameters

This section begins by presenting the OLS estimates of EASI food demand function. Eight separate budget-share equations are estimated as a function of log households' real expenditures and other demographic and economic variables (household size and proportions of various demographic groups, age, education, employment and gender of the household head, transfer and disaster dummies, distance to the nearest market from the household location, urban dummy, and survey years dummies).[14] The definitions of variables used in various estimations are shown in appendix A2.1.

[14] Distance to the nearest market is included to get a rough measure of the households' accessibility to a variety of goods and services. The urban dummy is included to measure the urban preferences and tastes of urban consumers, and the availability of goods (particularly manufactured goods) on the demand for foods. Dummies of survey rounds are included to control for possible seasonal effects. Dummies for female-headed households and education of the household head are included as potential taste-shifters.

Table 2.6: OLS estimates of food consumption patterns in pooled sample

Specifications	Main staple (1)	Starchy staple (2)	Veg & fruit (3)	Meat (4)	Fish (5)	Food diversity (6)
Log of real expenditure	0.028***	-0.006**	-0.028***	0.035***	0.008***	0.018***
	(0.008)	(0.006)	(0.004)	(0.003)	(0.004)	(0.005)
Household size	0.007***	-0.000	-0.005***	-0.004***	-0.008***	0.008***
	(0.003)	(0.000)	(0.002)	(0.002)	(0.000)	(0.003)
pro_m0_4	-0.029	0.005	-0.046**	0.108***	0.027	0.108***
	(0.029)	(0.002)	(0.020)	(0.025)	(0.018)	(0.023)
pro_m5_9	-0.005	0.000	-0.035	0.017	0.024	0.095***
	(0.060)	(0.006)	(0.020)	(0.024)	(0.015)	(0.022)
pro_m10_14	0.049	0.002	-0.050**	-0.047*	0.011	0.022
	(0.033)	(0.003)	(0.023)	(0.029)	(0.015)	(0.028)
pro_m15_55	-0.082***	-0.005	-0.032**	-0.034*	-0.012	-0.047**
	(0.037)	(0.007)	(0.018)	(0.019)	(0.012)	(0.022)
pro_m55_70	0.317	0.008	-0.020	-0.053***	-0.016	-0.042*
	(0.026)	(0.013)	(0.019)	(0.018)	(0.017)	(0.025)
pro_f0_4	-0.048	0.008	-0.042**	0.074***	0.030*	0.109***
	(0.030)	(0.007)	(0.027)	(0.021)	(0.017)	(0.023)
pro_f5_9	-0.025	0.006	-0.028	-0.002	0.026	0.105***
	(0.015)	(0.004)	(0.021)	(0.023)	(0.018)	(0.012)
pro_f10_14	0.038	0.007	-0.047**	-0.058**	0.030*	0.054**
	(0.037)	(0.006)	(0.017)	(0.024)	(0.018)	(0.024)

Specifications	Main staple (1)	Starchy staple (2)	Veg & fruit (3)	Meat (4)	Fish (5)	Food diversity (6)
pro_f15_55	-0.069**	0.002	0.008	0.010	0.024*	0.080***
	(0.034)	(0.006)	(0.016)	(0.017)	(0.014)	(0.020)
pro_f55_70	-0.060	0.027	-0.021	-0.029	0.009	0.043*
	(0.080)	(0.018)	(0.019)	(0.021)	(0.016)	(0.021)
Age of head	0.009***	0.000***	-0.000	0.000***	0.000**	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Education of head	-0.006***	-0.000*	0.007***	0.008***	-0.000	0.004***
	(0.002)	(0.000)	(0.002)	(0.003)	(0.000)	(0.002)
Head is employed	0.014*	-0.006	-0.007	-0.009*	-0.000	0.027***
	(0.008)	(0.006)	(0.002)	(0.004)	(0.003)	(0.004)
Head is male	-0.018***	-0.006	-0.017***	-0.019*	0.007*	0.028***
	(0.007)	(0.004)	(0.006)	(0.015)	(0.004)	(0.006)
Transfer	0.054***	0.004	-0.016***	-0.020***	-0.010***	0.008
	(0.007)	(0.001)	(0.005)	(0.005)	(0.004)	(0.014)
Disaster	-0.025***	0.003***	-0.019***	-0.016***	-0.012***	-0.008
	(0.005)	(0.002)	(0.005)	(0.004)	(0.004)	(0.002)
Distance to the market	-0.007***	-0.000***	0.000	-0.003***	0.004***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Urban	-0.070***	-0.005***	0.024***	0.045***	-0.022***	0.015***
	(0.006)	(0.002)	(0.001)	(0.002)	(0.005)	(0.004)
Year 2000	-0.092***	0.028*	0.045**	-0.068***	-0.024***	-0.052***
	(0.029)	(0.015)	(0.017)	(0.015)	(0.009)	(0.017)

Specifications	Main staple (1)	Starchy staple (2)	Veg & fruit (3)	Meat (4)	Fish (5)	Food diversity (6)
Year 2007	-0.120***	0.007***	0.045***	-0.092***	-0.016***	-0.068***
	(0.018)	(0.004)	(0.009)	(0.008)	(0.003)	(0.012)
constant	0.097	0.046***	0.567***	-0.337***	-0.040*	0.482***
	(0.036)	(0.018)	(0.040)	(0.035)	(0.022)	(0.049)
Observations	9000	9000	9000	9000	9000	9000
R-squared	0.224	0.035	0.060	0.117	0.070	0.081

Notes: Robust standard errors appear in brackets. Asterisks denote significance: *** p < 0.01, ** p < 0.05, * p < 0.1. Columns (1)-(5) show OLS estimates for five food groups and column (6) shows OLS estimate of food diversity. In all specifications, the control variable are: log of per capita food expenditure at the household level, household size, proportions of households by age-sex groups (for instance, pro_m0_4 shows the ratio of the number of boys in the age group 0-4 to the household size and pro_f0_4 shows the ratio of the number of girls to the household size in the same age group and other age-sex groups variable has been defined in the similar ways and age group 71-90 are an omitted category for both males and females), age of household head in years, highest grade attained by the head of the household, dummy variables for whether the household head is male and employed, dummy for whether household received any cash or kind transfer from the government, a dummy variable for whether household experienced any natural disasters, a continuous variable of the distance from the household to the nearest market in kilometres, a dummy variable for whether household is an urban area, and dummy variables of the survey year 2000 and 2007 with dummy 1997 is an omitted category.

Some of the key coefficients presented in Table 2.6 can be interpreted as follows. The log of real expenditure is positively and significantly associated with the main staple, meat, and fish budget shares, and negatively and significantly associated with the starchy staple and vegetables budget shares. The positive coefficient on the main staple share (0.028) indicates that even though households are becoming more affluent in terms of total household resources, measured by total household expenditure, they still consume more rice, which is the largest proportion of the main staple. Specifically, a 1% increase in the log of real expenditure is on average associated with a 2.8% increase in consumption of main staple. The coefficient of the urban dummy is highly statistically significant, which implies that rural people consume more rice than urban people. This is consistent with the fact that people working in agricultural production require more carbohydrates and energy, because there are household resource constraints on purchasing high-value commodities such as meat and fish. Household size is positively and significantly associated with the demand for staple foods, and negatively and significantly associated with the demand for vegetables and fruits, meat, and fish. This is because larger households have to meet a minimum caloric requirement, which is usually realised through the consumption of basic staple foods such as rice. Household size is positively and significantly associated with this main staple. Consumption of high-value goods such as vegetables and fruit, meat and fish, is partially affected by the number of members in the households. As the family size increases, the family is less likely to consume high-value foods mainly because of the pressure on budget. Except for starchy foods, natural disasters reduce all food consumption significantly. The distance to the nearest market has a negative and highly statistically significant relationship to the consumption of the main staple and meat, and positively related to fish consumption. As expected, the education of the household head is positively and significantly associated with the consumption of vegetables and fruit, meat, and food diversity, and negatively associated with the main staple.

2.7.4.2 Instrumental variable estimates of food consumption parameters

It is important to bear in mind the IFLS is a seven-day recall survey of food purchases by all household members. In the IFLS survey, the consumption module carefully collects information on food consumption from both market purchases and self-produced goods. In addition, the survey also captures in kind or cash transfers and any purchases of rice from the poor program or market food operations. As the identical questionnaire is used in each survey year, these factors are less likely to cause bias estimates of food consumption parameters.

Nevertheless, it is possible that correlated measurement errors in total food consumption could cause the food expenditure elasticity estimates to be biased. By construction, the food group’s share is based on the total household’s allocation towards food expenditure. In the case of correlated measurement errors in the response and explanatory variables, it is not clear whether the upward bias from correlated measurement errors is greater than the standard downward attenuation bias from the measurement error in total food consumption. Hence, the direction of net bias from the expenditure elasticity estimates would depend on the relative magnitude of the correlation between the measurement errors and the variance of measurement error in household food consumption. To address the issue of such endogeneity of total food consumption expenditures, I use the IV method. Strauss et al (2004) and Skoufias, Tiwari, and Zaman (2011) used non-land productive assets and its squares as an instrument for ascertaining per capita household expenditure in the estimation of household welfare by Strauss (2004) and Skoufias et al (2011), in the Indonesian context, has used an index of assets in the estimation of income elasticity of micronutrients. Olivia and Gibson (2012) used log of wages of the household head from an IFLS sample to estimate an Engel curve in Indonesia. Following this research, I use the log of household assets and the log of wages of the household head as IVs for log of real expenditure. The relevance of the instrument is tested by the first-stage regression. In all food groups, the chosen instruments are highly correlated with endogenous food expenditure. For all food consumption shares, the *F*-statistic of excluded instruments is highly significant as its value is considerably greater than 10.[15] The Sargan test of over-identifying restrictions clearly indicates that the null hypothesis that over-identifying restrictions are valid cannot be rejected. This implies that identifying restrictions does not explain the unexplained part of household consumption share equations, strongly suggesting the instrument's validity. Table 2.6 presents IV estimation of food shares.[16] The IV estimates support the OLS estimation for all food consumption groups. Although the IV estimates are generally highly statistically significant, they appear to be higher than OLS. This suggests that upward bias from the correlated errors may be lower than the standard downward attenuation bias in the OLS estimates. The first-stage results corresponding to Table 2.7 is shown in Table A2.6 in the appendix.

Table 2.7: Instrumental variable (IV) estimates of food consumption patterns in pooled sample

	Staple (1)	Starchy (2)	Veg (3)	Meat (4)	Fish (5)	Food Diversity (6)
Log of real expenditure	-0.162***	-0.003	0.047***	0.177***	0.011**	0.121***
	(0.014)	(0.002)	(0.007)	(0.010)	(0.005)	(0.009)
Household size	0.017***	0.000	-0.006***	-0.014***	0.001**	-0.005***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
pro_m0_4	0.168***	0.003	-0.110***	-0.064**	0.015	-0.013
	(0.035)	(0.005)	(0.020)	(0.028)	(0.015)	(0.023)
pro_m5_9	0.180***	0.001	-0.084***	-0.134***	0.013	-0.017
	(0.034)	(0.005)	(0.020)	(0.026)	(0.015)	(0.023)
pro_m10_14	0.193***	0.001	-0.090***	-0.166***	0.006	-0.061**
	(0.038)	(0.005)	(0.020)	(0.031)	(0.015)	(0.025)
pro_m15_55	0.018	-0.003	-0.067***	-0.090***	-0.013	-0.084***
	(0.027)	(0.004)	(0.017)	(0.020)	(0.011)	(0.019)
pro_m55_over	0.044	0.005	-0.028	-0.056**	-0.015	-0.031
	(0.031)	(0.008)	(0.018)	(0.023)	(0.014)	(0.021)
pro_f0_4	0.139***	-0.003	-0.095***	-0.074***	0.017	-0.013
	(0.034)	(0.005)	(0.020)	(0.027)	(0.015)	(0.024)
pro_f5_9	0.145***	0.001	-0.071***	-0.137***	0.013	-0.009
	(0.034)	(0.005)	(0.020)	(0.026)	(0.015)	(0.024)
pro_f10_14	0.186***	-0.000	-0.090***	-0.179***	0.023	-0.043*
	(0.036)	(0.005)	(0.020)	(0.028)	(0.015)	(0.024)
pro_f15_55	0.078***	-0.002	-0.038**	-0.080***	0.021*	0.007
	(0.027)	(0.005)	(0.017)	(0.020)	(0.011)	(0.018)
pro_f55_over	0.005	-0.001	-0.032*	-0.031	0.011	0.029
	(0.031)	(0.006)	(0.018)	(0.022)	(0.013)	(0.020)
Age of head	0.002***	0.000***	-0.000***	-0.001***	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Education of head	-0.001***	0.000	0.000***	0.001***	-0.000	-0.000***

[15] The *F*-statistic is a joint test of whether all excluded instruments (the variables in *z* which are not in *x*) are significant or not.
[16] IV estimation and IV tests are conducted using Stata ivreg2 command in Stata 14 version.

	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Employment of head	0.033***	-0.001	-0.012***	-0.027***	-0.002	0.003
	(0.008)	(0.001)	(0.004)	(0.006)	(0.003)	(0.005)
Male head	0.025***	-0.001	-0.026***	-0.043***	0.004	-0.004
	(0.008)	(0.001)	(0.004)	(0.006)	(0.003)	(0.005)
Transfer	0.019***	0.002	0.000	0.004	-0.009***	0.017***
	(0.007)	(0.001)	(0.003)	(0.005)	(0.003)	(0.004)
Disaster	0.014*	0.005***	0.002	-0.014**	0.010***	-0.023***
	(0.008)	(0.001)	(0.004)	(0.006)	(0.003)	(0.005)
Distance to the market	0.004***	-0.000***	0.000	-0.003***	0.003***	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Urban	-0.051***	-0.003***	0.017***	0.037***	-0.022***	-0.012***
	(0.004)	(0.001)	(0.002)	(0.003)	(0.002)	(0.003)
Year 2000	0.222***	0.025**	-0.084***	-0.299***	-0.025**	-0.216***
	(0.033)	(0.012)	(0.019)	(0.027)	(0.011)	(0.021)
Year 2007	0.355***	0.009*	-0.140***	-0.456***	-0.024*	-0.320***
	(0.036)	(0.005)	(0.019)	(0.027)	(0.013)	(0.024)
Constant	2.171***	0.055**	-0.346***	-1.958***	-0.076	-0.796***
	(0.162)	(0.025)	(0.086)	(0.121)	(0.057)	(0.111)
Observations	10,160	10,160	10,160	10,160	10,160	10,160
F-Statistic (excluded instruments)	26.3	29.5	22.9	60.3	156.9	19.2
Prob>F	0.000	0.000	0.000	0.000	0.000	0.000
Sargan test of overid (Prob>chi2)	0.58	0.37	0.21	0.47	0.35	0.52
Instruments	A+B	A+B	A+B	A+B	A+b	A+B

Notes: Robust standard errors appear in brackets and clustered at the community level. Asterisks denote significance: *** p < 0.01, ** p < 0.05, * p < 0.1. Columns (1)-(6) show IV estimates for five main food groups and column (6) shows IV estimate of food diversity. In all specifications, the control variable are: log of real expenditure at the household level, household size, proportions of households by age-sex groups (for instance, pro_m0_4 shows the ratio of the number of boys in the age group 0-4 to the household size and pro_f0_4 shows the ratio of the number of girls to the household size in the same age group and other age-sex groups variable has been defined in the similar ways and age group 71-90 are an omitted category for both males and females), age of household head in years, highest grade attained by the head of the household, dummy variables for whether the household head is male and employed, dummy for whether household received any cash or kind transfer from the government, a dummy variable for whether household experienced any natural disasters, a continuous variable of the distance from the household to the nearest market in kilometres, a dummy variable for whether household is an urban area, and dummy variables of the survey year 2000 and 2007 with dummy 1997 is an omitted category. The instrument A is the log of household's non-land productive assets and the instruments B is log of wages of the household head.

2.7.4.3 EASI real expenditure semi-elasticities across survey years

Table 2.8 shows the real expenditure elasticities for three survey waves. Though not statistically significant in all cases, the shares of the staple food and starchy food groups' elasticities are less than 1 for all the survey waves, confirming the Engel law. There are no statistically significant results for the vegetable and fruit group due to changes in real expenditures. In the 1997 data, only the dried food share is found to capture Engel's curvature of fourth degree polynomials in real expenditure, whereas in the 2000 data, the starchy food share and other food shares' curvature of the Engel curve capture the same higher order polynomials in real expenditure. In 2007, both the dried share and condiments share showed such relationships. In the appendix, Table A2.4 presents the results of the real expenditure semi-elasticities in pooled sample. Overall, there are no conclusive results for the shapes of Engel curves across the years for any specific food group.

Table 2.8: Shape of the Engel curve: EASI real expenditure semi-elasticities across survey years

	Staple	Starchy	Veg	Meat	Fish	Condiments
	(1)	(2)	(3)	(4)	(5)	(6)
1997						
Log of real expenditure	0.607***	0.805***	-7.106**	-1.249	1.300	1.208
	(0.050)	(0.127)	(3.035)	(2.739)	(1.745)	(0.750)
Log of real expenditure, square	-1.263***	-0.100	0.809**	0.192	-0.166	-0.133
	(0.453)	(0.083)	(0.344)	(0.330)	(0.199)	(0.087)
Log of real expenditure, cube	0.066***	0.005	-0.041**	-0.012	0.009	0.006
	(0.024)	(0.004)	(0.017)	(0.018)	(0.010)	(0.004)
Log of real expenditure, quartic	-0.001***	-0.000	0.001**	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-33.169***	-2.336	23.482**	2.780	-3.679	-3.934
	(12.198)	(2.380)	(10.002)	(8.454)	(5.715)	(2.407)
Observations	6,729	6,729	6,729	6,729	6,729	6,729
R-squared	0.142	0.009	0.035	0.002	0.003	0.033
2000						
Log of real expenditure	-0.263***	-2.855	-1.315	13.731***	2.253	-4.067***
	(0.072)	(2.709)	(4.781)	(3.001)	(1.528)	(0.888)
Log of real expenditure, square	2.819***	0.299	0.076	-1.485***	-0.272	0.470***
	(0.321)	(0.277)	(0.496)	(0.325)	(0.166)	(0.096)
Log of real expenditure, cube	-0.137***	-0.014	-0.001	0.070***	0.015*	-0.024***
	(0.015)	(0.013)	(0.023)	(0.016)	(0.008)	(0.005)

Log of real expenditure, quartic	0.002***	0.000	-0.000	-0.001***	-0.000**	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	83.392***	10.151	7.460	-46.989***	-7.010	12.899***
	(10.235)	(9.890)	(17.198)	(10.282)	(5.239)	(3.030)
Observations	10,141	10,141	10,141	10,141	10,141	10,141
R-squared	0.039	0.003	0.036	0.110	0.029	0.039
2007						
Log of real expenditure	-0.983***	-1.042	4.563	3.821***	-3.848***	-2.831
	(0.167)	(0.751)	(3.045)	(1.403)	(0.753)	(1.782)
Log of real expenditure, square	1.432***	0.106	-0.453	-0.430***	0.371***	0.296*
	(0.203)	(0.069)	(0.284)	(0.136)	(0.072)	(0.166)
Log of real expenditure, cube	-0.063***	-0.005*	0.019*	0.021***	-0.016***	-0.013**
	(0.008)	(0.003)	(0.012)	(0.006)	(0.003)	(0.007)
Log of real expenditure, quartic	0.001***	0.000*	-0.000*	-0.000***	0.000***	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	49.881***	3.846	-16.512	-12.322**	14.595***	9.946
	(8.614)	(3.030)	(12.160)	(5.375)	(2.908)	(7.132)
Observations	12,878	12,878	12,878	12,878	12,878	12,878
R-squared	0.045	0.006	0.031	0.087	0.034	0.012
Price controls	yes	yes	yes	yes	yes	yes
Demographics controls	yes	yes	yes	yes	yes	yes

Structural variable controls	yes	yes	yes	yes	yes	yes

Notes: Robust standard errors appear in brackets and clustered at the community level. Asterisks denote significance: *** p < 0.01, ** p < 0.05, * p < 0.1. Columns (1)-(6) show OLS estimates in EASI specifications for eight different food groups across survey years 1997, 2000, and 2007. In all specifications, a specific food group share is regressed on the fourth order polynomials of real food expenditure (linear, quadratic, cubic, and quintic food expenditures) where nominal food expenditure is made into real deflating by Stone Price Index. The control variables are: price controls (price indices for eight food groups), demographic controls (household size, proportions of households by different age-sex groups (for instance, pro_m0_4 shows the ratio of the number of boys in the age group 0-4 to the household size and pro_f0_4 shows the ratio of the number of girls to the household size in the same age group and other age-sex groups variable has been defined in the similar ways and age group 71-90 are an omitted category for both males and females), age of household head in years, highest grade attained by the head of the household, dummy variables for whether the household head is male and employed)), structural variables (a dummy variable for whether household experienced any natural disasters and a continuous variable of the distance from the household to the nearest market in kilometres).

2.7.4.4. EASI compensated price semi-elasticities

Table 2.9 reports the compensated budget-share semi-elasticities. The main diagonal shows Marshallian own-price, budget-share semi-elasticities. The signs of the own-price elasticities for all food groups are as expected and all of the own-price effects are large and statistically highly significant. The own-price semi-elasticity for the staple food share is 0.115 in 1997, which implies that a staple food-share price increase of 10% would be associated with a decrease in staple food consumption by 1.15 percentage points. The magnitude of own-price semi-elasticities for the staple food budget share is higher in 2000 and 2007. This may have happened because of the financial crisis in 1997 and 1998, during which households have to cut their staple food budgets to cope with the situation. This has policy implications, particularly for the poor people in rural areas. In rural areas, people working in agricultural production require more calories and the higher price of food may hurt their calorific fulfilment in the harvest season. The budget-share elasticities for other food groups show similar, significant results. The findings are consistent with those of similar previous studies, such as Pangaribowo and Tsegai (2011), Navamuel, Morollon, and Paredes (2014), Bouis (1990), Abdulai et al. (1999), Ahmed and Shams (1994), and Alfonzo and Peterson (2006).

Several cross-price effects are also large and statistically different from zero, suggesting that substitution effects are important. For example, the ‘starchy foods’ budget share compensated ‘staple foods” price semi-elasticity is -0.044, implying that an increase in the price of staple foods is associated with a significant increase in the budget share for starchy foods. Overall, all cross-price elasticities are found to be inelastic and compared to their own-price elasticities, the size of the cross-price elasticities are smaller. This implies that consumers are more sensitive to own good price changes than to cross-price changes. Table A2.5 shows the results of EASI compensated budget share semi-elasticities with respect to price in pooled sample.

Table 2.9: EASI compensated price semi-elasticities across three IFLS surveys

	Staple	Starchy	Veg	Meat	Fish
	(1)	(2)	(3)	(4)	(5)
1997					
P (Staple)	-0.115***	-0.002	-0.022***	-0.029***	-0.004**
	(0.011)	(0.001)	(0.006)	(0.008)	(0.002)
P (Starchy)	0.011***	-0.019***	-0.014***	0.004	-0.006***
	(0.004)	(0.002)	(0.003)	(0.003)	(0.001)
P (Veg)	-0.027***	-0.005***	-0.106***	-0.019***	-0.011***
	(0.009)	(0.002)	(0.008)	(0.007)	(0.003)
P (Meat)	-0.037***	0.002	-0.019***	-0.125***	-0.004**
	(0.006)	(0.001)	(0.005)	(0.009)	(0.002)
P (Fish)	-0.007	-0.006***	-0.007*	-0.001	-0.044***
	(0.004)	(0.001)	(0.003)	(0.003)	(0.002)
Constant	-0.056	0.030	-0.077	-0.492***	-0.051*
	(0.147)	(0.019)	(0.098)	(0.136)	(0.028)
Observations	391	391	391	391	391
R-squared	0.709	0.401	0.688	0.764	0.726
2000					
P (Staple)	-0.118***	-0.002***	-0.006***	-0.029***	-0.010***
	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)
P (Starchy)	-0.002	-0.028***	-0.000	-0.007***	-0.003***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
P (Veg)	-0.028***	-0.003***	-0.110***	-0.009***	-0.006***
	(0.003)	(0.001)	(0.002)	(0.002)	(0.001)
P (Meat)	-0.016***	-0.003***	-0.003***	-0.105***	-0.008***
	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)
P (Fish)	-0.000	-0.001***	-0.008***	-0.027***	-0.053***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
Constant	-0.357***	-0.036***	-0.275***	-0.226***	-0.054**
	(0.037)	(0.008)	(0.018)	(0.033)	(0.022)
Observations	3,002	3,002	3,002	3,002	3,002
R-squared	0.678	0.622	0.775	0.707	0.593
2007					
P (Staple)	-0.134***	-0.003**	-0.017***	-0.021***	-0.009***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
P (Starchy)	-0.002	-0.027***	0.007***	-0.010***	-0.003***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
P (Veg)	-0.018***	-0.000	-0.059***	-0.015***	0.002
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
P (Meat)	-0.019***	-0.003**	-0.008***	-0.113***	-0.009***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
P (Fish)	-0.007***	-0.002***	-0.006***	-0.011***	-0.054***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	-0.469***	-0.033	-0.014	-0.462***	-0.158***
	(0.035)	(0.025)	(0.029)	(0.030)	(0.019)

Observations	3,632	3,632	3,632	3,632	3,632
R-squared	0.766	0.499	0.513	0.735	0.580

Notes: Numbers in the parentheses are robust standard errors, corresponding to the compensated price semi-elasticities. Error are clustered at the community level. Each column represents a separate regression using a household level fourth order real expenditure polynomials and demographic controls. The demographic controls include household head's age, educational attainment, employment status, gender of head, and whether household receives any transfers from the government. The total number of observations (approximately 25,000 households) are from the pooled sample from IFLS 1997, IFLS 2000, and IFLS 2007.

2.7.4.5 Welfare comparisons between the poor and non-poor groups in food consumption

Estimates of the OLS regressions describe the relationship between changes in food budget shares and the log of real expenditures per capita with other household level controls. In order to adjust inflation with the nominal expenditures, the regressions include community-level fixed effect, as the price index has been constructed from community facilities surveys.

Comparing Table 2.10 and Table 2.11, the most interesting result is that whereas poor people consume more calorie-rich foods, well-off people consume less of the staple foods. In terms of high-value foods (e.g. vegetables and fruit, meat, and fish), the consumption patterns of rich people are quite different than those of poor people.

Table 2.10: Poor group food consumption patterns (log of real expenditure is below median)

	Staple	Starchy	Veg	Meat	Fish	Food diversity
	(1)	(2)	(3)	(4)	(5)	(6)
Log of real expenditure	0.095***	-0.004**	-0.033***	0.002	-0.000	0.023***
	(0.004)	(0.002)	(0.004)	(0.003)	(0.002)	(0.005)
Household size	-0.004***	0.001**	0.002***	0.001	0.001***	0.003***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Age of head	0.002***	0.000***	-0.000	-0.001***	-0.000	0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Education of head	-0.002***	0.000	0.001***	0.001***	-0.000**	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Employment of head	0.014*	-0.001	-0.007	-0.014**	0.005	0.016**
	(0.007)	(0.002)	(0.006)	(0.005)	(0.003)	(0.006)
Male head	-0.042***	0.001	-0.016***	0.002	0.001	-0.002
	(0.007)	(0.002)	(0.005)	(0.005)	(0.003)	(0.006)
Transfer	0.028***	0.004	-0.008	-0.001	-0.008**	0.027***
	(0.009)	(0.002)	(0.005)	(0.006)	(0.004)	(0.007)
Disaster	-0.006	0.001	0.036***	0.023***	-0.000	-0.015**
	(0.009)	(0.002)	(0.007)	(0.007)	(0.006)	(0.007)
Distance to the market	0.001	-0.000	0.002***	-0.002***	0.005***	0.002***
	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)
Urban	-0.012***	-0.001	0.006***	0.010***	0.001	0.009***
	(0.002)	(0.000)	(0.001)	(0.002)	(0.001)	(0.002)
Year 2000	-0.165***	0.036**	0.034**	-0.027	-0.016*	-0.058***
	(0.026)	(0.017)	(0.017)	(0.021)	(0.009)	(0.020)
Year 2007	-0.196***	0.009**	0.050***	-0.056***	-0.018***	-0.135***

	(0.010)	(0.004)	(0.009)	(0.007)	(0.004)	(0.011)
Constant	-0.933***	0.063**	0.574***	0.119***	0.062***	0.380***
	(0.054)	(0.027)	(0.052)	(0.042)	(0.024)	(0.062)
Observations	5,232	5,232	5,232	5,232	5,232	5,232
R-squared	0.179	0.022	0.059	0.040	0.057	0.151

Notes: Robust standard errors appear in brackets. Asterisks denote significance: *** p < 0.01, ** p < 0.05, * p < 0.1. Columns I-V show OLS estimates for five food groups and column (6) shows OLS estimate of food diversity in EASI budget share forms. Estimation is restricted for the samples of poor households, defined as the households whose per capita food consumption is below the median of per capita food expenditures. All five food group shares and food diversity is regressed on log of per capita food expenditure at the household level, household size, age of head in years, the highest educational attainment in terms of grade completed of the household head, indicator variables for whether household head is employed and male, a dummy variable for whether household received any cash or kind transfer from the government, a dummy variable for whether household experienced any natural disasters, a continuous variable of the distance from the household to the nearest market in kilometres, a dummy variable for whether household is an urban area, and dummy variables of the survey year 2000 and 2007 with dummy 1997 is an omitted category.

Table 2.11: Non-poor group food consumption patterns (log of real expenditure is above median)

	Staple	Starchy	Veg	Meat	Fish	Food diversity
	(1)	(2)	(3)	(4)	(5)	(6)
Log of real expenditure	-0.036***	0.001	-0.009*	0.060***	0.004	-0.050***
	(0.007)	(0.002)	(0.005)	(0.007)	(0.004)	(0.008)
Household size	0.005***	0.000	-0.001**	-0.005***	0.002***	0.003***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)
Age of head	0.001***	0.000***	0.000	-0.001***	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Education of head	-0.001***	-0.000	0.000***	0.001***	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Employment of head	0.004	0.001	0.005	-0.009	-0.004	0.008
	(0.012)	(0.002)	(0.005)	(0.009)	(0.006)	(0.007)
Male head	-0.004	-0.004***	-0.015***	-0.035***	0.001	0.016***
	(0.009)	(0.001)	(0.005)	(0.008)	(0.005)	(0.006)
Transfer	0.038***	0.001	-0.007*	-0.012*	-0.012***	-0.011**
	(0.008)	(0.001)	(0.004)	(0.007)	(0.004)	(0.005)
Disaster	-0.001	0.007***	-0.010**	-0.013*	0.013***	0.000
	(0.009)	(0.001)	(0.004)	(0.007)	(0.004)	(0.006)
Distance to the market	0.004***	-0.000***	-0.001***	-0.002***	0.002***	-0.001*
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Urban	-0.045***	-0.004***	0.011***	0.034***	-0.015***	-0.005
	(0.005)	(0.001)	(0.003)	(0.005)	(0.003)	(0.003)
Year 2000	-0.307***	-0.002	0.140**	-0.019	0.104***	0.238***
	(0.097)	(0.004)	(0.063)	(0.073)	(0.037)	(0.070)

Year 2007	-0.307***	0.005*	0.070***	-0.069	0.069***	0.245***
	(0.086)	(0.003)	(0.012)	(0.067)	(0.013)	(0.056)
Constant	1.298***	0.004	0.104	-0.849***	-0.064	1.112***
	(0.130)	(0.035)	(0.080)	(0.124)	(0.056)	(0.122)
Observations	3,731	3,731	3,731	3,731	3,731	3,731
R-squared	0.116	0.024	0.039	0.135	0.041	0.066

Notes: Robust standard errors appear in brackets. Asterisks denote significance: *** p < 0.01, ** p < 0.05, * p < 0.1. Columns (1)-(6) show OLS estimates for five food groups and column (6) shows OLS estimate of food diversity in EASI budget share forms. Estimation is restricted for the samples of non-poor households, defined as the households whose per capita food consumption is above the median of per capita food expenditures. All five food group shares and food diversity is regressed on log of per capita food expenditure at the household level, household size, age of head in years, education is the highest educational attainment in terms of grade completed of the household head, indicator variables for whether household head is employed and male, a dummy variable for whether household received any cash or kind transfer from the government, a dummy variable for whether household experienced any natural disasters, a continuous variable of the distance from the household to the nearest market in kilometres, a dummy variable for whether household is an urban area, and dummy variables of the survey year 2000 and 2007 with dummy 1997 is an omitted category.

Finally, I assess the welfare of individuals who are living just above the food poverty line. To accomplish this, I evaluate the economic significance of the EASI demand system with a cost-of-living experiment. In Indonesia, rice is not generally subject to value-added tax or sales tax, which is typically 10% for goods such as luxury houses or apartments. Let us consider the cost of living associated with being subject to 10% sales tax on rice for the people who face a zero price vector at the base level, so that $p_0 = 0_J$ and $P_1 = [ln0.10\ 0\ 0\ 0\ 0\ 0\ 0\ 0]$:

$$C(p_1, u, z, \varepsilon) - C(p_0, u, z, \varepsilon) = ln0.10w^{rice} + \frac{ln0.10^2\left(\sum_{l=0}^{L} z_l a_l^{rice,rice} + b^{rice,rice} y\right)}{2} = \frac{ln0.10^2\left(\sum_{l=0}^{L} z_l a_l^{rice,rice} + b^{rice,rice} x\right)}{2} \quad (16),$$

where $a_l^{rice,rice}$ and $b^{rice,rice}$ are the own-price elements of A_l and B. The comparison price vector is 0_J where $y = x$, and therefore there is only one budget-share and six parameters are required to estimate the cost-of-living index. After a tax increase, the consumer surplus for individuals who live just above the poverty line experiences a 2% decline. This implies that they are very vulnerable to any shocks, which may take them to the lower level of the poverty line of the people who are currently living slightly below the poverty line.

2.7 Conclusion and limitations

This study examines the changing patterns of food consumption in Indonesian households using three rounds (1997, 2000, and 2007) of longitudinal IFLS for eight aggregated food groups: staple foods, starchy foods, vegetables and fruits, meat, fish, dried foods, condiments, and other. I employ EASI to estimate food demand semi-elasticities and their determinants in Indonesian households and observe the following.

First, the estimated Engel curves are found to be irregular in shape across survey years under the approximate EASI model. Some food demand functions are close to linear, some are quadratic, and some show as possibly cubic. This implies that food consumption patterns have empirical evidence up to rank 3. Most likely, the estimation captures much of the unobserved heterogeneity in consumption, which is evident from the estimated Engel curves.

