ESSAYS IN TECHNOLOGY ADOPTION AND CHILD DEVELOPMENT USING SEMIPARAMETRIC ECONOMETRICS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of the Ohio State University

By
Gracious M. Diiro, BS/MS/MA
Graduate Program in Agricultural, Environmental and Development Economics

The Ohio State University
2013

Dissertation Committee:
Dr. Abdoul G. Sam, Advisor
Dr. Joyce Chen
Dr. David Kraybill
This dissertation comprises three chapters focusing on application of semiparametric and structural econometrics to understanding child development, and technology adoption decisions in India and Uganda.

The first chapter examines the heterogeneous effects of maternal labor market participation on the nutritional status of children under five in Rural India. Child malnutrition remains a serious public health concern in developing countries and is puzzlingly high in India, despite her recent achievements in economic progress. Almost half of the children under five in India are stunted, and about 42% are underweight, which presents a major challenge to human development and sustained economic growth. A number of studies have analyzed the determinants of child malnutrition, proposing various interventions to reduce its prevalence. These studies have, however, paid little attention to the role of female labor market participation in determining undernutrition in developing countries. Available evidence shows that women in these economies are increasingly engaging in wage employment to earn additional income for their households. Although maternal income is expected to improve child nutritional outcomes, reduced maternal time with the child may hamper child health. This essay uses a nationally representative dataset from India to investigate the causal effects of maternal participation in labor markets on child nutrition (the standardized height-for-age). The essay differs from previous research in two main aspects: First, it allows for heterogeneity in effects of maternal work on child nutritional status, which may arise due to differences in unobserved genetic and nongenetic attributes.
of children. Secondly, it is the first study to estimate the causal effect of maternal work on child nutrition in India. The results show that maternal labor market participation in India leads to improved nutritional status among children at the bottom of the distribution of height-for-age. However, the effect is not significant for an average child, suggesting that the mean estimate understates the effect of maternal work. The quantile estimates provide evidence of large heterogeneity in the effect of mother’s work on child nutrition. In particular, the results suggest that it is the children in the lower tail of the height-for-age distribution who experience more sizable ‘nutritional premiums’ due to maternal labor market participation; the effects are small for the children in the rest of the distribution.

The second chapter explores empirically if rural households in Uganda leverage their nonfarm earnings to overcome credit constraints and invest in high-yielding maize seed varieties. Unlike previous studies in this realm, this essay does not assume exogeneity of farm and nonfarm earnings, and uses a recently developed semiparametric estimator of binary outcomes that also accommodates endogenous regressors straightforwardly to estimate the effect of nonfarm income on technology adoption decisions. The results show that nonfarm income has a positive and significant effect on adoption of improved maize seed. Specifically, the findings show that a one standard deviation increase in nonfarm income for the average farmer raises the likelihood of adoption by about 47%. Thus, increased recourse to high yielding crop seeds in Uganda can be enhanced by promoting mechanisms that encourage income diversification in the rural areas.

The third chapter is based on my joint research with Abdoul Sam, and it applies a new semiparametric estimator for binary-choice single-index models, which uses parametric information in the form of a known link function and nonparametrically corrects it. The proposed estimator achieves significant bias reduction and efficiency gains relative to competing estimators. Empirical application shows that fertilizer adoption in Uganda is
significantly spurred by the number of agricultural extension visits, male head of household and average education level of the household, and farm and non-farm revenue.
Dedicated to my dear wife Dr. Teddy Namulema Diiro, our children: Maria Diiro, Xavier Diiro, Victor Diiro (RIP) and Adrian Diiro, and my mother, Ms. Veronica Namuhooya
I would like to thank all those who have helped and inspired me during my doctoral studies at the Ohio State University. This dissertation would not have been a success without the support I received from distinguished individuals and institutions.

My sincere appreciation goes to my committee chair, Professor Abdoul G. Sam, for his enduring guidance, encouragement and advice during my doctoral study and dissertation research at the Ohio State University. It is his generous commitment to work, and expertise that have made this dissertation a success. Professor Sam was always available whenever I needed his advice. My interests in the field of applied econometrics have been inspired greatly by his insightful research ideas, and excellent knowledge and skills in econometrics that he selflessly shares with his students. I am lucky to have been one of his student advisees, and I am confident that the knowledge I have learned from working with him will significantly benefit my future professional career.

I also want to thank my committee members Professor Joyce Chen and Professor David Krabill. This dissertation has benefited greatly from their illuminating contributions and feedback. Special thanks go to Prof. David Kraybill for the financial support I received, through the department of Agricultural, Environmental and Development Economics (AEDE), during my final year of study, and during summer holidays.

I also want to thank the professors from the department of AEDE and the department of Economics for their valuable training during my doctoral studies. In particular, I am delighted to have interacted with Dr. Mario Miranda, Dr. Pok-Sang Lam, Dr. Stephen
Cosslet, Dr. Hajime Miyazaki, Dr. Robert De Jong, Dr. Ian Sheldon, Dr. Lung-fei Lee, and Dr. Daeho Kim. My dissertation has benefited greatly from the knowledge I attained from attending their classes. My profound thanks also go to my classmates and friends in the AEDE program for their friendship and support during my doctoral studies. The following deserve special mention: Shuying Shen (Emiley), Ganita Bhupal, Rosemary Isoto, Jenny Trump, Kate Farrin, Xiaohui Tian and Matt Gnagey. I am also grateful to my friend Tatsuro from the department of Economics, Ohio State, for his kindness and friendship.

I would like to thank the institutions that funded my doctoral studies at the Ohio State University. I received generous support from Fulbright for three years, which enabled me realize my academic ambitions. I also want to thank the department of AEDE, not only for funding my final year, but also for admitting me into their PhD program. It has been a rewarding opportunity to study in this department. I greatly value the high quality and the flexibility of the coursework of the AEDE PhD program. I have benefitted from AEDEs strong working relationship with the department of Economics at Ohio State, where I have taken several theoretical courses that have enabled me design a field of specialization for achieving my career goal.

Finally, I must express my deep gratitude to Teddy, my wife, and our children for their continued patience, support and encouragement during the PhD program. Completing my doctoral studies has been greatly enhanced by love and care I received from each of them. I was always amazed by my wife’s willingness to proof read my work. I also want to thank my mother for being one of my best counselors.
1998 .............................. Advanced Level of Education (Uganda Certificate of Advanced Education), Kiira College Butiki, Uganda

2002 .............................. B.Sc. in Agriculture (Economics), Makerere University Kampala, Uganda

2006 .............................. MSc. Agricultural Economics, Makerere University Kampala, Uganda

2006-2007 ........................ Research Officer, National Agricultural Research Systems, Uganda

2007-2009 ........................ Research Associate, CIMMYT/Sasakawa Impact Evaluation project in Uganda

2008-present ...................... Assistant Lecturer, Department of Agricultural Economics, Makerere University, Kampala Uganda

2011 .............................. MA. Economics, The Ohio State University, Columbus, OH, USA

2012-Present ...................... Graduate Teaching Associate, The Ohio State University
FIELDS OF STUDY

Major Field: Agricultural, Environmental and Development Economics

Specialization: Econometrics and Development Economics
CONTENTS

Abstract .............................................................................................................. ii
Dedication ........................................................................................................... iv
Acknowledgments ............................................................................................... vi
Vita ....................................................................................................................... viii
List of Tables ..................................................................................................... xii
List of Figures .................................................................................................... xiii

Chapter............................................................................................................. Page
1 Heterogeneous Effects of Maternal Labor Market Participation on Nutritional Status of Children: Empirical Evidence From Rural India ............... 1