Second, Indonesian consumers are found to be highly responsive even with small changes in the magnitudes of the own price elasticities. This is evident from the estimation that the magnitude of cross-price elasticities is lower than own-price elasticities. This has an important policy implication, as it suggests that a small change in the price of staple food could significantly impact how much it is consumed, and this in turn could affect the minimum energy requirement particularly for the poor segments of the society.

Third, some non-economic variables such as household size, education, distance to the market from the household and natural disaster significantly affect food demand patterns in Indonesia.

The main limitation of this study is that the IFLS data does not contain information about quantity of food consumption, which is critical for analysing nutrition and welfare at the household level. Nor have consumer tastes and preferences been taken into consideration, as between 1997 and 2007, some important changes may have taken place in consumer preferences. Moreover, this study has not exploited the dynamics of food consumption, although the IFLS data are longitudinal. Hence, a fruitful avenue of future research would be analysis of inter-temporal food demand and consumption, the impact of tastes and preferences on food consumption, and the impact of urbanisation on changing food consumption patterns in developing countries, including Indonesia.

References

Abdulai, A., D. Jain, and A. Sharma (1999): "Household food demand analysis in India," Journal of Agricultural Economics, 50 (2), 316–327.

Abdulai, A. (2002): "Household demand for food in Switzerland. A quadratic almost ideal demand system," Available at https://ideas.repec.org/a/ses/arsjes/2002-i-1.html

Ahmed, N., M. Brzozowski and T. F. Crossley (2016): "Measurement errors in recall food consumption data," The Institute for Fiscal Studies working paper no. WP06/21. Available at https://www.ifs.org.uk/wps/wp0621.pdf

Ahmed, A. and Y. Shams (1994): "Demand elasticities in rural Bangladesh: An application of the AIDS model," *Bangladesh Development Studies*, 22, 1–25.

Ahuja, Vinod, Benu Bidani, Francisco Ferreira, and Michael Walton. (1997). *Everyone's miracle: Revisiting poverty and inequality in East Asia.* World Bank, Washington, DC.

Alfonzo, L., and H., Peterson (2006): "Estimating food demand in Paraguay from household survey data," *Agricultural Economics*, 34 (3), 243–257.

Atkinson, A. B., J., Gomulka, and N. H., Stern (1990): Spending on alcohol: evidence from the Family Expenditure Survey 1970–1983," *Economic Journal*, 100 (402), 808–827.

Attanasio, O. P., and V., Lechene (2014): "Efficient responses to targeted transfers," *Journal of Political Economy*, 122 (1), 178–222.

Attanasio, O., and M. Angelucci (2012): "The demand for food of poor urban Mexican households: Understanding policy impacts using structural models," *American Economic Journal: Economic Policy*, 5 (1), 146–178.

Attanasio, O., E., E. Battistin and A. Mesnard (2009): "Food and cash transfers: Evidence from Colombia," CEPR Discussion Paper No. DP7326. Available at SSRN: https://ssrn.com/abstract=1433922

Attanasio, O., V. Di Maro, V., Lechene and D., Phillips (2013): "The effect of increases in food prices on consumption and welfare in rural Mexico," *Journal of Development Economics*, 104 (C), 136–151.

BPS (2013). *Statistical Yearbook of Indonesia 2013.* Statistics Indonesia, Jakarta.

Baltagi, B. H. (2005). *Econometric analysis of panel data*, 4th edn. John Wiley & Sons: Chichester, West Sussex.

Banks, J., R. Blundell, and A. Lewbel (1997): "Quadratic Engel curves and consumer demand," *Review of Economics and Statistics*, 79 (4), 527–539.

Bellemare, Marc F., Christopher B. Barrett, and David R. Just (2016): "The welfare impacts of commodity price volatility: Reply," *American Journal of Agricultural Economics* forthcoming.

Blundell, R., and T. M. Stoker (2005): "Heterogeneity and aggregation," *Journal of Economic Literature*, 43 (2), 347–391.

Blundell, R., A. Duncan, and K. Pendakur (1998): “Semiparametric Estimation and Consumer Demand,” *Journal of Applied Econometrics*, 13 (5), 435–461.

Blundell, R., P. Pashardes and G. Weber (1993): “What do we learn about consumer demand patterns from micro data?” *American Economic Review,* 83 (3), 570–597.

Blundell, R., X. Chen, and D. Kristensen (2007): “Semi-nonparametric IV estimation of shape-invariant Engel curves,” *Econometrica,* 75 (6), 1613–1669.

Bouis, H. (1990): “Evaluating demand for calories for urban and rural populations in the Philippines: Implications for nutrition policy under economic recovery,” *World Development,* 18 (2), 281–299.

Briggs, A., and S. Chowdhury (2014): “Economic development, food demand and the consequences for agricultural resource requirements (Indonesia),” 58th National AARES Conference, Canberra, ACT. Australian Agricultural and Resource Economics Society (AARES).

Cameron, Lisa. (1999). Indonesia: A quarterly review. *Bulletin of Indonesian Economics Studies*. 35 (1), 3–41.

Clements, K.W., and S. Selvanathan (1994): “Understanding consumption patterns,” *Empirical Economics,* 19 (1), 69–110.

Deaton, A. (2016): “Measuring and understanding behavior, welfare, and poverty,” *American Economic Review,* 106 (6), 1221–1243.

Deaton, A., and M. J. Irish (1984): “Statistical models for zero expenditures in household budgets,” *Journal of Public Economics,* 23 (1–2), 59–80.

Deaton, A. S., and J. Muellbauer (1980): “An almost ideal demand system,” *American Economic Review,* 70 (3), 312–336.

Deaton, A., J. Ruiz-Castillo, and D. Thomas (1989): “The influence of household composition on household expenditure patterns: theory and Spanish evidence,” *Journal of Political Economy,* 97 (1), 179–200.

Deaton, A. and Zaidi, S. (2002). *Guidelines for constructing consumption aggregates for welfare analysis*. LSMS Working Paper No. 135, World Bank: Washington DC.

Deaton, A., and J. Muellbauer (1980a): “An Almost Ideal Demand System,” *American Economic Review,* 70 (3), 312–26.

Deaton, Angus S. and Muellbauer, John. (1980b). *Economics and consumer behavior.* Cambridge University Press: Cambridge.

Eichengreen, B. and Gupta, P. (2009): The two waves of service sector growth. National Bureau of Economic Research Working Paper No. 14968. Available at http://www.nber.org/papers/w14968.pdf

Elias, Stephen, and Noone, Clare (2011). The growth and development of the Indonesian economy. *Bulletin of the Reserve Bank of Australia*, December Quarter, 33–44. Available at http://www.rba.gov.au/publications/bulletin/2011/dec/pdf/bu-1211-4.pdf

Fallon, P. R., and R. E. B. Lucas (2002): "The impact of financial crises on labor markets household incomes, and poverty: A review of evidence," *World Bank Research Observer*, 17 (1), 21–45.

Frankenberg, E., J. Smith, and D., Thomas (2003): "Economic Shocks, Wealth and Welfare," *Journal of Human Resources*, 38 (2), 280–321.

Gorman, W. M. (1981). Some Engel curves. In *Essays in the theory and measurement of consumer behaviour in Honor of Sir Richard Stone*, ed. A. Deaton, A. Cambridge University Press: Cambridge.

Guilmoto, C. Z., and G. W. Gones (2016): *Contemporary demographic transformations in China, India, and Indonesia*. Demographic Transformation and Socioeconomic Development, Vol. 5, Springer. DOI: 10.1007/978-3-319-24783-0_5

Hutasuhut, M., H-S. Chang, G. Griffith, C. J. O'Donnell, and H. Doran (2002): "The demand for beef in Indonesia: Implication for Australian agribusiness," *Australian Agribusiness Review* 10 (4), 1–10.

Jensen, H., and J. Manrique (1998): "Demand for food commodities by income groups in Indonesia," *Applied Economics*, 30 (4), 491–501.

Jorgenson, Dale W., Lawrence J. Lau, and Thomas M. Stoker (1980): "Welfare comparison and exact aggregation," *American Economic Review,* 70 (2), 268–72.

Jorgenson, Dale W., Lawrence J. Lau, and Thomas M. Stoker (1982): "The transcendental logarithmic model of aggregate consumer behaviour," In *Advances in Econometrics, Vol. 1*. ed. R. L. Basmann and G. Rhodes. JAI Press: Greenwich, pp. 97–238.

Júarez-Torres, M. (2015): "The Impact of Food Price Shocks on Consumption and Nutritional Patterns of Urban Mexican Households," International Association of Agricultural Economists 2015 Conference, August 9–14, 2015, Milan, Italy.

Kakwani, N. (1977): "On the estimation of Engel elasticities from grouped observation with application to Indonesian data," *Journal of Econometrics*, 6 (1), 1–19

Kebede, B. (2003). *Intra-household distribution of expenditures in rural Ethiopia: A demand system approach*. Wellbeing in Developing Countries (WeD) and ESRC Research Group, and Centre for the Study of African Economies (CSASE), Working Paper Series 2003-08, Oxford University.

Levinsohn, J., S. Berry, and J. Friedman (2003): "Impacts of the Indonesian economic crisis: Price changes and the poor," in *Managing currency crises in emerging markets*, ed. by Michael Dooley and Jeffrey Frankel, 393–424. Chicago, University of Chicago Press.

Lewbel, A. (1991): "The rank of demand systems: Theory and Nonparametric Estimation," *Econometrica*, 59 (3), 711–730.

Lewbel, A., and K. Pendakur (2009): “Tricks with Hicks: The EASI Demand System,” American *Economic Review*, 99 (3), 827–863.

Meghir, C., and J. M. Robin (1992): “Frequency of purchase and the estimation of demand systems,” *Journal of Econometrics*, 53(1–3), 53–85.

Magana-Lemus D., A. Ishdorj and C. P. Rosson (2013): “Welfare impacts of increasing food prices in Mexico: An application of unrestricted Engel curves and LA/EASI demand system,” Southern Agricultural Economics Association, Working Paper No. 143057.

McFadden, D., and K. Richter (1991) "Stochastic Rationality and Revealed Stochastic Preference," in J. Chipman, D. McFadden, K. Richter (eds) *Preferences, Uncertainty, and Rationality*, Westview Press, 161–186.

Moeis, J. Prananta. (2003). Indonesian food demand system: An analysis of the impact of the economic crisis on household consumption and nutritional intake. Unpublished Doctor of Philosophy’s Dissertation. The Faculty of Columbian College of Art and Sciences, George Washington University.

Moschini, G. (1995): “Unit of measurement and the Stone index in demand system estimation,” *American Journal of Agricultural Economics*, 77 (1), 63–68.

Niimi, Y. (2005). An analysis of household responses to price shocks in Vietnam: Can unit values substitute for market prices? PRUS Working Paper No. 30. Poverty Research Unit at University of Sussex: Susssex**.**

Navamuel, E. L., F. R. Morollón, and D. Paredes (2014): “City size and household food consumption: demand elasticities in Spain**,”** *Applied Economics*, 46 (14), 1624–1641.

Olivia, S., and J. Gibson (2012): “Measuring Economic Change in Indonesia-As Seen from Space,” Available at https://crawford.anu.edu.au/acde/events/adew-2013/papers/Olivia,S__B2.pdf

Olivieri, A. (2014). *Estimation of consumption choices with the EASI demand system: Application to Italian data*. Prometeia Associazione per le Previsioni Econometriche: Bologna.

Pangaribowo, E. H., and D. Tsegai (2011): “Food Demand Analysis of Indonesian Households with Particular Attention to the Poorest,” ZEF-Discussion Papers on Development Policy No. 151, University of Bonn.

Pendakur, K. (2009): "EASI Made Easier", in *Quantifying Consumer Preferences*, (Contributions in Economic Analysis Series), edited by D. Slottje, London: Emerald Group Publishing, pp 179–206.

Philips, L. (1983). *Applied consumption analysis: Advanced Textbooks in Economics*, 2nd edn. Elsevier Science Publishing Company, Inc., Amsterdam, The Netherlands.

RAND. 2010. IFLS4 public release. Available at http://www.rand.org/labor/FLS/IFLS.html (Accessed 22 December, 2010).

Ray, A. N. (1999): "Food consumption patterns and nutrition in urban Java households: The discriminatory power of some socioeconomic variables," *The Australian Journal of Agricultural and Resource Economics*, 43 (3), 359–383.

Ravallion, M., and M., Lokshin (2007): "Lasting Impacts of Indonesia's Financial Crisis," *Economic Development and Cultural Change*, 56 (1), 27–56.

Skoufias, E., Tiwari, S., and Zaman, H. (2011). *Can we rely on cash transfers to protect dietary diversity during food crises? Estimate from Indonesia.* World Bank Policy Research Working Paper No. 5548, World Bank. Available at http://dx.doi.org/10.1596/1813-9450-5548 (Accessed 11 October, 2016).

Skoufias, E., A., Suryahadi, and S., Sumarto (2000): "Changes in Household Welfare, Poverty and Inequality during the Crisis," *Bulletin of Indonesian Economic Studies*, 36 (2), 97–114.

Song, Z., Li, L. and Ma, C. (2013), *The EASI demand system: Evidence from China household*, MPRA Paper No. 48435. Available at https://mpra.ub.uni-muenchen.de/48435/1/MPRA_paper_48435.pdf

Strauss, John, Kathleen Beegle, Agus Dwiyanto, Yulia Herawati, Daan Pattinasaramy, Elan Satriawan, Bondan Sikoki, Sukamdi, and Firman Witolear (2004): *Indonesian living standards before and after the financial crisis*. Singapore: Institute of Southeast Asian Studies

Strauss, J., Witoelar. F., Sikoki, B., and Wattie, A. M. (2009). *The fourth wave of Indonesian Family Life Survey: Overview and field report, Volume 1*. WR-675/1-NIA/NICHD, RAND Corporation, Washington DC.

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

Tabor, S. R, A. Klaus, and B. Adinugroho. (1989): "Food crop demand in Indonesia: A system approach," *Bulletin of Indonesian Economic Studies*, 25 (2), 31–51.

TelI, T., and S. R. Johnson (1987): "Demand system from cross section data: An experiment for Indonesia," Working Paper 87-WP 24.

Timmer, P. C., and H. Alderman (1979): "Estimating consumption parameters for food policy analysis," *American Journal of Agricultural Economics*, 61 (5), 982–987.

Thomas, D., and E. Frankenberg (2007): "Household responses to the financial crisis in Indonesia: Longitudinal evidence on poverty, resources, and well-being (pp. 517–560)," in *Globalization and poverty*, ed. A. Harrison (ed.), University of Chicago Press for National Bureau of Economic Research.

United Nations (2016): Department of Economic and Social Affairs, Population Division. Available at http://www.un.org/en/development/desa/population/

Varian, H. R. (1992): *Microeconomic analysis*, (3rd edn), Norton: New York and London.

Vu, L. H., and P. W. Glewwe (2011): "Impact of Rising Food Prices on Poverty and Welfare in Vietnam," *Journal of Agricultural and Resource Economics*, 36 (1), 14–27.

Widarjono, A. (2012): “Food and nutrient demand in Indonesia,” PhD Dissertation, Oklahama State University.

Wood, B., C. H. Nelson, and L. Nogueira (2012): “Poverty effects of food price escalation: The importance of substitution effects in Mexican households*,”* *Food Policy*, 37 (1), 77–85.

Woods, J. E. (1979): “A Note on Hicks’ Composite Commodity Theorem,” *Journal of Economics,* 39 (1/2), 185-188.

Appendix A

Table A2.1: Definitions of variables

Variable name	Definition	Units	Household/ Community level
Log of real expenditures	Log of annual total food expenditure deflated by Stone Index	Rupiah	Household
HH size	Number of persons in a household (HH)	Persons	Household
Food group share	Expenditure on a particular food group divided by HH size	Rupiah	Household
Prices of food groups	Stone price index of various food groups	Index	Community
Asset	Annual assets excluding farm assets and non-farm business assets	Rupiah	Household
Age of head	Age of HH head	Years	Household
Education of head	Highest grade completed by HH Head (0 to 7 scale, 0 indicates no grade completion; 7 indicates highest grade completion)	Grade scale	Household
Employment of head	Whether HH head employed	Dummy (HH head employed=1)	Household
Transfer	Whether HH received any assistance from the government	Dummy (Transfer received=1)	Household
Disaster	Whether HH experienced any disaster after the last survey	Dummy (Experienced disaster=1)	Community
Distance to the market	Distance to the nearest market	Kilometre	Community
Distance to the capital	Distance to the provincial capital	Kilometre	Community
Urban	Whether HH located in urban area	Dummy (Urban=1)	Household
Year 1997	Survey year 1997	Dummy (year 1997=1)	Household
Year 2000	Survey year 2000	Dummy (year 2000=1)	Household
Year 2007	Survey year 2007	Dummy (year 2007=1)	Household
pro_m0_4	Number of male persons in age group 0-4 divided by HH size	Proportion	Household
pro_f0_4	Number of female persons in age group 0-4 divided by HH size	Proportion	Household

pro_m5_9	Number of male persons in age group 5-9 divided by HH size	Proportion	Household
pro_f5_9	Number of female persons in age group 5-9 divided by HH size	Proportion	Household
pro_m10_14	Number of male persons in age group 10-14 divided by HH size	Proportion	Household
pro_f10_14	Number of female persons in age group 10-14 divided by HH size	Proportion	Household
pro_m15_55	Number of male persons in age group 15-55 divided by HH size	Proportion	Household
pro_f15_55	Number of female persons in age group 15-55 divided by HH size	Proportion	Household
pro_m55_70	Number of male persons in age group 55-70 divided by HH size	Proportion	Household
pro_f55_70	Number of female persons in age group 55-70 divided by HH size	Proportion	Household
pro_m71_90	Omitted category	Proportion	Household
pro_f71_90	Omitted category	Proportion	Household
Food diversity	Sampson Index to measure food diversity (Index value takes from 0 to 1 and as the Index value closes to 1, food diversity increases)	Index	Household

Source: IFLS household level and Community Facility Surveys.

A2.2 EASI implicit Marshallian demand function

EASI implicit Marshallian demand function can be expressed with the following parametric cost function:

$$ln\, C(p,u,z,\varepsilon) = u + \sum_{j=1}^{J} m^j(u,z)\, ln\, p^j + \frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{J}\sum_{h=1}^{H} a^{jk} z_h\, ln\, p^j\, ln\, p^k + \frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{J}\sum_{h=1}^{H} b^{jk}\, ln\, p^j\, ln\, p^k\, u + \frac{1}{2}\sum_{j=1}^{J} \varepsilon^j\, ln\, p^j \quad \text{(A21)}$$

Where J=1,…, J indicates goods, z is a H-vector of demographic variables, p is a J-vector of prices, u is utility and ε represents unobserved individual heterogeneity. The log of costs function is assumed to be concave, increasing, differentiable and homogeneity of degree one in prices and is monotonically increasing and differentiable in utility, u.

Let us define $m^j(u, z)$ as:

$$m^j(u,z) = \sum_{r=1}^{R} b_r^j\, u^r + \sum_{h=1}^{H} g_h^j\, z_h + \sum_{h=2}^{H} d_h^j\, z_h u \qquad \text{(A22)}$$

Applying Shephard's lemma ($\frac{\delta \ln C(.)}{\delta \ln p^j} = w^j$), the share of expenditure in good j, Hicksian budget-share equations, becomes:[17]

$$w^j = \sum_{r=1}^{R} b_r^j\, u^r + \sum_{h=1}^{H} g_h^j\, z_h + \sum_{h=2}^{H} d_h^j\, z_h y + \sum_{j=1}^{J} \sum_{k=1}^{J} \sum_{h=1}^{H} a^{jk}\, z_h lnp^k + \sum_{k=1}^{J} b^{jk} lnp^k\, y + \varepsilon^j \qquad \text{(A23)}$$

From Equation 3, we can write the implicit utility ($y = u$) as follows:

$$y = u = \frac{lnx - \sum_{j=1}^{J} w^j lnp^j + \frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{J}\sum_{h=1}^{H} a^{jkh} z_h lnp^j lnp^k}{1 - \frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{J} b^{jk} lnp^j lnp^k} \qquad \text{(A24)}$$

[17] Following Pendakur (2009), we can provide natural interpretation of Shephard's lemma as follows. Suppose a household is faced with 10% increase in the price of a staple food, say, rice, and rice accounts for half of the budget, then the cost increase for the household is 10%*50%=5%. For a small price change, we usually do not adjust our consumption choices. Rather, we just require 5% extra money to buy the same commodity as before. Henceforth, the proportionate change in cost due to a small price increase is equal to the budget share of the commodity.

The above Equation 6 can be compactly expressed in a matrix form as:

$$y = g(w, p, x, z) = \frac{x - p^{'} w - \sum_{l=0}^{L} z_l p^{'} A_l p/2}{1 - p^{'} Bp/2} \qquad \text{(A25)}$$

Equations (3) and (4) define the EASI implicit Marshallian demand system that possesses several desirable features in common with traditional demand systems, such as AIDS and QUAIDS. First, the model is amenable to estimation via iterative linear methods. Second, there are linear price effects that may depend upon observable factors ($a^{jk}(z)$). Third, the functions $m^j(u, z)$ can be independent of y, linear in y like AIDS and quadratic in y as in QUAIDS. Fourth, the utility function in Equation 4 is expressed in terms of observables.

The implicit utility defined as y in Equation 4 encompasses the properties of the log real expenditures (log of Stone index deflated nominal expenditures). It is equal to a cardinal measure of utility u and is affine in nominal expenditures x. When the logs of all prices are zero under the assumption of base period price (equals one), it is equal to x. When B becomes zero, y is exactly equal to log of nominal expenditures (x) deflated by Stone price index ($p^{'} w$). The real expenditure is highly correlated with the term $x - p^{'} w$.

A2.3 Separability of preferences

The Hicks' composite commodity theorem provides guidance on how to classify commodities into various groups. The theorem implies that if the prices of a group of commodities change in the same proportion, that group will act as a single commodity (Woods, 1979). As empirical economists regularly deal with commodity aggregation, the usefulness of this theorem is evident. However, as prices of closely related commodities vary in different directions, it is difficult to form commodity groups difficult using Hick's composite commodity theorem (Deaton and Muellbauer, 1980a).

Alternatively, it is possible to use Varian's (1992) functional separability condition in preferences to form a group for the goods. The preference ordering follows the property that

$$(w, z) \succ (w', z) \text{ if and only if } (w, z') \succ (w', z')$$

for all consumption bundles w, z, w', and z'.

This condition means that if w is preferred to w' for some choices of the other goods, then w is preferred to w' for all choices of the other goods. Put differently, the preferences the w-goods are independent of the z-goods and the utility function for w and z can be expressed as $u(w, z) = U(v(w), z)$, where $U(v, z)$ is an increasing function of v. The overall utility from w and z is a function of the sub-utility of w, $v(w)$ and the level of consumption of the z-goods. Solving the sub-utility maximisation problem would give demand function for the w-goods only as a function prices of the w-goods and expenditure on w- goods in the demand function of good i, expressed as $w_i = h_i(x_i, p_i)$. The prices of the other goods are only relevant if they determine the expenditure of the current good. Weak separability ensures the existence of sub-utility demand functions.

Here, it is assumed that utility functions of the different food groups are weakly separable. This allows us to classify the commodities into sub-groups in a way that preferences can be described within the sub-group regardless of the amount demanded in the other food groups. This implies that the consumer can purchase the bundle of goods in any group irrespective of consumption in the other groups. Weak separability implies the existence of sub-group demand functions.

Table A2.4: Shape of the Engel curve in pooled sample: Real expenditure elasticities

	Staple (1)	Starchy (2)	Veg (3)	Meat (4)	Fish (5)	Condiments (6)
Log real expenditure	2.016***	-1.165***	-2.475***	7.448***	-0.093	-1.055***
	(0.747)	(0.327)	(0.713)	(0.716)	(0.364)	(0.311)
Log real expenditure, squared	-0.141*	0.118***	0.240***	-0.778***	0.001	0.108***
	(0.073)	(0.032)	(0.069)	(0.072)	(0.036)	(0.030)
Log real expenditure, cube	0.004	-0.005***	-0.010***	0.035***	0.000	-0.005***
	(0.003)	(0.001)	(0.003)	(0.003)	(0.002)	(0.001)
Log real expenditure, quartic	-0.000	0.000***	0.000***	-0.001***	-0.000	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-9.627***	4.282***	9.739***	-26.212***	0.710	3.881***
	(2.827)	(1.254)	(2.733)	(2.659)	(1.368)	(1.178)
Observations	29,748	29,748	29,748	29,748	29,748	29,748
R-squared	0.031	0.004	0.017	0.057	0.007	0.002
Price controls	yes	yes	yes	yes	yes	yes

Demographic controls	yes	yes	yes	yes	yes	yes
Structural variable controls	yes	yes	yes	yes	yes	yes

Notes: Robust standard errors appear in brackets and clustered at the community level. Asterisks denote significance: *** p < 0.01, ** p < 0.05, * p < 0.1. Columns (1)-(6) show OLS estimates in EASI specifications for six food groups in pooled samples of the survey years 1997, 2000, and 2007. In all specifications, a specific food group share is regressed on the fourth order polynomials of real food expenditure (linear, quadratic, cubic, and quintic food expenditures) where nominal food expenditure is made into real deflating by Stone Price Index. The control variables are: price controls (price indices for eight food groups), demographic controls (household size, proportions of households by different age-sex groups (for instance, pro_m0_4 shows the ratio of the number of boys in the age group 0-4 to the household size and pro_f0_4 shows the ratio of the number of girls to the household size in the same age group and other age-sex groups variable has been defined in the similar ways and age group 71-90 are an omitted category for both males and females), age of household head in years, highest grade attained by the head of the household, dummy variables for whether the household head is male and employed)), structural variables (a dummy variable for whether household experienced any natural disasters and a continuous variable of the distance from the household to the nearest market in kilometres).

Table A2.5: EASI compensated price share semi-elasticities in pooled sample

	Staple (1)	Starchy (2)	Veg (3)	Meat (4)	Fish (5)	Condiments (6)
Compensated price semi-elasticities						
P (Staple)	0.127***	-0.003***	-0.016***	-0.028***	-0.019***	-0.013***
	(0.003)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
P (Starchy)	-0.003*	0.027***	0.002	-0.008***	-0.005***	-0.001
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
P (Veg)	-0.010***	-0.002***	0.076***	-0.019***	-0.007***	-0.011***
	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)
P (Meat)	-0.021***	-0.003***	-0.012***	0.108***	-0.019***	-0.008***
	(0.002)	(0.000)	(0.001)	(0.002)	(0.001)	(0.001)
P (Fish)	-0.007***	-0.002***	-0.005***	-0.012***	0.058***	-0.004***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)

P (Dried)	-0.006***	-0.002***	-0.002*	-0.025***	-0.008***	-0.003***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
P (Condiments)	-0.003	-0.001***	-0.016***	0.000	-0.007***	0.041***
	(0.002)	(0.000)	(0.001)	(0.001)	(0.001)	(0.002)
P (other)	-0.021***	-0.003***	-0.013***	-0.021***	-0.013***	-0.009***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Structural variables semi-elasticities						
Disaster	0.031***	0.000	-0.043***	0.070***	-0.024***	-0.004***
	(0.005)	(0.001)	(0.003)	(0.004)	(0.003)	(0.002)
Distance to the market	0.003***	0.000***	0.002***	0.001**	-0.002***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-0.248***	0.008	0.169***	-0.074***	0.231***	0.169***
	(0.031)	(0.007)	(0.017)	(0.020)	(0.017)	(0.018)
Observations	2,747	2,747	2,747	2,747	2,747	2,747
R-squared	0.744	0.565	0.668	0.815	0.653	0.630
Real expenditure polynomial controls	yes	yes	yes	yes	yes	yes
Demographics controls	yes	yes	yes	yes	yes	yes

Notes: Numbers in the parentheses are robust standard errors, corresponding to the compensated price semi-elasticities. Each column represents a separate regression using household level fourth-order real expenditure polynomials and demographic controls. The demographic controls include household head's age, educational attainment, employment status, gender, and whether household receives any transfers from the government. The total number of observations (approximately 25,000 households) are from the pooled sample from IFLS 1997, IFLS 2000, and IFLS 2007.

Table A2.6: First-stage regressions for Table 2.7

	Log of real expenditure	Log of real expenditure
	(1)	(2)
Log of asset	0.103***	
	(0.005)	
Log of wages		0.003***
		(0.000)
Household size	0.072***	0.071***
	(0.003)	(0.003)
pro_m0_4	1.093***	1.092***
	(0.106)	(0.106)
pro_m5_9	0.882***	0.885***
	(0.106)	(0.106)
pro_m10_14	0.729***	0.724***
	(0.116)	(0.117)
pro_m15_55	0.436***	0.429***
	(0.091)	(0.091)
pro_m55_over	0.156	0.152
	(0.105)	(0.105)
pro_f0_4	0.982***	0.981***
	(0.107)	(0.107)
pro_f5_9	0.815***	0.814***
	(0.109)	(0.109)
pro_f10_14	0.802***	0.798***
	(0.123)	(0.122)
pro_f15_55	0.543***	0.538***
	(0.090)	(0.090)
pro_f55_over	0.125	0.115
	(0.101)	(0.100)
Age of head	0.003***	0.003***
	(0.001)	(0.001)
Education of head	0.000	0.000
	(0.000)	(0.000)
Employment of head	0.115***	0.119***
	(0.026)	(0.026)
Male head	0.223***	0.222***
	(0.023)	(0.023)
Transfer	-0.063***	-0.063***
	(0.021)	(0.021)
Disaster	0.167***	0.166***
	(0.023)	(0.023)
Distance to the market	0.004***	0.005***
	(0.001)	(0.001)
Urban	0.020	0.014

	(0.014)	(0.014)
Year 2000	1.908***	1.904***
	(0.088)	(0.088)
Year 2007	2.664***	2.669***
	(0.016)	(0.016)
Constant	9.880***	10.731***
	(0.123)	(0.104)
Observations	10,160	10,160
R-squared	0.789	0.789

Notes: Robust standard errors appear in brackets and clustered at the community level. Asterisks denote significance: *** p < 0.01, ** p < 0.05, * p < 0.1. Column (I) regresses endogenous log of per capita food expenditure on exogenous log of per capita assets (an instrument)-(VIII) on household size, proportions of individuals in different age-sex groups (for instance, pro_m0_4 shows the ratio of the number of boys in the age group 0-4 to the household size and pro_f0_4 shows the ratio of the number of girls to the household size in the same age group and other age-sex groups variable has been defined in the similar ways and age group 71-90 are an omitted category for both males and females), age of household head in years, highest grade attained by the head of the household, dummy variables for whether the household head is male and employed)), structural variables (a dummy variable for whether household experienced any natural disasters and a continuous variable of the distance from the household to the nearest market in kilometres), a dummy variable for whether household received any government transfers, a dummy for urban households, and dummies for survey years 2000 and 2007

Chapter 3 Essay 2: Migration and Food Consumption Patterns in Indonesia

3.1 Introduction

The impact of migration on reshaping consumption patterns of migrant-sending households has gained much attention over the last two decades.[18] While much of the focus in the literature has been on the impact of migration on overall consumption patterns of the migrant's households, the impact of migration on different food group's consumption on migrant-sending-households is less understood but extremely important for policy-makers. Applying the instrumental variable and matching technique to micro data in Indonesia, this chapter addresses the question of how an internal migration affects food consumption diversity in migrant sending households compared to non-migrant sending households.

Indonesia is an interesting country for this comparison as there is an increasing trend in the proportion of lifetime interprovincial migrants (Muhidin, 2014; Hugo, 2012) as well as a drastic change in food consumption pattern (Hasanah, et al. 2017). Development economists and policy makers mostly agree that having more diversified and higher levels of consumption can lead to higher household welfare. Therefore, this study may have policy implications for health and nutrition aspects of welfare due to enhanced and diversified food consumption.

[18]Studies have documented the relationship between migration and consumption in a household: Taylor and Mora, 2006; Taylor et. al., 1996; Chami, Fullenkamp, and Jahjah, 2003; Durrand and Massey, 1996; de Brauw, Mueller and Woldehama, 2013; Zarat-Hoyos, 2004; Deb and Seck, 2009; Hasanah, Medolia, and Yerokhin, 2017; Olowa and Awoyemi, 2011.

Household's well-being in developing economies are generally affected by resource constraints, insufficient infrastructure, and recurrent natural disasters. There are several ways in which households' may respond to sustain their own welfare when faced with these constraints. The first type of response is related to family strategies to cope with income fluctuations and liquidity constraints (Morduch, 1995; Dercon, 2002). Some effective strategies could be income diversification, saving in good times, and labour supply adjustments. The second type of response is internal migration. Although people in developing countries (e.g. Ethiopia) sometimes face negative government policy on labour movement shocks such as banning migration (Dercon, 2002), others place special attention on the role of movement across regions and nations to increase the household's welfare (Rosenzweig and Stark, 1989; Stark, 1991; Yang and Martinez, 2005). Groger and Zylberberg (2015) argued that internal migration provides an effective shock-coping instrument for agricultural households in developing countries. The final response to adjust with the new situation is borrowing. Compared with all other possible strategies by the households, internal migration is found to be very cost effective and effective to deal with adverse situations.

In the literature, several studies have found that internal migration in recent times has played an important role in changing people's well-being and in reshaping development in Asian countries. Using secondary data from Bangladesh, China, Vietnam and the Philippines, Anh (2003) showed that migration has provided opportunities for one to escape poverty and to improve on one's livelihood and well-being. She further argued that controlling the movement of people is counterproductive. Afsar (2*003*) pointed out that migrants' remittances increase the area of land cultivation. Ping (2003) argued that migration has contributed to China's development, and in particular, to the rapid growth of cities such as Beijing, Shanghai, Guangzhou and Shenzhen. Andersson (2002) claimed that internal migration could bring numerous benefits to Bolivia, given its low population density, poverty and mountainous regions, by providing services in rural areas.

It is estimated that the number of internal migrants largely exceeds that of international migrants, with internal migration recorded at about three per cent of global population as of 2010 (United Nations, 2013). According to Bell and Muhidin (2011), each of the countries of China, Indonesia, and South Africa, more than 5 million people move between geographic zones over a five-year interval. The estimated lifetime migration intensity is of less than 10% between the states of India, and between the provinces of China and of Indonesia.[19]

Indonesia experiences a lot of internal migration, where one in two people migrated across the regions at least once in their lifetime (Deb and Seck, 2009). From the Indonesia Census of 2010, recent migrants numbered more than 5 million and were dominated by the working-age (i.e. population ages 15 to 64) demographic group (Badan Pusat Statistik, 2011). About 40% of migrants moved for job purposes, either transferring or looking for a job, and about 7% per cent of migrations were related to schooling (Badan Pusat Statistik, 2009).