1.1 Introduction .................................................................................................. 1
1.2 Background and Literature Review ............................................................. 6
  1.2.1 Social Employment Programs in India ................................................. 6
  1.2.2 Literature Review ............................................................................... 7
1.3 Econometrics ............................................................................................... 12
  1.3.1 Model Specification .......................................................................... 12
  1.3.2 Quantile Treatment Effects Estimation ............................................. 18
    1.3.2.1 Local Average Treatment Effects Framework (LATE) .............. 20
1.4 Data and Summary Statistics ..................................................................... 23
1.5 Estimation Results ....................................................................................... 26
  1.5.1 Conditional Mean Regression Estimates ........................................... 26
  1.5.2 Conditional Quantile Regression Estimates ....................................... 28
1.6 Conclusions ................................................................................................. 30
1.7 References ................................................................................................... 32
1.8 Tables and Figures ...................................................................................... 43
2 Semiparametric Analysis of Agricultural Technology Adoption in Uganda: The Role of Nonfarm Earnings ........................................ 52

2.1 Introduction ........................................ 52
2.2 Background and Literature Review .................. 57
   2.2.1 Significance of Maize and Technological Progress in Uganda .. . 57
   2.2.2 Literature Review ................................ 58
2.3 Econometrics ........................................ 63
   2.3.1 Model Specification ............................. 63
   2.3.2 A Semiparametric Binary Choice Model with Endogenous Co-
       variates .................................................. 68
2.4 Data and Summary Statistics .......................... 70
2.5 Estimation Results ................................... 72
2.6 Conclusions ......................................... 76
2.7 References .......................................... 77
2.8 Tables and Figures ................................... 87

3 Parametrically-Guided Semiparametric Estimation of Binary-Choice Models: an Application to Fertilizer Adoption in Uganda ............... 91

3.1 Introduction ......................................... 91
3.2 Ichimura’s Semiparametric Estimator of Binary Data .................. 93
3.3 Parametrically-Guided Single-Index Model .................... 94
3.4 Empirical Application to Fertilizer Adoption in Uganda .......... 97
   3.4.1 A Brief Literature Review on the Determinants of Fertilizer
       Adoption in SSA ......................................... 99
   3.4.2 Empirical Model and Data .......................... 101
   3.4.3 Empirical Results ................................ 104
3.5 Conclusions ......................................... 105
3.6 References .......................................... 107
3.7 Tables and Figures ................................... 112

Bibliography ............................................. 114
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Summary Statistics</td>
</tr>
<tr>
<td>1.2</td>
<td>First Stage Regression Estimates</td>
</tr>
<tr>
<td>1.3</td>
<td>Mean Effects of Maternal Work on Child Nutritional Status</td>
</tr>
<tr>
<td>1.4</td>
<td>Quantile Treatment Estimates (LATE)</td>
</tr>
<tr>
<td>2.1</td>
<td>Summary Statistics</td>
</tr>
<tr>
<td>2.2</td>
<td>First-Stage Regressions</td>
</tr>
<tr>
<td>2.3</td>
<td>Second Stage Regressions</td>
</tr>
<tr>
<td>3.1</td>
<td>Summary Statistics of the Variables Included in the Model</td>
</tr>
<tr>
<td>3.2</td>
<td>Results from PGSIM and Probit Procedures</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Kernel Density Estimates of Height-for-Age, Pooled Sample</td>
</tr>
<tr>
<td>1.2</td>
<td>Kernel Density Estimates of Height-for-Age, by Mother Category</td>
</tr>
<tr>
<td>1.3</td>
<td>Distribution of Height-for-Age of Children of Working and Nonworking Mothers</td>
</tr>
<tr>
<td>2.1</td>
<td>Average Structural Function</td>
</tr>
</tbody>
</table>
Chapter 1
HETEROGENEOUS EFFECTS OF MATERNAL LABOR MARKET PARTICIPATION ON NUTRITIONAL STATUS OF CHILDREN: EMPIRICAL EVIDENCE FROM RURAL INDIA

1.1 Introduction

Nutrition security is currently at the helm of the policy agenda for fostering social and economic development across many developing economies (United Nations Childrens Fund, 2009; Leonhauser, 2013; Haddad et al., 2003). Malnutrition of children under five remains a major public health concern in many of these countries (Schmidt, Jorgensen, and Michaelsenk, 1995). Inadequate nutrition during early life inhibits children’s physical and cognitive development, lowering their productivity later in life (Black et al., 2008; Reading, 2008). Child malnutrition is also reported to be a cause of more than one-third of child deaths in developing countries (Liu et al., 2012; Black et al., 2008).

Nearly one half of the children under five in developing countries are malnourished (undernourished) (WHO, 2012).\(^1\) The trend is even more puzzling in South Asia, particularly in India, where the number of undernourished children doubles that in Sub-Saharan African (SSA) countries (Marwaha et al., 2011). Of the 159 million children under the

\(^1\)The terms ‘malnutrition’ and ‘undernutrition’ are often used loosely and interchangeably. Malnutrition refers to all deviations from adequate and optimal nutritional status, including energy undernutrition and over-nutrition (obesity is a form of malnutrition). The term ‘undernutrition’ as used to refer to generally poor nutritional status, but also implies underfeeding” (Shetty, 2003).
age of five in India, about 48% are stunted, 42% are underweight, and 19.8% are wasted (Tarozzi and Mahajan, 2007; WHO, 2012). High differentials in child malnutrition are also reported between rural and urban areas, and across states.

The literature on child health has put forth several explanations for child malnutrition in developing countries, ranging from gender discrimination (low status of women), poor sanitation, lack of adequate knowledge on health/nutrition, mother’s education, and poor maternal health. This study focuses on a factor that has not been rigorously analyzed in developing countries: maternal time allocation between employment and household health production. Women in developing countries have dual roles in their households. They are the primary caregivers to the children and the rest of the household members, and also engage in paid work activities to earn additional income for their families (Jacoby, 1995; Glick, 2002; Glick and Sahn, 1998; Bianchi, 2000). There is evidence that women in developing countries are increasingly entering the labor markets. In India, female labor market participation has increased from 30% in 2000 to 33% in 2005 in the rural areas and from 15% to 18% in the urban areas (India Human Development Survey, 2005). In societies such as India where women are traditionally confined to work at home or on a family farm (Mehrotra and Kapoor, 2009), their growing significance in labor markets is likely to induce transformative changes in intrahousehold decision making and allocation of household resources and time, which directly influences production of child health (Desai and

---

2 In SSA countries 38% and 18% of the 158 million children under five years of age are stunted and underweight, respectively (WHO, 2012). The high prevalence of malnutrition in India does not match her recent economic progress. For example, GDP per capita is substantially high in India than in the SSA economies. In India, GDP per capita (US$), based on purchasing power parity, increased from 1,527.74 in 2000 to 3,650.17 in 2012. GDP per capita in the SSA has increased from 1,431.12 to 2,362.85 over the same period of time (The World Bank, 2013).

3 Stunting, underweight, and wasting, respectively, account for about 50.7%, 45.6% and 20.7% among children under five residing in the rural areas, and about 39.6%, 32.7% and 16.9% among their urban counterparts (Tarozzi and Mahajan, 2007; WHO, 2012).

4 E.g. Coffey, Khera, and Spears (2013); Webb and Block (2003); Tarozzi and Mahajan (2007).
However, the effects of maternal work decisions on child health outcomes in India have yet to be empirically explored.