Migration is a common action for many Indonesian who are unable to gain employment in their local region or home town (Hugo, 2001). In the fifth wave of Indonesian Family Life Survey (IFLS, 2014) sample, 35% of individuals migrated for work purposes, followed by 15% for marriage, 10% for education and training, and 9% to get closer to the family. Similar results were obtained from the IFLS fourth wave (IFLS, 2007). Here, I focus on Indonesia as it has a high regional migration intensity and food consumption is expected to be affected considerably by migration.[20]

[19] Migration intensity is defined as the total number of internal migrants in a given time period as a percentage of the population at risk.

[20] There are a number of channels through migration may influence food consumption patterns. Firstly, migration may bring about additional information on healthy diet practices that can enhance a household’s knowledge about micronutrient values of food. Secondly, positive income effects due to migrant’s remittances may be realised through

Among empirical studies there is little consensus on how migration affects welfare. Although Mendola (2008) found that internal migrant households are poorer than non-migrant households in Bangladesh, Beegle, Weerdt and Dercon (2011) found that migrants have 36% higher consumption than non-migrants. Studying internal migration from Indonesia, Deb and Seck (2009) noted that migration enhances consumption. Despite the inconclusive empirical relationship between household migration and well-being and the uncertainty of the returns from migration, the number of people moving from their birthplace has increased over the years. The human development report 2009 of UNDP recorded that internal migration reached 700 million people worldwide (UNDP, 2009).

In this paper, I investigate two research questions. First, I test the hypothesis that, for a given location, food consumption differs between migrant-sending households and non-migrant-sending households.[21] In other words, I consider Indonesian's migration flows and intensity to investigate whether out migration has a statistically significant impact on food consumption patterns of the migrant-sending-households, and if both per capita food consumption and food group consumption differ between migrants and non-migrants. Second, I test the hypothesis that migrant households consume more carbohydrate-dense foods relative to protein-rich foods than non-migrant-sending households.

The distance from the migrant-households residence of origin to the destination may strongly correlate with the migrant's movement across provinces. By exploiting the IFLS data, I can obtain the log of distance from the respondent's residence to the destination at the time of each interview.

increased per capita expenditure on food and health related products. Finally, migration may affect food security through human capital formation.

[21] A household that stays in one place but one or two family members migrate and send back money or helps original households in some other ways.

It is plausible that distance from the centre of the community where households reside to the destination should not directly affect food consumption. Thus distance may be a promising candidate as an instrument. An instrumental variable approach recognizes that migrants have both different observable and unobservable characteristics. As migration is a choice variable and it's not randomly assigned, so migrant does not constitute a random sample. The decision to migrate could be highly selective in a household, causing selection bias to emerge, which makes it difficult to compare a migrant household with a non-migrant household. The main concern to measure consumption gains from migration is that migrant households also differ from non-migrant households in terms of unobserved characteristics. Using IFLS data, I can examine differences in pre-migration consumption to check whether there is positive or negative selection on the unobservables, conditional of the observed characteristics in the data. Antman (2012) argues that fixed effects may circumvent selection issues without relying on dubious instrument. Comparison of migrants with not-very-similar non-migrants may result in another source of bias McKenzie et al. (2010) claim that matching estimator do this comparison semi-parametrically and minimize that bias.

This study contains two main contributions. First, using rich longitudinal IFLS data and utilising distance from the migrant's birthplace to the migrant's current place of residence destination as an instrument, this study has an exclusive focus on household welfare expressed in terms of food consumption, which may have been affected by internal migration. Specifically, it adds to the literature of distinguished distance-based prediction of migration research by McKenzie, Gibson, and Stillman (2010), Ham, Li, and Reagan (2005), and Batista, Seither, and Vicente (2016). However, a key departure of this paper from the literature is that it looks at the effect of migration on changing food consumption diversification, while the literature considers mainly the effect of migration on earnings or overall consumption. Second, selection bias can be a challenging issue in migration research. To address this, I use propensity score matching (PSM) (Rosenbaum and Rubin, 1983) to estimate the effect of migration on food consumption. This study also adds to the growing literature on matching methods (Atkin, 2016; de Brauw, Mueller, and Woldehanna, 2013; McKenzie et al., 2010; Molini, Pavelesku and Ranzani, 2016) on the importance of migration to food consumption (as an overall welfare measure) of the people in developing countries.

Applying IFLS 2007 and IFLS 2014, I find three main empirical characteristics about the relationship between migration and food consumption. First, applying Propensity Score Matching (PSM) and Instrumental Variable (IV) approach to account for endogeneity of endogeneity of migration, the results indicate that migration has a positive impact on overall per capita food consumption. Second, migration appears to cause shifting of consumption from carbohydrate-rich food to protein-rich food. Third, migration has significant and noticeable impact on diversifying food consumption.

The structure of this chapter is as follows. Section 3.2 describes the review of related literature; Section 3.3 highlights the migration decision and migration-food consumption linkages; Section 3.4 presents data and methodology; Section 3.5 presents results and discussion; and Section 3.6 does the robustness check; and Section 3.7 ends with conclusions and policy issues.

3.2 Literature review

In the past few decades, there has been much attention on investigating the impact of internal migration on economic development following the earliest work on migration from Sjaastad (1962), Todaro (1969), and Harris and Todaro (1970). Sjaastad (1962) studied costs and returns to migration in a resource allocation framework and viewed that migration cannot be explained in isolation from private costs and monetary and non-monetary returns. His work also emphasised the role of gross rather net migration as more relevant for studying the impact of migration. The standpoint of Harris–Todaro model is that potential migration is a function of rural–urban wage differentials and migration takes place when urban returns are greater than rural returns and will continue until net returns for marginal migrants equals zero.

Nguyen and Winters (2011) examined the relationship between migration and food consumption patterns in Vietnam using household panel data in 2004 and 2006. They found that short-term migration has a positive impact on per capita food expenditure and food diversity, which implies that facilitating short-term migration may improve food security in Vietnam. Karamba, Quinones and Winters (2011) explored the link between migration and food consumption patterns in Ghana based on a cross-sectional living standards survey in 2005/2006. Their results indicate that overall, migration does not significantly affect total food expenditure per capita and has a minimal impact on food consumption patterns. However, investigations for the selected household samples of interest (where networks and culture of migration is well-established) show that only in high-out-migration regions (40% or more households have migrants), migrant-sending-households seem to shift consumption to less nutritious food such as sugar and beverages.

Atkin (2016) investigated the caloric costs of culture in Indian migrants using two cross-sections of a national survey. He found that interstate migrants consume fewer calories per rupee of food expenditure than non-migrants or locals. He also documented that caloric intake gap of locals and migrants depends on the intensity of the migrants' origin birth-of-state preferences. Deb and Seck (2009) measured the returns to migration in Indonesia and Mexico and found that migration can improve socioeconomic status via the increase in consumption, but is harmful to the health and emotional well-being of migrants.

Although the idea that migration can improve household welfare has strong intuitive appeal, the empirical findings on the welfare improving effects of migration are inconclusive. For example, Litchfield and Waddington (2003) found that migrant households have statistically significant living standards in terms of consumption levels than non-migrants. However, there seem to be no differences between migrants and non-migrants when other measure of non-monetary metric of welfare such as education is considered. Unlike Litchfield and Waddington (2003), Boakye-Yiadom (2008) accounted for the non-random selection of migrants and concluded that urban migration could improve consumption of internal migrants. Ackah and Medvedev (2010) estimated that conditional on remittances, migration significantly and positively increases household's consumption for the households who sent migrants to urban areas. They accounted for the non-random selection of migration using district level migration rate as an exclusion restriction. Using the same set of data, Adams et al. (2008) found that an increase in remittances reduces poverty and increases inequality. They took into account the issues of selection by exploiting variations in migration networks among several ethno-religious groups. Karamba et al. (2011) investigated the impact of internal migration on food consumption in Ghana and find that migration does not affect overall food consumption per capita and reduces consumption of meat, fish, vegetables, and fruits. They found a positive effect of migration on consumption only in regions with intensive rate of migration. Finally, Adams and Cuecuecha (2013) showed that households in receipt of remittances spend marginally less on food and marginally more on non-food items such as education, housing, and health.

Using longitudinal South Indian village data, Rosenzweig and Stark (1989) studied the movement of women due to marital arrangements in a distinct route. They found that marital migration decreases the variability of household food consumption significantly. In Vietnam, de Brauw and Harigaya (2007) documented household consumption growth and found that seasonal migration can contribute 5.2 percentage points of annual expenditure growth and migration accounts for a 3 percentage point increase in headcount poverty.

Vidyattama (2014) applied the growth model to inter-provincial migration and regional growth in Indonesia between 1975 and 2005 and found that migration played no role in regional convergence during that period. However, in-migration within poorer provinces had a positive impact on economic growth in Indonesia. Other evidences are also suggestive to boost regional economic growth in the growing economies (Resosudarmo and Vidyattama, 2006; Ozgen et al., 2010; Mountford, 1997; Chami et al., 2005). While these studies concentrated on the impact of in-migration on regional economic growth, my study is an attempt to estimate the impact of out-migration on the consumption growth in the migrant-sending households.

A central methodological concern with the existing literature that studies the impact of migration on household's outcome (e.g. consumption or investment) is that migrants are not randomly selected across households. Therefore, the observed relationship between migration and food consumption could be driven by a third unobserved variable. For instance, households that have members with unobserved network-building capacity could have more migrants, receive more remittances, and thus have higher consumption per capita.

A large number of studies have investigated the returns to migration. In these studies, the main estimation issue is the problem of self-selection (Grogger and Hanson, 2011). Previous studies have used FE in panel data (Beegle, De Weerdt, and Dercon, 2011), selection correction methods (Barham and Boucher, 1998), matching (Gibson and McKenzie, 2010), IVs (McKenzie and Rapoport, 2007; Yang 2011) and natural policy experiments (McKenzie, Gibson and Stillman, 2010; Bryan, Chowdhury, and Mobarak, 2014)) to estimate the causal impact of migration. My study is an application of both IV and matching techniques to address for selection and omitted variable bias and my work is closely related to McKenzie, Gibson and Stillman (2010).

3.3 Migration decision and migration-food consumption linkages: conceptual issues

The basic tenet of consumer theory is that people choose among various consumption possibilities to maximise utility. Migration can be conceptualised by consumer theory with two adjustments. First, migration is typically undertaken by all members of a household; therefore, in empirical research, the household is defined as the decision-making unit of the migration process and the relevant motivating factors for migration, such as consumption of goods and services, are defined for the entire household unit (Wallace, DeLorme, and Kamerschen, 1997). Second, the household's net utility is obtained from the modified consumption bundles resulting from migration after adjustment for any disutility associated with alternative food bundles and new location characteristics.

It is assumed that the migration decision is mainly a household's decisions rather individual's choices. The decision to migrate is equivalent to choosing a different location and consumption refers to location-specific selection of food bundles. However, if a migrated individual strongly prefers food bundles from his origin, he does not bring his preferences in the new location and he prefers to consume original food bundles. The other important points are that migrated individuals may send remittances to their residence of origin, share health knowledges with them, and may motivate them to diversify food consumption. These are very crucial factors to motivate a person in a household to migrate than just consumption itself. Location-specific factors, such as proximity to the market and district, migration network, and climatic conditions, affect food choices considerably.

The decision to migrate by a household may take place in three stages (Mincer, 1978; Rossi, 1980). The first stage deals with whether a person considers migration or not. The possible reasons to move out from the residence of origins are job loss, employer relocates, dissatisfaction to the current place. The second stage deals with the collection of relevant information about the potential destinations. The main drivers in this stage are formal and informal networks, characteristics of new location, and job prospects. In the final stage, a destination is selected which gives the best possible benefits after comparing opportunities from the alternative locations. In the last stage, opportunity costs and relative attractiveness are considered to have final selection of the place.

There are several reasons why migration may affect food consumption. First of all, the increase in income resulting from remittances sent by migrants to the family of origin induces the purchase of more goods and services, and thus, leads to increase in overall consumption for the migrant-sending households. Secondly, due to exposure to additional household's knowledge on health and nutrition, migrant households may spend more on the food items, which are more healthy and nutritious (Karamba et al., 2011). Thirdly, migration may have indirect impact on reducing food insecurity of the migrant-sending-household through remittances when migrant-household's encounter natural disasters or any other shocks (Lucas and Stark, 1985; Hoddinott, 1992). Fourthly, due to migration, food consumption expenditure of the migrant-sending households may decrease due to shrinking of the size of the household. Finally, farm production in the rural areas may be hampered due to out-migration, specifically in the cropping seasons.

3.4 Data and methodology

3.4.1 Data

The data are taken from the recent two IFLS waves: IFLS 2007 and IFLS 2014. The IFLS is a household-level, ongoing longitudinal, socioeconomic, and health survey. According to the first IFLS wave in 1993, the sample of households represent about 83% of the Indonesian population living in 13 of the nation's 26 provinces, and the same number of provinces are surveyed across all the subsequent surveys, including IFLS 2014.[22] The survey collects data from the individuals, their families, their households, and their neighbouring or local communities (Strauss, Witoelar, and Sikoki, 2016).

The IFLS 2014 was fielded in late 2014 and early 2015 and became publicly available in April 2016; 16,204 households and 50,148 individuals were interviewed. The IFLS 2007 was fielded in late 2007 and early 2008 and the data were released to the public in April 2009; 13,535 households and 44,103 individuals were interviewed (Strauss, Witoelar, Sikoki, and Wattie, 2009). The two panel data sets are stacked to create a combined household-level data set. Overall, about 29,000 households are available in the combined data set for the final analysis.

The IFLS contains an extensive consumption module that records expenditure in Indonesian rupiah made on 37 individual foods. The individual food items were aggregated into eight food groups: staples, vegetables and fruit, meat and fish, dairy products, dried foods, condiments/spices, beverages and other foods. For each food item, the survey covers consumption expenditure during the past 7 days, including the estimated values of self-produced consumption, any food assistance received from the government program, and other sources. Total food expenditure is calculated as the annual consumption expenditure on the eight food groups. Per capita food group expenditure is then constructed by dividing total expenditure on each food group by the number of household

---

[22] Four provinces on Sumatra (North Sumatra, West Sumatra, South Sumatra, and Lampung), all five of the Javanese provinces (DKI Jakarta, West Java, Central Java, DI Yogyakarta, and East Java), and four provinces covering the remaining major island groups (Bali, West Nusa Tenggara, South Kalimantan, and South Sulawesi). Although there are currently 34 provinces in Indonesia, the number of IFLS provinces (equal to13) is the same across all surveys to ensure comparability across panel households.

members. Finally, the share of each food group is calculated by dividing the expense on each food group by total food expenditure. The construction of food groups is shown in Appendix table B3.1.

Administratively, Indonesia is divided into 34 provinces (*provinsi*).[23] The provinces are divided into districts/cities (*kabupaten* or *kota*). There are 405 regencies and 97 cities. Further, districts/cities are divided into 6543 subdistricts (*kecamaten*). The smallest geographical unit is the village (*desa* or *kelurahan).* There are 75,244 villages in Indonesia. This analysis is based on migration at the provincial level (the largest geographical and administrative unit).

This paper explains migration at the household level. The IFLS survey defines a migrant as a household member who at the age of 12 or higher moves outside the locality of residence for more than 6 months.[24] There are also some return migrants in the households; however, this study considers only out-migrants who are away from the household and have not yet returned. A household would be considered a migrant-sending household (i.e. a migrant household) if there is at least one member out-migration recorded in both 2007 and 2014 and that member has not returned at the time of the survey. (Households with current migrants will be referred to as migrant households unless otherwise defined.) For the same locality, if none of the members has been recorded as a migrant in the household, it is defined as a non–migrant-sending household (i.e. non-migrant households).

Table 3.1 shows recent trends in people movement based on the IFLS dataset. It can be observed that even though migration has slowed overall, there is an increasing trend of moving longer distances. While across-district movement declined by 8.5% within the same province between

[23] North Kalimantan is a new province created in 2012 after the 2010 census.

[24] Locality of residence is defined at the community level (i.e. EA in the surveys).

2000 and 2010, the movement across districts between different provinces increased from 50.8% to 53.5%.

Table 3.1: Recent internal migration trends in Indonesia

	Population 5 years and older that moved during the past 5 years				
	2000	2010	% change 2000-2010	2000	2010
Inter-district migrants	10,703,029	9,791,734	-8.5%	5.9	4.6
Intra-provincial migrants	5,262,790	4,555,956	-13.4%	2.9	2.1
Inter-provincial migrants	5,440,239	5,235,778	-3.8%	3.0	2.4
Inter-provincial migrants as % of inter-district	50.8%	53.5%			
Intra-corridor migrants	3,715,660	3,325,121	-10.5%	2.1	1.6
Inter-corridor migrants	1,724,579	1,910,657	10.7%	1.0	0.9
Inter-corridor migrants as % of inter-provincial	31.7%	36.5%			

Notes: The data are collected from the 2000 and 2010 censuses in Indonesia. Under the master plan for Acceleration and Expansion of Indonesia Economic Development, Indonesia is divided into six economic corridors: Sumatra, Java, Kalimantan, Sulawesi, Bali and Nusa Tenggara, and Maluku and Papua. Each corridor's own economic activities induce migration flows. From the geographical point of view, inter-economic corridor (inter-corridor) and inter-island migration are synonymous. *Source*: Sukamdi and Mujahid (2015)

Table 3.2 provides the sample's summary statistics in the data set. About 43% of households in the combined 2007 and 2014 survey data are migrant households. Of the households in the full sample, about 42% are rural households and about 43% are rural migrant households.

Table 3.2: Sample summary statistics

	No. households	Proportion of full sample households	Proportion of migrant sample households	Proportion of rural sample households	Proportion of rural migrant sample households
Full sample	29,457	1.000	0.4593	0.4240	0.4091
Households moved for work sample	4,242	0.1440	0.3135	0.3396	0.8302
Households moved for marriage sample	2,785	0.0945	0.2058	0.2229	0.5451

Notes: The table shows the proportion of sample households used in paper. All proportions are with respect to the full sample of households for each row category. The total number of households is 29,457, which is the sum of sample in IFLS 2007 and IFLS 2014 surveys.

3.4.1.1 Comparison of food consumption behaviour of migrants and non-migrants

Table 3.3 displays the behaviour of migrants and non-migrants within the same district. This analysis largely follows the set of controls used by Subramanian and Deaton (1996), Karamba *et al.* (2011), Thomas and Frankenberg (2007), and others under the Indonesian context. Each row of household-level variables is regressed on the migrant household dummy and on a district FE and the coefficient on the migrant indicator variable reported. Compared with other households in the district, migrant households have 21.7% higher food expenditure and 20.33% higher total expenditure.

Table 3.3: Summary statistics (mean) and differences between migrant and non-migrant households

	Mean (full sample)	Migrant difference (full sample)	Migrant difference (HH member moved for work sample)
	(1)	(2)	(3)
Log household per capita food consumption	10.9731	0.3537***	0.0588***
	(0.8786)	(0.0103)	(0.0173)
Log household per capita expenditure	17.0482	0.2033***	0.0424**
	(0.9702)	(0.0115)	(0.0203)
Log household food expenditure	14.3211	0.2170***	0.1193***
	(1.8159)	(0.0216)	(0.0338)
Log household size	1.5351	-0.2241***	-0.1044***
	(0.6280)	(0.0074)	(0.0140)
Proportion males 0-4	0.0403	0.0266***	-0.0113***
	(0.0932)	(0.0011)	(0.0019)
Proportion females 0-4	0.0375	0.0247***	-0.0117***
	(0.0903)	(0.0010)	(0.0018)
Proportion males 5-9	0.0367	0.0266***	-0.0063***
	(0.0932)	(0.0011)	(0.0014)
Proportion females 5-9	0.0342	0.0247***	-0.0043***
	(0.0818)	(0.0010)	(0.0014)
Proportion males 10-14	0.0335	-0.0211***	-0.0003
	(0.0851)	(0.0009)	(0.0012)
Proportion Females 10-14	0.0327	-0.0209***	-0.0015
	(0.0847)	(0.0009)	(0.0011)

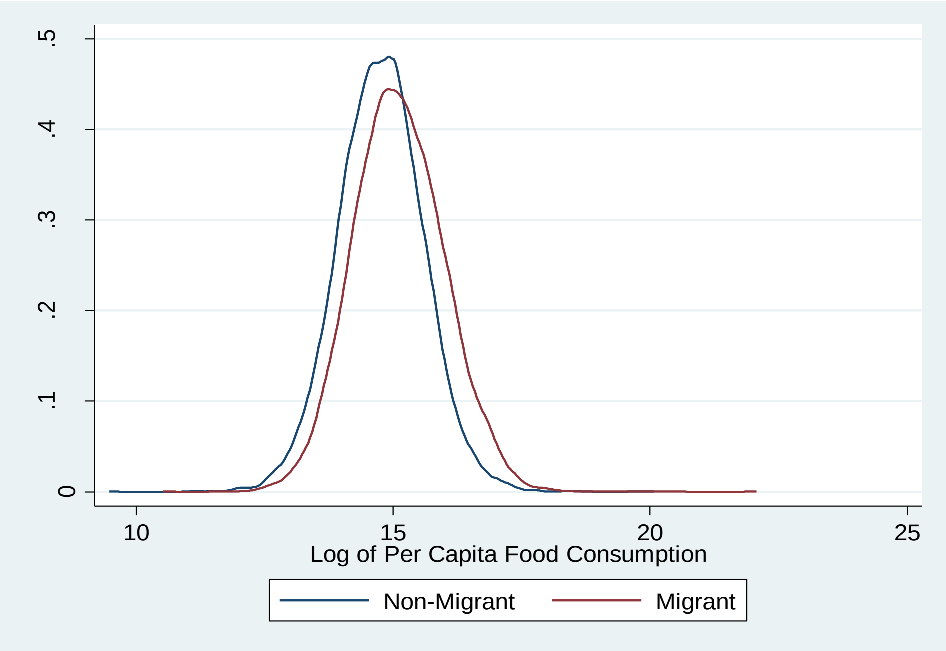
	Mean (full sample)	Migrant difference (full sample)	Migrant difference (HH member moved for work sample)
	(1)	(2)	(3)
Proportion males 15-55	0.3210	0.0516***	0.0646***
	(0.2091)	(0.0024)	(0.0048)
Proportion females 15-55	0.3210	0.0311***	-0.0193***
	(0.1924)	(0.0022)	(0.0042)
Proportion males 55+	0.0427	-l420.0192	0.0003
	(0.0929)	(0.0010)	(0.0015)
Proportion females 55+	0.0544	-0.0309***	-0.0061***
	(0.1149)	(0.0012)	(0.0015)
Age of HH head	43.1603	-10.5349***	0.8617
	(23.7058)	(0.3828)	(0.6230)
Age of HH head squared	2,424.7460	-775.3103**	49.5185
	(18,425.6400)	(305.5496)	(520.5111)
Education of HH head	5.1929	0.3593***	0.5370***
	(2.6431)	(0.0440)	(0.0617)
Female-headed household	0.0698	0.0188***	0.0008
	(0.2549)	(0.0030)	(0.0050)
HH head married	0.8682	-0.1683***	-0.0925***
	(0.3381)	(0.0056)	(0.0111)
Employment of HH head	0.4745	0.0551***	0.0402***
	(0.4993)	(0.0058)	(0.0095)
Religion of HH head			
Islam	0.4434	0.0524***	0.0094
	(0.4968)	(0.0058)	(0.0092)
Hindu	0.0223	0.0024	0.0039
	(0.1478)	(0.0017)	(0.0027)
Protestant	0.0221	0.0122***	0.0013
	(0.1471)	(0.0017)	(0.0031)
Ethnicity			
Javanese	0.2187	0.0176***	0.0437***
	(0.4134)	(0.0048)	(0.0079)
Sundanese	0.0634	0.0013	0.0126***
	(0.2437)	(0.0028)	(0.0046)
Minang	0.0234	0.0141***	-0.0037
	(0.1514)	(0.0018)	(0.0031)
Batak	0.0230	0.0164	0.0025
	(0.1499)	(0.0017)	(0.0033)
Sasak	0.0217	0.0022	0.0032
	(0.1460)	(0.0017)	(0.0027)

	Mean (full sample)	Migrant difference (full sample)	Migrant difference (HH member moved for work sample)
	(1)	(2)	(3)
HH experienced any natural disaster	0.2247	0.0045	0.0110
	(0.4174)	(0.0049)	(0.0077)
Log of household asset	17.5879	0.0312***	-0.1472***
	(1.7768)	(0.0230)	(0.0359)
Log of distance to the destination (km)	3.1195	-122.2644***	46.3725***
	(1.9629)	(0.7854)	(1.7042)
Region (Base: DKI Jakarta)			
North Sumatra	0.0715	0.0300***	-0.0056
	(0.2577)	(0.0030)	(0.0051)
West Sumatra	0.0442	0.0120***	-0.0087
	(0.2056)	(0.0024)	(0.0039)
South Sumatra	0.0444	0.0132***	-0.0246***
	(0.2059)	(0.0024)	(0.0037)
Lampung	0.0392	0.0024	0.0240***
	(0.1942)	(0.0022)	(0.0049)
DKI Jakarta	0.0684	0.0016	0.0240***
	(0.2528)	(0.0029)	(0.0049)
West Java	0.1506	-0.0181***	-0.0043
	(0.3577)	(0.0041)	(0.0064)
Central Java	0.1210	-0.0178***	0.0039
	(0.3261)	(0.0037)	(0.0058)
DI Yogyakarta	0.0525	-0.0100***	-0.0081**
	(0.2230)	(0.0025)	(0.0038)
East Java	0.1403	-0.0338***	0.0008
	(0.3473)	(0.0040)	(0.0060)
Bali	0.0472	-0.0072***	0.0124***
	(0.2121)	(0.0024)	(0.0039)
West Nusa Tenggara	0.0661	-0.0020	-0.0090**
	(0.2486)	(0.0029)	(0.0044)
South Kalimantan	0.0448	-0.0007	-0.0149***
	(0.2068)	(0.0024)	(0.0035)
South Sulawesi	0.0473	-0.0033	-0.0049
	(0.2124)	(0.0024)	(0.0037)

Notes: Column 1 shows the mean of each household-level variable in each row. Column 2 displays the estimated coefficients on a migrant dummy when the variable is regressed on a province-year FE and migrant-status dummy. Column 3 displays the coefficient for the same regression but for the members in the households who moved to another location for work. Standard deviations are shown in parentheses in column 1 and robust standard errors are shown in the parentheses in columns 2-3. In all regressions, errors are clustered at the community level.
***Significant at 1% level; ** Significant at 5% level; *Significant at 10% level

Figure 3.1 below shows per capita food consumption based on the household migration status.

Figure 3.1: Per capita food consumption by household migration status



The above diagram implies that there is a subtle difference between per capita food consumption of migrant households and non-migrant households. On average, per capita consumption is higher for migrant households and the difference in consumption is observed throughout the distribution. The kernel density of the logarithm of per capita food consumption for migrant households has been shifted to the right relative to that of non-migrant households.

Figure B3.1 in appendix also shows the kernel density of log of per capita food consumption for each year in the panel data. Per capita food consumption is expressed in a log form as it tends to follow log-normal distribution. The figure B3.1 implies that that per capita food consumption is well approximated by the assumption of log-normal distribution. The same figure also shows that the distribution of food consumption has shifted to the right over the two periods, indicating a general increase in living standards for the panel households.

3.4.2 Empirical methodology

Model 1: The migration decision model

Following Nguyen, Raabe and Grote (2015), I estimate a non-linear probability model that links the household migration status in 2014 and 2007 to both household and community characteristics in 2007. The model is stated below as Equation 1:

$$\Pr\left(M_{ij,2014,2007}\right) = F(X_{ij,2007}, Z_{j,2007}, FE_{Prov}) \qquad (1)$$

The dependent variable expresses the probability that household i in community j is a migrant household in 2007 and/or 2014. The dummy variable, $M_{ij,2014,2007}$, is equal to 1 if household i in village j had at least one migrant household member in 2007 and/or 2014. The vector $X_{ij,2007}$ of household i in community j shows a set of observable household's characteristics in 2007. The vector $Z_{j,2007}$ includes a set of observable community characteristics. Finally, province fixed effects, FE_{Prov}, unobserved province heterogeneity in out-migration.

Model 2: The impact of migration on food consumption: Propensity Score Matching (PSM)

Following Heackman and Navarro-Lozano (2004), Equation 2 below quantifies and compares outcome (i.e. food consumption) of households with migrants against that of comparable households with non-migrants.

$$ATT = E(Y_{1i} - Y_{0i})|M = 1) = E(Y_{1i}|D = 1) - E(Y_{0i}|D = 1) \qquad (2)$$

where ATT refers to the average treatment effect on the treated, which measures the impact of migration on the outcome of migrant households. M is a dummy variable that equals 1 if the household has at least one migrant and zero, otherwise. Y_{1i} and Y_{0i} refres the outcome of household i with and without migrants, respectively.

It is not feasible to find the outcome of the migrant household, where no one is actually migrated ($Y_{0i}|D = 1$) which is unobserved. This chapter employs PSM to solve this problem. I employ migration decision model in Equation 1 to estimate propensity score. I use Nearest-Neighborhood and Kernel matching methods. These methods would estimate the outcome that migrant household would have had no household members ever migrated.

One of the standard problem of PSM method is that it only controls for selection on unobservable, but cannot account for unobserved variables that may affect the probability of migration and the outcome variable. I use the following Equation 3 to address this type of endogeneity.

$$ATT = [Y^1_{2014} - Y^1_{2007}|X_{2007}, M = 1] - [Y^0_{2014} - Y^0_{2007}|X_{2007}, M = 0] \qquad (3)$$

Model 3: The impact of migration on food consumption: OLS and IV

I test the hypothesis that the food consumption pattern of migrant-sending households is different from that of non-migrant-sending households living in the same locality. The empirical specification follows Atkin (2016), Subramanian and Deaton (1996), and de Brauw and Giles (2008):

$$Foodcon^{(k)}_{it} = \beta_0 + \beta_1 migrant_{it} + X'_{it}\alpha + \lambda_i + \delta_t + p * t + \varepsilon_{it} \qquad (4)$$

where $Foodcon^{(k)}_{it}$ indicates three different measures of food consumption (k) in household i at time t: log of per capita food consumption, budgetary shares of six food groups (staple foods, vegetables and fruits, meat and fish, dairy products, dried foods, and spices/condiments), and a measure of food diversity. $migrant_{it}$ is a binary variable that is equal to 1 if at least one member of the household migrated at time t and 0 otherwise. X_{it} is a vector of other household-level controls that contain demographic and characteristics variables. Conditional on household-level controls X_{it}, the main parameter of interest is β_1: if positive (negative), it implies that migrant households consume more (less) than non-migrant households on average. λ_i indicates household-specific unobserved time-invariant variables that affect food consumption for any given level if there are out-migrants in the household. $\delta_t$captures the common time-specific effect. I further include a vector of province-year interactions dummies, $p * t,$ to absorb any province-specific macroeconomic shocks. ε_{it} refers to household-specific error term.

3.4.2.1 Endogeneity of migration

Three major endogeneity concerns are: *omitted variable bias*, *reverse causality*, and *self-selection*. These issues are explained below.

One of the core threats to identification (i.e. causality from migration to food consumption) is that migrants are not usually randomly spread across labour markets. This is mainly because migrants select their locations in function of the characteristics of the local labour market of destination. If such characteristics (unobserved) are correlated with the outcome of interest and cannot be controlled for in the main estimating equation, then *omitted variable or simultaneity bias* might arise. It might also be plausible that migrants choose regions with lower density of migration or more job prospects. This could create a spurious correlation between migration and outcomes.

Decisions on migration, education, labour supply, and other economic choices are made simultaneously by the household. Therefore, any household characteristics that might influence migration decisions could also affect household consumption. Most importantly, many household's characteristics are unobservable (e.g. motivation or ability or effort). These issues make it difficult for us to estimate the effect of migration and its outcome as these unobservables could bias our OLS estimates.

The main challenge in estimating the impact of migration on food consumption is the possibility of unobserved characteristics of individuals in the households that affect the migration decision also affect food consumption (Molini, Pavelesku, and Ranzani, 2016). For instance, individuals with higher unobserved abilities are more likely to migrate so as to earn more, and thus, consume more. Irrespective of migration, the same unobserved factors that induce them to migrate may also have an effect on food consumption. If this is the case, then simple OLS comparison of the outcomes of migrants and non-migrants may overestimate the gains in food consumption due to migration. One may also argue that natural disasters (e.g. floods or earthquakes) can force individuals to migrate and at the same time may reduce their consumption. This leads to a spurious correlation between migration and food consumption.

Relating to our estimating Equation 1, there is a concern that the migration variable is endogenous. For instance, migrant-sending households may have unobserved characteristics or unobserved heterogeneity that could affect food consumption expenditure. This induces omitted variable bias due to the potential correlation between error terms and regressors. Some unobservable individual characteristics that affect migration may also affect both the levels and shares of household food expenditure. For example, parents who place greater value on their children's nutrition may be motivated to migrate as a family or send their children away to earn more money. Consequently, they may allocate more to food consumption and/or to food diversity. Karamba et al. (2011) noted that migrant households are likely to take risks and to purchase new possessions. In this case, the OLS estimate is likely to overestimate the impact of migration on food consumption and absorb both the impact of migration and the positive traits of migrants.

The second estimation issue is *reverse causality*. For instance, individuals who have a higher propensity to consume may choose to migrate to areas with more opportunities, and thus, consumption may reverse cause migration. This is likely to create an upward bias estimate of the migration variable.

The third estimation issue is *self-selection*, which causes selection bias. Because migrants may differ fundamentally from non-migrants in terms of observables and unobservable characteristics, they self-select and selection bias arises. Thus, it is not possible to determine what would happen to non-migrant households if they were to migrate, from observing the experience of migrant households. For instance, more educated individuals may migrate to exploit higher wages, leading to positive selection, and individuals with strong ethnical and cultural backgrounds may less be likely to migrate, leading to a negative selection.