A priori, child health effects of maternal labor market participation are ambiguous. On one hand, the wage effect of maternal labor market participation may generate positive outcomes of child health. Additional income from maternal work is likely to improve children’s nutrition through increased expenditure on food and health (Glick, 2002; Qian, 2008; Smith et al., 2003). On the other hand, maternal work participation may have a negative effect on child health because it is likely to reduce the quantity and quality of maternal care to the children. For example, a child may become malnourished if a working mother lacks time and energy to adequately breastfeed or prepare nutritionally balanced meals or to utilize the available public interventions for improving health of children (Cawley and Liu, 2012; Glick, 2002; Desai, Chase-Lansdale, and Michael, 1989; Powell, 1984; Smith et al., 2003). Potential market care-giving substitutes may not be readily available especially in rural areas, and may be expensive for poor households residing in urban areas to afford them. As a result, many mothers, especially those in the rural areas, rely on the nonmarket child care substitutes (such as relatives or older children), who may not provide quality child care (Glick and Sahn, 1998; Mcguire and Popkin, 1990).

A number of studies (e.g. Cawley and Liu, 2012; Liu et al., 2009; Scholder, 2008; Glick and Sahn, 1998) have analyzed the effects of maternal work on child health outcomes in both developed and developing countries, but one important limitation persists: lack of estimates that allow the effect to vary across the conditional distribution of child

5 A mother’s decision to work involves tradeoffs between the negative consequences of reductions in quality (quantity) of child care and the income from work (Glick, 2002).

6 Maternal employment and associated earnings lead to increased allocations of household resources to children’s (Behrman and Skoufias, 2006), not only through increased women empowerment within households but also because women are more likely to shift marginal resources to their children (Gettler, 2010).
health outcomes. Although mean estimates reported in the literature provide an informative summary of the health effects of maternal work, they are likely to mask substantial heterogeneity in the effects across the distribution of children in the sample. In the presence of heterogeneous effects, the mean estimate may overstate or understate the effect of the treatment (Buchinsky, 1998; Bilias and Koenke, 2001). A large (significant) value of the mean estimate of the effect of maternal work on child nutrition, for example, may mask the marginal effects on the undernourished children, rendering the inferences incorrect and misleading for policy prescriptions. This essay extends the literature by allowing for heterogeneity in the effects of maternal work participation on child nutrition conditional on observable factors that may influence health of a child. Characterizing the heterogeneous impact of mother’s work on the entire distribution of child nutritional status helps to identify the groups of children that are most affected by maternal work decisions, which is a critical ingredient in designing more effective and targeted policy interventions to reduce undernutrition. For example, implementers of public programs such as the social employment and child development schemes in India are likely to have more interest in the effect of maternal work on the lower tail of the distribution of nutritional status of children than its effect on the mean of the distribution.

The essay also allows for endogeneity of maternal work that may arise from unobservable factors that affect both mother’s decision to participate in the labor market and health

---

7 A review of selected studies is presented in the literature review section that follows.

8 Heterogeneity in child health may arise from differences in genetic, environmental and socio-economic factors. While heterogeneity due to observed characteristics can be easily captured by controlling for these factors during estimation, much of the heterogeneity may be attributed to unobserved factors. Empirical studies on genetic heritability such as Raychaudhuri et al. (2003) in South Asia show that heritability for height is about 62%.

9 Estimates of mean treatment effects provide very limited information about the distribution of the dependent variable (Firpo, 2007; Koenker and Bassett, 1978; Buchinsky, 1998), and may overstate or understate the effect of the treatment because mean estimates are prone to influence by outliers (Buchinsky, 1998; Bilias and Koenke, 2001).
status of a child. Evaluation of maternal work decisions on child health status is likely to be affected by simultaneity bias. For instance, a child with chronic illnesses requires more care than a healthy one, implying that mothers with unhealthy children will probably be less likely to participate in the labor markets. In this case, we can observe a positive association of work and nutrition but with reverse causality, running from child health to maternal work rather than the reverse—the true effect of work on health could be zero or even negative (Glick, 2002). In addition, differences in unobserved idiosyncratic preferences and tastes of mothers are also likely to influence their time allocation decisions. In India, customs and norms dictate that women stay at home to do domestic work (Desai and Jain, 1994). This suggests that a mother who goes against the restrictive societal norms to enter the labor market is likely to be driven by unobserved attributes (such as preferences, motivation and ability). On the other hand, a mother may decide not to work because she enjoys spending time cooking or playing with their children. Others may not enter the labor market because they do not like physical work, the most common form of employment in the informal sector. This essay thus estimates the distributional effects of maternal employment using an instrumental variables quantile regression approach by Abadie, Angrist, and Imbens (2002) to relax the homogeneity assumption imposed by the mean regressions, and to control for endogeneity of the treatment variable, maternal work in this case. The essay exploits the exogenous local variations in social employment programs, implemented by the central government of India, to identify the effect of maternal labor market participation on child nutritional status.\footnote{A brief review of government supported employment schemes in India is provided in the literature section.}

The rest of the essay is organized as follows: The next section presents a brief review of literature on the determinants of child health in developing countries. The third section outlines the conceptual model and estimation methods, describing the instrumental variable quantile regression methods and the estimator adopted in this essay. The fourth section
describes the data used, and the selected characteristics of the children and households in the sample. The fifth section presents the estimation results. The last section of the essay concludes.

1.2 Background and Literature Review

1.2.1 Social Employment Programs in India

The government of India has, for the past several years, implemented a number of wage employment programs aimed at creating gainful employment for the rural poor households in the country (Government of India, 2010, 2006). Examples of schemes implemented in the recent years include the National Rural Employment Guarantee Act (NREGA) and the Universal Rural Employment Program, locally referred to as the Sampoorna Grameen Rozgar Yojana (SGRY) (Government of India, 2010, 2011). NREGA was enacted in 2005, and it aims at improving rural livelihoods through provision of stable employment in the rural areas. The scheme provides a legal guarantee for at least 100 days of work every financial year to adult members of any household willing to do unskilled manual public work at the statutory minimum wage notified for the program. NREGA also promotes female employment by allocating about one-third of work opportunities to women of working age.

Prior to launching the NREGA, the government of India implemented the SGRY scheme, which was launched in 2001 by merging two other schemes: the Employment Assurance Scheme (EAS) and Jawahar Gram Samridhi Yojana (JGSY) (Government of India, 2006).11 The scheme aimed at creating need based rural infrastructure at the village level, and had special provisions for women, scheduled castes, scheduled tribes and parents of

11The EAS was initiated in 1993 to create additional employment opportunities during the period of acute shortage of wage employment through manual work for the rural poor living below the poverty line. The JGSY was launched in 1989 by merging National Rural Employment Program (NREP) and Rural Landless Employment Guarantee Program (RLEGP) (Government of India, 2006).
children withdrawn from hazardous occupations. The employment schemes described above were initiated by the central government of India but implemented by state governments while observing the guidelines outlined by the central government. The programs are mainly implemented at district panchayats, intermediate panchayats and gram panchayats, which are selected on the basis of poverty indicators (Government of India, 2006, 2010).