3.4.2.2 Estimators

To address endogeneity of migration due to omitted variables, reverse causality, and self-selection, I employ the IV method. The instrument generates what we hope are plausibly exogenous variations in the endogenous variable, i.e. migration, such that it affects the outcome variable, i.e. food consumption, only through the migration variable. As well, another assumption is that the instrument has to be orthogonal to the error term.[25] In other words, an instrument should be relevant (strongly correlated with the explanatory variable) and exogenous (may not be correlated with the outcome variable other than through the explanatory variable). The first stage estimates the relationship between the endogenous variable (migration) and the instrument (e.g. the migration network) and isolates the error-free component of migration that is correlated with the

[25] The instruments are a subset of exogenous variables. The exogenous variables that are not used as instruments are said to be exogenous covariates (Angrist and Pischke, 2009).

error term. In the second-stage, the error-free predicted migration is used to estimate its coefficient.[26]

To account for the potential problem of endogeneity of migration, I use two types of instrument for migration: migration networks and distance from the migrant's residence of origin to the destination. Several well-established branch-of-migration studies have used migration networks or district-level migration rates as instruments for migration (de Brauw, Mueller, and Woldehanna, 2013; Carrington et al., 1996; Karamba *et al.*, 2011; Molini, Pavlesku, and Ranzani, 2016; Nguyen and Winters, 2011; Stark and Taylor, 1991; Winters et al., 2011). I follow this literature to construct the migrant networks as an instrument for current migration. Arguably, the migration networks are likely to influence migration, but not directly affect household-level food consumption. In particular, the district-level migration rate (i.e. migration networks) is simply the percentage of adult labourers who participate in migration, excluding the household being analysed. Following a unique contribution to migration literature by McKenzie et al. (2010), I use distance from the migrant's birth place to the destination as another instrument for migration.[27] This distance is recorded in the migration module of the IFLS questionnaire. In particular, I use the log of the distance from migrants the previous place of residence to the destination place as an exclusion restriction for migration in food consumption estimation.

[26] In Karamba et al. (2011), the relationship between endogenous variables and instruments is explained precisely.

[27] McKenzie et al. (2010) used distance from the New Zealand consulate in Tonga as an instrument for migration when investigating the impact of migration from of Tongan people to New Zealand. In a recent paper, Batista, Seither, and Vicente (2016) have used household size, natural catastrophes, and networks link with shortest path link as an instrument for migration they found that network instrument with the shortest path link is strong predictor of migration.

The distance between the migrant's origin and destination locations is significantly correlated with migration (the instrument is relevant), with first stage excluded *F*-statistics of 23 for Indonesia as a whole and 30 when the sample is restricted to East Java. It seems plausible that distance from the locality of the migrant's residence to the potential destination should not directly affect food consumption in Indonesia. However, a concern for an identification of a food consumption regression on migration is the possibility that this distance affects earnings and thereby affects consumption within a province. In Indonesia, individuals living in the outer province have a different earning potential than individuals from the mainland province, where households are located closer to the migration destination or closer to the workplace. Thus, this instrument is more likely to be valid when the sample is restricted to the main islands.

One of the threats to identification when 'migration network' is used as an instrument that it can be both related to migration and to the income level of the community. For instance, it may be true that poorer regions are more likely to send migrants therefore have more extensive migrant networks. Then migrant network would be both related to migration decision of the household and consumption level of the household. To address this issue, log of households' annual income is used as a control variable and a separate regression is performed. Table B3.4 in the appendix presents the results when income is controlled for both in OLS and IV regressions.[28]

Another threat to identification may arise when 'distance' is used as an instrument in food consumption-migration regression. Specifically, the distance between the migrant's birth place and the destination could be self-selected based on unobserved migrant household's characteristics. To address this issue, households fixed effect is applied in the regression as

[28] There are seven main sources of income recorded in IFLS data: crop income, employment, livestock, other rental, agricultural rent, self-employment and transfers. I have only used public transfer and private transfer is not included in income calculation as many observations are missing.

distance is time-invariant. As two years' survey data (IFLS 2007 and IFLS 2014) have been used, first difference has been applied in the estimation which is equivalent to fixed effect estimation.

To account for the third endogeneity issue of migration (i.e. potential self-selection), a standard econometric approach of Angrist and Pischke (2009) is used. I try to control for as many household characteristics as possible, as they determine decisions to migrate, and how far they migrate. Angrist's (1998) conditional independence assumption implies that conditional on all observed characteristics, migrants and non-migrants are comparable.[29] Conditional on X_{it}, variation in migration status comes solely from the fact that a potential migrant may not ultimately move to a new location.

The fundamental problem in identifying the impact of migration on food consumption is the lack of a counterfactual. The construction of counterfactuals requires comparing a migrant household with an 'identical' household that has not migrated. In reality, an identical household does not exist. One response to this is to use Propensity Score Matching (PSM) (see Rosenbaum and Rubin, 1983), where a migrant household is compared with a non-migrant household with the same propensity to migrate. Matching is a non-experimental evaluation technique that has recently attracted enormous research interest. Proponents of PSM claim that this method can replicate experimental benchmarks when used appropriately (Dehejia and Wahba, 2002; Dehejia, 2005). To estimate propensity scores, I use STATA programs developed by Becker and Ichino (2002), which implement the algorithm suggested by Dehejia and Wahba (2002).

The intuition and estimation procedure of matching can be described as follows. First, I start by estimating a Probit regression on the migrant dummy. In this Probit regression, I use the following

[29] Angrist (1998) pointed out that controlling for covariates can make CIA more plausible.

control variables: age, sex, and educational attainment of the household head; log of per capita household expenditure; household's location (rural or urban); and log of household assets and migration networks. The Probit regression generates a propensity score for each household. Based on these scores, a migrant household is matched with a non-migrant household in the same locality. The result of the matching procedure is based on the two generated samples, one sample of each migrant and non-migrant groups. Based on the observable characteristics that are controlled in the Probit regression, the two samples are statistically indistinguishable. This suggests that the matched pairs have almost the same probability of receiving the treatment (migration). However, one group receives the treatment (migrated) and the other does not (not migrated). This procedure generates a pseudo-random sample and based on these samples, households are randomly assigned to both the treatment and control groups. Hence, any resulting differences between the treatment and control groups would reflect the treatment effect and the differences are not influenced by any pre-existing individual characteristics.

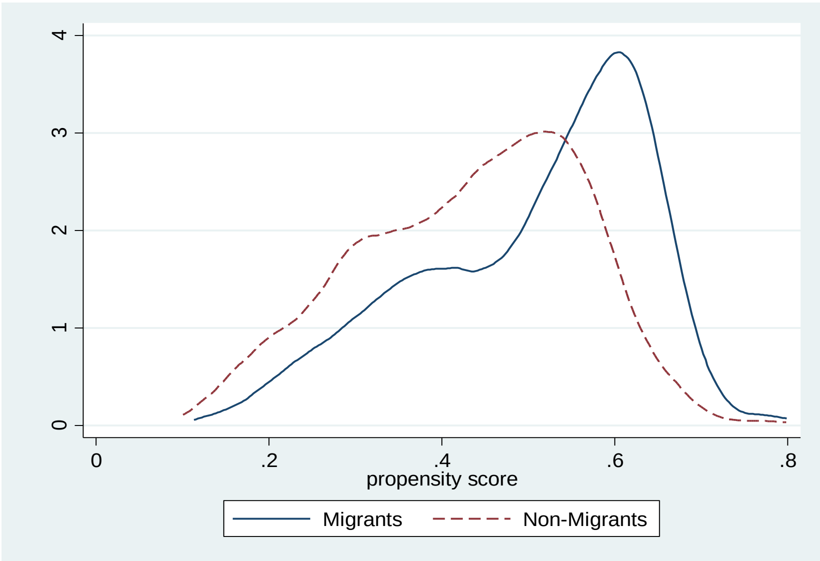
3.5 Regression results and discussion

3.5.1 Migration and food consumption patterns: PSM and OLS

This section starts with presenting balance tests for Propensity Score Matching (PSM) which is presented in appendix Table B3.3. The table displays the mean, the standardized bias, reduction in bias and the *t-test* of the selected covariates of the treatment and control group. The results show a significant improvement in balance. The reduction of bias for household size, age of head, and employment status of the head have reduced by more than 85% after matching. Except for sex of head, other covariates also show the reduction from 28% to 50%.

Figure 3.2 shows the distribution of propensity scores of migrants and non-migrants households. The results of the Probit estimation are used to predict propensity scores and display the kernel density in Figure 3.2. Overlapping in the propensity scores is ensured by plotting kernel densities of the propensity scores among both migrants and non-migrants.

Figure 3.2 Estimated density of propensity to migrate, by migrant status, in 2014, Indonesia



The graph in Figure 3.2 displays the estimated density of the predicted probabilities that a non-migrant household is indeed a non-migrant household and the estimated density of the predicted probabilities that a migrant household is indeed a migrant household. Neither plot shows too much probability mass near 0 or 1, nor do the two estimated densities have most of their respective masses in overlapping regions, so there is no indication that the overlap assumption is violated.[30]

The results of the impact of migration on changes in consumption, estimated from Equation 3, are presented in Table 3.4. This table documents the estimates for both Ordinary Least Squares (OLS) and Average Treatment Effect on the Treated (ATT) derived from the method of Propensity Score

[30] To make this claim, I consulted with the Stata command *teffects overlap* in version Stata 14.

Matching (PSM). An advantage of the model represented by Equation 3 is that it is possible to control for individual differences through the vector X, which includes variables such as age and educational attainment. Three types of matching algorithms are applied to get the robust estimates of counterfactuals from the PSM method. *Nearest-neighbor (NN)* matching picks up counterfactual household for each migrant household based on closest proximity. *Caliper* matching is applied where the closest neighbour is far away. Finally, *Kernel matching*, is used which matches each treated unit (migrant households) to a weighted sum of comparison unit (non-migrant households) with the highest weight assigned to units with closer scores. For efficiency, this chapter uses tighter *caliper* set at 0.001.

To provide evidence that the OLS results are not being driven by household-level unobserved characteristics, I employ covariate matching (difference-in-difference matching) methods in Equation 3 as matching methods attempt to control for selection bias. The determinants of migration from the Probit estimation is shown in Table B3.2 in the Appendix and it has been observed that both migration network and distance to the destination are quite strong predictors of migration. It is possible to identify within household differences in food consumption when this method is used. The PSM estimates show that being a migrant household significantly increases the range of expenditure on staple food from 5.7% to 34.1%.

Table 3.4: Impact of migration on food consumption: PSM and OLS

Outcome variable (log of annual expenditure)	PSM estimates (ATT)				OLS estimates	
	Matching Algorithm	Difference	SE	t-test	B	SE
Staple foods	NN Caliper=0.001 with replacement	0.341**	0.102	3.34	0.105***	0.008
	NN Caliper=0.001 without replacement	0.212***	0.014	15.14		
	Kernel	0.057***	0.002	28.5		
Veg and fruits	NN Caliper=0.001 with replacement	0.361*	0.154	2.34	0.285***	0.018
	NN Caliper=0.001 without replacement	0.313*	0.112	2.79		

	Kernel	0.456**	0.115	3.96		
Meat and fish	NN Caliper=0.001 with replacement	0.189***	0.018	10.5	0.108**	0.018
	NN Caliper=0.001 without replacement	0.174**	0.054	3.22		
	Kernel	0.117*	0.051	2.29		
Dairy products	NN Caliper=0.001 with replacement	-0.217	0.118	-1.83	-0.084	0.045
	NN Caliper without replacement	-0.322	0.301	1.07		
	Kernel	0.151	0.116	1.30		
Condiments	NN Caliper=0.001 with replacement	0.188***	0.006	31.33	0.201***	0.017
	NN Caliper=0.001 without replacement	0.203***	0.012	16.92		
	Kernel	0.261*	0.103	2.53		
Food diversity	NN Caliper=0.001 with replacement	0.231***	0.012	19.25	0.154***	0.008
	NN Caliper=0.001 without replacement	0.452**	0.102	4.43		
	Kernel	0.356**	0.107	3.32		

Notes: The first column show the log of annualized expenditures for five main food groups-staple foods, vegetable and fruits, meat and fish, dairy products, and condiments. The final variable in the first column in food diversity. The second column shows the different matching algorithms. The third to fifth column shows the PSM estimates and the derived average treatment on the treated (ATT) effects differences, standard errors and t-tests, respectively. The final sixth and seventh column show coefficients and standard errors for OLS estimates.

***Significant at 1% level; ** Significant at 5% level; *Significant at 10% level

Table 3.4 shows the results for five main food groups and food diversity. Both OLS and PSM estimates suggest that being a migrant in the households positively affects food consumption except for the dairy products. Migration supports migrant households to have increased food access and enhanced food diversity.

3.5.2 Migration and food consumption patterns: OLS and IV

This section presents results for the analysis of per capita food consumption and budget shares of six food groups with regards to the impact of migration on food consumption patterns for the regression specified in Equation 4 in Section 2.5. I first report the ordinary least squares (OLS) estimates. I then report the various instrumental variable (IV) estimates for a single specification and sample. Finally, I present the estimates from the household's fixed effect (FE) model. Under the assumption that the unobserved heterogeneity of migrants (for example, migrants foresight) is constant over time, the household fixed-effect (FE) is conducted, thus removing the household-specific constant components. The household FE would also absorb a host of other time-invariant geographic differences such as distance from the community to the destination of migrants that might affect food consumption. Still there is a possibility that there are some unobserved time-varying factors in the households. To account for the remaining unobserved time-varying factors, an instrumental variable approach is applied so that this can partially remove impact of time-varying confounding factors to estimate the impact of migration on food consumption.

Tables 3.5 presents the coefficient estimates for these three estimation methods, where the dependent variable is log of per capita food consumption. To estimate Equation 4 (shown in Table 3.5), I include a vector of household demographics and household-head characteristics variables of the head of the household (ethnicities and religion), an indicator for urban households, dummies for interactions of the survey years and provinces, and other controls (log of household assets and natural disaster). The controls for household demographics include household size as well as proportions of household members that fall into five sex-age brackets: 0-4, 5-9, 10-14, 15-55, and over 55. The controls for household-head characteristics are age in years and educational attainment, indicator variables for female-headed household, employment, and marital status. The demographics and characteristics variables control for the possibility that as compared with other households in the community, migrants may choose less physically intensive jobs and have different demographic structures. As ethnicities and religion may be cultural determinants of food preferences, I also include these controls. When females control household resources, they favour basic needs and children's welfare. Rogers (1996) found that female-headed households consume more expensive and protein-dense foods. The geographical location also plays an important role in consumption. For instance, urban households may consume more junk foods than homemade foods. In developing countries, because of religious beliefs, people sometimes strongly prefer a specific food. For example, although Christians and Muslims eat animal products, the latter do not eat pork and the Hindu majority avoids beef in India (Atkin, 2016). Finally, province and year dummies are included to control for regional food preferences and to absorb provincial-level macro-shocks.

The discussion begins with the result presented in Column 1 in Table 3.5. Even after controlling for all other characteristics relevant to household food consumption (e.g. welfare), the dummy variable that indicates if a household has at least one out-migrant is positive and statistically significant in the specifications (OLS, IV, and FE). The estimated coefficient (0.126) shown in Column 1 of Table 3.5 suggests that migrant households' average welfare (in terms of food consumption) is higher than average welfare of non-migrant households by 0.126 log-point. This estimate implies that migrant-household consumption is on average 13.4% higher than non-migrant-household consumption.[31] This welfare increase is similar with the descriptive statistics shown in Table 3.3 and also fairly consistent with the findings of Beegle et al. (2011) for Tanzania and Karamba et al.'s (2011) study of migration and food consumption in Ghana. In this study, the coefficient estimates of migration in Indonesia fall in middle of the pack in terms of the magnitude of the finding from, on the one hand, Beegle et al. (2011), in which the coefficient of migration estimates is substantially larger, and, on the other hand, Karamba et al. (2011), in which the coefficient of migration estimates is substantially smaller. It is observed that in all regressions, the larger household size significantly depresses per capita food consumption. In particular, an 11.3 percentage point decrease in welfare due to larger household size is observed in the OLS regression (the estimated coefficient of household size is -0.113, shown in Column I in Table 3.5). This result

[31] The semi-elasticity of a dummy variable (migrant) coefficient in a log-level model is calculated as $(e^{\beta} - 1)$. If we plug the estimated coefficient 0.126, the migrant coefficient becomes $(exp^{0.126} - 1 = 1.134282)$ 0.134282, which is about 13.4%.

is quite reasonable in that the distribution of resources allocated to food consumption reduces when there are more household members to feed. This finding is also consistent with Subramanian and Deaton's (1996) study of demand for food and calories in India. The educational attainment and employment of the household head have the expected impact and both variables significantly and positively affect per capita food consumption. Somewhat surprisingly, there is no impact of the age of the head of household on food consumption.

Table 3.5: Migration and food consumption patterns. (Dependent variable: log of households' per capita food consumption)

	(1)	(2)	(3)	(4)	(5)
	OLS	IV	IV	IV	FE
Migrant	0.126***	0.115***	0.152***	0.114***	0.109***
	(0.013)	(0.021)	(0.092)	(0.021)	(0.013)
Household demographics					
Household size	-0.113***	-0.123***	-0.104***	-0.114***	-0.116***
	(0.004)	(0.004)	(0.008)	(0.004)	(0.004)
Proportion males 0-4	0.583***	0.590***	1.100***	0.591***	0.576***
	(0.108)	(0.108)	(0.385)	(0.108)	(0.111)
Proportion females 0-4	0.911***	0.909***	0.804***	0.909***	0.924***
	(0.109)	(0.109)	(0.149)	(0.109)	(0.112)
Proportion males 5-9	0.824***	0.819***	0.466	0.819***	0.856***
	(0.109)	(0.109)	(0.285)	(0.109)	(0.112)
Proportion females 5-9	1.081***	1.086***	1.449***	1.086***	1.068***
	(0.094)	(0.095)	(0.281)	(0.095)	(0.098)
Proportion males 10-14	0.548***	0.554***	1.018***	0.555***	0.548***
	(0.124)	(0.124)	(0.367)	(0.124)	(0.129)
Proportion females 10-14	0.592***	0.599***	1.136***	0.600***	0.549***
	(0.109)	(0.110)	(0.403)	(0.110)	(0.112)
Proportion males 15-55	0.726***	0.724***	0.621***	0.724***	0.740***
	(0.111)	(0.111)	(0.149)	(0.111)	(0.114)
Proportion females 15-55	0.756***	0.750***	0.322	0.750***	0.798***
	(0.111)	(0.111)	(0.334)	(0.111)	(0.114)
Proportion males 55+	0.797***	0.802***	1.170***	0.803***	0.795***
	(0.093)	(0.093)	(0.285)	(0.093)	(0.096)
Proportion females 55+	0.247**	0.250**	0.418**	0.250**	0.229*
	(0.121)	(0.121)	(0.185)	(0.121)	(0.126)
HH head characteristics					
Age	-0.001	-0.001	-0.009	-0.001	-0.000
	(0.001)	(0.001)	(0.006)	(0.001)	(0.001)
Education	0.022***	0.022***	0.026***	0.022***	0.021***
	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)
Female	0.053	0.052	0.038	0.052	0.058
	(0.036)	(0.036)	(0.044)	(0.036)	(0.037)
Married	-0.146***	-0.147***	-0.258***	-0.148***	-0.117***

	(1) OLS	(2) IV	(3) IV	(4) IV	(5) FE
	(0.031)	(0.031)	(0.086)	(0.031)	(0.032)
Employed	0.531***	0.530***	0.506***	0.530***	0.556***
	(0.036)	(0.036)	(0.044)	(0.036)	(0.036)
Ethnicities					
Javanese	-0.144***	-0.144***	-0.152***	-0.144***	-0.144***
	(0.014)	(0.014)	(0.017)	(0.014)	(0.015)
Sundanese	-0.032	-0.032	-0.058*	-0.032	-0.034*
	(0.020)	(0.020)	(0.030)	(0.020)	(0.021)
Minang	0.149***	0.150***	0.231***	0.150***	0.152***
	(0.028)	(0.028)	(0.066)	(0.028)	(0.028)
Batak	0.034	0.035	0.106	0.035	0.039
	(0.035)	(0.035)	(0.066)	(0.035)	(0.035)
Religion					
Islam	-0.120***	-0.121***	-0.150***	-0.128***	-0.119***
	(0.031)	(0.031)	(0.041)	(0.031)	(0.032)
Hindu	0.011	0.010	-0.057	0.010	0.009
	(0.042)	(0.042)	(0.066)	(0.042)	(0.043)
Location: urban	-0.218***	-0.218***	-0.260***	-0.218***	-0.225***
	(0.013)	(0.013)	(0.034)	(0.013)	(0.013)
Constant	14.906***	14.914***	15.541***	14.915***	14.873***
	(0.109)	(0.109)	(0.470)	(0.109)	(0.112)
Province*year dummies	yes	yes	yes	yes	yes
Other controls	yes	yes	yes	yes	yes
IV Coefficient from the first stage regression	-	0.124***	0.162***	0.122***	-
Regression statistics					
Instruments	-	Log distance	Migration network	Both	-

	(1) OLS	(2) IV	(3) IV	(4) IV	(5) FE
F-test of excluded instruments	-	26.28	17.29	15.78	-
Hansen J statistic (overid test)	-	-	-	6.39	-
P-value, J statistic		-	-	0.50	
Hausman (*p*-value)	-	0.14	0.55	0.48	0.56
Observations	12,000	12,000	12,000	12,000	12,000
No. clusters	316	316	316	316	316

Notes: The dependent variable for columns 1-5 is log of per capita food consumption. The key independent variable is migrant: a migrant-household dummy equal to 1 when the household contains at least one migrant and 0 otherwise. All specifications include province-year FEs, household size and other demographics, household-head characteristics, ethnicities of the head of the household, religion, and urban dummy. Other controls are dummy for whether or not a household experienced any natural disasters in the past and log of household asset. The log of distance from the residence of origin to the destination and migration network have been used as an instrument for endogenous variable, migrant. In particular, migration is instrumented with the log of distance to the destination in column 2, with the migration network at the community level at column III, and with both the log of distance to the destination and migration network in column IV. OLS is applied in column 1, IV is applied from columns 2 to 4, and FE is applied in column 5. This FE is indeed the first difference, as only two waves of panel survey (2007 and 2014) are used. The robust standard errors are shown in parentheses and clustered at the community (i.e. EA) level.

***Significant at the 1% level; ** Significant at the 5% level; *Significant at the 10% level

The OLS estimates of the effect of migration on food consumption potentially suffer from several issues. First, there is a threat that both higher consumption and intensity of migration are associated with some other household-level factor, such as culture or belief of consuming a specific food that is either unobservable or excluded from the estimation. If this is the case, the OLS estimates of the relationship between migration and consumption could merely reflect a coincidental association. It is also plausible that the likelihood of healthy and fresh food consumption in the rural areas has less motivational impact on migration (reverse causation). If either of these happen (omitted factors or reverse causality), then one would expect that the estimated relationship would change when different variables, related to migration but not directly related to either food consumption or to the lurking “unknown” variable, were used as are potential candidates as an instrument to migration (Pritchett and Summers, 1996).

The empirical association between migrant networks and the probability of migration has been established in the literature. Following this literature, I use two types of instruments for migration: migration network and log of distance to the destination. Networks are likely to support potential migrants in terms of information about job opportunities, initial living, and other direct (e.g. monetary) and indirect assistance. In this paper, the migration network at the community level is defined as a ratio of the number of migrants in each community (i.e. EA), excluding the household in question, to the number of adult workers in the community, again excluding the household in question. This captures the density of migration in the neighbourhood of the household (see for instance Nguyen and Winters, 2011). For the second instrument for migration, I follow McKenzie et al. (2010). I use distance from the migrant’s birth place to the destination as an instrument for migration to food consumption.

The F statistic for these excluded instruments (shown in Table 3.5) ranges from 15.78 to 26.28. This implies that both log of distance and migration network are significant predictors of potential migration. For all outcomes, the p value of 6.39 for the Hansen J (over-identification test) statistic do not lead us to reject the hypothesis that the instruments are orthogonal to the second-stage error term at the usual 5% significance level.

The results of IV estimation of Equation 1 are presented from column 2 to column 4 in Table 3.5. Column 2 uses log of distance from the household to the destination as the instrument. Compared to the OLS result, the semi-elasticity of food consumption of migration is smaller (0.11 versus 0.12) and the estimate is more precise. By contrast, in column 3 it can be seen that the estimated impact of migration using migration network as an instrument for migration is larger (0.15 versus 0.12) and highly statistically significant. Finally, in column 4 it can be observed that the impact of migration on food consumption is smaller (0.11 versus 0.12) when both instruments are used; the impact is again found to be smaller and precise.

These IV results suggest that using only the variable which is highly correlated with migration) produces estimates of the impact of migration on food consumption as if the estimates produced by using migration itself. The Hausman test investigates if the null hypothesis that the coefficients estimated using migration (OLS) and using only migration related to another variable (IV) are equal. For each instrument, the Hausman test that the OLS and IV estimates are equal is never rejected at the 5% significance level. The last column of Table 3.5 presents the households' FE results, which show that migration leads to 0.11 percentage point increase in per capita food consumption for migrant-sending households. The result is more similar to IV estimates than to OLS estimates. This may partly imply that that the findings are not driven by unobserved heterogeneity of the households.

Apart from the analysis of migrants' impact on log of per capita food consumption, an empirical analysis of the impact of migration on shares of food consumption for different food groups is also conducted (shown in Table 3.6, Table 3.7 and Table 3.8: OLS, IV, and FE results, respectively). In all budget share regressions, I additionally control for log of per capita total food expenditure. The reason for controlling this variable is because, as households gain more access to resources, they tend to shift their consumption patterns, probably towards more protein- and nutrient-rich foods. Failure to control this variable would lead to biased estimates of the impact of migration on food expenditure. As a result, the difference between migrant and non-migrant expenditure may be misleading, even if controlling for the other important covariates. It is observed that the standard errors for IV estimates are systematically lower than FE estimate except migration coefficient. This situation may arise due to presence of measurement errors in the data or there still may have some omitted variables that cannot be controlled for in the migration-consumption stated in Equation 4. Table B3.3 exhibits the OLS results for all food groups in pooled sample and Table B3.4 shows the IV results for all food groups in pooled sample.

Table 3.6: Migration and food consumption patterns. (Dependent variable: Budget shares of staple foods, and vegetables and fruits)

	(1)	(2)	(3)	(4)	(5)	(6)
	Staples	Staples	Staples	Vegetables and fruits	Vegetables and fruits	Vegetables and fruits
Estimator	OLS	IV	FE	OLS	IV	FE
Migrant	-0.011***	-0.015	-0.010***	0.003*	-0.060	0.004**
	(0.003)	(0.098)	(0.003)	(0.001)	(0.052)	(0.001)
Log of per capita food con	-0.043***	-0.037***	-0.042***	-0.002**	0.001	-0.002**
	(0.002)	(0.006)	(0.002)	(0.001)	(0.003)	(0.001)
HH Head characteristics						
Age of head	0.001***	-0.001	0.001***	0.000***	-0.000	0.000***
	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)
Education of head	-0.002***	-0.002***	-0.002***	0.001***	0.001***	0.001***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Female head	0.006	0.004	0.006	-0.000	-0.002	-0.000
	(0.007)	(0.008)	(0.007)	(0.004)	(0.004)	(0.004)
Married head	0.070***	0.057***	0.071***	0.031***	0.023***	0.032***
	(0.005)	(0.013)	(0.005)	(0.003)	(0.007)	(0.003)
Head employed	-0.001	-0.008	-0.003	-0.004	-0.008	-0.002
	(0.008)	(0.010)	(0.008)	(0.005)	(0.005)	(0.005)
Ethnicities						
Javanese	-0.023***	-0.023***	-0.022***	0.005***	0.005***	0.005***
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
Sundanese	-0.011***	-0.015***	-0.011***	-0.005**	-0.007***	-0.005**
	(0.004)	(0.005)	(0.004)	(0.002)	(0.003)	(0.002)
Minang	0.003	0.013	0.002	-0.010***	-0.005	-0.010***
	(0.006)	(0.011)	(0.006)	(0.003)	(0.006)	(0.003)
Batak	0.001	0.010	0.001	-0.006*	-0.001	-0.006*
	(0.008)	(0.011)	(0.008)	(0.004)	(0.006)	(0.004)

	(1)	(2)	(3)	(4)	(5)	(6)
	Staples	Staples	Staples	Vegetables and fruits	Vegetables and fruits	Vegetables and fruits
Estimator	OLS	IV	FE	OLS	IV	FE
Religion						
Islam	-0.011	-0.014*	-0.010	-0.008**	-0.009**	-0.008**
	(0.007)	(0.008)	(0.007)	(0.003)	(0.004)	(0.003)
Hindu	0.009	0.000	0.010	0.007	0.002	0.007
	(0.010)	(0.013)	(0.010)	(0.005)	(0.007)	(0.005)
Location: urban	0.045***	0.041***	0.045***	-0.000	-0.003	-0.000
	(0.003)	(0.005)	(0.003)	(0.001)	(0.002)	(0.001)
Constant	0.716***	0.708***	0.702***	0.097***	0.092***	0.098***
	(0.039)	(0.041)	(0.039)	(0.020)	(0.022)	(0.020)
Demographic controls	yes	yes	yes	yes	yes	yes
Province*round dummies	yes	yes	yes	yes	yes	yes
Other controls	yes	yes	yes	yes	yes	yes
Observations	12,559	12,559	12,559	12,559	12,559	12,559
R-squared	0.202	0.110		0.064		
Instruments	-	Migration network *HH size and log of distance	-	-	Migration network *HH size and log of distance	-
First stage *F*		32.78			28.87	
Endogeneity: *F* *p*-value		0.122			0.230	
Overid test: χ^2 p-value		0.39			0.75	
Coefficient of IV from the first-stage reg	-	0.019***	-	-	0.068***	-

Notes: The dependent variable for columns 1-3 is the budget share of staple foods and the dependent variable for columns 4-6 is the budget share of vegetables and fruit. The key independent variable is migrant, a migrant-household dummy equal to 1 when the household has at least one migrant and 0 otherwise. All specifications include province-year FEs, household size and other demographics, household-head characteristics, ethnicities of the head of the household, religion, and urban dummy. Other controls are dummy for whether a household experienced any natural disasters in the past and log of household asset. The log of distance from the residence of origin to the destination and migration network have been used as an instrument for the endogenous variable, migrant. In particular, migration is instrumented with the log of distance to

	(1)	(2)	(3)	(4)	(5)	(6)
	Staples	Staples	Staples	Vegetables and fruits	Vegetables and fruits	Vegetables and fruits
Estimator	OLS	IV	FE	OLS	IV	FE

the destination and an interaction of migration networks with household size in column 2 and in column 5. OLS is applied in column 1, IV is applied in column 2, and FE is applied in column 3 for the regression of budget shares of staple foods and the corresponding regressions methods are applied in columns 4-6 for the regression of budget shares of vegetables and fruits. This FE is indeed the first difference, as only two waves of panel survey (2007 and 2014) are used. The robust standard errors are shown in parentheses and clustered at the community (i.e. EA) level.

***Significant at 1% level; ** Significant at 5% level; *Significant at 10% level

Table 3.7: Migration and food consumption patterns. (Budget shares of 'meat and fish' and 'dairy products')

	(1)	(2)	(3)	(4)	(5)	(6)
	Meat and fish	Meat and fish	Meat and fish	Dairy products	Dairy products	Dairy products
Estimators	OLS	IV	FE	OLS	IV	FE
Migrant	0.003*	-0.060	0.005**	0.003**	-0.040	0.003**
	(0.001)	(0.052)	(0.001)	(0.001)	(0.043)	(0.001)
Log of per capita food consumption	-0.002**	0.001	-0.002**	0.010***	0.013***	0.010***
	(0.001)	(0.003)	(0.001)	(0.001)	(0.003)	(0.001)
Age of head	0.000***	-0.000	0.000***	0.000***	-0.000	0.000***
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Education of head	0.001***	0.001***	0.001***	0.001***	0.002***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female head	-0.000	-0.002	-0.000	-0.002	-0.003	-0.002
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Married head	0.031***	0.023***	0.032***	0.016***	0.011*	0.017***
	(0.003)	(0.007)	(0.003)	(0.003)	(0.006)	(0.003)
Head employed	-0.004	-0.008	-0.002	-0.008**	-0.011***	-0.007**
	(0.005)	(0.005)	(0.005)	(0.003)	(0.004)	(0.003)
Ethnicities						
Javanese	0.005***	0.005***	0.005***	0.008***	0.008***	0.008***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Sundanese	-0.005**	-0.007***	-0.005**	0.002	0.000	0.002
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Minang	-0.010***	-0.005	-0.010***	0.001	0.005	0.001
	(0.003)	(0.006)	(0.003)	(0.003)	(0.005)	(0.003)
Batak	-0.006*	-0.001	-0.006*	-0.004	-0.001	-0.004
	(0.004)	(0.006)	(0.004)	(0.003)	(0.005)	(0.003)
Religion						
Islam	-0.008**	-0.009**	-0.009**	-0.004	-0.005	-0.004

	(1)	(2)	(3)	(4)	(5)	(6)
	Meat and fish	Meat and fish	Meat and fish	Dairy products	Dairy products	Dairy products
Estimators	OLS	IV	FE	OLS	IV	FE
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
Hindu	0.007	0.002	0.007	0.002	-0.002	0.002
	(0.005)	(0.007)	(0.005)	(0.005)	(0.006)	(0.005)
Location: urban	-0.000	-0.003	-0.000	-0.014***	-0.016***	-0.014***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)
Constant	0.097***	0.092***	0.098***	-0.121***	-0.125***	-0.119***
	(0.020)	(0.022)	(0.020)	(0.018)	(0.019)	(0.018)
Coefficient of IV from the first-stage regression	-	0.075**	-	-	0.048**	-
Observations	12,559	12,559	12,559	12,559	12,559	12,559
R-squared	0.064			0.121	0.056	
Instruments		Migration network *HH size and log of distance	11,030		Migration network *HH size and log of distance	11,030
First stage *F*		21.18			14.23	
Endogeneity: *F* *p*-value		0.18			0.26	
Overid test: χ^2 p-value		0.33			0.40	

Notes: Dependent variable for columns 1-3 is the budget shares of meat and fish and dependent variable for columns 4-6 is the budget shares of dairy products. The key independent variable is migrant, a migrant-household dummy equal to 1 when the household contains at least one migrant and 0 otherwise. All specifications include province-round FEs, household size and other demographics, household-head characteristics, ethnicities of the head of the household, religion, and urban dummy. Other controls are dummy for whether a household experienced any natural disasters in the past and log of household asset. The log of distance from the residence of origin to the destination and migration network have been used as an instrument for endogenous variable, migrant. In particular, migration is instrumented with the interaction of migration networks with household size and log of distance to the destination in column 2 and in column 5. OLS is applied in column 1, IV is applied in column 2 and FE is applied in column 3 for the regression of meat and fish budget shares. This FE is indeed the first difference, as only two waves of panel survey (2007 and 2014) are used. The robust standard errors shown in parentheses and clustered at the community (i.e. EA) level.