1.2.2 Literature Review

This subsection gives a brief review of empirical literature on the determinants of child health in developing countries, with special focus on South Asia. The literature is summarized into four main strands, first discussing studies that relate parental education to child health, then looking at gender bias and woman’s status, health infrastructure and environmental factors, and ending with income and labor allocation decisions. A large volume of recent literature has explored the causal mechanisms of education on health of children, concluding that the health returns to women’s education are indisputable. In India, Govindasamy and Ramesh (1997) finds that a higher level of maternal education results in improved child survival. They argue that compared to uneducated mothers or those with little education, mothers with higher education tend to seek and utilize more preventive health care for children. Similarly, Mishra and Retherford (2000) find a positive relationship between the mother’s education and child nutrition. Kravdal (2004) investigates the effect of community-level education of women on child health, and finds that education of individual mothers and that of the community generate improved child health outcomes. Singh-Manoux et al. (2008) also report lower mortality among children of parents with more years of formal education. Elsewhere, Webb and Block (2003) examine the functional distinction between mother’s nutritional knowledge and formal education in household production of child health in Indonesia. They find that maternal nutritional knowledge improves short-term health of a child (weight-for-height) but formal schooling
exerts a larger positive effect on long-term health (child’s height-for-age) than nutrition knowledge.\textsuperscript{12}

Another hypothesis often raised in the literature is that low status of women and discrimination against females in south Asia (Mehrotra and Kapoor, 2009) may explain poor nutritional status in the region. For example, Smith et al. (2003) analyze the effect of decision making power of women relative to their partners in households, and the effect of gender equality in the community on child nutrition in developing countries including India.\textsuperscript{13} Their findings for India show that increasing relative decision making power of women in households can reduce child stunting and wasting. They also show that the number of underweight children under 3 years could be reduced by 13.4 million (about 13 percentage points reduction) if males and females had equal status in society. They argue that a woman’s ability to make decisions at home and in her community not only affects the care she receives (and thus her own nutritional well-being) but it also enables her to provide better care and nutrition for her children. Similarly, a recent study by Coffey, Khera, and Spears (2013) reveals that children of lower ranking daughters-in-law in rural joint households in India are shorter, on average, than children of higher ranking daughters-in-law.\textsuperscript{14} They argue that a low ranking daughter-in-law has lower decision making power in the household compared to a high ranking daughter-in-law. Dancer, Rammohan, and

\textsuperscript{12}They argue that maternal education contributes to short-run child outcomes through nutritional knowledge, and that paternal education contributes independently to long-run (but not short-run) child nutrition.

\textsuperscript{13}Relative power of a woman in the household was captured using an index constructed from four variables: employment status of the woman, the woman’s age at first marriage, the percent difference in education and age of the woman and her partner. The index for societal gender equality was constructed using three variables measured at the community level: difference between boys and girls under five with respect to weight for-age Z-scores, and age-adjusted vaccination score, and the difference in years of education of adult women and men.

\textsuperscript{14}“In India, many families live in patrilocal joint households, that is, households in which adult sons live with their parents after marriage, and their wives join them in their parents’ homes. The wives of older brothers are assigned higher social rank in the in-laws’ households than the wives of younger brothers” (Coffey, Khera, and Spears, 2013).
Smith (2008) analyze the relationship between child nutrition and gender differences in survival probabilities of children in Indonesia. They find that boys have lower probability of survival in their first year of life than girls, but the boys who survive grow significantly taller than girls. Several other studies also report discrimination against girls with respect to child feeding, antenatal and post natal care, and treatment seeking in India.\textsuperscript{15}

The third strand of literature links child health to the environment where the child lives, and health care infrastructure. Poor sanitation and lack of safe water remain a major challenge to public health in India (Kumar, Kar, and Jain, 2011; Kremer, 2010). For example, about 67\% of India’s population does not have toilets whereas 21\% has no access to safe water (United Nations Children’s Fund, 2013), culminating in severe diarrheal illnesses especially among children (Checkley et al., 2008). Diarrhea and other chronic intestinal diseases limit children’s ability to absorb and use food nutrients, resulting in severe malnutrition and death (Humphrey, 2009; Checkley et al., 2008). The United Nations Children’s Fund (2013) estimates that diarrhea causes death of about 400,000 children under five years in India every year. However, there is evidence that better hygiene and sanitation in households can substantially improve child health status in India. For instance, Jalan and Ravallion (2003) use propensity score matching methods to analyze the effects of piped water on the prevalence of diarrhea in India. They find that the prevalence and duration of diarrhea among children under five in rural India are significantly lower, on average, for a family with piped water than for an identical family without piped water. Van der Klaauw and Wang (2011) estimate the determinants of mortality and report that improvements in

\textsuperscript{15}Selected examples include Mishra, Roy, and Retherford (2004) report gender discrimination (in favor of boys) in childhood feeding, immunization coverage, treatment-seeking, and nutritional status. Bharadwaj and Lakdawala (2013) find evidence towards gender discrimination in prenatal investments–mothers visit antenatal clinics and receive tetanus shots more frequently when pregnant with a boy. Jayachandran and Kuziemko (2011) model breastfeeding under son-biased fertility preferences and find that breastfeeding duration increases with birth order, and is lowest for daughters and children without older brothers because their parents try again for a son. Similarly, Barcellos, Carvalho, and Lleras-Muney (2012) find that boys receive significantly more childcare time, and are more likely to be breastfed longer, and to be given vaccinations and vitamin supplementation than girls.
indoor air quality and access to safe water can substantially reduce child mortality in India. In particular, they find that 17 more children out of 1,000 live births survive until age five if all families switch to clean cooking fuels in a separate kitchen. They also report that the presence of a doctor in the village significantly increases the probability of child survival after the first birthday. Kumar and Vollmer (2013) also report that improvements in hygiene and sanitation in rural India can reduce the risk of contracting diarrhea by 2.2 percentage points.

The fourth strand of literature, and the most relevant to this essay relates parental income and time allocation to child health. According to the hypothesis of ‘maternal altruism’, children benefit more if income is earned or received by a woman than if it is in a man’s hands (Behrman and Skoufias, 2006). Women are believed to be natural child care givers, and tend to shift the scarce household resources more toward the welfare of young children (Gettler, 2010). A number of studies support this hypothesis. For example, in China, Qian (2008) finds that increases in female income improves survival rates (for girls) and increases education attainment (for girls and boys) whereas increases in male income worsens survival rates for girls and has no effect on boy’s educational attainment. Using pension eligibility for men and women as instrumental variables, Duflo (2003) analyzes the impacts of old age pension on child health in South Africa, and finds that a pension received by women increased health of children more than that received by men. The study estimates that pensions received by a woman increased weight for height of girls by 1.19 standard deviations, and their height-for-age by 1.16 standard deviations. However, no effect was found for pensions received by men. Shuhaimi and Muniandy (2012) find higher prevalence of severe wasting among children of unemployed mothers than those of their employed counterparts in Malaysia.\footnote{They find a positive correlation between maternal hours of wage work and short term health status (weight and BMI) of children but no correlation with their long-term health status.} Their findings also show that working mothers
generally introduce complementary feeding at an earlier age of a child than nonworking mothers. This could be one of the reasons for the observed differences in the heath status of children of the two types of mothers. However, some studies find deleterious effects of maternal work on child health production whereas others report no effects. For example, Toyama et al. (2001) examine children under five in Indonesia, and find that children of working mothers are at a higher risk of stunting. They also note that households of working mothers experienced food shortages more often than those of nonworking mothers. Similar findings are reported by Maddah et al. (2007) in a study of 1,319 children between 3 and 5 years of age in Northern Iran. On the other hand, Berman et al. (1997) finds no effect of maternal work and earnings on health expenditure in Haryana State in India. Glick and Sahn (1998) use assets, unearned income, price of food, cost of medical care as instrumental variables to estimate the causal effects of mother’s employment on child health in west Africa, and find that additions to maternal labor income yields larger increases in child height than do equivalent additions to other (non-mother) household income. They, however, find a negative net effect of maternal employment implying that the negative health consequences of reductions in quality (quantity) of child care offset the positive effects of additions to income. Their results, however, need to be interpreted with caution because most of the instrumental variables they use may not be exogenous to child health status.

The work in this essay differs from the above studies in two different ways. First, unlike the previous studies in this realm, this essay does not assume that the effects of maternal labor market participation are the same for the whole distribution of child nutritional status. In this essay, the health effects of maternal work are estimated on the distribution of child nutrition using a quantile treatment effects estimator. Second, to my knowledge, this study is the first one in South Asia to analyze the causal effect of maternal work on child nutrition using a nationally representative dataset. The findings from this study will provide new empirical information to formulate and implement appropriate social policies for improved
maternal and child welfare in India. Lessons learned from this study can also be extended to other countries in South Asia.

### 1.3 Econometrics

#### 1.3.1 Model Specification

The study adopts the Roy (1951) model to conceptualize the maternal decisions to participate in labor markets in India. It is assumed that women in India are rational decision makers and therefore only enter the labor market if doing so maximizes their expected utilities. Their participation in the labor market is thus driven by the difference between the expected benefits from working in nondomestic activities and those derived from working at home (including work within the home). The model for work participation by a mother of child $i$ is specified as:

$$w^*_i = \gamma s_i + \varepsilon_i \quad (1.3.1)$$

$$w_i = I(\gamma s_i - \varepsilon_i \geq 0) \quad (1.3.2)$$

where, $w^*_i$ is a latent variable denoting the difference between her expected utility from engaging in paid work and that derived if she decides not to work (or doing domestic work), $w_i$ is a dummy for maternal labor market participation: $w_i = 1$ if $w^*_i > 0$ and $w_i = 0$ if $w^*_i \leq 0$, $s_i \equiv (x_i, z_i)$ is a vector of observable factors that affect maternal work decisions, $z_i$ is a vector of instrumental variables, $x_i$ is a vector of other exogenous covariates, $\gamma$ is a vector of parameters, and $\varepsilon_i$ denotes the effect of unobservable factors.

This study further assumes that child human capital is, in part, determined by the mother’s work status. Specification of the model of child health outcomes in this essay follows the standard paradigm of the household utility maximization framework (Becker,
The study assumes a reduced form health production function that involves combining representations of the endogenous maternal labor force activity with exogenous explanatory variables (Schultz and Tansel, 1997; Glewwe, 1999; Glick and Sahn, 1998). The model for child health can be specified as below:

\[ c_i = c(x_i, w_i; \theta) + v_i, \quad i = 1, 2, ...n, \]  

(1.3.3)

where, \( c_i \) is the current health status of a given child \( i \), \( x_i \) is as defined above, \( \theta \) is a vector of parameters, and \( v_i \) is a random error term denoting the effect of the unobservable characteristics.

In this study, health status of a child is captured using their height-for-age Z-score (HAZ), a measure of chronic malnutrition. HAZ reflects the accumulated investment over the life of child (Deaton, 2007; Steckel, 2008). Height is commonly used in development studies to predict human capital development (Bozzoli, Deaton, and Quintana-Domeque, 2009) and economic productivity of an individual (Judge and Cable, 2004; Case and Paxson, 2008; Haddad et al., 2003). The problem of relying on the short term indicators (weight-for-age (WAZ) and weight-for-height (WHZ)), is that they may not capture the true health status of a child because they are affected by both observed and unobserved transitory shocks such as short term illnesses and stress at the time of measurement (Sahn and Stifel, 2002a). The HAZ scores for children in the sample were computed following the WHO child growth standards: observed height of the child minus the median height in the reference population normalized by the standard deviation of height in the reference populations (WHO and de Onis, 2006). The WHO norms originate from well-nourished (well fed, exclusively breastfed) populations of children around the world.

---

17 The household maximizes utility subject to prices, budget and time constraints, and the health production function. The maximization process jointly determines health outputs, health inputs, and demands for other goods, including leisure (hence the labor supply of the mother and others is determined as well) (Schultz, 1994; Glick and Sahn, 1998).
The study proposes to use maternal proximity to work opportunities (captured by a dummy variable for presence of government supported employment programs in the village of residence) as an instrumental variable \( z_i \) to address endogeneity of maternal labor market participation in the child health production function. Presence of employment programs can increase the indirect benefits of work in two ways. First, the employment programs create local labor opportunities in the communities, increasing the chances of finding a job, especially if there are more job seekers than available jobs. Secondly, growth of local work opportunities may induce women to join the labor market since it is more convenient for them to work in their locality, and if increased labor demand raises wages in the sector. A possible threat to exogeneity of this instrumental variable is that the work opportunities generated by the social employment programs are accessed by both men and women. These programs are therefore expected to increase labor demand for men as well, who through a combination of cultural and physical factors may be the most favored for manual labor jobs in their local communities. Consequently, presence of work programs is likely to influence child health through labor market participation choices of both mother and father of the child. To address this concern, a dummy variable for labor market participation decisions of the father, and the number of other working members in the household (other than the mother and father) are included in the empirical model, such that mother’s work participation is the most important remaining channel through which implementation of work programs may affect child health. Furthermore, it is unlikely that employment programs have a direct effect on child health through correlation with other state-level social safety net programs because the guidelines and procedures for implementing the employment schemes in the states are provided by the central government.

The covariates \( x_i \) included in the model consist of several factors that influence child health including health inputs, health endowments of the child, individual characteristics of the child and those of the parents, and environmental conditions. These factors have
been drawn from previous research studies on child health in developing countries (e.g. Glewwe (1999); Duflo (2003); Dancer, Rammohan, and Smith (2008); Sahn and Alderman (1997); Vandell and Ramanan (1992); Thomas and Strauss (1992); Morales, Aguilar, and Calzadilla (2004); Webb and Block (2003); Borooah (2005); Glick and Sahn (1998)). Four variables are included in the model as proxies for inputs invested in the production of child health in the household. These include a dummy for child immunization, years of formal education of the mother, number of antenatal visits attended by the mother during pregnancy with the child, and maternal health knowledge. Whether a mother believes that "first breast milk" (colostrum) is good for a new born baby is used to proxy for a mother’s health knowledge. Further, the study controls for the effect of mother’s access to health related information using the mother’s exposure to media, captured by a dummy variable indicating whether the mother listens to radio or watches television or reads newspapers on a regular basis. This study hypothesizes that increased access to media generates positive child health outcomes because media channels are expected to supply the public with vital health related information, which is likely to increase a mother’s knowledge of household health production. The effect of female empowerment and autonomy is captured using an index computed as the mean of four dummy variable for woman’s involvement in decision-making in the household. These include whether a woman participates in making decisions about the type of food to cook, what to do when a child is sick, how to spend the income earned in the household, and whether she participates in grocery shopping in the household. The index takes values from 0 to 1, corresponding to no bargaining power and higher bargaining power, respectively. Consonant with recent literature on female autonomy and health in India (e.g. Coffey, Khera, and Spears (2013)), female autonomy is expected to induce positive child health outcomes through its effect on increased female resource control

---

18 The first breast milk produced during pregnancy, rich antibodies that boost baby’s immunity against infections.
within the household. The number of children below 14 years in the household are also included in the model to control for competition for the scarce resources among children in the household.