***Significant at 1% level; ** Significant at 5% level; *Significant at 10% level

Table 3.8: Migration and food consumption patterns. (Budget shares of dried foods and condiments/spices)

Variables	(1) Dried foods OLS	(2) dried foods IV	(3) dried foods FE	(4) condiments/spices OLS	(5) condiments/spices IV	(6) condiments/Spices FE
Migrant	-0.003**	-0.096*	-0.003**	-0.002	-0.044	-0.001
	(0.001)	(0.054)	(0.001)	(0.001)	(0.051)	(0.001)
Log of per capita food consumption	-0.005***	0.000	-0.005***	-0.018***	-0.016***	-0.018***
	(0.001)	(0.003)	(0.001)	(0.001)	(0.003)	(0.001)
Age of head	-0.000***	-0.001**	-0.000***	0.000***	0.000	0.000***
	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)
Education of head	0.000**	0.001**	0.000**	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female head	0.016***	0.014***	0.016***	0.001	0.000	0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Married head	-0.009***	-0.020***	-0.009***	0.048***	0.043***	0.048***
	(0.003)	(0.007)	(0.003)	(0.003)	(0.007)	(0.003)
Head employed	-0.019***	-0.024***	-0.018***	-0.000	-0.003	-0.001
	(0.004)	(0.006)	(0.004)	(0.005)	(0.005)	(0.004)
Ethnicities						
Javanese	-0.002	-0.002	-0.002	0.003*	0.003	0.003*
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Sundanese	-0.002	-0.004	-0.002	-0.010***	-0.011***	-0.009***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)
Minang	-0.008***	0.000	-0.008***	0.016***	0.020***	0.016***
	(0.003)	(0.006)	(0.003)	(0.003)	(0.006)	(0.004)
Batak	-0.020***	-0.013**	-0.020***	0.014***	0.017***	0.014***
	(0.004)	(0.006)	(0.004)	(0.004)	(0.006)	(0.004)
	0.002	-0.001	0.002	0.012***	0.011***	0.013***
Religion	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)
Islam	-0.008**	-0.016**	-0.008**	0.010**	0.006	0.010**

Variables	(1) Dried foods OLS	(2) dried foods IV	(3) dried foods FE	(4) condiments/spices OLS	(5) condiments/spices IV	(6) condiments/Spices FE
	(0.004)	(0.006)	(0.004)	(0.004)	(0.006)	(0.004)
Hindu	-0.004***	-0.007***	-0.004***	0.023***	0.021***	0.022***
	(0.001)	(0.003)	(0.001)	(0.001)	(0.002)	(0.001)
Location: Urban	0.170***	0.162***	0.170***	0.272***	0.268***	0.264***
	(0.020)	(0.023)	(0.020)	(0.023)	(0.024)	(0.022)
Coefficient of IV from the first-stage regression	-	0.098	-	-	0.002**	-
Observations	12,559	12,559	12,559	12,559	12,559	12,559
R-squared	0.046			0.207	0.148	
Number of cluster	316	316	316	316	316	316
F(excluded instruments)	14.21	19.50	12.23	25.50	28.32	38.54
Prob>F	0.000	0.000	0.000	0.000	0.000	0.000

Notes: Robust standard errors are shown in parentheses and clustered at the community level. As migration suffers from an endogeneity problem, migration is instrumented with the log of distance from the migrants' households to the destination and the interaction of migration network with household size district-level rate of migration as a measure of household network. In testing the explanatory power of the chosen instrument, the first-stage result indicates that the instrument is reasonably strong with F-statistics larger than 10. The under-identification test indicates that the model is identified, as the null hypothesis from the Kleibergen Paap rk LM statistic is not accepted. The over-identification test implies that the null hypothesis should not be rejected and that the instruments are valid. The test suggests that the instrument is uncorrelated with the error term, and the excluded instruments are correctly excluded from the estimating equation.

***Significant at 1% level; ** Significant at 5% level; *Significant at 10% level

Both the OLS and the IV result suggest that migration has a strong impact on budget shares of different food group consumption. There appears to be substantial shift from rice, corn, and wheat consumption towards vegetables and fruits, dairy products, and meat and fish food group consumption.

3.5.3 Do migrant-sending-households shift consumption from carbohydrate-rich foods to protein- or vitamin-rich foods?

I test further whether migrants consume more carbohydrate-rich foods than non-migrants in the neighbourhood of the same community. I present a simpler specification of constructing carbohydrate-rich food over protein-rich foods as follows:

$$\frac{carbohydrate_{it}}{carbohydrate_{it} + protein_{it}} = \alpha_0 + \alpha_1 migrant_{it} + X_{it}'\Omega + \lambda_i + \delta_t + p\ x\ t + \mu_{it} \tag{5}$$

where $\frac{carbohydrate_{it}}{carbohydrate_{it}+protein_{it}}$ refers to the household's relative preference for carbohydrate foods over protein-rich foods. The positive coefficient of migrant, α_1, implies that migrants spends more on carbohydrate goods than on protein goods. The negative coefficient of migrant indicates that migrants spend more on protein-rich foods than on carbohydrate-rich foods.

I regress carbohydrate's share of total household spending, $\frac{carbohydrate_{it}}{carbohydrate_{it}+protein_{it}}$, on migrant dummy and other household-level controls, X_{it}, used in the regression in Table 3.5. This regression also includes province year FEs. The regression results are summarised in Table 3.9. Both in all migrants sample and work migrants sample, the coefficient of migrant is statistically significant and negative. In particular, in all migrant samples, migrant-sending households consume 2.1% less carbohydrate-dense foods than non-migrant-sending households and in work samples, those households consume 2.3% less carbo-dense foods. The estimated coefficients in the households where individuals are away from the residence for 1-3 years and 4-6 years are analogous in magnitude and directions and also highly statistically significant. However, for the corresponding coefficients, no relationships has been observed of the migrants for work sample. Table B3.5 in the appendix B presents OLS estimates with different specifications for the same dependent variable. The results in the appendix table B3.5 is consistent with the findings in table 3.8, though magnitudes are quite different.

Table 3.9 OLS regression of migration on carbohydrate-rich food. (Dependent variable: carbohydrate's expenditure share: $\frac{carbohydrate_{it}}{carbohydrate_{it}+protein_{it}}$)

	(1) All migrants sample	(2) Work migrants sample
Migrant	-0.021***	
	(0.003)	
Work migrant		-0.023***
		(0.006)
Migration 1 to 3 years	-0.035***	-0.051***
	(0.003)	(0.009)
Migration 4 to 6 years	-0.022***	-0.042
	(0.002)	(0.039)
Migration 7 to 9 years	-0.020	-0.034
	(0.033)	(0.039)
Migration 10 plus years	-0.021	-0.033
	(0.033)	(0.039)
Log of per capita food exp.	-0.088***	-0.084***
	(0.006)	(0.006)
Log of household size	0.002	0.010
	(0.009)	(0.011)
Proportion of male aged 0 to 4	-0.125**	-0.107
	(0.055)	(0.066)
Proportion of male aged 5 to 9	-0.062	-0.055
	(0.058)	(0.071)
Proportion of male aged 10 to 14	0.069	0.153**
	(0.063)	(0.078)
Proportion of male aged 15 to 55	-0.003	0.018
	(0.051)	(0.063)
Proportion of male aged 56 to 70	0.138**	0.154*
	(0.068)	(0.082)
Proportion of female aged 0 to 4	-0.122**	-0.098
	(0.055)	(0.066)
Proportion of female aged 5 to 9	0.017	0.018
	(0.059)	(0.072)
Proportion of female aged 10 to 14	0.063	0.060
	(0.062)	(0.080)

	(1) All migrants sample	(2) Work migrants sample
Proportion of female aged 15 to 55	0.015	0.053
	(0.050)	(0.061)
Proportion of female aged 56 to 70	0.052	0.097
	(0.059)	(0.072)
Age of head	-0.001	-0.001
	(0.000)	(0.000)
Education of head	-0.007***	-0.007***
	(0.001)	(0.002)
Female head	0.007	0.006
	(0.014)	(0.017)
Employment of head	0.005	-0.005
	(0.023)	(0.026)
North Sumatra	0.030**	0.024
	(0.014)	(0.015)
West Sumatra	0.060***	0.051***
	(0.017)	(0.019)
South Sumatra	0.062***	0.058***
	(0.017)	(0.019)
Lampung	0.109***	0.119***
	(0.017)	(0.020)
West Java	0.067***	0.072***
	(0.012)	(0.014)
Central Java	0.097***	0.087***
	(0.013)	(0.015)
Yogyakarta	0.058***	0.062***
	(0.019)	(0.022)
East Java	0.046***	0.043***
	(0.013)	(0.015)
Bali	0.089***	0.078***
	(0.018)	(0.021)
West Nusa Tenggara	0.095***	0.094***
	(0.014)	(0.015)
South Kalimantan	0.026*	0.043**
	(0.015)	(0.017)
South Sulawesi	0.014	0.006
	(0.015)	(0.017)
Constant	1.477***	1.423***
	(0.087)	(0.101)
Ethnicity controls	yes	yes
Religion controls	yes	yes
Province year FE	yes	yes
Observations	11,690	4,242
R-squared	0.100	0.095

Notes: The dependent variable for columns 1-2 is carbohydrate food share over protein-rich food. The key independent variable is migrant, a migrant-household dummy equal to 1 when the household contains at least one migrant and 0 otherwise. All specifications include province year FEs, household size and other demographics, household-head characteristics, ethnicities of the head of the household, religion, and urban dummy. Other controls are dummy for whether or not a household experienced any natural disasters in the past and log of household asset. Migration 1-3 years is a dummy variable which is defined as at least one member of the household moved from the home from one year to three years; migration 4-6 years, migration 7-9 years, and migration 10 plus years are defined similarly; these coefficients are compared with reference to the locality of non-migrant-households. The robust standard errors shown in parentheses and clustered at the community (i.e. EA) level.

***Significant at 1% level; ** Significant at 5% level; *Significant at 10% level

3.5.4 Impact of migration on food diversity[32]

Table 3.8 exhibits the impact of migration on food diversity. The OLS coefficient of migration on food diversity is highly statistically significant. The corresponding IV and FE estimates support OLS finding. However, IV coefficient estimate (of the effect of migration on food diversity) is larger than the OLS estimate. This indicates that there may be a measurement error in migration. Overall, migration appears to diversify food consumption significantly. More diversification of food implies that migrant households are highly concerned on micronutrient dense foods which have health implications.

Table 3.10: Impact of migration on food diversity

Variables↓	(1) Food Diversity OLS	(2) Food Diversity IV	(3) Food Diversity FE
Migrant	0.012***	0.154***	0.010***
	(0.003)	(0.007)	(0.003)
Log of per capita food consumption	-0.020***	-0.012	-0.021***
	(0.002)	(0.008)	(0.003)
Household size	0.005***	0.007***	0.005***
	(0.001)	(0.003)	(0.001)
Pro of male aged 0 to 4	0.174***	0.257***	0.175***
	(0.022)	(0.084)	(0.023)
Pro of male aged 5 to 9	0.125***	0.100***	0.125***
	(0.023)	(0.035)	(0.023)
Pro of male aged 10 to 14	0.068***	0.001	0.064***
	(0.025)	(0.069)	(0.025)
Pro of male aged 15 to 55	-0.100***	-0.047	-0.096***

[32] To measure food diversity, this paper uses the Simpson index of food diversity: Simpson index $= 1 - \sum_i w_i^2$, where w_i is the budget share of food group i. This index ranges in value from 0 to 1, where 0 indicates no diversification of food consumption and 1 indicates the highest diversification. The more the index closes to 1, the greater the diversification (see Nguyen and Winters, 2011 for an explanation of this index).

Variables↓	(1) Food Diversity OLS	(2) Food Diversity IV	(3) Food Diversity FE
	(0.020)	(0.056)	(0.021)
Pro of male aged 55-70	-0.113***	-0.038	-0.108***
	(0.033)	(0.080)	(0.032)
Pro of female aged 0 to 4	0.170***	0.257***	0.170***
	(0.022)	(0.088)	(0.022)
Pro of female aged 5 to 9	0.133***	0.110***	0.134***
	(0.022)	(0.034)	(0.023)
Pro of female aged 10 to 14	0.082***	0.003	0.084***
	(0.024)	(0.080)	(0.024)
Pro of female aged 15 to 55	0.118***	0.175***	0.119***
	(0.020)	(0.059)	(0.020)
Pro of female aged 55-70	0.137***	0.164***	0.134***
	(0.026)	(0.038)	(0.026)
Age of head	0.001***	-0.000	0.001***
	(0.000)	(0.001)	(0.000)
Education of head	0.002***	0.003***	0.002***
	(0.001)	(0.001)	(0.001)
Female head	-0.048***	-0.050***	-0.046***
	(0.008)	(0.010)	(0.008)
Married head	0.219***	0.201***	0.220***
	(0.009)	(0.020)	(0.009)
Head is employed	-0.022*	-0.031**	-0.019
	(0.012)	(0.015)	(0.012)
Ethnicities			
Javanese	-0.016***	-0.016***	-0.015***
	(0.003)	(0.004)	(0.003)
Sundanese	-0.029***	-0.033***	-0.028***
	(0.005)	(0.007)	(0.005)
Minang	-0.002	0.010	0.000
	(0.006)	(0.014)	(0.007)
Batak	0.006	0.018	0.009
	(0.009)	(0.015)	(0.009)
Religion			
Islam	0.020**	0.015	0.022**
	(0.008)	(0.010)	(0.009)
Hindu	0.005	-0.007	0.008
	(0.011)	(0.016)	(0.011)
Urban	0.030***	0.025***	0.030***
	(0.003)	(0.006)	(0.003)
Constant	0.669***	0.658***	0.662***
	(0.045)	(0.048)	(0.045)
Coefficient of IV from the first-stage regression	-	0.124***	-
Observations	12,559	12,559	12,559
R-squared	0.426	0.333	
Province-year FE	yes	yes	yes

Variables↓	(1) Food Diversity OLS	(2) Food Diversity IV	(3) Food Diversity FE
First stage *F*		19.26	
Endogeneity: *F* *p*-value		0.24	
Sargan Overid test: χ^2 p-value		0.32	

Notes: Dependent variable for columns 1-3 is food diversity. The key independent variable is migrant, a migrant-household dummy equal to 1 when the household contains at least one migrant and 0 otherwise. All specifications include province-year FEs, household size and other demographics, household-head characteristics, ethnicities of the head of the household, religion, and urban dummy. Other controls are dummy for whether a household experienced any natural disasters in the past and log of household asset. The log of distance from the residence of origin to the destination and migration network were used as an instrument for endogenous variable, migrant. OLS is applied in column 1, IV is applied in column 2, and FE is applied in regression column 3. The FE is indeed the first difference, as only two waves of panel survey (2007 and 2014) are used. The robust standard errors are shown in parentheses and clustered at the community (i.e. EA) level. Age group from 71 to 90 is an omitted category. ***Significant at 1% level; ** Significant at 5% level; *Significant at 10% level

3.6 Robustness checks

Column I of Table 3.11 summarises the estimates of the impact of migration on food consumption in the baseline sample. The main parameter of interest is the coefficient on migration, β_1, which shows that conditioning on food expenditure, and other various level of household controls, whether or not migrant's household food consumption is different from that of non-migrants. The migrant household's per capita food consumption is 15.1% more food than that of non-migrants in the neighbouring community with no controls.

The remaining columns of Table 3.11 (column 2 to column 7) conduct a variety of robustness checks. As mentioned in section 3, about 35% of households move for work purposes. Column 2 regresses log of per capita food consumption on migrant with all controls. The migrant household's per capita food consumption is 5.7% more food than that of non-migrants in the neighbouring community with all controls and less than baseline estimates (15.1%), with a 95% confidence interval between 0.06% and 0.08%. The coefficient of migrant attenuates by 9.4% with all controls than no controls. As mentioned in section 3, about 35% of households move for work purposes. In column 3, a regression of per capita consumption on migration is implemented for the sample of household members moving for work or searching for employment. It is assumed that the unobserved heterogeneity is of less concern here as key household-level variables are controlled in this estimation. The migrant coefficient is still significantly positive for this sample. Nonetheless, the coefficient is attenuated by 12.9%. As types of employment are not checked for and it is plausible that the current out-migrants may be unemployed for a while, that limits the remittance sent to the parents, thereby exerting less influence over household spending decisions on food.

Table 3.11: Comparing food consumption patterns of migrants and non-migrants (Dependent variable: log of per capita food consumption)

Specifications	Baseline	All migrants sample	Work migrants sample	Total exp. controls	Real food exp. adjustment	Migration instrumented	Raskin program included
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Migrant	0.151***	0.057***	0.022***	0.119***	0.072***	0.073***	0.044***
	(0.006)	(0.002)	(0.009)	(0.007)	(0.006)	(0.004)	(0.003)
observations	29,000	29,000	4,000	29,000	29,000	29,000	29,000
Within R-square	0.705	0.410	0.712	0.481	0.678	0.550	0.918
Ethnicity controls	no	yes	yes	yes	yes	yes	yes
Total exp. controls	no	yes	no	yes	no	no	no
Real food exp. adjustment	no	yes	no	no	yes	no	no
Demographics controls	no	yes	yes	yes	yes	yes	yes
HH characteristics controls	no	yes	yes	yes	yes	yes	yes
Province year FE	no	yes	yes	yes	yes	yes	yes

Notes: The dependent variable is log of per capita food consumption. The seven-day food purchases were converted into annual terms by the formula: (food consumption of the seven days*365)/7. All other quantitative variables are also annualised. The key independent variable is $migrant_i$ is a binary variable equal to 1 if at least one member of the household migrated in 2007 and in 2014 and 0 otherwise. Except baseline, all specifications include province-year FE and controls for household size, household demographics, and household's characteristics. Column 1 runs OLS regression of log of per capita food consumption on migration with no controls and column 2 runs the OLS with all controls using the full sample. Column 3 confines regression for the households that move for work purposes. Column 4 restricts regression controlling to the sum of food and non-food expenditure. Column 5 confines regression with real food expenditure adjustment and real food expenditure adjustment made by converting nominal food expenditure into real food expenditure, deflating by the consumer price index (CPI) data from the Bank of Indonesia for 2014 and 2007=100 serves as a base year. In column 6, migration is instrumented on the migration networks and distance from the birth place to the current destination of residence. Finally, column 7 shows the estimates controlling for government rice for the poor program (Raskin) in Indonesia in both 2007 and 2014. ***significant at 1% level.

Column 3 in Table 3.11 restricts the same regression for the sample of household that moved for work reasons. The types of work is not considered in the estimation. In a developing country context, an individual usually moves to the new place primarily for work purposes. So, to draw a plausible comparison of food consumption, two groups are made: household members are moving into intra-province (non-migrant) and they are moving into inter-province (migrant). The result is still statistically significant and the effect of migration on per capita food consumption is positive. Columns 4 and 5 include alternative controls of the polynomials of log of per capita food consumption. Column 4 conducts regressions of migration on food consumption patterns including controls of third-order polynomials in log per capita total expenditure on all goods. In case of total expenditure controls, the effect of migration on per capita food consumption, is smaller in magnitude, 3.2% smaller than baseline specification (11.9% compared to 15.1%). Column 5 runs regression controlling for real food expenditure, where nominal values are adjusted with annual CPI data. The coefficient of migrant (0.072) is positive and statistically significant and the coefficient is different and lower in size by 7.9% from the baseline regression (15.1%) (See Table 3.11). This supports the hypothesis that migrant households consume less amount of food per capita after controlling for real food expenditure.

Column 6 of Table 3.11 runs the IV regression, where migrant is instrumented by both migration networks and distance of the household between birth origin and destination. An instrumental variable is applied because of the following concern: there may be correlated measurement error in the data sets, as per capita consumption and migration is calculated from the same of set of consumption modules. Migration may also suffer from another type of endogeneity problem. For instance, a shock that increases demand for per capita food consumption, such as changing type of work, may also increase migration. This will cause upward bias on the coefficients of migration. Finally, column 7 documents the estimates from the regression of per capita food consumption on migration, including the government's food subsidisation program, the rice for the poor program (*Raskin* program). It is assumed that migrant-sending households may have restricted access to the government subsidy program or new rules may be introduced to provide access to the public food distribution program.

Although the coefficient on migrants (0.044) is highly statistically significant, somewhat surprisingly, the estimate attenuates by more than 10.7%.

Table 3.12 exhibits the regressions of migration on log of per capita food consumption for three estimation methods for all migrants sample: OLS, IV, and covariate matching. The effect of migration on log of per capita consumption is positive and significantly different from zero across three estimation methods (OLS coefficient: 0.089; IV coefficient: 0.075; and covariate matching coefficient: 0.051). This implies that the estimated coefficient of migration may not be driven by third unobserved factors.

Table 3.12: Comparing food consumption patterns of migrants and non-migrants: whole sample (Dependent variable: log of per capita food consumption)

	(1) OLS	(2) IV	(3) Covariate matching
Migrant	0.089***	0.075***	0.051***
	(0.006)	(0.012)	(0.002)
Demographics controls	Yes	Yes	Yes
HH head characteristics controls	Yes	Yes	Yes
Ethnicity controls	Yes	Yes	Yes
Religion controls	Yes	Yes	Yes
Province-year FE	Yes	Yes	Yes
Coefficient of IV from the first-stage regression	-	0.035***	-

Notes: Dependent variable is log of per capita food consumption for columns 1-3. Independent variable is migrant, measured as the number of out-migrants in a household. Column 1 shows estimates of the OLS regression of log of per capita food consumption on migrant, and other covariates. Column 2 is the IV regression with the same specification as OSL, where migration networks are used as instruments for migration. The matching variables used in column 2 are: household size, age, education, and employment status of the household head.

The *Wall Street Journal* (2015) reported that rice is a dominant commodity in the Indonesians' food basket and annual per capita rice consumption is higher than almost any other country. Average Indonesians consume 15 times more rice than Americans. Indonesians often say a meal without rice means one hasn't eaten. The government is trying to get away its people from pilling rice on their plates and has initiated a rice-reduction campaign "One Meal, No Rice."

(The Wall Street Journal, 2011). Given that rice is a very important staple in Indonesian's diet, I conduct a robustness check whether migration could reduce rice consumption. Table 3.13 does this robustness check. I pick staple share as the dependent variable as rice is the largest share (more than 70%) in staple food consumption which is tantamount to regress rice as the dependent variable. I run this exercise in subsamples on East Java and North Sumatra, which are the highest rice consumption provinces in the dataset. The results support our conjecture that migration may reduce the share of staple food consumption significantly, although the size of these reductions are not substantial in either OLS or IV regression. The estimates of covariate matching also show the similar results.

Table 3.13: Comparing food consumption patterns of migrants and non-migrants: subsamples (Dependent variable is the budget share of staple food)

	(1) OLS	(2) IV	(3) Covariate matching
Migrant	-0.019***	-0.035***	-0.058***
	(0.006)	(0.012)	(0.002)
Demographics controls	Yes	Yes	Yes
HH head characteristics controls	Yes	Yes	Yes
Ethnicity controls	Yes	Yes	Yes
Religion controls	Yes	Yes	Yes
Province-round FE	Yes	Yes	Yes
Coefficient of IV from the first-stage regression	-	0.021***	-

Notes: Dependent variable is budget share of staple food for columns 1-3. Independent variable is migrant, measured as the number of out-migrants in a household. Column 1 shows estimates of the OLS regression of log of per capita food consumption on migrant, and other covariates. Column 2 is the IV regression with the same specification as OSL, where migration networks are used as IV for migration. The matching variables used in column 3 are: household size, age, education, and employment status of the household head.

3.7 Conclusions and future directions of migration and food consumption research

This study is an attempt to determine to what extent migration alters food consumption for migrant households in Indonesia. This endeavour is important to bridge the existing gap in the migration literature and to provide a better understanding to aid the policy makers in Indonesia to design cost-effective policies for regional migration. Furthermore, the latest global fuel, food, financial, and economic crises intensify the need to understand the relationship between internal migration and food consumption. This chapter employs the household-level panel data sets of IFLS 2007 and IFLS 2014 to determine the relationship between internal migration and food consumption in Indonesia. Both propensity score matching (PSM) and an instrumental variable (IV) have been used to predict food consumption patterns due to internal migration.

This chapter finds three pieces of empirical evidences linking migration and food consumption. First, migration is found to be positively and significantly associated with migrant sending household's per capita food consumption. Our OLS results indicate that on average, migrant households per capita food consumption has increased by 12.6% compared to the households living in the same locality with no migrants. The similar finding is found when instrumental variable and households fixed effect model have been used to predict food consumption. This finding is robust to alternative specifications. This finding has implications for the policy makers to design policies to relocate people where population density is high and food security is a concern. Second, migration is found to have a substantial and statistically significant impact on different food group shares. However, migration is said to have varied association with budgetary shares of various food groups. For instance, while migration causes to reduce staple food consumption significantly, its magnitude is not noticeable. In case of budget shares of vegetables and fruits, the coefficient of migration is positive and marginally significant. In case of budgetary shares of meat and fish and dairy products migration has a minimum impact, though sign is as expected. A related result is that there is a minimum shift of carbohydrate-rich food towards vitamin- and protein-rich foods at the margin. Finally, on average, migration is found to be positively and significantly associated with greater food diversity. In particular, migration is found to have increased food diversity by 1.2%. This finding has an important implications for health and nutrition for the migrant sending households.

However, analysing Indonesian Family Life Survey (IFLS) data has several limitations to study migration and food consumption. First, migration module of IFLS contains no direct information on remittances, on which poor households in the developing countries are often dependent to purchase variety of foods. Second, IFLS does contain no consumption information for the migrants in the destination, so comparison of consumption of original households with destination households are not possible. Third, there is no quantitative information on food consumption. Hence, meaningful study on linking migration to nutrition and welfare is difficult. Finally, the findings are sensitive to the regional context of migration. Further research is required to generalise the findings for the developing countries. Fruitful avenue of more research would be urbanization, migration, and food consumption linkage in a dynamic setting.

References

Antman, F. M. (2012): “Gender, educational attainment, and the impact of parental migration on children left behind," Journal of Population Economics, 25 (4), 1187-1214.

Angrist, J. D. (1998): “Estimating the labor market impact on voluntary military service using social security data on military applicants,” *Econometrica*, 66 (2), 249–288.

Angrist, J. D., and J. S. Pischke (2009): *Mostly harmless econometrics: An empiricist’s companion*. Princeton University Press, Princeton.

Anh, D. N. (2003): “Migration and poverty in Asia: with reference to Bangladesh, China, the Philippines and Viet Nam,” ad Hoc Expert Group Meeting on Migration and Development, organised by the Economic and Social Commission for Asia and the Pacific, Bangkok, 27–29 August.

Afsar, R. (2003): “Dynamics of poverty, development and population mobility: the Bangladesh case,” ad Hoc Expert Group Meeting on Migration and Development, organized by the Economic and Social Commission for Asia and the Pacific, Bangkok, 27–29 August.

Andersson, L. (2002): “Rural–urban linkages in Bolivia: advantages and disadvantages,” Instituto de Investigaciones Socio-Económicas, Universidad Catolica Boliviana. Available at http://www.childmigration.net/files/LAndersen.pdf.

Atkin, D. (2016): “The caloric costs of culture: Evidence from Indian migrants: Dataset. *American Economic Review*, 106 (4), 1144–1181.

Badan Pusat Statistik (2009): *Profil migran berdasarkan hasil SUPAS 2005*. Badan Pusat Statistik, Jakarta.

Badan Pusat Statistik (2011): *Statistical yearbook of Indonesia*. Badan Pusat Statistik, Jakarta.

Batista, C., J. Seither, and P. C. Vicente (2016): “Migration, Political Institutions, and Social Networks,” Available at http://conference.iza.org/conference_files/AMM_2016/seither_j24043.pdf

Becker, S., and A. Ichino (2002): “Estimation of Average Treatment Effects Based on Propensity Scores,” *Stata Journal*, 2 (4), 358–377.

Beegle, K., J. De Weerdt, and S. Dercon (2011): "Migration and economic mobility in Tanzania: Evidence from a tracking survey," *Review of Economics and Statistics*, 93 (3), 1010–1033.

Bell, M., and S. Muhidin (2011): "Comparing internal migration between countries using Courgeau's *k*," in *Population dynamics and projection methods,* ed. J. Stillwell, and Clarke, Springer, Netherlands.

de Brauw, A., and T. Harigaya (2007): "Seasonal migration and improving living standards in Vietnam," *American Journal of Agricultural Economics*, 89 (2), 430–447.

de Brauw, A., and J. Giles (2016): "Migrant Opportunity and the Educational Attainment of Youth in Rural China," *Journal of Human Resources*, 50 (2), 317–372.

de Brauw, A., V. Mueller, and T. Woldehanna (2013): "Motives to Remit: Evidence from Tracked Internal Migrants in Ethiopia," *World Development* 50, 13–23.

Bryan, G., S. Chowdhury, and A. Mobarak (2014): "Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh," *Econometrica*, 82, 1671–1748.

Carrington, W., E. Detragiache, T. Vishnawath (1996): "Migration with endogenous moving costs," *American Economic Review* 86 (4), 909–930.

Chami, R., C. Fullenkamp, and S. Jahjah (2005): "Are immigrant remittance flows a source of capital for development?" *IMF Staff Papers,* 52 (1), 55–81.

Deb, P., and P. Seck (2009): "Internal migration, selection bias and human development: Evidence from Mexico," Human Development Research Paper 2009/31, United Nations Development Program (UNDP).

Dehejia, R. (2005): "Practical Propensity Score Matching: A Reply to Smith and Todd," *Journal of Econometrics,* 125 (1–2), 355–364.

Dehejia, R., and S. Wahba (2002): "Propensity Score Matching Methods for Non-Experimental Causal Studies," *Review of Economics and Statistics*, 84 (1), 151–161.

Dercon, S. (2002): "Income risk, coping strategies and safety nets," *World Bank Research Observer*, 17 (2), 141–166.

Durand, J., and D. S. Massey (1992) "Mexican Migration to the United States: A Critical Review," *Latin American Research Review*, 27, 3-42.

Ham, J. C., X. Li, and P. B. Reagan (2005): "Propensity score matching, a distance-based measure of migration, and the wage growth of young men," Federal Reserve Bank of New York Staff Reports, No. 212.

Harris, J. R. and M. P. Todaro (1970): "Migration, unemployment and development: A two-sector analysis," *American Economic Review*, 60 (1), 126–142.

Hasanah, A., S. Mendolia and O. Yerokhin (2017): "Labour Migration, Food Expenditure, and Household Food Security in Eastern Indonesia," Economic Record, VOL. 93, Special Issue, June, 122–143.

Hoddinott, J. (1992): "Rotten kids or manipulative parents: are children old age security in Western Kenya?" *Economic Development and Cultural Change*, 40 (3), 545–565.

Hugo, G. (2001): "Internal and international population mobility: Implications for the spread of HIV/AIDS. Executive summary and recommendations," in Hugo, G. (2001). *Population mobility and HIV/AIDS in Indonesia*. Available from http://www.ilo.org/wcmsp5/groups/public/---asia/---ro-bangkok/---ilo-jakarta/documents/publication/wcms_141794.pdf

Hugo, G. (2012): "Changing patterns of population mobility in Southeast Asia," in L. Williams & M. P. Guest (Eds.), *Demographic change in Southeast Asia: Recent histories and future directions*, Ithaca: Cornell Southeast Asia Program Publications.

Karamba, W. R., E. J. Quiñones and P. Winters (2011): "Migration and food consumption patterns in Ghana," *Food Policy*, 36 (1), 41–53.

Lucas, R. E. B., and O. Stark (1985): "Motivations to remit: evidence from Botswana," *Journal of Political Economy*, 93, 901–918.

McKenzie, D., J. Gibson, and S. Stillman (2010): "How important is selection? Experimental vs. non-experimental measures of the income gains from migration," *Journal of the European Economic Association*, 8 (4): 913–945.

Mendola, M. (2008): "Migration and technological change in rural households: Complements or substitutes?" *Journal of Development Economics*, 85(1–2), 150–175.

Mincer, J. (1978): "Family migration decisions," *Journal of Political Economy*, 86, 759-773.

Molini, V., and D. Pavelesku and M. Ranzani (2016): "Should I Stay or Should I Go? Internal Migration and Household Welfare in Ghana," World Bank Policy Research Working Paper No. 7752. Available at SSRN: https://ssrn.com/abstract=2812300

Morduch, J. (1995): "Income smoothing and consumption smoothing," *Journal of Economic Perspectives*, 9 (3), 103–114.

Mountford, A. (1997): "Can a brain drain be good for growth in the source economy?" *Journal of Development Economics*, 53 (2), 287–303.

Muhidin, S. (2014): "Migration patterns: people on the move," in H. Hill (Ed.), *Regional dynamics in a decentralized Indonesia,* Institute of Southeast Asian Studies: Singapore.

Nguyen, M. C., and P. Winters (2011): "The impact of migration on food consumption patterns: The case of Vietnam," *Food Policy,* 36 (1), 71–87.

Olowa, O. W., and T. T. Awoyemi (2011): "Remittances and Household Expenditure in Rural Nigeria," Journal of Rural Economics and Development, 20 (1), 30-43.

Ozgen, C., P. Nijkamp, and J. Poot (2010): "The effect of migration on income growth and convergence: Meta-analytic evidence," *Papers in Regional Science,* 89 (3), 537–561.

Ping, H. (2003): "China migration country study," Paper presented at the Regional Conference on Migration, Development and Pro-Poor Policy Choices in Asia, organised by the Bangladesh Refugee and Migratory Movements Research Unit, Bangladesh/DFID UK, Dhaka, 22-24 June. Available at http://www.eldis.org/vfile/upload/1/document/0903/Dhaka_CP_3.pdf

Rogers, B. L. (1996): "The implications of female household headship for food consumption and nutritional status in the Dominican Republic," *World Development*, 24 (1), 113–128.

Rosenzweig, M. R., and O. Stark (1989): "Consumption smoothing, migration and marriage: Evidence from rural India," *Journal of Political Economy,* 97 (4), 905–926.

Rosenbaum, P., and D. Rubin (1983): "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 70 (1), 41–55.

Resosudarmo, B., and Y. Vidyattama (2006): "Regional income disparity in Indonesia: A panel data analysis," *ASEAN Economic Bulletin*, 23 (1), 41-44.

Rossi, P. H. (1980): "Why families move," Sage Publications, Beverly Hills, CA.

Sjaastad, L. A. (1962): "The costs and returns of human migration," *Journal of Political Economy*, 70 (5), 80–93.

Stark, O. (1991). *The migration of labour*. Blackwell, Cambridge.

Stark, O., and J. E. Taylor (1991): "Migration incentives, migration types: the role of relative deprivation," *Economic Journal*, 101 (408), 1163–1178.

Strauss, J., F. Witoelar and B. Sikoki (2016): *The fifth wave of the Indonesia Family Life Survey (IFLS5): Overview and field report, Vol. 1.* WR-1143/1-NIA/NICHD. Rand Labor & Population.

Strauss, J., F. Witoelar, B. Sikoki, and A. M. Wattie (2009): *The fourth wave of the Indonesia Family Life Survey (IFLS4): Overview and field report, Vol. 1.* WR-675/1-NIA/NICHD. Rand Labor & Population.