The study uses three variables to capture the effect of health endowments of a child. These include a dummy variable for birth weight of the child (1 if the child was big at birth, 0 otherwise), maternal child rearing experience measured as the difference between age of the mother and that of the child, and mother’s height. The mother’s height (in centimeters) was included in the model to control for the effect of genetics and unobserved family background. In addition, dummies for the caste category of the mother are included to capture the effect of social and cultural norms on child health outcomes. Four commonly used caste categories in India are included, that is, scheduled caste, scheduled tribe, other backward tribes (OBT) and Brahmin (Desai, Vanneman, and National Council of Applied Economic Research, 2008).

Furthermore, individual and household factors are included in the model to control for the effects of attributes specific to the child and those of their households. The individual variables include age of the child (measured in years), gender (1 if male, 0 otherwise) and birth order (1 if firstborn, 0 otherwise). Household variables include household poverty and labor participation by other members of the household other than the mother. Household poverty is a dummy variable indicating whether the household is below the poverty level, and is included to control for the effect of household wealth. Rich households are expected to have healthier children because they can afford to spend more on important health inputs such as nutrient supplements and medical care for their children than poor households.

Furthermore, the number of other working members in a household (apart from the mother

---

19 Height of mother is a form of health endowment and it captures the effect of unobserved family background and the genetics of the mother. It may also be used as an indicator of the long-run overall poverty in the family, and in this case human capital investment in the mother, which manifests itself in the height Z-scores (Thomas, Strauss, and Henriques, 1991).
and the father), and a dummy variable for paternal participation in wage employment is included in the model to control for the effect of father’s income on child health. Fathers are the main bread winners in households, and their income is expected to improve the welfare of children in the household. On the other hand, the number of other working members in a household is hypothesized to have a negative effect on child health due to reductions in the quantity of child care in the household. In addition, including paternal work participation and the number other working members in a household are important for controlling for the likely “spillover effects” of government supported work programs on child health via the father and other members who may access work through these programs. A dummy variable indicating whether a mother’s marital home is close to her natal family is also included in the empirical model to control for the additional child care support she may receive from her parents and relatives. Her natal home is close to her marital home if it takes less than one hour to travel between the two homes.

The study controls for the effect of environmental conditions in which the child lives using two variables: a dummy variable for household access to and use of protected water source, and an index for hygiene and sanitation in the household. The index was computed as the mean of seven indicators of hygiene and sanitation in the household, including having a toilet, washing hands after using the toilet, washing hands with soap, cooking in a ventilated kitchen, having good drainage around the home to reduce stagnant water during the rainy season, proper disposal of animal excreta (no animal or human excreta around the home), keeping animals in designated structures (that is, people do not share house with animals). The index takes values from 0 to 1, corresponding to poor hygiene and better hygiene.²⁰

²⁰The survey required enumerators to report their observations on selected topics at the time of interview, and the indicators of the sanitation and hygiene were one of them.
1.3.2 Quantile Treatment Effects Estimation

This study analyzes the causal effect of maternal labor participation on child nutritional status using standard approaches for estimating treatment effects models. A number of estimators for analyzing effects of endogenous variables on distributions of outcome variables have been proposed. Firpo (2007) develops a two step semiparametric estimator for quantile treatment effects under the identifying restriction that selection to treatment is based on observable characteristics. The procedure involves nonparametric estimation of the propensity score in the first step followed by computation of the difference between the solutions of two separate minimization problems in the second step. The main limitation of this approach is that it assumes that selection bias is only due to observables, which may not be the case in the current context because there are likely to be inherent attributes that influence both maternal work choices and health of nutritional status of a child.

Chernozhukov and Hansen (2005) propose a conditional quantile treatment effects estimator for analyzing models with endogenous treatments, without imposing functional form restrictions. This approach is applicable to models with continuous outcome variables and does not impose restrictions on the distribution of the endogenous and the instrumental variables. It admits both continuous and discrete forms of endogenous variables and the instruments. They show that the estimator is $\sqrt{n}$ consistent and asymptotically normal under appropriate regularity and identification conditions.

Another appealing method is proposed by Imbens and Newey (2009), similar to earlier works by Chesher (2003) and Ma and Koenker (2006). This approach is based on a triangular system of equations. It requires continuity of the endogenous variable, and relies on monotonicity of the reduced form equation in the scalar error term. It also imposes a restrictive independence condition that the instrumental variable is independent of both the structural and the reduced form error term. To assume a triangular structure in unobservables may be very strong in the current study because the structural and the reduced form
errors are likely to contain common attributes, such as genetic and nongenetic environment effects.\textsuperscript{21} This estimator is also not plausible for our study because we focus on a binary endogenous regressor, maternal work participation.

The quantile effects estimators discussed above share one main limitation. They assume that the instrumental variable affects the behavior of all treated units (subjects) in the sample. Imbens and Angrist (1994), however, shows that noncompliers are likely to exist in the presence of endogeneous treatments such as maternal work status in this essay. In the current context, some mothers may not work even if they had access to work opportunities. In the presence of noncompliance problem, the average treatment effect (ATE) parameters estimated using standard procedures including those described above would identify the intention to treat effects (ITT) and not the causal effect (Imbens and Angrist, 1994). This essay opts for the local average treatment effect (LATE) framework proposed by Abadie, Angrist, and Imbens (2002) to circumvent the noncompliance problem. The Abadie, Angrist, and Imbens (2002) quantile estimator is a generalization of the local average treatment effects (LATE) framework of Imbens and Angrist (1994).\textsuperscript{22} The LATE estimators have been shown to provide unbiased parameter estimates of the average treatment effect (ATE) in presence of non-compliance because they restrict the computation of the ATE to the subpopulation of ‘compliers’ (Imbens and Angrist, 1994).\textsuperscript{23} This method estimates policy parameters in models with continuous outcomes, a binary treatment, and a

\textsuperscript{21}In our context, it would require that there exists an additional latent factor which influences current height-for-age but does not affect maternal work decisions, and is orthogonal to the structural error term.

\textsuperscript{22}The Abadie, Angrist, and Imbens (2002) quantile treatment estimator is akin to the recent approach by Frolich and Melly (2008). The difference between these two estimators is that former estimates conditional treatment effects whereas the later estimates unconditional treatment effects. Conditioning on covariates provides detailed description of quantiles of potential outcomes for the compliers in the population.

\textsuperscript{23}In the LATE framework, the treatment effect can only be identified for the compliers (a subpopulation of treated individuals whose behavior is affected by the instrument) because the always takers and the never-participants cannot be induced to change status by the hypothetical movement of the instrument.
binary instrumental variable, making it relevant to the current study, which focuses on a bi-
mary endogenous variable (maternal labor market participation) and a binary instrumental
variable (for maternal access to work opportunities).