Sukamdi and G. Mujahid (2015): *Internal migration in Indonesia.* UNFPA Indonesia Monograph Series No. 3. Available at http://indonesia.unfpa.org/application/assets/publications/FA_Isi_BUKU_Monograph_Internal_Migration_ENG.pdf

Subramanian, S., and A. Deaton **(**1996) "The Demand for Food and Calories," *Journal of Political Economy*, 104 (1), 133–62.

Taylor, J. E. and J. Mora (2006): "Does Migration Reshape Expenditures in Rural Households? Evidence from Mexico," *World Bank Policy Research Paper 3842,* Washington, DC: World Bank.

Taylor, J. E., D. S. Massey, J. Arango, G. Hugo, A. Kouaouci, and A. Pellegrino (1996): "International Migration and Community Development," *Population Index* 62 (3), 397-418

Thomas, D., and E. Frankenberg (2007): "Household responses to the financial crisis in Indonesia: Longitudinal evidence on poverty, resources, and well-being," in Ann Harrison (ed.), *Globalization and poverty* (pages 517–560), National Bureau of Economic Research and University of Chicago Press. Available from URL: http://www.nber.org/chapters/c0116.pdf

Todaro, M. P. (1969): “A model of labor migration and urban unemployment in less developed countries,” *American Economic Review*, 59(1), 138–148.

UNDP (2009): Human Development Report 2009 – Overcoming Barriers: Human Mobility and Development. United Nations Development Program, New York.

United Nations (2013): Trends in international migrant stock: The 2013 revision. (United Nations database). Available at http://www.un.org/en/development/desa/population/migration/data/estimates/docs/OriginMIgrantStocks_Documentation.pdf

Vidyattama, Yogi (2014/2016): “Inter-provincial migration and 1975–2005 regional growth in Indonesia,” *Papers in Regional Science*, 95 (S1).

Winters, P., A. de Janvry, and E. Sadoulet (2001): “Family and community networks in Mexico–US migration,” *Journal of Human Resources* 36 (1), 159–184.

Yang, D. (2011): “Migrant remittances,” *Journal of Economic Perspectives*, 25 (3), 129–152.

Yang, D., and C. Martinez (2005): *Remittances and poverty in migrants' home areas: Evidence from the Philippines*. Background paper for World Bank's International Migration and Development Research Program: The World Bank. Available from https://www.cbd.int/financial/charity/philippines-remittance.pdf.

The World Bank (2017): “Urban Poverty and Mobility in Indonesia,” Available at https://olc.worldbank.org/.../Matthew-Urban_Poverty_SummerSchool_2017-07-07.pdf.

Wallace, S. B., C. D. DeLorme, and D. R. Kamerschen (1997): “Migration as a Consumption Activity”, *International Migration*, 35 (1), 37-58.

Zarate-Hoyos, G. A. (2004): “Consumption and Remittances in Migrant Households: Toward a Productive Use of Remittances,” *Contemporary Economic Policy*, 22 (4), 555-565.

Appendix B

Table B3.1: Construction of food groups

Group	Food group	Food items
1	Staple	Rice, corn, sago/flour, cassava, tapioca, dried cassava, and other staples like potatoes, sweet potatoes and yams
2	Vegetables and fruit	Green vegetables (kankung (water spinach), spinach etc.); mustard greens, tomatoes, cabbage, green beans and the like; fruits like papaya, mango, banana and the like
3	Meat and fish	Beef, mutton, chicken, water buffalo meat and the like; chicken, duck and the like; fresh fish, oysters, shrimp, squid and the like; salted fish and smoked fish
4	Dairy products	Eggs, fresh milk, canned milk, powdered milk and the like
5	Dried foods	Noodles, macaroni, chips and the like; cookies, beards and crackers
6	Condiments/spices	Salt, sweet and salty soy sauce, shrimp paste; chili sauce, tomato sauce and the like; shallot, garlic, chili, candle nuts, coriander and the like; Javanese brown sugar, butter, cooking oil, coconut oil, peanut oil, corn oil, palm oil and the like
7	Beverages	Drinking water, granulated sugar, coffee, tea, cocoa, soft drinks etc.
10	Other foods	Tofu, tempe and other side dishes; jerky, shredded beef, canned meat, sardine and the like

Note: Constructed from the IFLS consumption modules in survey years 2007 and 2014.

Table B3.2: Probit estimates of the determinants household migration

Determinants↓	Dependent variable: migrant
Log of per capita food consumption	0.190***
	(0.028)
Household size	5.819***
	(0.024)
Proportion males 0-4	0.282***
	(0.294)
Proportion Females 0-4	-0.877***
	(0.303)
Proportion males 5-9	-0.574***
	(0.316)
Proportion females 5-9	0.080***
	(0.249)
Proportion males 10-14	0.367***
	(0.326)
Proportion Females 10-14	.625***
	(0.293)
Proportion males 15-55	.635***
	(0.008)
Proportion females 15-55	-1.555***
	(0.306)
Proportion males 55-70	0.189***
	(0.245)
Proportion females 55-70	0.400
	(0.338)
Urban	0.137
	(0.097)
Age of head	0.012
	(0.017)
Education of head	-0.009
	(0.009)
Female-headed household	0.332**
	(0.140)
Marital status of head	0.281
	(0.352)
Employment of head	-0.192
	(0.224)
Religion: Islam	-0.070
	(0.086)
Religion: Hindu	-0.246**
	(0.119)
Ethnicities	
Javanese	-0.041
	(0.041)
Sundanese	-0.139**

	(0.058)
Minang	-0.062
	(0.090)
Batak	0.428***
	(0.099)
Log of household asset	-0.047***
	(0.011)
Disaster	-0.071*
	(0.043)
Log of Distance	-0.030***
	(0.001)
Migration network	0.850***
	(0.003)
Log distance from the centre of the community to district headquarter	-0.203*** (0.042)
Community road condition (1=good; 0=bad)	-0.154
	(0.122)
Constant	-6.349***
	(2.158)
Observations	11,196

Notes: The robust standard errors are shown in parentheses. Dependent variable is a migrant, dummy variable taking the value of 1 if a household has at least one migrant in survey years 2007 and 2014. The Probit regression is applied, with key household level controls and exogenous instruments. ***denotes statistical significance at 1% level; **denotes statistical significance at 5% level; and *denotes statistical significance at 10% level.

Table B3.3: Covariate balance before and after matching

Variables	(1) Mean-treated	(2) Mean-control	(3) % bias	(4) % reduction in bias	(5) *t*
Household characteristics:					
Household size	3.44	3.38	0.5	95.2	0.09
Age of head	39.32	38.51	6.2	88.1	0.67
Sex of head	0.87	0.85	14	2.1	1.12
Education of head	11.3	10.8	15	34	1.18
Head is employed	0.60	0.58	9.1	85.9	0.75
Community characteristics:					
Community road condition (1=good; 0=bad)	0.32	0.36	-12.3	49.5	-1.16
Log distance from the centre of the community to the district headquarter	0.26	0.25	4.5	28	0.97

Table B3.4: Robustness check for Table 3.5 (Dependent variable-log of per capita food consumption)

	(1) OLS	(2) IV	(3) IV	(4) FE
Migrant	0.131***	0.128***	0.122**	0.227
	(0.011)	(0.019)	(0.028)	(0.238)
Log of households' income	0.024*	0.024*	0.023*	0.024*
	(0.013)	(0.013)	(0.013)	(0.013)
Household size	-0.121***	-0.121***	-0.121***	-0.121***
	(0.011)	(0.011)	(0.011)	(0.011)
Proportion males 0-4	0.055	0.053	0.084	0.051
	(0.310)	(0.306)	(0.313)	(0.319)
Proportion males 5-9	0.627**	0.630**	0.579*	0.634*
	(0.309)	(0.307)	(0.315)	(0.326)
Proportion males 10-14	0.518	0.521	0.462	0.525
	(0.336)	(0.334)	(0.350)	(0.349)
Proportion males 15-55	0.974***	0.973***	0.989***	0.972***
	(0.307)	(0.303)	(0.305)	(0.308)
Proportion males 55-70	0.338	0.336	0.370	0.334
	(0.347)	(0.342)	(0.347)	(0.354)
Proportion females 0-4	-0.045	-0.047	-0.007	-0.050
	(0.303)	(0.298)	(0.311)	(0.320)
Proportion females 5-9	-0.081	-0.079	-0.123	-0.076
	(0.301)	(0.299)	(0.303)	(0.313)
Proportion females 10-14	0.734*	0.737*	0.675*	0.742*
	(0.388)	(0.383)	(0.401)	(0.409)
Proportion females 15-55	0.465*	0.464*	0.479*	0.463*
	(0.278)	(0.275)	(0.278)	(0.278)
Proportion females 55-70	0.334	0.335	0.319	0.336
	(0.313)	(0.309)	(0.309)	(0.312)
Age of head	-0.002	-0.002	-0.003	-0.002
	(0.002)	(0.002)	(0.003)	(0.003)
Education of head	0.035***	0.035***	0.037***	0.035***
	(0.008)	(0.008)	(0.009)	(0.009)
Employment of head	0.081	0.082	0.081	0.082
	(0.096)	(0.094)	(0.095)	(0.094)
Female head	-0.048	-0.048	-0.060	-0.047
	(0.133)	(0.132)	(0.135)	(0.133)
Married head	-0.402***	-0.400***	-0.427***	-0.398**
	(0.153)	(0.150)	(0.163)	(0.170)
Religion Islam	-0.294*	-0.294*	-0.289*	-0.294*
	(0.177)	(0.175)	(0.175)	(0.176)
Religion Hindu	-0.052	-0.053	-0.050	-0.053
	(0.195)	(0.192)	(0.193)	(0.193)
Javanese head	-0.077*	-0.077*	-0.080*	-0.076*
	(0.046)	(0.045)	(0.045)	(0.046)
Sundanese head	0.071	0.072	0.066	0.072
	(0.066)	(0.065)	(0.065)	(0.067)
Minang head	0.082	0.081	0.101	0.080
	(0.099)	(0.098)	(0.103)	(0.102)
Batak head	0.014	0.013	0.027	0.012
	(0.193)	(0.192)	(0.192)	(0.199)
Constant	10.807***	10.802***	10.893***	10.795***
	(0.350)	(0.351)	(0.381)	(0.413)
Observations	985	985	985	985
R-squared	0.329	0.329	0.325	0.329
Coefficient of IV from the first-stage	-	0.133***	0.115**	-

regression

Notes: The dependent variable for columns 1-4 is log of per capita food consumption. The key independent variable is migrant: a migrant-household dummy equal to 1 when the household contains at least one migrant and 0 otherwise. All specifications include province-year FEs, household size and other demographics, household-head characteristics, ethnicities of the head of the household, religion, and urban dummy. Other controls are dummy for whether or not a household experienced any natural disasters in the past and log of household asset. In column 2, log of distance from the residence of origin to the destination is used as an instrument for migration and in column 3, migration network have been used as an instrument for endogenous variable, migrant. In column 4, household fixed effect has been applied. The robust standard errors are shown in parentheses and clustered at the community (i.e. EA) level.

***Significant at 1% level; ** Significant at 5% level; *Significant at 10% level

Table B3.5: OLS in pooled sample for all food groups

Dep Var.:	(1) staple	(2) veg	(3) meat	(4) dairy	(5) dried	(6) spices	(7) bev	(8) other
Migrant	-0.0127*** (0.00282)	0.00268 (0.00140)	-0.00229 (0.00192)	0.00396** (0.00142)	-0.00283* (0.00122)	-0.00345* (0.00140)	0.0157*** (0.00378)	-0.00115 (0.000665)
lpcfoodexp	-0.0453*** (0.00213)	-0.00118 (0.00111)	0.0142*** (0.00133)	0.0126*** (0.000990)	- (0.00120)	-0.0184*** (0.00134)	0.0514*** (0.00314)	- (0.000613)
lhhsize	0.0147*** (0.00378)	0.00652*** (0.00188)	0.0314*** (0.00252)	0.0120*** (0.00174)	-0.00255 (0.00163)	0.00384 (0.00209)	-0.0646*** (0.00567)	-0.00139 (0.000929)
Male 0-4	0.0107 (0.0234)	-0.00267 (0.0126)	0.000531 (0.0157)	0.161*** (0.0122)	0.0403*** (0.00918)	0.00131 (0.0114)	-0.220*** (0.0303)	0.00934 (0.00562)
Male 5-9	-0.00378 (0.0233)	-0.0121 (0.0127)	0.0152 (0.0161)	0.0500*** (0.0107)	0.0334*** (0.00911)	0.00395 (0.0116)	-0.0988** (0.0318)	0.0121* (0.00551)
Male 10-14	0.0152 (0.0228)	-0.0130 (0.0122)	0.0145 (0.0166)	0.0126 (0.0104)	0.0414*** (0.0111)	0.000590 (0.0113)	-0.0784* (0.0317)	0.00697 (0.00561)
male 15-55	-0.000362 (0.0206)	-0.0161 (0.0112)	-0.0189 (0.0136)	-0.00672 (0.00863)	-0.00172 (0.00771)	-0.00278 (0.0105)	0.0449 (0.0271)	0.00167 (0.00484)
Male 55-70	0.0501 (0.0268)	-0.0109 (0.0138)	-0.0207 (0.0171)	-0.00316 (0.0112)	0.00190 (0.00957)	0.00172 (0.0145)	-0.0180 (0.0408)	-0.000873 (0.00630)
Female 0-4	-0.0136 (0.0232)	-0.00165 (0.0128)	-0.00511 (0.0155)	0.130*** (0.0119)	0.0346*** (0.00905)	0.0107 (0.0117)	-0.167*** (0.0307)	0.0117* (0.00556)
Female 5-9	0.00739 (0.0236)	0.0142 (0.0126)	0.00302 (0.0159)	0.0395*** (0.0112)	0.0379*** (0.00929)	0.00647 (0.0118)	-0.119*** (0.0312)	0.0103 (0.00564)
Female 10-14	-0.00707 (0.0226)	0.00424 (0.0128)	0.00894 (0.0148)	0.00137 (0.00978)	0.0236* (0.00964)	0.00895 (0.0117)	-0.0467 (0.0320)	0.00671 (0.00530)
Female 15-55	0.00765 (0.0199)	0.0187 (0.0109)	0.0140 (0.0132)	0.0114 (0.00848)	0.0171* (0.00754)	0.0245* (0.0104)	-0.103*** (0.0263)	0.0100* (0.00468)
Female 55-70	0.0331 (0.0263)	0.0429** (0.0142)	0.0131 (0.0169)	0.0149 (0.0116)	0.00930 (0.0100)	0.0778*** (0.0137)	-0.210*** (0.0349)	0.0187** (0.00657)
Age	0.000157 (0.000176)	0.000124 (0.000086	0.000359** (0.000118)	0.000162 (0.000087	- (0.000071	0.000153 (0.000087	- (0.000243)	0.000143** (0.000041
Education	- (0.000520)	0.00107*** (0.000266)	0.00148*** (0.000339)	0.00174*** (0.000243)	0.000557* (0.000235)	- (0.000261)	-0.00127 (0.000724)	0.000229 (0.000123)

Female	0.0146*	0.00354	-0.00538	0.000801	0.0144***	0.00503	-0.0339***	0.000938
	(0.00723)	(0.00363)	(0.00485)	(0.00382)	(0.00381)	(0.00397)	(0.00991)	(0.00169)
Married	0.0797***	0.0289***	0.0440***	0.00766**	-0.00943**	0.0538***	-0.216***	0.0115***
	(0.00547)	(0.00308)	(0.00378)	(0.00271)	(0.00307)	(0.00320)	(0.0102)	(0.00143)
Employed	0.00198	-0.00436	-0.00784	-0.00992**	-0.0186***	0.00123	0.0412**	-0.00375*
	(0.00762)	(0.00460)	(0.00410)	(0.00332)	(0.00443)	(0.00444)	(0.0130)	(0.00188)
Islam	-0.0101	-0.00745*	0.000742	-0.00418	0.00175	0.0122***	0.00250	0.00454**
	(0.00705)	(0.00324)	(0.00429)	(0.00314)	(0.00297)	(0.00331)	(0.00995)	(0.00161)
Hindu	0.00948	0.00681	-0.0196***	0.00166	-0.00813*	0.00955*	0.000378	-0.000129
	(0.00971)	(0.00485)	(0.00590)	(0.00457)	(0.00381)	(0.00444)	(0.0129)	(0.00189)
Javanese	-0.0230***	0.00515***	-0.0358***	0.00813***	-0.00208	0.00284	0.0336***	0.0111***
	(0.00298)	(0.00151)	(0.00209)	(0.00157)	(0.00135)	(0.00150)	(0.00414)	(0.000707)
Sundanese	-0.0148***	-0.00525*	-0.0366***	0.00297	-0.00117	-0.0110***	0.0581***	0.00776***
	(0.00403)	(0.00211)	(0.00280)	(0.00195)	(0.00185)	(0.00197)	(0.00611)	(0.000991)
Minang	-0.000739	-0.0106***	-0.0107**	0.00173	-0.00771**	0.0138***	0.0158	-0.00150
	(0.00641)	(0.00286)	(0.00396)	(0.00308)	(0.00260)	(0.00345)	(0.00835)	(0.00134)
Batak	0.00794	-0.00678	0.0202***	-0.00679	-0.0210***	0.0165***	-0.00924	-0.000819
	(0.00783)	(0.00352)	(0.00514)	(0.00349)	(0.00357)	(0.00373)	(0.0111)	(0.00208)
Constant	0.629***	0.0689***	-0.125***	-0.128***	0.144***	0.232***	0.0806	0.0983***
	(0.0323)	(0.0165)	(0.0206)	(0.0144)	(0.0165)	(0.0182)	(0.0467)	(0.00917)
Obs	12625	12625	12625	12625	12625	12625	12625	12625
Prov*Yr FE	yes	yes	yes	yes	yes	yes	yes	yes

Notes: The dependent variables for columns 1-8 are the budget shares of eight food groups. The key independent variable is migrant: a migrant-household dummy equal to 1 when the household contains at least one migrant and 0 otherwise. All specifications include province-year FEs. The control variables are: household size and other demographics, household-head characteristics, ethnicities of the head of the household, religion, and urban dummy. Other controls are dummy for whether or not a household experienced any natural disasters in the past and log of household asset. The robust standard errors are shown in parentheses and clustered at the community (i.e. EA) level.

***Significant at 1% level; ** Significant at 5% level; *Significant at 10% level

Table B3.6: IV in pooled sample for all food groups

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	staple	veg	meat	dairy	dried	spices	bev	other
Migrant	-0.0141***	0.00272	-0.00413	0.00822***	-0.00184	-0.00494*	0.0156*	-0.00146
	(0.00420)	(0.00223)	(0.00284)	(0.00227)	(0.00193)	(0.00217)	(0.00617)	(0.00100)
Log of per capita food con	-0.0452***	-0.00118	0.0143***	0.0124***	-0.00500***	-0.0183***	0.0514***	-0.00838***
	(0.00214)	(0.00111)	(0.00133)	(0.000993)	(0.00121)	(0.00134)	(0.00315)	(0.000611)
Log of household size	0.0147***	0.00652***	0.0315***	0.0120***	-0.00256	0.00386	-0.0646***	-0.00138
	(0.00378)	(0.00188)	(0.00252)	(0.00175)	(0.00163)	(0.00209)	(0.00567)	(0.000929)
Male aged 0-4	0.0116	-0.00269	0.00155	0.159***	0.0397***	0.00214	-0.220***	0.00952
	(0.0234)	(0.0126)	(0.0157)	(0.0122)	(0.00923)	(0.0114)	(0.0304)	(0.00564)
Male aged 5-9	-0.00410	-0.0121	0.0148	0.0510***	0.0336***	0.00363	-0.0989**	0.0120*
	(0.0233)	(0.0127)	(0.0161)	(0.0107)	(0.00911)	(0.0116)	(0.0318)	(0.00551)
Male aged 10-14	0.0145	-0.0129	0.0135	0.0149	0.0419***	-0.000201	-0.0784*	0.00680
	(0.0229)	(0.0122)	(0.0166)	(0.0104)	(0.0111)	(0.0114)	(0.0318)	(0.00564)
Male aged 15-55	0.000123	-0.0161	-0.0183	-0.00812	-0.00205	-0.00229	0.0450	0.00177
	(0.0206)	(0.0113)	(0.0137)	(0.00865)	(0.00772)	(0.0105)	(0.0271)	(0.00484)
Male aged 55+	0.0507	-0.0110	-0.0200	-0.00480	0.00152	0.00229	-0.0180	-0.000754

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	staple	veg	meat	dairy	dried	spices	bev	other
	(0.0268)	(0.0138)	(0.0171)	(0.0113)	(0.00957)	(0.0145)	(0.0409)	(0.00631)
Female aged 0-4	-0.0127	-0.00167	-0.00403	0.128***	0.0340***	0.0116	-0.167***	0.0118*
	(0.0233)	(0.0129)	(0.0156)	(0.0120)	(0.00909)	(0.0117)	(0.0308)	(0.00558)
Female aged 5-9	0.00709	0.0142	0.00265	0.0404***	0.0381***	0.00617	-0.119***	0.0103
	(0.0236)	(0.0126)	(0.0159)	(0.0112)	(0.00929)	(0.0118)	(0.0312)	(0.00564)
Female aged 10-14	-0.00795	0.00427	0.00784	0.00390	0.0242*	0.00806	-0.0468	0.00652
	(0.0227)	(0.0128)	(0.0148)	(0.00983)	(0.00970)	(0.0118)	(0.0322)	(0.00534)
Female aged 15-55	0.00818	0.0187	0.0146	0.00991	0.0168*	0.0251*	-0.103***	0.0101*
	(0.0199)	(0.0110)	(0.0133)	(0.00850)	(0.00755)	(0.0104)	(0.0264)	(0.00468)
Female aged 55+	0.0333	0.0429**	0.0133	0.0144	0.00919	0.0779***	-0.210***	0.0187**
	(0.0263)	(0.0142)	(0.0169)	(0.0116)	(0.0100)	(0.0137)	(0.0349)	(0.00657)
Age	0.000146	0.000125	0.000345**	0.000195*	-0.000190**	0.000141	-0.000901***	0.000141***
	(0.000179)	(0.0000877)	(0.000118)	(0.0000890)	(0.0000727)	(0.0000886)	(0.000246)	(0.0000421)
Education	-0.00313***	0.00107***	0.00149***	0.00173***	0.000553*	-0.000666*	-0.00127	0.000231
	(0.000520)	(0.000266)	(0.000339)	(0.000243)	(0.000235)	(0.000261)	(0.000724)	(0.000123)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	staple	veg	meat	dairy	dried	spices	bev	other
Female	0.0146*	0.00354	-0.00544	0.000943	0.0144***	0.00498	-0.0339***	0.000928
	(0.00723)	(0.00363)	(0.00485)	(0.00382)	(0.00381)	(0.00397)	(0.00991)	(0.00169)
Married	0.0795***	0.0289***	0.0437***	0.00826**	-0.00929**	0.0536***	-0.216***	0.0115***
	(0.00548)	(0.00309)	(0.00378)	(0.00272)	(0.00309)	(0.00320)	(0.0102)	(0.00144)
Employed	0.00190	-0.00436	-0.00794	-0.00968**	-0.0185***	0.00115	0.0412**	-0.00377*
	(0.00763)	(0.00461)	(0.00410)	(0.00333)	(0.00443)	(0.00444)	(0.0130)	(0.00188)
Islam	-0.0101	-0.00745*	0.000684	-0.00405	0.00178	0.0121***	0.00249	0.00453**
	(0.00705)	(0.00324)	(0.00429)	(0.00314)	(0.00297)	(0.00331)	(0.00995)	(0.00162)
Hindu	0.00935	0.00681	-0.0198***	0.00202	-0.00805*	0.00942*	0.000365	-0.000155
	(0.00971)	(0.00486)	(0.00590)	(0.00457)	(0.00382)	(0.00443)	(0.0129)	(0.00190)
Javanese	-0.0230***	0.00515***	-0.0358***	0.00817***	-0.00207	0.00283	0.0336***	0.0111***
	(0.00298)	(0.00151)	(0.00209)	(0.00157)	(0.00135)	(0.00150)	(0.00414)	(0.000706)
Sundanese	-0.0149***	-0.00525*	-0.0366***	0.00309	-0.00114	-0.0111***	0.0581***	0.00775***
	(0.00404)	(0.00211)	(0.00280)	(0.00195)	(0.00185)	(0.00197)	(0.00611)	(0.000990)
Minang	-0.000596	-0.0106***	-0.0105**	0.00132	-0.00780**	0.0139***	0.0158	-0.00147
	(0.00641)	(0.00287)	(0.00396)	(0.00309)	(0.00260)	(0.00345)	(0.00836)	(0.00134)

	(1) staple	(2) veg	(3) meat	(4) dairy	(5) dried	(6) spices	(7) bev	(8) other
Batak	0.00806	-0.00679	0.0203***	-0.00713*	-0.0211***	0.0166***	-0.00922	-0.000794
	(0.00784)	(0.00352)	(0.00515)	(0.00350)	(0.00355)	(0.00374)	(0.0111)	(0.00207)
constant	0.629***	0.0689***	-0.125***	-0.129***	0.144***	0.232***	0.0807	0.0983***
	(0.0323)	(0.0165)	(0.0206)	(0.0144)	(0.0165)	(0.0182)	(0.0467)	(0.00918)
Coefficient of IV from the first-stage regression	0.021**	0.035**	0.042**	0.054**	0.012**	0.024***	0.122**	0.101*
Obs.	12,625	12,625	12,625	12,625	12,625	12,625	12,625	12,625
Prov*Yr FE	yes	yes	yes	yes	yes	yes	yes	yes

Notes: The dependent variables for columns 1-8 are the budget shares of eight different food groups. The key independent variable is migrant: a migrant-household dummy equal to 1 when the household contains at least one migrant and 0 otherwise. All specifications include province-year FEs, household size and other demographics, household-head characteristics, ethnicities of the head of the household, religion, and urban dummy. Other controls are dummy for whether or not a household experienced any natural disasters in the past and log of household asset. An interaction of log of distance from the residence of origin to the destination with migration network have been used as an instrument for endogenous variable, migrant. The robust standard errors are shown in parentheses and clustered at the community (i.e. EA) level.

***Significant at 1% level; ** Significant at 5% level; *Significant at 10% level

Table B3.7: OLS regression of migration on carbohydrate-rich food. (Dependent variable: carbohydrate's expenditure share: $\frac{carbohydrate_{it}}{carbohydrate_{it}+protein_{it}}$)

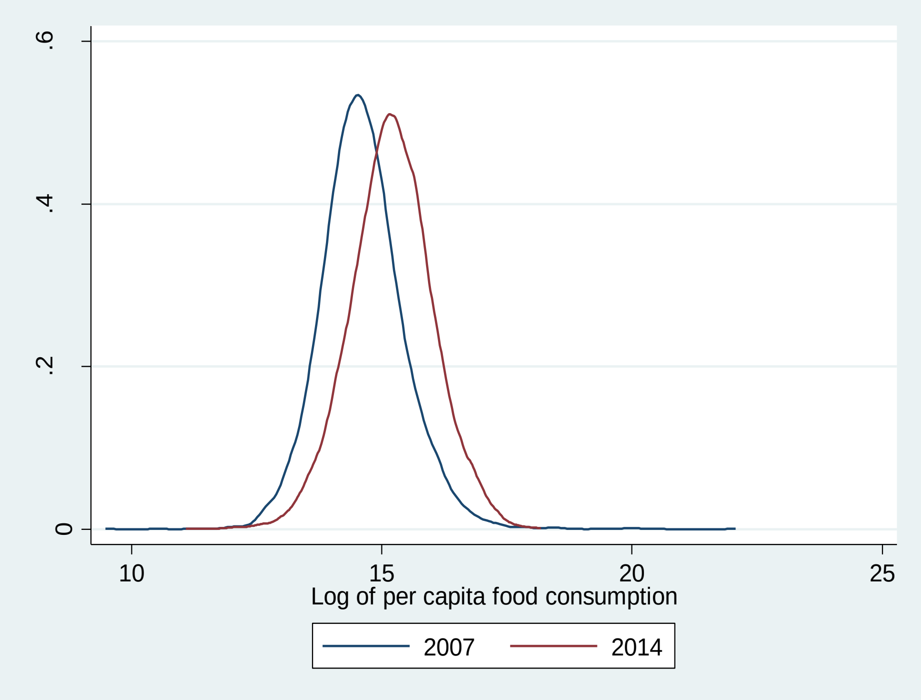
Specification	(1) Full sample	(2) Short migrant	(3) Long migrant	(4) Work migrant
Migrant	-0.026***			
	(0.005)			
Short migrant		-0.031***		
		(0.006)		
Long migrant			-0.026***	
			(0.011)	
Work migrant				0.025***
				(0.008)
Log of per capita food exp.	-0.092***	-0.085***	-0.085***	-0.086***
	(0.004)	(0.006)	(0.006)	(0.006)
Log of household size	-0.022***	0.005	0.005	0.006
	(0.008)	(0.011)	(0.011)	(0.011)
Pro of male aged 0 to 4	-0.152***	-0.120*	-0.120*	-0.120*
	(0.043)	(0.071)	(0.071)	(0.071)
Pro of male aged 5 to 9	-0.075*	-0.058	-0.058	-0.057
	(0.043)	(0.075)	(0.075)	(0.075)
Pro of male aged 10 to 14	0.039	0.144*	0.142*	0.144*
	(0.046)	(0.081)	(0.081)	(0.081)
Pro of male aged 15 to 55	0.000	0.079	0.080	0.079
	(0.038)	(0.068)	(0.068)	(0.068)
Pro of male aged 55-70	0.116**	0.147*	0.148*	0.148*
	(0.051)	(0.087)	(0.087)	(0.087)
Pro of female aged 0 to 4	-0.152***	-0.107	-0.107	-0.106
	(0.043)	(0.071)	(0.071)	(0.071)
Pro of female aged 5 to 9	-0.036	0.023	0.023	0.024
	(0.044)	(0.076)	(0.076)	(0.076)
Pro of female aged 10 to 14	-0.022	0.082	0.079	0.083
	(0.045)	(0.084)	(0.084)	(0.084)
Pro of female aged 15 to 55	0.009	0.062	0.062	0.063
	(0.037)	(0.066)	(0.067)	(0.066)
Pro of female aged 55-70	0.015	0.073	0.073	0.075
	(0.048)	(0.085)	(0.085)	(0.085)
Age of head	-0.001*	-0.001*	-0.001*	-0.001*
	(0.000)	(0.000)	(0.000)	(0.000)
Education of head	-0.009***	-0.008***	-0.008***	-0.008***
	(0.001)	(0.002)	(0.002)	(0.002)
Female head	0.008	0.051**	0.052**	0.051**
	(0.014)	(0.022)	(0.022)	(0.022)

Specification	(1) Full sample	(2) Short migrant	(3) Long migrant	(4) Work migrant
Married head	0.086***	0.077***	0.077***	0.077***
	(0.016)	(0.019)	(0.019)	(0.019)
North Sumatra	0.045***	0.044**	0.044**	0.045**
	(0.014)	(0.018)	(0.018)	(0.018)
West Sumatra	0.088***	0.093***	0.093***	0.094***
	(0.023)	(0.029)	(0.029)	(0.029)
South Sumatra	0.063***	0.057***	0.057***	0.057***
	(0.014)	(0.019)	(0.019)	(0.019)
Lampung	0.100***	0.113***	0.113***	0.114***
	(0.014)	(0.020)	(0.020)	(0.020)
West Java	0.056***	0.050***	0.050***	0.050***
	(0.011)	(0.015)	(0.015)	(0.015)
Central Java	0.103***	0.098***	0.098***	0.098***
	(0.011)	(0.016)	(0.016)	(0.016)
Yogyakarta	0.060***	0.059**	0.059**	0.060**
	(0.015)	(0.024)	(0.024)	(0.024)
East Java	0.058***	0.041***	0.041***	0.042***
	(0.010)	(0.015)	(0.015)	(0.015)
Bali	0.078***	0.070**	0.070**	0.070**
	(0.023)	(0.034)	(0.034)	(0.034)
West Nusa Tenggara	0.101***	0.091***	0.091***	0.091***
	(0.012)	(0.017)	(0.017)	(0.017)
South Kalimantan	0.020	0.033*	0.033*	0.034*
	(0.012)	(0.018)	(0.018)	(0.018)
South Sulawesi	0.006	0.006	0.005	0.006
	(0.013)	(0.019)	(0.019)	(0.019)
Constant	1.493***	1.310***	1.314***	1.308***
	(0.065)	(0.100)	(0.100)	(0.100)
Ethnicity controls	yes	yes	yes	yes
Religion controls	yes	yes	yes	yes
Province-year FE	yes	yes	yes	yes
Observations	11,690	5,956	5,956	5,956
R-squared	0.117	0.099	0.099	0.099

Notes: The dependent variable for columns 1-4 is carbohydrate food share over protein-rich food. The key independent variable is migrant, a migrant-household dummy equal to 1 when the household contains at least one migrant and 0 otherwise. All specifications include province year FEs, household size and other demographics, household-head characteristics, ethnicities of the head of the household, religion, and urban dummy. Other controls are dummy for whether or not a household experienced any natural disasters in the past and log of household asset. Short migrant is defined as at least one member of the household moved from the home for one year to five years and long migrant is defined as at least one member of the household moved from the home for more than five years. The robust standard errors shown in parentheses and clustered at the community (i.e. EA) level.

***Significant at 1% level; ** Significant at 5% level; *Significant at 10% level

Figure B3.1: Kernel density estimate of log of per capita food consumption



Chapter 4 Essay 3: Education and Food Consumption Patterns

4.1 Introduction and background

The positive association between education and living standards is a well-established fact in social sciences (Case, 2006, Duflo, 2001, Psacharopoulos, 1994, Case and Deaton, 1999, Lucas, 1988). However, what are the channels for this association are still under debate. In recent times, there has been an increasing interest to investigate how education may affect choices in food consumption. The focus on food consumption is important, since in developing countries, food is closely related to welfare and poverty (Todaro and Smith, 2012 and Goulet, 1997). Furthermore, studying the impact of education on food consumption may help to uncover food behaviour for the design national nutritional policies. In a recent study, Wantchekon, Klasnja and Novta (2015) have noted that education can have an intense transformational impact on individuals and communes. Another study by Moreira and Padrao (2004) found that education is one of the key elements to influence shifting consumption towards healthy food groups like vegetables and fruits. However, these studies have not ascertained whether the relationship is causal.

While the monetary returns to education (for instance, impact of education on earnings) for both developing and developed countries, including Indonesia's, are well documented, causal studies on nonmonetary returns to education (for example, food consumption) are scarce. In this chapter, I examine the impact that education may have on food consumption. In particular, I test if more educated household head tend to make a healthy food choice or unhealthy food choice. The findings may have policy implications for developing countries to invest more in education if it can be established that better educated individuals would choose to consume healthier foods.

Indonesia has made great advances in many areas, including larger investment in the education and infrastructural development. The Indonesian school structure is massive and diverse. Education is the principal focus to the Indonesian Government's development program. Since the onset of the economic crisis, the spending on education has increased significantly in Indonesia. In real terms, spending on education has doubled between 2000 and 2006. Government expenditure on education was greater than on other sector in 2007. The expenditure on education as a share of GDP is about 3.4 percent as a share of GDP in the recent years. After China, India and USA, the Indonesian education system is the fourth largest in the world and third largest in Asia region with more than 50 million students and 250,000 schools (World Bank, 2014). It is the largest economy in the Southeast Asia with average GDP growth is between 5% and 6.5% for more than a decade. Currently, the country has 9 years of compulsory schooling (CIA World Fact Book, 2014). Moreover, Indonesia's enrolment rates both at primary and secondary levels have been increased dramatically over the last few decades.[33]

[33] Since the 1970's both the primary and secondary enrolment rates have increased dramatically in Indonesia (Economist, 2014). One reason for this huge increase in enrolment has been identified by Duflo (2001) that between 1973 and 1978 more than 61,000 primary schools built in Indonesia under the major school construction program, the Sekolah Dasar INPRES program.[33] About ten years later (initiated in 1973) government implemented compulsory

The recent literature on the returns to education underscores that returns vary across individuals, and are correlated with the education. In terms of an equation, $Y = \alpha + \beta S + \varepsilon$ (where *Y* is log of food consumption and S is years of schooling), β can be interpreted as random coefficient and is potentially correlated with S. For example, the individual at the margin between two different levels of education may have different returns from all the infra-marginal individuals (Carneiro, Lokshin, and Umapathi, 2016). This study investigates food consumption returns to education in Indonesia under the assumption that β may vary across individuals and correlated with *S*.

There are several reasons why I focus on the impact that education has on food consumption. The first reason is due to survival and existence: food consumption is fundamental for the existence of life regardless of the level of education attained by an individual. Hence from the policy standpoint on welfare, consumption is more relevant than earnings as labour income would ultimately translate into consumption.

Because lifetime consumption is smoother than income, consumption resembles more like a log normal distribution than income itself. Battistin, Blundell, and Lewbel (2009) observe that consumption expenditures across households are more log normally distributed while

---

education for primary school children (7-12 years). Consequently, primary school participation rate rose to 92 percent in 1993 compared to 79 percent 10 years before.[33] Again in 1994, the country expanded compulsory education to 9 years for every Indonesian in the 7-15 age group. Since 2009, the government has allocated one-fifth of its yearly budget in education.

significant departures from log normality is found in income data. They note that consumption within a cohort has several implications for welfare and econometric modelling. Likewise income data are noisier than consumption. It is anticipated that education has both lifetime productive returns and have had more role on consumption smoothing that can serve as a better measure of welfare than other educational outcomes, such as permanent income and earnings (Fulford, 2014).

Moreover, it is possible that a person does not have any earnings in the short run. For instance, a labourer who works in the agricultural sector sometimes suffers in seasonal unemployment and ends up with no earnings. However, he has to consume regularly, which makes consumption proportional to lifetime resources and reflects living standard throughout the year (see, Deaton and Grosh, 1998; Musgrove, 1978 and 1979; Paxson, 1992 and 1993; Wolpin, 1982).[34]

The second reason is that it is plausible that higher carbohydrate consumption due to lack of proper nutrition knowledge may result in health-risk; for example, among the poor segments in developing countries, higher price shock may force them to live below the required amount of calories (Abdulai and Aubert, 2004).

[34] A remarkable quotation about 200 years back by Anthelme Brillat -Savarin noted in Anand and Sen (1998), "Tell me what you eat," and I will tell you what you are."

The final reason is that everyone in the households consume, although not all members earn and have similar levels of education. Head of the household invest in children and schools in the expectation of accruing potential income in the household at the cost of current postponed adult consumption. As such, education has a prospective consumption returns.

Applying IFLS 2014 data to a semi-parametric model, this study finds that individuals who have received upper secondary school or higher levels of education, on average, consume 31.5 percentage points' more healthy foods than those who have lower secondary school education or less. In terms of unhealthy food consumption, more-educated individuals, on average, consume 22.8% less unhealthy foods than less-educated individuals.

This study contributes to the food consumption literatures in two ways. First, this study is unique in the sense that it attempts to provide the first quasi-experimental evidence of the impact of education on food consumption patterns in Indonesia, while the literature has mainly focused on the impact of education on earnings or health or schooling for the next generation. In particular, this study investigates the role of education on choosing healthy food group or unhealthy food group by exploiting an exogenous variation of schooling – time required to attend school – to construct an IV. Second, it adds to the literature on the consumption returns to education by attempting to estimate the causal effect of education, while the literature has mainly focused on the correlation between education and food consumption (see, for example, Michael, 1975; Fulford, 2014; Bhandari, 2008; and Alem and Soderbom, 2012).

4.2 Literature review

Three interrelated groups of literatures are dominant in explaining the impact of education on sociol-economic outcome: earnings, health, and growth. A vast number of literature investigates the relationship between education and earnings, including compulsory schooling and earnings ((Angrist and Krueger, 1991; Stephens and Yang, 2014); returns to schooling from Sibling data (Ashenfelter and Krueger, 1994; Butcher and Anne, 1994; schooling and selectivity bias (Garen, 1984, education, ability and earnings)).

Koc and Kippersluis (2015) investigated Discrete-Choice-Experiment (DCE) of educational disparities on making food consumption and found that health knowledge differentials play a greater role in education disparity in food in Netherlands. Cutler and Lleras-Muney (2010) finds that there is a strong disparity in healthy behaviours like diet choice across education groups. Haines, Guilkey and Popkin (1988) examined the food consumption decisions as a two-step process decomposing food groups into low fat milk versus high fat milk and high fat low fibre bread group. Their findings suggest that decision to consume a specific food within a food group is statistically significantly different from how much to consume for more broadly defined food groups.

Fulford (2014) examined the returns to education in India and found that an additional year of education brings 4 percent more consumption of male cohorts with no extra consumption for female cohorts. A related study in Ethiopia by Alem and Soderbom (2012) found that a significant percentage of households adjust food consumption due to large price shock. Yen, Lin, and Davis (2008) explored the linkage between consumer knowledge and meat consumption both at home and away from home and found that dietary knowledge reduces beef and pork consumption both at home and away from home and men consumes more meat and fish than women. Abdulai and Aubert (2004) conducted parametric and nonparametric analysis of calorie consumption in the presence of behavioural heterogeneity and measurement error using panel data from Tanzania and concluded that higher food prices could reduce calorie demand significantly, and as such, it would be important to allocate targeted food subsidies for poor households. Cain et al. (2010) conducted an empirical study using household level consumption expenditure data from India and concluded that the amount of inequality generated by the education of household heads were much greater than the sum of all other household characteristics. In particular, they found that education accounted for 8% and 9% of inequality in 1993 and 2004 in rural areas and the respective figures are 25% and 28% in urban areas in the same years. They also found that when inequality decomposition is carried out without incorporating occupation into main specification, education could explain 53% of the increase in the Gini coefficient in rural areas and 57% of the increase in the Gini coefficient in urban areas. Overall, their finding implies that education may play a role in increasing income inequality.

Studies also found that consumption expenditures are strongly correlated with education. In a recent study on Nepal, Fafchamps and Shilpi (2014) have found that there is a strong statistical association between male education and household’s welfare even after controlling for educational attainment within their birth cohort. In another benchmark study, Michael (1975) has found that the education elasticity of goods is -0.07 and of services is 0.19, meaning that an additional year of schooling shifts the spending patterns toward services. A number of experimental studies have linked diseases to the choices of food consumption. Therefore, health and development practitioners have been concerned about understanding which factors could influence consumption patterns of the population (Fraser et al. 2000, Chait et al. 1993, Potter, 1997, and Denke, 1997).

Ricchiuto, Tarasuk, and Yatchew (2006) characterized the role of household's socioeconomic status on choosing a particular food in Canadian households. They concluded that irrespective of household size, income and composition, higher education was associated with the purchase of larger quantities of vegetables, milk and food products. Interestingly, households with post-secondary education purchased 6% more fruit and vegetables than those have fewer than 9 years of schooling. Some other studies also obtained similar findings that higher income and higher education are associated with consuming more vegetables and fruits (Nayga et al., 1999; Groth et al., 2001, Perez, 2002).

Duflo (2001) found that the economic returns (i.e. earnings) from an additional year of education in Indonesia ranges from 6.8% to 10.6%. Though she generated huge variations of schooling by exogenous district-level changes of the number of schools constructed by the Indonesian government, other local factors (for instance, district level teacher-student ratio) could threatened the exclusionary restriction of her instrument. Purnastuti, Salim, and Joarder (2015) examined returns to schooling in Indonesia using IFLS 2007 and an IV approach. Their

OLS estimates show that the returns to schooling is 4.36 percent for males and 5.26 percent for females. However, the relationship between education and earnings is not significant when IV is used. A similar work by Comola and Mello (2010) for the Indonesian labor market had shown that the returns to education from 9.49% to 10.32%, although their work did not address the identification issues that are standard when estimating the returns to education. Dumauli (2015) examined the private returns to education in Indonesia accounting for sample selection and endogeneity issues. The household FE estimates indicate the returns to education fell from 10.8% to 5% between 1986 and 2007. This may be a reason for why college enrolment rate in Indonesia has stagnated during this period.

The theoretical foundations of schooling as a formation of human capital and its impact on monetary returns are quite strong in the literatures. However, the impact of education on non-monetary returns are scarce in the empirical literatures, though studying education's impact on non-pecuniary outcomes are very indispensable at its own right. This paper exploits quasi-natural experiment to estimate the food consumption returns to education in Indonesia which is the main departure from the existing studies that investigate the returns to education.

4.3. Conceptual framework of linking education to consumption

Individuals in the household make their education decisions by comparing potential returns with the costs of education. Individuals are assumed to be utility maximizers of becoming graduate if the expected utility from graduating is greater than the expected utility from not graduating. The household is assumed to derive utility from consuming health food items and unhealthy foods. Healthy and unhealthy consumption goods could be purchased or self-produced in the households. It is assumed that the consumption of healthy foods increases utility and the consumption of unhealthy food decreases utility.

The household utility is maximized when marginal benefits of healthy consumption is equal to the marginal monetary cost of healthy food consumption. Likewise, household is said to have utility maximiser when marginal utility individuals derive from the consumption of unhealthy goods is equal to the monetary (health) cost of unhealthy consumption.

There are a number of ways that education might affect food consumption. First, education may enhance the capacity of understanding of nutritional aspects of variety of foods that may lead to consume the healthy foods. Second, more highly educated people may have higher earnings potentials, which may lead to greater access to varied food groups in the market. Third, more highly educated individuals may spend more time on media and newspaper that has a coverage of food and nutrition and thus may have more knowledge of how unhealthy food choices would be threatening to health. Fourth, more highly educated individuals may build up healthier dietary habits more quickly than lower educated individuals. Fifth, they may have broader cultural views of food consumption that lead to diversity of food consumption. Finally, compared to other demographic factors such age and gender, education is a policy variable that is responsive to government interventions. If government is concerned with the health aspects of individuals, she may introduce nutrition education in the schools.

4.4 Data and descriptive statistics

The empirical analysis in this paper draws on the publicly available household level panel datasets of IFLS. I use data from the fifth wave of the IFLS fielded from September 2014 to March 2015. The IFLS 2014 is an ideal data choice for our case as Indonesia has been gradually transforming into decentralized economy followed by deregulation and government has been allocating more finance to expand for since 2000. In particular, implementation of compulsory education policies in 1974 and 1984 approved by the government are likely to facilitate cohort-specific individuals to complete full time education. For a detailed description of the IFLS 2014 survey see Strauss, Witoelar, and Sikoki (2016).

IFLS collects a wide range of information at the individual, household and community level. The IFLS sample is drawn from 321 randomly selected villages, covering 13 Indonesian provinces and representing 83% of the country's population. The last survey is carried out in 2014. The sub-sample I use consists of household head aged 15-70, and who have reported non-missing food consumption and schooling information. The dependent variables in our analysis are: log of per capita healthy food consumption and log of per capita unhealthy food consumption at the household level.[35] The final sample contains about 13000 households.

[35] Following Usfar and Fahmida (2011), I have constructed healthy and unhealthy food groups using IFLS consumption module: i) the main staples, vegetables and fruits, meat and animal products, and fish constitute a healthy food group; and ii) the dried foods, condiments, and other foods constitute an unhealthy food group. Both food groups have been converted into annualized per capita food group at the household level.

Table 4.1 presents descriptive statistics for the main variables used in this study. It shows that individuals with higher secondary or more levels of education have, on average, 0.33 log points higher than those with less than higher secondary education. They have 8.23 extra years of schooling. Graduates from the higher secondary schools or more are likely to come from families with better educated parents and have fewer household members in the family. Higher secondary or more educated individuals are more likely to live in urban areas than rural areas and are less likely to consume unhealthy foods. They also tend to live in the proximity of high schools than lower educated persons.

Table 4.1: Summary statistics for the treatment and control groups

	Higher secondary or more (Treatment group)	Less than higher secondary (Control group)
	N = 9945	*N* = 4978
Log of per capita healthy food consumption	14.557 (0.872)	14.231 (0.861)
Log of per capita unhealthy food consumption	14.374 (0.966)	14.892 (0.934)
Years of education	14.787 (3.348)	6.508 (1.856)
Household size	5.046 (3.115)	5.817 (3.135)
Age	39.098 (11.866)	44.632 (12.867)
Sex	0.853 (0.353)	0.822 (0.382)
Employment	0.721 (0.448)	0.683 (0.665)
Married	0.863 (0.343)	0.955 (0.206)
Muslim	0.852 (0.354)	0.914 (0.278)
Catholic	0.021 (0.144)	0.009 (0.095)
Protestant	0.059 (0.235)	0.037 (0.188)
Other	0.066 (0.249)	0.038 (0.193)
Javanese	0.392 (0.488)	0.465 (0.498)
Sundanese	0.108 (0.311)	0.136 (0.343)
Minang	0.065 (0.247)	0.045 (0.208)
Other	0.434 (0.495)	0.352 (0.477)
Fathers education	6.123 (2.067)	5.202 (2.425)
Mothers education	4.112 (2.101)	4.011 (2.001)
Distance to school (minutes)	16.145 (12.185)	16.321 (12.079)
Distance to health post (km)	5.009 (8.357)	6.183 (10.461)
Rural household	0.249 (0.432)	0.462 (0.498)
North Sumatra	0.080 (0.271)	0.074 (0.261)
West Sumatra	0.046 (0.210)	0.041 (0.198)
South Sumatra	0.048 (0.214)	0.046 (0.210)
Lampung	0.028 (0.166)	0.045 (0.208)
Jakarta	0.078 (0.269)	0.059 (0.235)
Central Java	0.090 (0.286)	0.136 (0.342)
Yogyakarta	0.066 (0.249)	0.040 (0.196)
East Java	0.108 (0.310)	0.149 (0.356)
Bali	0.063 (0.244)	0.039 (0.194)
West Nusa Tenggara	0.088 (0.284)	0.062 (0.241)
South Kalimantan	0.043 (0.204)	0.044 (0.205)
South Sulawesi	0.047 (0.212)	0.048 (0.214)
Rural	0.249 (0.432)	0.462 (0.498)

Source: Calculated from the IFLS 2014 and sample is restricted to the non-missing schooling and distance to the school.

4.5 Methodology

The estimation approach in this paper is carried in three steps. First, I set up the model of consumption returns to education. Second, I explicate the endogeneity issues of education. Third, I explain of building up the scenario of IV.

4.5.1 A semiparametric selection model

Estimating marginal food consumption returns to education is a key parameter of interest in this study. In other words, estimation of the food consumption returns to education is one of the central focus for the policy makers to evaluate cost and benefit of the policy (e.g. educational expansion policy of the government) (Carneiro, Heckman, and Vytlacil 2011). Appendix figure C4.6 exhibits density of returns to education intuitively. The mean marginal returns to education can be estimated by the following equation:

$$\ln Y = \alpha + \beta * S + \varepsilon \qquad (1),$$

where lnY is the log of per capita food consumption, S is a dummy variable indicating if an individual (i.e. household head) has had a high school education, β is the returns to schooling of the household head (which may differ among individuals), and ε is the residual. The coefficient β would be positive if an individual chooses healthy food groups and it would be negative if he chooses an unhealthy food group. It is assumed here that knowledge about diet-health relationship induces what to consume. The effect of graduating from high school on food consumption may be confounded by self-selection. I will address this issue by using an IV or implementing model that corrects for selection.

I follow Carneiro, Lokshin, and Umapathi's (2016) model of potential outcomes applied to education. Consider a model with two levels of schooling:

$$\ln Y_1^{(k)} = \alpha_1 + X\beta_1 + U_1 \quad (2)$$

$$\ln Y_0^{(k)} = \alpha_0 + X\beta_0 + U_0 \quad (3)$$

$$\mathrm{S} = 1 \; if \; Z\lambda - U_s > 0 \quad (4)$$

$\ln Y_1^{(k)}$ is the log of per capita food consumption with k equal to healthy or unhealthy food consumption if the head of the household have completed higher secondary education or more, $\ln Y_0^{(k)}$ is the log or per capita food consumption if the head of the household have completed less than secondary education, X is a vector of observable characteristics which affect food consumption, and U_1 and U_0 are the error terms, Z is the other vector of household characteristics affecting schooling. From now onwards, I drop superscript *k* for convenience.

Equation 11 can be rewritten as:

$$\mathrm{S} = 1 \; if \; P(Z) > V \quad (5)$$

where $P(Z) = F_{U_s}(Z\lambda)$ and $V = F_{U_s}(U_s)$ and F_{U_s} is a cumulative distribution function of U_s. V is assumed to be uniformly distributed.

The observed food consumption can be written as:

$$\ln Y = S\, lnY_1 + (1 - S) \ln Y_0 \qquad (6)$$

Food consumption returns to schooling can be expressed as:

$$\ln Y_1 - lnY_0 = \alpha_1 - \alpha_0 + X\,(\beta_1 - \beta_0) + U_1 - U_0 \qquad (7)$$

Note from Equation 14 that consumption returns to education vary across individuals with different *X'*s and different U_1 and U_2. This implies heterogeneity in returns to education and the Equation 14 also aids to make a distinction between the returns for average and marginal individuals.

The marginal treatment effect (MTE) is the main object to be estimated here. The MTE can be expressed as follows:

$$\mathrm{M}TE(x, v) = E(lnY_1 - lnY_0 | X = x, V = v) = \alpha_1 - \alpha_0 + x(\beta_1 - \beta_0) + E(U_1 - U_0 | X = x, V = v) \qquad (8)$$

The MTE measures the consumption returns to education for individuals with different levels of observables (*X*) and unobservables (*V*). Hence, MTE offers a characterization of heterogeneity in returns. $(\beta_1 - \beta_0)$ could be positive or negative. If parents were highly educated, children would be more educated as well. In this case, we expect higher consumption returns to schooling, which implies that $(\beta_1 - \beta_0)$>0. If parents were less educated, children would be less educated as well. In this case, we expect lower consumption returns to education. Likewise, *V* may be interpreted as positive unobserved ability. Individuals with high values of *V* are less likely to enrol in school than those with low values of *V*. If this is the case, $E(U_1 - U_0|X = x, V = v)$ would tell us how the returns to education vary with unobserved ability.

Several other parameters of interest can be constructed as weighted averages of the MTE. Heckman and Vytlacil (2005) show that the following important parameters can be derived from MTE:

$$ATE(x) = \int MTE\ (x, v) \int f_{v|x}\ (v|x)dv \qquad (9)$$

$$ATT(x) = \int MTE\ (x, v) \int f_{v|x}\ (v|x, S = 1)dv \qquad (10)$$

$$ATU(x) = \int MTE\ (x, v) \int f_{v|x}\ (v|x, S = 0)dv \qquad (11)$$

Where *ATE(x)* is the average treatment effect, *ATT(x)* is average treatment on the treated, *ATU(x)* is average treatment on the untreated (conditional on *X*=*x*, and $f_{v|x}(v|x)$ is the density of *V* conditional on *X*.

Another policy related parameter of interest can be constructed as follows:

$$\mathrm{P}RTE\ (x) = \int MTE\ (x, v) \int f_{v|x'}(v|x, S(z) = 0, S(z') = 1)dv \qquad (12)$$

Where PRTE(x) is the policy related treatment effect which measures the average returns to schooling for those induced to change their schooling decision in response to a specific policy assuming policy shifts Z from Z=z to Z=z'.

To estimate MTE, I use the method of local IV that imposes no distributional assumptions on the unobservables of the model, apart from the assumptions that X and Z are independent of error terms.

There are two steps in the procedure for estimating MTE. The first step is to estimate a regression of the outcome lnY, on X and P (propensity scores). This step can be written as follows:

$$\mathrm{E}(lnY|X, P) = E[\alpha_0 + X\beta_0 + S(\alpha_1 - \alpha_0) + SX\ (\beta_1 - \beta_0) + U_0 + S(U_1 - U_0)|X, P$$

$$= \alpha_0 + X\beta_0 + P(\alpha_1 - \alpha_0) + PX\ (\beta_1 - \beta_0) + E(U_1 - U_0|S = 1, X, P)P$$

$$= \alpha_0 + X\beta_0 + P(\alpha_1 - \alpha_0) + PX\ (\beta_1 - \beta_0) + K\ (\mathrm{P}) \qquad (13)$$

K(P) is a flexible function of P. We will estimate it using a non-parametric method, such as, local linear regression. The regression in (19) is partially linear, where *X* and *XP* are partially linear, and the function *K (P)* is non-parametrically estimated.

The second step is to take derivative of (9) with respect to *P* to get MTE:

$$\mathrm{M}TE(x,v) = \frac{\delta E(Y|X,P)}{\delta P} = X(\beta_1 - \beta_0) + K'(P) \qquad (14)$$

Therefore, the local IV estimator for the Equation 1 and 2 just requires regressing *ln(Y)* on *X* and *P* and taking the partial derivative of the estimated regression function with respect to *P*.

4.5.2 Endogeneity of education

If the error term in Equation 1 (the food consumption equation) were orthogonal to the regressors, the OLS estimates of the parameters in Equation 1 will be unbiased and consistent. However, these OLS estimates, which capture the impact of education on food consumption, may could be biased could be biased, particularly in the case of self-reported schooling. It is noted in paper by Lleras-Muney (2005) that the OLS estimates of education may be biased due to measurement error in education. Likewise, Angrist and Krueger (1999) conclude that with no controls, measurement error can shrink returns to education by about 10 percent. Griliches (1997, 1979) underscores that measurement errors in education would lead a downward bias in the estimates of schooling on earnings. Card (1999) points out that errors in reported schooling may be mean-regressive as higher educated individuals cannot state positive errors in schooling and lower educated individuals cannot state negative errors in schooling.

Equation 1 may also suffer from endogeneity problem due to omitted variable. For example, ability or current health condition of a household head may be correlated with both education and food consumption. Or, there may arise endogeneity due to reverse causation. For instance, healthy food choices may increase life expectancy which would lead to persuade to take schooling at a longer time.

4.5.3 Validity of the instrumental variable

I deal with the endogeneity of education by employing IV and MTEs. To do so, I instrument years of schooling with the time required (measured in minutes) to attend school from home (distance to the school). Distance to the nearest school as an instrument is first used by Card (1993) and successively has been used by other researchers (Kane and Rouse, 1995, Kling, 2001, Currie and Moretti, 2003, Cameron and Taber, 2004, Carnerio, Heckman and Vytlacil, 2011, and Carnerio, Lokhsin, Ridao-Cano, and Umapathi, 2016). When compulsory education laws comes into effect (e.g. Compulsory 6 years schooling law in 1974 and compulsory 9 years schooling law in 1984 in Indonesia), the decision to take further education is less about tuition and more about location. This is particularly true for developing countries where there is inadequate infrastructure, transportation, and schools.

In order for distance to the school to be a valid instrument, it needs to satisfy two conditions: i) it should have a strong correlation with years of schooling and ii) it should have no direct linkage with outcome variable (food consumption per capita). In this connection, two main concerns need to be mentioned. First, households and schools may not be located randomly across localities in Indonesia. Second, though distance is measured while a student completed education, distance is not measured at the time of schooling decision. As it is measured at the time of survey year, there is likely reporting error in minutes or it's an approximation like Card (1995). The main problem with this sort of approximation is that educated families or individuals may have already shifted to the place where there are more private and public schools and good infrastructures. This could entail reverse causality in the first stage relationship.

A consistent estimate of the impact of education on food consumption can be obtained if there would be obtained a variable that affects education but not food consumption. In this case, one needs to identify a causal determinant of education that can be legitimately excluded from the food consumption equation. Distance from the community to the nearest secondary schools may be used as a valid instrument to education as it presumably affect food consumption only through education, and it may not influence other household level determinants of food consumption. For instance, individuals who grow up in an area without secondary schools face a higher cost of attending schools. This higher cost reduces investments in further education, particularly for the children with low-income families. This implies that raising up near secondary schools may have a larger impact on the education outcomes at least in terms of enrolment, particularly for the children from economically disadvantaged families.

There are some reasons why individuals who grow up in close proximity to schools may have higher consumption than others, controlling for education, location, and parental education. First, Families that put a strong emphasis on education may choose to reside at a nearby schools. Children of such family background may have higher earning potentials the labour market that would translate into higher consumption. Second, the presence of schools at nearby communities reduces transaction cost, may be due to easy driving distance, to attend in the schools. Hence, families have some scopes to allocate more on consumption. Finally, it is possible that nearby schools are also located in close proximity to industries and markets. Hence, both higher labour market earnings potentialities and greater access to goods may result in higher consumption. So, all of the points imply that distance to schools from the community is likely to serve as a valid instrument to education.

Table 4.2 indicates that distance to the nearby higher secondary school in the community is a strong predictor of schooling enrolment in the higher secondary level. I run a logit regression where the dependent variable is a dummy variable that indicates (i.e. is equal to 1) if the individual has attended in higher secondary schools or more. I control for age, employment and marital status of the household head, religions, ethnicities, parental education, distance to the nearest school, and distance to the nearest health post as a proxy for location geographies, and rural dummy as a proxy for area.

Table 4.2: Higher secondary school decision model – average marginal derivative

	(1) Coefficients	(2) Average Derivative
Distance to school in minutes	-0.223***	-0.020***
	(0.030)	(0.006)
Age	0.036**	-0.015**
	(0.002)	(0.001)
Employment	0.172	0.172
	(0.164)	(0.164)
Married	-1.312*	-0.019*
	(0.744)	(0.744)
Protestant	0.640***	0.042***
	(0.157)	(0.003)
Catholic	0.914*	0.014*
	(0.503)	(0.008)
Other religions	0.618*	0.618*
	(0.330)	(0.330)
Javanese	-0.567***	-0.567***
	(0.172)	(0.172)
Sundanese	-0.762***	-0.762***
	(0.253)	(0.253)
Minang	0.302	0.302
	(0.352)	(0.157)
Father higher education	0.824***	0.231***
	(0.040)	(0.003)
Mother higher education	0.441***	0.154**
	(0.003)	(0.127)
Rural	-0.918***	-0.092***
	(0.168)	(0.101)
Distance to health post in km	-0.019	-0.025
	(0.008)	(0.005)
Constant	1.256	1.112
	(0.836)	(0.225)

Location fixed effect	yes
Test for joint significance of instruments: Chi-square/p-value	11.63/0.0011

Notes: This table presents the coefficients and average marginal derivatives from a logit regression of higher secondary school attendance (an indicator variable that is equal to 1 if an individual has ever attended higher secondary school or more and equal to 0 if he has never attended higher secondary school but graduated from lower secondary school) on household level observables. Type of location is controlled for using province dummy variables. Column 1 presents coefficients of logit regression where only distance to the secondary school is used an IV. In Column 2 average derivatives (computed at the mean values of X) are presented and instruments include distance to secondary school and interactions with all the Xs. Reference category for religion is Muslim and for ethnicities is other ethnic group. Robust Standard errors are shown in parenthesis and clustered at the community level.

***significant at 1 percent level; **significant at 5 percent level and; *significant at 10 percent level.

4.6 Empirical results

4.6.1 Standard estimates of the food consumption returns to education

In my empirical model, I divide schooling into two categories: i) completed lower secondary or below, and ii) completion of upper secondary or higher. While this division bands together different levels of schooling, it simplifies the model and has been considered in the literature (e.g., Willis and Rosen, 1979).[36] The transition to upper secondary schooling is of interest in the Indonesian context given its current effort to expand secondary education. Higher secondary schooling corresponds to 10 or more years of completed education in Indonesia. The calculation

[36] Other schooling might have taken as a categorical variables such as no schooling, elementary, high school, college, and tertiary education. Impact of each level of schooling on food consumption at the household level is difficult to establish as no individual food consumption data are available in the surveys.

of length of schooling has been presented in detail in tables C4.1-C4.3 and in note C4.2 in appendix C.

I use the distance (in minutes) from the household to the nearest secondary school in the community as an instrument for schooling. The distance is self-reported by the respondent who is an adult 15 years or older and the distance in recorded in section DL in Book IIIA in the household survey.

There are two endogeneity concerns that come with using the distance of the household to the school as an instrument. First, the distance between school and house can be choice as households may locate themselves closer to school in order to make it easier for their children to go to school. Second, distance may be related to wellbeing of the household if schools are located at central locations and distant households are in general poorer. There are three ways I have dealt with these sources of simultaneity. First, the location of the school may be exogenous when a set of controls are rich enough to capture those sources of endogeneity. I have controlled a detailed set of household and regional characteristics to minimize reverse causality. Specifically, I have used distance to the nearest health post and provincial dummies as a proxy for location characteristics and find that the distance to the nearest health post does not predict school choice. Second, the well-off households may choose live in the close proximity of the schools than poorer households. To minimize this problem of endogeneity, I have used log of household head earning as a control which is shown in the Appendix Table C4.4 and the impact of education on healthy food consumption is found to be statistically insignificant in an instrumental variable regression when log of earnings is added as a control. Third, nonetheless one could concern that instrument is weak. Hence, I have additionally used the estimated propensity score and the interaction of a set of controls and the distance as an instrumental variable to minimize the threat that the results are driven by the confounded by the other household level unobserved factors.

Table 4.3 presents results of standard OLS and IV estimates. The key variable of interest is schooling: takes 1 for household head who has completed higher secondary education or more, and takes 0 for less than higher secondary schooling. I use the log of per capita healthy food consumption and the log of per capita unhealthy food consumption as dependent variables. All specifications in Table 4.3 includes the following controls: household size, age of the household head in years, indicator variables for whether head of the household is employed and married, indicators for main religions (Protestant, Catholic, other religions and Muslim is an omitted category), indicators for main ethnicities (Javanese, Sundanese, Minang, and 'Other ethnicities' are omitted), fathers and mothers education measured with the highest grade completed parental education, distance to the nearest health post in kilometres, an indicator for rural households.

It is observed that individuals with higher secondary school education consume 31.5% more healthy foods than the individuals with lower secondary school education. The corresponding IV estimate is 33%. On the other hand, the reduction of unhealthy food consumption due to upper secondary education from our OLS and IV estimates are 22.8% and 26.9% respectively. In both types of food consumption, IV estimate produces larger coefficient than OLS. It may be attributed to the measurement error in the data or the differences in the magnitude may depend on the choice of instrument.

Table 4.3: OLS and IV estimates of food consumption returns to higher secondary schooling

	(1)	(2)	(3)	(4)
	OLS	IV	OLS	IV
	Healthy food	Healthy food	Unhealthy food	Unhealthy food
Schooling	0.315***	0.330***	-0.228***	-0.269**
	(0.050)	(0.043)	(0.050)	(0.552)
Household size	-0.061***	-0.054***	-0.090***	-0.088***
	(0.007)	(0.008)	(0.009)	(0.012)
Age	-0.004	-0.007	-0.012***	-0.004
	(0.003)	(0.004)	(0.003)	(0.005)
Employment	0.192***	0.203***	0.210***	0.163**
	(0.040)	(0.052)	(0.044)	(0.068)
Married	0.645*	0.210	-0.777***	-0.523
	(0.334)	(0.504)	(0.183)	(0.330)
Protestant	0.108***	0.172**	-0.377***	-0.115**
	(0.026)	(0.055)	(0.132)	(0.051)
Catholic	-0.072	0.038	-0.381**	0.044
	(0.208)	(0.258)	(0.190)	(0.254)
Other religions	-0.125	0.060	-0.515***	0.036
	(0.248)	(0.310)	(0.196)	(0.283)
Javanese	0.040***	0.003***	-0.014**	-0.030**
	(0.010)	(0.001)	(0.007)	(0.011)
Sundanese	-0.063	-0.081	0.043	0.298*
	(0.099)	(0.157)	(0.092)	(0.175)
Minang	0.091	0.116	-0.056	-0.437
	(0.217)	(0.297)	(0.149)	(0.302)
Fathers education	0.017***	0.015***	-0.014***	-0.020***
	(0.002)	(0.007)	(0.001)	(0.003)
Mothers education	0.027***	0.018***	-0.015***	-0.023***
	(0.002)	(0.008)	(0.002)	(0.007)
Rural	0.011	0.459	-0.264***	1.024
	(0.048)	(0.371)	(0.053)	(1.055)
Distance to health post (km)	-0.001	0.001	0.002	0.010
	(0.002)	(0.005)	(0.002)	(0.011)
Location controls	yes	yes	yes	yes
F-test of excluded instruments		11.0		13.11
Observation	8000	8000	8000	8000
R-squared	0.167	0.158	0.301	0.37

Note: This table reports the coefficients for OLS and 2SLS IV for regression of log of per capita healthy food consumption (column I and column II) and log of per capita unhealthy food consumption (column III and column IV) on higher secondary education (higher secondary schooling (a dummy variable that is equal to 1 if an individual has completed upper secondary school and above and equal to 0 if he has completed schooling below upper secondary level). Both OLS and IV regressions control for household demographics, religion, ethnicities, parental education, rural household, distance to the health post and location. Excluded instruments are distance from the office of the community head to the nearest secondary school and interactions with parental education, religion and age. Type of location is controlled using province dummies. Reference categories are other religions for religion, other ethnicities for ethnicities. Standard errors are in parenthesis and clustered at the community level.