1.3.2.1 Local Average Treatment Effects Framework (LATE)
This essay follows the standard literature on treatment effects to define the counterfactual
states for child health and maternal labor market participation (Abadie, 2003; Abadie, An-
grist, and Imbens, 2002; Imbens and Angrist, 1994). Two potential health outcomes are
defined for each child in the sample: the nutritional status when a mother participates in
the labor market \(c_1^i\) and the nutritional status when she does not participate \(c_0^i\). For
any given child, the causal impact of maternal work participation is naturally defined as
\(c_1^i - c_0^i\). But since the realizations of the two potential health outcomes are mutually
exclusive, we cannot simultaneously observe both outcomes for a given child. Only the
average outcome, \(c_i = w_i c_1^i + (1 - w_i) c_0^i\), is observed. As explained above, maternal labor
market participation is influenced by a binary instrumental variable–presence of govern-
ment supported employment programs in the village, giving rise to two counterfactuals for
maternal work participation. Maternal work status is \(w_1^i\) when a mother has access to the
work opportunities \(z_i = 1\) and is \(w_0^i\) when she does not have access to work \(z_i = 0\). The
work status indicator can be expressed as \(w_i = z_i w_1^i + (1 - z_i) w_0^i\). In this framework, the
population of interest is that of all compliers (that is, the (unobserved) subpopulation with
\(w_1^i > w_0^i\) (Abadie, Angrist, and Imbens, 2002; Firpo, 2007; Frolich and Melly, 2008; Im-
bens and Angrist, 1994; Abadie, 2003).\(^{24}\) This surmounts to the group of children whose

\(^{24}\) According to Abadie, Angrist, and Imbens (2002), identification of average treatment effects, and ATE on
the treated is difficult when the treatment variable is endogenous.
mothers would participate in the labor market if they had access to work opportunities but would not participate otherwise.\(^{25}\)

Given the binary instrumental variable \(z_i\), the \(\tau\)-quantile of the outcome \(c_i\) conditioned on exogenous variables \((x_i)\) and endogenous (to child nutrition) work choices \((w_i)\) for compliers is specified in (1.3.4) below.

\[
q_{\tau}(c|x, w, w^1 > w^0) = x_i \beta^\tau + w_i \alpha^\tau
\]

(1.3.4)

where \(\alpha^\tau\) and \(\beta^\tau\) are the unknown parameters of the model at \(\tau^{th}\) quantile of the conditional distribution of height of the child, and \(\alpha^\tau\) represents the conditional quantile treatment effect—the difference in the conditional \(\tau\)-quantiles of \(c^1_i\) and \(c^0_i\) for compliers.\(^{26}\)

The key identifying assumptions of this approach are that of monotonicity (to rule out defiers) and a conditional independence of the binary instrumental variable (to identify the quantile treatment effects for the compliers).\(^{27}\) Under these assumptions, Abadie, Angrist, and Imbens (2002) show that \(\alpha^\tau\) (the conditional QTE) for the compliers can be consistently estimated by a weighted sum of check functions specified in (1.3.5).\(^{28}\)

\[
(\beta^\tau, \alpha^\tau) = \arg\min_{\beta, \alpha} \mathbb{E}[k_i (x, w, z) \rho_{\tau}(c_i - x_i \beta - w_i \alpha)]
\]

(1.3.5)

\(^{25}\)Depending on the reaction of \(w_i\) on an external intervention of \(z_i\), four subpopulations of children can be derived (Rubin, 1974). The first are the compliers (individuals whose treatment status can be manipulated by the experiment (instrument) at hand) with \(w^1_i > w^0_i\) or \(w^1_i = 1\), and \(w^0_i = 0\). Others are the never-taker with \(w^1_i = w^0_i = 0\), always-takers with \(w^1_i = w^0_i = 1\), and the defiers with \(w^1_i < w^0_i\). Since only one of the potential treatment indicators is observed, we cannot identify which one of these four groups any particular child belongs to.

\(^{26}\)Note that 1.3.4 identifies differences in quantiles on the potential outcomes, \(c^1_i\) and \(c^0_i\), and not the quantile of the difference, \(c^1_i - c^0_i\), which is harder to identify (Abadie, Angrist, and Imbens, 2002; Abadie, 2003). If we observed the event \(w^1_i > w^0_i\), then (1.3.4) could be estimated by standard quantile regression using the subpopulation of compliers. But, in effect, the binary variable \(I(w^1_i > w^0_i)\) is an omitted variable, and \(z_i\) is the available instrument for \(w_i\).

\(^{27}\)Monotonicity ensures that the treatment effects can only be identified for the compliers because always-takers and never-participants cannot be induced to change treatment status by hypothetical movement of the instrument.

\(^{28}\)The check function ensures that (i) all \(\rho_{\tau}\) are positive, and (ii) the scale is according to the probability \(\tau\).
where \( p_\tau(\lambda) = \begin{cases} \tau \times \lambda & \text{if } \lambda \geq 0 \\ (\tau - 1) \times \lambda & \text{if } \lambda < 0 \end{cases} \) is the check (sign) function, \( k_i(x, w, z) \) is weighting function that is to identify the subpopulation of compliers, and is defined in (1.3.7) below.\(^{29}\)

\[
k_i(x, w; z) = 1 - \frac{w_i(1 - z_i)}{1 - p(z_i = 1|x_i)} - \frac{(1 - w_i)z_i}{p(z_i = 1|x_i)}
\]  

(1.3.7)

The weight function takes a value of one for potential work participants (when \( w_i = z_i \)) and a negative value otherwise. Negative values of the weighting function render the optimization problem in (1.3.5) nonconvex, making it difficult to implement. To derive a convex optimization problem, this essay follows the procedure suggested by Abadie, Angrist, and Imbens (2002); Frolich and Melly (2008) of taking conditional expectations of objective function (given the outcome variable, treatment variable and instrumental variable).

\[
k_i^+(u) = E[k(x, w, z)|u] = 1 - \frac{w_i(1 - s_i(u))}{1 - p(z_i = 1|x_i)} - \frac{(1 - w_i)s_i(u)}{p(z_i = 1|x_i)}
\]  

(1.3.8)

where \( u = (c, x, w) \), and \( s_i(u) = E(z_i|u) = p(z_i = 1|c, x, w) \).

A positively weighted check function in (1.3.9) is thus derived by replacing \( k_i \) in (1.3.5) with positive weight function, \( k_i^+(u) \), defined in (1.3.8).

\[
(\beta^\top, \alpha^\top) = \arg \min_{\beta, \alpha} \frac{1}{n} \sum_{i=1}^{n} \left( 1[k_i^+(u_i \geq 0)]k_i^+(u_i) \rho_\tau(c_i - x_i \beta - w_i \alpha) \right)
\]  

(1.3.9)

The indicator function \( 1[k_i^+(u_i \geq 0)] \) ensures that only observations with nonnegative weights are used in the optimization.

\(^{29}\) According to Abadie, Angrist, and Imbens (2002); Abadie (2003), the joint independence assumption \((c_i^1, c_i^0, w_i^1, w_i^0) \perp z_i|x_i)\) implies that any observed relationship between the treatment \( (w_i) \) and the instrument \((z_i)\) has a causal interpretation in the subpopulation of compliers (see lemma 2.1 for details). The subpopulation of compliers is not identified since \( w_i^1 \) and \( w_i^0 \) are mutually exclusive. This problem can be solved by weighting the moments by \( k_i(\cdot) \) in an average sense as shown in (1.3.6) below, and it implies that the parameters defined as the solution to a moment conditions involving \((c_i, w_i, x_i)\) are identified (Abadie, Angrist, and Imbens, 2002; Abadie, 2003; Frolich and Melly, 2008):

\[
E[k_i \rho_\tau(c_i - x_i \beta - w_i \alpha)] = p(w_i > w_i^0)E[\rho_\tau(c_i - x_i \beta - w_i \alpha)|w_i > w_i^0]
\]  

(1.3.6)
In practice, the estimator is implemented in two steps. The first step involves estimation of the conditional probabilities (propensity scores) of the instrumental variable, \( \pi(x_i) = p(z_i = 1|x_i) \) and \( s_i(\mu) \), which are used to construct a nonparametric estimate of the weight function in (1.3.8). Using a rich set of community level data, this essay applies a nonparametric local linear logistic regression procedure suggested by Frolich and Melly (2008) to predict the propensity score \( \pi(x_i) \). The positive weight function (1.3.8) is then generated by non-parametric regression of its sample analog in (1.3.7) on the outcome variable (nutritional status of the child), the observable factors \( x_i \) and the treatment (maternal labor market participation). In the second step, the fitted values of the weight function \( \hat{k}_i^+ \) are used to estimate the conditional quantile parameters \( \alpha^\tau \) and \( \beta^\tau \) in equation (1.3.9).