***significant at 1 percent level; **significant at 5 percent level and; *significant at 10 percent level.

Appendix Table C4.4 conducts a robustness check corresponding to the results presented in Table 4.3. In doing so, the definition of schooling (i.e. an indicator variable) is the same as defined in table 4.3, but log of household head annualized per capita earnings is added into the specification. The OLS coefficient of schooling for healthy food becomes smaller (0.198 versus 0.315) although highly statistically significant. The IV coefficient on schooling for healthy food becomes larger, but not statistically significant. It has happen due to the fact that IV estimates are very sensitive to the choice of instrument. On the other hand, OLS and IV point estimates (-0.136 and -0.153) from regressing unhealthy food consumption on schooling are much smaller, though highly significant, compared to the estimates (-0.228 and -0.269) presented in table 4.3. Appendix table C4.5 conducts another robustness check, where log of earnings has been dropped and education is measured as years of completed education by the head of the household, and other controls are the same as table C4.4. The signs of OLS estimates from both healthy and unhealthy food group consumption regressions are as expected. However, the magnitudes of the estimates are quite smaller compared to estimates in table C4.4 in both regressions. The signs of the corresponding IV coefficients (see table C4.5) are as expected.

However these coefficients are not statistically significant. The difference between OLS and IV coefficients may be attributed to measurement errors in schooling.

4.6.2 Estimates of the average and marginal treatment effects

Table 4.4 exhibits the average food consumption returns to higher secondary education for different groups of individuals. The return in terms of healthy food consumption to higher secondary school for a random person (ATE) is 10.2%. The return for the individuals who were enrolled in the higher secondary school (ATT) is marginally higher (11.3%). The return for the individuals who did not attend higher secondary school had they attended there (ATU) is 7.1%. The average return for those induced to attend higher secondary school for a particular policy shift (PRTE) is 13%.[37] An estimate for the return to a marginal student (AMTE/MPRTE) is 12.3%.

Table 4.4: Estimates of average returns to higher secondary schooling (Dependent variable: log of per capita healthy food consumption)

Parameter	Semi-Parametric Estimate	Normal Selection Model
ATT	0.113	0.105
	(-0.003, 0.023)	(-0.004, 0.238)
ATE	0.102	0.065
	(0.021, 0.254)	(-0.013, 0.122)
ATU	0.701	0.017
	(-0.301, 0.831)	(-0.091, 0.176)
PRTE	0.13	0.019
	(-0.046, 0.335)	(-0.018, 0.301)
MPRTE	0.123	0.012

[37] The particular policy exercise I have executed is: a 15% reduction of a distance to higher secondary school and find the parameter to understand the impact of an education expansion program.

	(-0.057, 0.228)	(-0.019, 0.291)

Note: This table records estimates of different consumption returns to higher secondary education for the semi-parametric and normal selection models: average treatment on the treated (ATT), average treatment effect (ATE), average treatment on the untreated, policy relevant treatment effect (PRTE), and the marginal policy relevant treatment effect (MPRTE). Bootstrapped highest posterior density 95% intervals are reported in parentheses.

4.6.3 Additional sensitivity results

Table 4.5 shows the additional sensitivity results of regressing education on healthy food consumption by employing IV using *P* (estimated propensity scores) and interactions of original IV (distance to the secondary school) and selected household observables to bring more variations education.

The OLS coefficient (0.172) of higher secondary education is positive and highly statistically significant (see table 4.5). This implies that individuals who have attended higher secondary schools, on average, consume 17.2% more healthy food than those who have been graduated from the lower secondary schools. When propensity scores (*P*) are used as an instrument to education, the coefficient is highly significant and very similar in magnitude to OLS estimate (0.189). When an interaction of distance to secondary schools with household observables (parental education, religion, ethnicities, and age) are used as instrument for education, the IV estimate gets quite large, though significant and expected sign is observed.

Table 4.5: Estimates of average returns to higher secondary schooling (Dependent variable: log of per capita healthy food consumption) - Additional sensitivity results

	OLS	IV (*Z*X* interactions)	IV using *P*
Higher secondary education	0.172***	0.253***	0.189***
	(0.022)	(0.025)	(0.049)

Note: This table reports the coefficients for OLS and 2SLS IV for regression of log per capita healthy food consumption on schooling (an indicator variable that is equal to 1 if an individual has ever attended higher secondary school or more and equal to 0 if he has never attended higher secondary school but graduated from lower secondary school), controlling for parental education, religion and location. Column 1 shows the OLS results, controlling for parental education, religion, ethnicities, and age. Column 2 exhibits IV estimates and excluded instruments are distance to secondary school and interactions with parental education, religion, ethnicities, and age. Column 3 records the IV estimates and excluded instrument is the estimated propensity scores. Type of location is controlled using province dummies. Muslim is an omitted category for Muslim and other ethnicities for ethnicities. Standard errors are shown in parenthesis and are robust to clustering at the community level. All coefficients are significant at 1 % level.

4.7 Conclusion

Indonesia has been very successful for its initiative to expand education since 1970s: The enrolment rates are closely universal for elementary schooling and are about 75% for secondary education. Applying very recent data from Indonesia, this study explores the impact of education on food consumption. In particular, I have investigated whether education has a role to pick up consumption bundles which has health implications.

I find that those who have completed higher secondary education or more substantially consume more healthy food and considerably reduces unhealthy food consumption. Specifically, individuals who have been graduated from upper secondary schools or higher educational institutions, on average, consume 31.5% higher healthy foods than those who have been graduated from the lower secondary schools or less. With respect to unhealthy food consumption, more-educated individuals, on average, consume 22.8% less unhealthy foods than less-educated individuals. This implies a large inequality in consuming healthy food bundles because of taking more education. So, it is important to design policies to expand education for all for at least up to higher secondary level in the context of Indonesia. This finding is important for better understanding of food-health gradient and human capital formation in a country like Indonesia.

However, one of the caveats of the above finding is that construction of healthy food group contains all staple foods and higher consumption of rice, which has the largest share in staple food, may not be always a worthy choice with respect to health and nutrition. Hence, it requires careful division of healthy and unhealthy food groups in the context of Indonesia. Without proper nutrition knowledge, the generalization of the result would be less practicable to analyse food consumption parameters due to education. Furthermore, I have used IV when calculating the treatment effect which may result in local treatment effect nonetheless. This issue is not well-studied in my education and consumption chapter.

References

Abdulai, A., and D. Aubert (2004): “Nonparametric and parametric analysis of calorie consumption in Tanzania,” *Food Policy,* 29 (2), 113–129.

Alem, Y., and M. Söderbom (2012): "Household-Level Consumption in Urban Ethiopia: The Effects of a Large Food Price Shock," *World Development*, Elsevier, 40(1), 146–162.

Anand, S., and Sen, A. (1998): “Consumption & Human Development Background paper,” Human Development Report 1998, Published for the United Nations Development Programme, Oxford University Press, New York.

Angrist, J. D., and A. B. Krueger (1991): “Does Compulsory School Attendance Affect Schooling and Earnings?” *Quarterly Journal of Economics,* 106 (4), 979–1014.

Angrist, J. D., and A. B. Krueger (1999): “Chapter 23 – Empirical Strategies in Labor Economics,” *Handbook of Labor Economics*, Volume 3A, 1277–1366.

Ashenfelter, O., and A. B. Krueger (1994): "Estimates of the Economic Return to Schooling from a New Sample of Twins," *American Economic Review,* 84 (5), 1157–73.

Battistin, E., R. Blundell, and A. Lewbel (2009): “Why Is Consumption More Log Normal than Income? Gibrat’s Law Revisited,” *Journal of Political Economy,* 117 (6), 1140–1154.

Bhandari, R., and F. J. Simth (2000): “Education and Food Consumption Patterns in China: Household Analysis and Policy Implications,” *Journal of Nutrition Education,* 32 (4), 214–224.

Butcher, K. F., and A. Case (1994): "The Effect of Sibling Composition on Women's Education and Earnings," *Quarterly Journal of Economics*, 109 (3), 531–563.

Cain, J. S., R. Hasan, R. Magsombol, and A. Tandon (2010): “Accounting for Consumption Inequality in India: Evidence from Household Expenditures,” *World Development,* 38 (3), 282–297.

Cameron, S., and C. Taber (2004): "Estimation of Educational Borrowing Constraints Using Returns to Schooling," *Journal of Political Economy*, 112 (1), 132–82.

Card, D. (1995): "Using Geographic Variation in College Proximity to Estimate the Return to Schooling," *Aspects of Labour Economics: Essays in Honour of John Vanderkamp*, ed. by Louis Christofides, E. Kenneth Grant and Robert Swindinsky. University of Toronto Press.

Card, D. (1999): "The Causal Effect of Education on Earnings," in *Handbook of Labor Economics*, ed. by Ashenfelter, O. & Card, D., Vol. 3A, 1801-1863, Amsterdam, Elsevier, North-Holland.

Carneiro, P., J. Heckman, and E. Vytlacil (2011): "Estimating Marginal Returns to Education," American *Economic Review*, 101 (6), 2754-2781.

Carneiro, P., J. Heckman, and E. Vytlacil (2003): "Understanding What Instrumental Variable Estimate: Estimating Average and Marginal Returns to Education," Working Paper, available at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.552.5201&rep=rep1&type=pdf

Carneiro, P., M. Lokshin, and N. Umapathi (2016): "Average and Marginal Returns to Secondary Schooling in Indonesia," *Journal of Applied Econometrics*, 30 (7), 1115–1143.

Case, A. C., and A. Deaton (1999): "School Inputs and Educational Outcomes in South Africa," *Quarterly Journal of Economics*, 114 (3), 1047–1084.

Case, A. C. (2005): "The Primacy of Education," 2005. Chapter 18 in *Understanding Poverty*, Abhijit Banerjee, Roland Benabou and Dilip Mookherjee (eds.), Oxford University Press.

Chait, A., J. D. Brunzell, M. Derke, D. Eisenberg, N. D. Ernst, F. A. Franklin, H. Ginsberg, T. A. Kotchen, L. Kuller, R. M. Mullis, M. Z. Nichaman, R. J. Nocolosi, E. J. Schaefer, N. J. Store, and W. H. Weidman (1993): "Rationale of the diet–heart statement of the American Heart Association," *Circulation*, 88, 3008–3027.

Comola, M., and L. de. Mello (2010): “Educational Attainment and Selection into the Labour Market: The Determinants of Employment and Earnings in Indonesia,” Paris School of Economics Working Paper No. 2010-06, Paris.

Currie, J., and E. Moretti (2003): “Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings,” *Quarterly Journal of Economics*, 118 (4), 1495–1532.

Cutler, D. M., and A. Lleras-Muney (2010): “Understanding differences in health behaviors by education,” *Journal of Health Economics*, 29 (1), 1–28.

Central Intelligence Agency (2014): The World Fact Book East and Southeast Asia: Indonesia. Available at https://www.cia.gov/library/publications/the-world-factbook/geos/print/country/countrypdf_id.pdf

Deaton, A., and M. Grosh (1998): “Consumption”, in Grosh, M. and P. Glewwe (eds.) *Designing Household Survey Questionnaires for Developing Countries: Lessons from Ten Years of the Living Standards Measurement Study*, Oxford University Press, New York, NY, for the World Bank.

Denke, M. A. (1995): “Cholesterol-lowering diets. A review of the evidence,” *Arch Int Med*, 155, 17–26.

Duflo, E. (2001): “Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment.” *American Economic Review*, 91(4), 795–813.

Dumauli, M. T. (2015): “Estimate of the private return on education in Indonesia: Evidence from sibling data,” *International Journal of Educational Development*, Elsevier, 42 (C), 14–24.

Fafchamps, M. and F. Shilpi, F. (2014): "Education and Household Welfare," *Economic Development and Cultural Change*, 63(1), 73–115.

Fraser, G. E., A. Welch, R. Luben, S. A. Bingham, N. E. Day (2000): “The Effect of Age, Sex, and Education on Food Consumption of a Middle-Aged English Cohort—EPIC in East Anglia,” *Preventive Medicine* 30, 26–34.

Fulford, S. (2014): “Returns to Education in India,” *World Development*, 59 (C), pp. 434–450.

Galama, T. and H. v. Kippersluis (2010): “A theory of socioeconomic disparities in health over the life cycle,” *RAND Working Paper no. 773*.

Garen, J. (1984): "The Returns to Schooling: A Selectivity Bias Approach with a Continuous Choice Variable," *Econometrica*, 52 (5), 1199–1218.

Goulet, D. (1997): “Development Ethics: A New Discipline,” *International Journal of Social Economics,* 24 (11), 1160–1171.

Griliches, Zvi (1977): “Estimating the Returns to Schooling: Some Econometric Problems,” *Econometrica*, 45 (1), 1–22.

Griliches, Zvi and W. Mason (1972): "Education, Income and Ability," *Journal of Political Economy*, 80 (3), S74–S103.

Grossman, M. (1972): “On the concept of health capital and the demand for health.” *Journal of Political Economy*, 80 (2), 223–255.

Groth, M., S. Fagt, and L. Brondsted (2001): “Social determinants of dietary habits in Denmark,” *European Journal of Clinical Nutrition,* 55, 959–966.

Harmon, C., and I. Walker (1995): “Estimates of the Economic Return to Schooling for the United Kingdom.” *American Economic Review,* 85 (5), 1278–86. http://www.indexmundi.com/facts/indonesia/school-enrollment (accessed on August 29, 2016).

Kane, T., and C. Rouse (1995), “Labor-Market Returns to Two- and Four-Year College,” *American Economic Review*, 85 (3), 600–614.

Kling, J. (2001): “Interpreting Instrumental Variables Estimates of the Returns to Schooling,” *Journal of Business and Economic Statistics*, 19 (3), 358–364.

Koc, H., and H. v. Kippersluis, H. (2015), "Thought for Food: Understanding Educational Disparities in Food Consumption", *Tinbergen Institute Discussion Paper* no. 15-034.

Lleras-Muney, A. (2005): “The Relationship between Education and Adult Mortality in the United States,” *Review of Economic Studies*, 72 (1), 189-122.

Lucas, R. E. (1988): “On the Mechanics of Economic Development,” *Journal of Monetary Economics*, 22 (1), 3–42.

Magdalyn, A. (2013): “The Rate of Returns to Education: The Case of Indonesia,” Master’s thesis paper, Institute of Social Studies. The Hague, The Netherlands.

Michael, R. T. (1975): "Education and Consumption," in *Education, Income and Human Behavior*, ed. by Juster, F. T., New York: Columbia University Press for National Bureau of Economic Research, 235–252.

Michael, R. T., and G. S. Becker (1973), “On the new theory of consumer behavior.” *The Swedish Journal of Economics*, 75 (4), 378–396.

Moreira, P. A. and P. D Padrao (2004): Educational and Economic Determinants of Food Intake in Portuguese Adults: A Cross-sectional Survey, BMC Public Health, 4: 58

Musgrove, P. (1978): “Consumer Behavior in Latin America,” *Brookings Institution*: Washington DC.

Musgrove, P. (1979): “Permanent Household Income and Consumption in Urban South America” *American Economic Review*, 69 (3), 355–368.

Nayga, R., B. Tepper, and L. Rosenzweig (1999): “Assessing the importance of health and nutrition related factors on food demand: a variable preference investigation,” *Appl Econ*, 31, 1541–1549.

Paxson, C. H. (1992): “Using Weather Variability to Estimate the Response of Savings to Transitory Income in Thailand,” *American Economic Review*, 82(1), 15–33.

Paxson, C. H. (1993): “Consumption and Income Seasonality in Thailand,” *Journal of Political Economy*, 101 (1), 39–72.

Perez, C. E. (2002): “Fruit and Vegetable Consumption,” *Health Reports*, 13 (3), Statistics Canada, Catalogue 82-003, Ottawa, ON.

Potter J. D. (Ed.) (1997): *Food, nutrition, and the prevention of cancer: a global perspective*, Washington, DC: World Cancer Research Fund and the American Institute for Cancer Research.

Pritchett, L. (2001): “Where has all the education gone?” *The World Bank Economic Review*, 15(3), 367–391.

Psacharopoulos, G. (1994): "Returns to investment in education: A global update," *World Development*, Elsevier, 22 (9), 1325–1343.

Psacharopoulos, G., and H. Patrinos (2004), “Returns to Investment in Education: A Further Update,” *Education Economics*, 12 (2), 111–134.

Purnastuti, L., R. Salim, and M. A. M. Joarder (2015): “The Returns to Education in Indonesia: Post Reform Estimates,” The Journal of Developing Areas 49 (3), 183-204.

Ricciuto, L., V. Tarasuk, and A. Yatchew (2006): “Socio-demographic influences on food purchasing among Canadian households,” *European Journal of Clinical Nutrition*, 60 (6), 778–790.

Stephens, M. Jr., and D. Y. Yang (2014): "Compulsory Education and the Benefits of Schooling," *American Economic Review*, 104 (6), 1777–1792.

Strauss, J., F. Witoelar, and B. Sikoki (2016) "The Fifth Wave of the Indonesia Family Life Survey (IFLS5): Overview and Field Report," WR-1143/1-NIA/NICHD.

Todaro, M. P., and S. C. SMITH (2012): *Economic Development*. Boston, Pearson Addison Wesley.

Usfar, A. A., and U. Fahmida (2011): “Do Indonesians follow its dietary guidelines?—Evidence related to food consumption, healthy lifestyle, and nutritional status within the period 2000–2010,” Asia Pacific Journal of Clinical Nutrition, 20 (3), 484–494.

The Economist (2014): “Education in Indonesia,” available at http://www.economist.com/news/asia/21636098-indonesias-schools-are-lousy-new-administration-wants-fix-them-schools.

World Bank (2014): “World Bank and Education in Indonesia,” available at http://www.worldbank.org/en/country/indonesia/brief/world-bank-and-education-in-indonesia.

Wantchekon, Klasnja and Novta (2015): "Education and Human Capital Externalities: Evidence from Colonial Benin", *Quarterly Journal of Economics*, 130 (2), 703–757.

Wolpin, K. I. (1982): "A New Test of the Permanent Income Hypothesis: The Impact of Weather on the Income and Consumption of Farm Households in India," *International Economic Review*, 23 (3), 583–594.

Willis, R., and S. Rosen (1979): “Education and Self-Selection,” *Journal of Political Economy*, 87(5), Pt2, S7-36.

Yen, S. T., B. Lin, and C. Davis (2008): “Consumer Knowledge and Meat Consumption at Home and Away From Home,” *Food Policy*, 33 (6), 631–639.

Appendix C

Table C4.1: Calculation of the length of schooling

Educational attainment	Cumulative length of schooling
No Schooling	0
Did not Complete/Have not yet Completed Primary School	3
Primary School	6
Packet A	6
Junior High School (General)	9
Junior High School (Vocational)	9
Packet B	9
Senior High School (General)	12
Senior High School (Vocational)	12
Packet C	12
Diploma I/II	14
Academy/Diploma III	15
University	17
Masters / PhD	20/22

Source: Magdalyn (2013) and Packet A, B, and C are the informal school

C4.2: Notes on calculating years of education

In IFLS surveys information on education are available in three separate modules. First, is in household roster module (Section AR in Book K) where question of education comes as: (i) highest level of schooling (like primary, secondary or tertiary level of education) attended by all household members listed in the surveys; (ii) highest grade ever completed by household members (like 1 represents first grade and 7 represents graduate). Second, there is a separate 'education module' (Section DL in Book IIIA) where extensive information is provided including first enrolment in the school, month and year of graduation, highest level of schooling, time to trip to school from the home, and tuition fees. Third, is from the community survey module where information on principals including teachers, students, per capita expenditure of the students in the schools, competitive examinations scores and tuition fees are available.

However, none of the modules have an information on years of schooling. So I compute the years of schooling from the highest level of schooling in household roster module. To compute years of completed education, I assign each level the number of years it takes to complete that level from shown in appendix A and B. For instance, if someone completes senior high school, he will have 12 years of completed formal education.

Table C4.3: Construction of years of schooling: detailed break-down

Types of Elementary School	Year of Schooling	Types of Junior High School	Years of Schooling	Types of Senior High School	Years of Schooling	D1, D2, D3/University	Years of Schooling
Elementary	6	Junior High General	3	Senior High General	3	College (D1, D2, D3)	1/2/3
Adult Education A	6	Junior High Vocational	3	Senior High Vocational	3	University (BA)	4
School for Disabled	6	Adult Education B	3	Adult Education C	3	University (MA)	2
Madrasah Elementary	6	School for Disabled	3	School for Disabled	3	University (PhD)	3
Kindergarten*	2	Madrasah Junior High school	3	Madrasah Senior High school	3	Open University	2
Other	6	Other	3	Other	3	Other	2

Notes: The data is taken from the Indonesian Ministry of Education; 6 years in primary school and 3 years in junior high school are compulsory by law in Indonesia (9 years of compulsory education); *Kindergarten is not a considered as a part of elementary school, its rather part of primary school with ages from 3 to 5. If someone completed junior high school or equivalent types, his total year of schooling becomes 11 years provided he completed Kindergarten as well.

Table C4.4: Impact of education on food consumption patterns: robustness check

Specifications→	(1) OLS Healthy food	(2) IV Healthy food	(3) OLS Unhealthy food	(4) IV Unhealthy food
schooling	0.198***	0.684	-0.136***	-0.153***
	(0.066)	(0.747)	(0.060)	(0.060)
log of earnings	0.124***	0.076	0.186***	-0.163
	(0.024)	(0.095)	(0.023)	(0.172)
household size	-0.055***	-0.048***	-0.110***	-0.107***
	(0.009)	(0.011)	(0.010)	(0.018)
age	-0.002	-0.003	-0.007**	0.016
	(0.003)	(0.007)	(0.003)	(0.012)
married	0.480	-0.192	-1.021***	-0.437
	(0.465)	(0.516)	(0.229)	(0.752)
Protestant	0.091	0.105	-0.032	-0.336
	(0.116)	(0.169)	(0.142)	(0.348)
Catholic	-0.033	0.117	-0.039	0.201
	(0.273)	(0.353)	(0.253)	(0.514)
Other religions	0.189	0.187	0.049	-0.191
	(0.122)	(0.143)	(0.109)	(0.272)
Javanese	-0.115*	-0.078	0.056	0.324*
	(0.060)	(0.100)	(0.060)	(0.194)
Sundanese	-0.120	-0.068	0.154**	0.588**
	(0.083)	(0.146)	(0.073)	(0.240)
Minang	0.038	0.094	0.111	-0.165
	(0.131)	(0.172)	(0.156)	(0.369)
Fathers' education	0.023**	0.003	-0.016	-0.045
	(0.011)	(0.019)	(0.011)	(0.036)
Mothers' education	0.021***	0.002	-0.015***	-0.031
	(0.001)	(0.312)	(0.002)	(0.030)
Distance to the health post (km)	0.001	0.002	0.001	0.003
	(0.003)	(0.003)	(0.002)	(0.005)
Constant	12.105***	13.514***	13.231***	16.516***
	(0.623)	(1.202)	(0.469)	(1.895)
Observations	5000	5000	5000	5000

Province dummies	yes	yes	yes	yes
Test for joint significance of instruments: *F-stat/p-value*		14.92/0.00		11.15/0.00

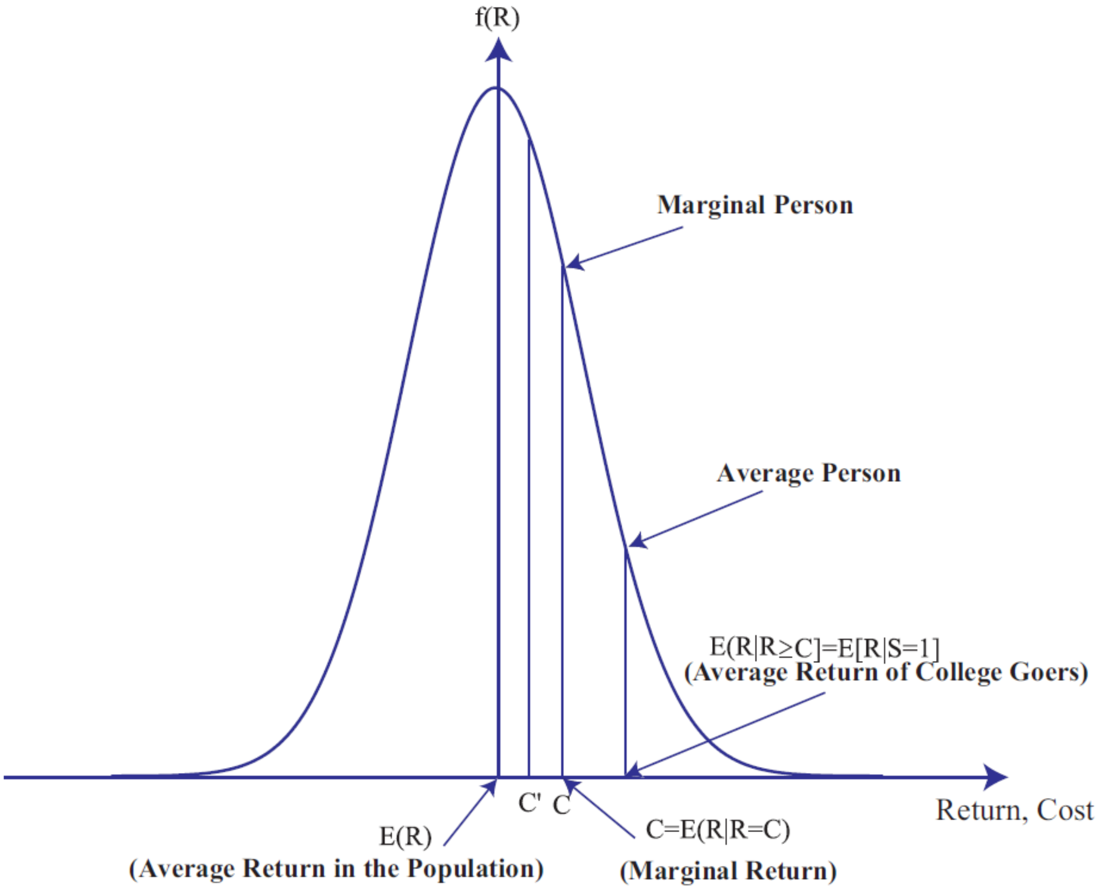
Notes: Column I and II in table C4.4 presents the coefficients for OLS and 2SLS IV, respectively, for regression of log of per capita healthy food consumption on schooling (an indicator variable that is equal to 1 if an individual has completed upper secondary school or more and 0 if he completed lower secondary or less), controlling for log of household head per capita annualized earnings, household size, age and marital status of household head, dummies for religions, dummies for ethnicities, parental education in terms of highest grade attainment, and distance to the health post in kilometres. Reference category for religion is Muslim and for ethnicities is 'other ethnic'. Column III and IV run the regression of log of per capita unhealthy food consumption on schooling and other controls as in column I and II regressions. Excluded instruments in columns II and IV are average distance to the secondary school from the community and interactions with parental education, religion, and age. The robust standard errors are shown in parentheses and clustered at the community level. *** Significant at 1 percent level; ** significant at 5 percent level; * significant at 10 percent level.

Table C4.5: Impact of education on food consumption patterns: further robustness check

Specifications→	(1) OLS Healthy food	(2) IV Healthy food	(3) OLS Unhealthy food	(4) IV Unhealthy food
education (years completed)	0.035***	0.327	-0.034***	-0.288
	(0.005)	(0.504)	(0.005)	(0.426)
household size	-0.067***	-0.051**	-0.095***	-0.088***
	(0.008)	(0.021)	(0.009)	(0.018)
age	-0.003	0.008	-0.010***	0.004
	(0.002)	(0.026)	(0.002)	(0.022)
Married head	0.620	0.471	-0.924***	-0.409
	(0.408)	(0.825)	(0.192)	(1.168)
Protestant	0.139	0.110	-0.043	-0.081
	(0.105)	(0.265)	(0.116)	(0.243)
Catholic	0.064	-0.389	-0.174	-0.431
	(0.198)	(1.121)	(0.204)	(0.953)
Other religion	0.193	-0.931	0.288*	-0.750
	(0.199)	(1.951)	(0.157)	(1.595)
Javanese	0.028	0.199	0.055	0.231
	(0.065)	(0.398)	(0.064)	(0.342)
Sundanese	-0.099	0.232	0.046	0.392
	(0.094)	(0.647)	(0.084)	(0.534)
Minang	0.095	-0.443	-0.087	-0.642
	(0.172)	(1.075)	(0.223)	(0.967)
Fathers education	0.020**	0.073	-0.022**	-0.047
	(0.009)	(0.144)	(0.009)	(0.121)
Mothers education	0.030***	0.083	-0.025***	-0.045***
	(0.001)	(0.093)	(0.002)	(0.012)
Distance to the health post	-0.002	0.004	0.001	0.007
	(0.002)	(0.011)	(0.002)	(0.010)
North Sumatra	0.189*	0.289	-0.286***	-0.121
	(0.102)	(0.352)	(0.104)	(0.326)
West Sumatra	-0.018	0.395	-0.015	0.361
	(0.174)	(0.840)	(0.216)	(0.780)
South Sumatra	-0.066	-0.185	-0.363***	-0.434**

	(0.120)	(0.254)	(0.122)	(0.213)
Lampung	-0.217	-0.028	-0.336**	-0.343
	(0.145)	(0.351)	(0.132)	(0.300)
West Java	-0.087	-0.333	-0.275***	-0.462
	(0.087)	(0.447)	(0.078)	(0.355)
Central Java	-0.316***	-0.270	-0.601***	-0.533**
	(0.106)	(0.230)	(0.106)	(0.231)
Yogyakarta	-0.508***	-0.704	-0.405***	-0.709
	(0.123)	(0.554)	(0.120)	(0.510)
East Java	-0.231***	-0.204	-0.342***	-0.316**
	(0.083)	(0.154)	(0.079)	(0.142)
Bali	-0.029	0.746	-0.434***	0.328
	(0.183)	(1.420)	(0.160)	(1.189)
West Nusa Tenggara	-0.104	-0.600	-0.473***	-0.841
	(0.115)	(0.856)	(0.118)	(0.686)
South Kalimantan	-0.307**	-0.489	-0.461***	-0.629*
	(0.135)	(0.401)	(0.156)	(0.328)
South Sulawesi	0.247**	-0.206	-0.258**	-0.582
	(0.111)	(0.784)	(0.112)	(0.653)
Constant	13.917***	11.265*	16.267***	13.060**
	(0.423)	(5.842)	(0.211)	(5.538)
Observations	5000	5000	5000	5000
Province dummies	yes	yes	yes	yes
Test for joint significance of instruments: *F-stat/p-value*		17.12/0.00		13.20/0.00

Figure C4.6: Density of Absolute Returns



.

Source: Carneiro, Heckman and Vytlacil (2003)

Chapter 5 Conclusions, policy implications, and directions of future research

5.1 Conclusions and policy implications

This dissertation endeavours to provide in-depth analysis of changing food consumption patterns in Indonesia. Specifically, this study comprises three essays to analyse food consumption patterns in the Indonesian economy, which has been gradually transitioning from a centralised to a decentralised economy in the past few decades. During this transition, the government of Indonesia has undertaken several economic policies that have reshaped the nation's living standards in terms of per capita food consumption. Some of those policies are directly related to boosting spatial share of GDP growth, encouraging both internal and international labour movement and investing more in human capital. It is expected that these policies have a significant impact on food consumption parameters in Indonesia.

This study makes use of the ongoing longitudinal, publicly available IFLS to analyse food demand patterns quantitatively. The sample of the IFLS household level survey is representative of about 83% of the Indonesian population and covers more than 30,000 individuals living in 13 of the 27 provinces. The IFLS contains very comprehensive consumption, education, and migration (plus many other) modules and there are also a complementary and contemporary community-level surveys that record valuable information in the local areas where households live. Investigating the hypotheses on food demand by employing IFLS data is anticipated to provide a better understanding of heterogeneous consumer consumption patterns in Indonesia.

The main findings are summarised below, with their implications. The results are documented sequentially for each of the three main essays.

First, as consumers become affluent over time, the estimated staple food share Engel curve can be approximated by quadratic function, the vegetable and fruit share Engel curve by a linear approximation, the meat share Engel curve by a cubic function, and the fish share Engel curve by an approximation of an upward line. The variations in the shapes of the estimated Engel function for different food groups supports the premise of the EASI system, that nonlinearities of the Engel curve are evident in food consumption patterns. Overall the findings of a nonlinear relationship between different food group shares and real per capita expenditure suggests that a substantial number of households may consume significantly below the threshold level of protein-rich food.

Second, with respect to traditional variables, the numerical values of own-price elasticities for almost all the food groups are larger in absolute terms than the cross-price elasticities. This indicates that consumers are more sensitive to price changes of own-good than price changes of other goods. With respect to structural variables, both natural disaster and average distance of the households in the community to the nearest market negatively and significantly affect food consumption in Indonesia. As distance to the nearest market implies access to and availability of foods, direct food transfers and subsidised food sales during natural catastrophes would be an appropriate policies to ensure food security of the households.

Third, somewhat surprisingly, the wealthier households do not appear to diversify their food consumption further when their income rises, whereas poorer households tend to diversify their food consumption significantly when their wealth increases.

Fourth, on average, the migrant-sending household's welfare in terms of per capita food consumption is larger (13.4%) than that of non-migrant sending households living in the same neighbourhood (10.7%).

Fifth, migrant-sending households seem to shift consumption minimally from carbohydrate-rich foods towards more nutrient-dense foods. This shift is probably induced by improved dietary knowledge in migrant households due to migration networks, remittances and access to knowledge of healthy food groups.

Sixth, household individuals who graduated from upper secondary high schools appear to choose more healthy food bundles and household individuals with less education are more likely to consume unhealthy foods. However, in this case, generalisation of the result is problematic as it is difficult to make a distinction between the healthy food group and the unhealthy food group.

5.2 Limitations and directions for future research

This thesis applies IFLS data to exploration of food consumption patterns in Indonesia. The main measure of welfare is household-level food consumption per capita. The IFLS records only expenditure on a particular food and no quantity information is available. This restricts the ability to conduct precise empirical analysis of welfare and nutrition. In further work, it would be valuable to apply the National Socioeconomic Survey of Indonesia (SUSENAS) data to the performance of Engel's law of food demand because the SUSENAS contains information on the quantity of food consumption.

Moreover, this study did not investigate the issue of structural changes in Indonesia between 1997 and 2014. Studying the impact of structural changes during this period may provide valuable information to add to the explanations provided here of household choices.

With respect to the first essay, the Engel curve methodology could be examined in a dynamic context. This might capture more precisely the tastes and preferences of food consumption of Indonesians over time.

In the second essay, this thesis considers out-migration only. From the IFLS data, it is possible to find the destination of the migrants, and compare consumption of migrants in both the origins and destinations, which would be useful for designing food security policies due to migration.

In the third essay, analysing the impact of education on food consumption patterns is critical. The educational attainment could encompass detailed national nutrition knowledge of food groups. Including only general education may obscure the relationship between education and the consumption of healthy foods. Local health practice centres could train community leaders about the importance of understanding the value of nutrition's in food and the IFLS community survey could then collect information in the future surveys.