As a benchmark, conventional instrumental variables estimates of the the effect of maternal work on child nutrition are computed. The estimates are generated by using fitted-values from the probit model of mother labor market participation as an instrumental variable for child nutritional status.

### 1.4 Data and Summary Statistics

The study utilizes data from the India Human Development Survey (2005) to analyze the causal effect of maternal work participation on nutritional status of children. The Indian Human Development survey (IDHS) IHDS is a survey of a nationally representative sample of 41,554 households drawn from 1,504 villages and 970 urban blocks across 31 states.

---

30 This essay recognizes the instrumental variable (presence of work program in a village) is not purely random. As earlier noted, placement of employment programs in India is determined mainly on the basis of socioeconomic indicators. In this essay, the probability of having a social employment program in a village is predicted using six community/village level social-economic indicators. These include existence of rural credit and savings associations, presence of trade unions, state level unemployment rates, distance to district headquarter, proportion of poor households in the village, and the proportion of households with electricity.

31 The fitted values were generated using optimal smoothing parameters derived using cross validation.
in India. The household questionnaire covered employment, income, education, consumption expenditure, health, and access to services for each household. An individual interview was conducted for a sample of ever-married women between 15 and 49 years old covering a variety of questions on education, birth history, anthropometric measurements, health inputs, sanitation, fertility and gender relations within the household. A mother-child matched dataset was constructed using the household and relationship identifiers in the household and individual level datasets. The sample for this study was restricted to children aged 0-5 residing in rural areas for two main reasons. First, detailed information on child health and survival was captured for only the previous two children born to a woman after the year 2000. In addition, the WHO (1986) recommends limiting the analysis of height measures to children between 0 to 5 years old, because children of this age group are the most vulnerable to environmental factors. Second, the community (village) questionnaire, which contains one of the instrumental variables (social employment program) was only administered to rural households. The size of the final analytic sample is 10,554 children born to 7,068 mothers in the rural areas. In this sample, the proportion of mothers who are engaged in paid work is about 27%. The sample also contains a substantial portion of working mothers with children below 5 years, making it relevant to analyze the relationship between maternal employment and child health. About 29.5% of the working mothers have a child under 2 years. The proportion of working mothers with a child between 2 and 5 years is about 24%.

Figures 1.1 and 1.2 show Epanechnikov kernel density estimates of height-for-age for children under five in India. It is clear that the health status of children in rural India is poor relative to the reference population. The average Z-score of height-for-age is -1.36 suggesting that an average child in India is more than one and quarter standard deviations shorter than a healthy child (of the same age) from the rest of the world. Moreover, about
54% of children in the sample exhibit stunted growth—they are more than 2 standard deviations below the mean of the reference population. Children of working mothers, are on average shorter than children of non working mothers. However, kernel density estimates show that the nutritional outcomes of children of working mothers may not be worse than children of working mothers. As can be seen in figure 1.2, the density function for working mothers has a slightly smaller mass in the lower tail, and a slightly larger mass in the center of the distribution than that of nonworking mothers. Figure 1.3 also compares the distribution of children of working and nonworking mothers on the quantiles of the Z-scores, and it shows that nutritional outcomes of the two types of children are quite similar in the middle quantiles (τ), between τ = 0.2 and τ = 0.8. Slight differences are observed in both tails of the distribution.

Summary statistics for selected attributes of the sample are presented in Table 2.1, for households with and without a working mother. The descriptive statistics show significant differences between children of the two categories of mothers (working mothers and nonworking mothers) with respect to several pre-determined characteristics. This suggests that mere comparison of unadjusted mean difference of the nutritional status of children of working mothers and nonworking would give misleading inferences because the two types of children are significantly different from each other, and thus incomparable without making adjustments. For example, the summary statistics reveal that a larger proportion of working mothers belong to poor households than their nonworking counterparts. About 49% of households with working mothers are poor (below the poverty line) compared to only 27.4% of the household of nonworking mothers. Further, households of non-working mothers reported substantially higher annual per capita income as well as monthly per capita expenditure relative to households of working mothers. However, monthly expenditure on food accounts for a larger proportion of the total household monthly expenditure in
households of working mothers compared to those of nonworking mothers. These statistics suggest that mothers who chose to work are mainly driven by the need to improve welfare in their household rather than improved economic opportunities in the labor market. Further, mothers who work are more involved in making decisions in the household than mothers who do not work. For example, 71% and 55% of the working mothers make decisions about type of food to cook in the households and participate in purchasing food, respectively, compared to the corresponding 61% and 39.7% of their non-working counterparts.

1.5 Estimation Results

Before estimating the quantile effects, it is important to examine the mean effects of maternal work participation on child nutrition as a basis for comparison with findings in the literature. This section first presents results of mean regressions, before discussing the estimates from the quantile regressions. Table 1.2 reports the estimates of the first stage probit model, which are used to generate predicted work status in the second stage. The results show a significant coefficient on the instrumental variable in the expected direction as discussed above, indicating that presence of work opportunities in the village increases the probability of female labor market participation.

1.5.1 Conditional Mean Regression Estimates

In the second step of the mean regression, the propensity score predicted from the first stage binary choice model is included as an instrumental variable in the parametric model of child health. The estimated results are presented in table 1.3. The coefficient on maternal work is not significant, suggesting that, on average, maternal labor market participation does not affect child nutritional status in India. Similar findings were reported by Glick and Sahn (1998).
The mean regression results, however, show that nutritional status of a child is affected by health inputs, health endowments, environmental factors and other household attributes. In particular, an additional maternal visit to an antenatal care facility improves health of a child by 0.05 standard deviations. Similarly, height-for-age of a child is estimated to improve by 0.03 standard deviations if a mother attained an additional year of formal schooling. The study findings also reveal that immunizing a child improves their stature by about 0.46 standard deviations. Furthermore, higher bargaining power of women within their households is estimated to improve child nutritional status by 0.41 standard deviations. Bargaining power in households gives women an opportunity to participate in making important decisions that directly or indirectly affect child health. Examples include making decisions about types of food to buy, types of food to cook, and how to spend the income earned in the household. The results further show that mothers who live near their parents tend to have healthier children, possibly because they are likely to benefit from social support in the form of child care that their parents may provide.

Access to safe water, hygiene and sanitation are the other important determinants of child health status. The results show that a child with access to safe water is about 0.19 standard deviations taller than their counterparts who do not have safe water. Similarly, children residing in households with better hygiene are about 0.25 standard deviations taller than the children who live in poor hygiene conditions. As expected, the mother’s height exerts a positive and highly significant effect on the child’s stature. Similarly, mothers with regular access to media (either a radio, or a television or print media) tend to have taller children. On the other hand, the number of resident children under 14 years is negatively related to the health of children under five in the household. More dependents in households strain the already scarce household resources, making it difficult to supply adequate levels of vital health inputs to children. The results also show that age of a child is negatively
related to their nutritional status, suggesting that an average child in India is significantly shorter than a comparable child in the WHO reference population.

1.5.2 Conditional Quantile Regression Estimates

The motivating question of this analysis is to understand how well the quantile effects, relative to the mean effects, characterise the health benefits derived by children in India when their mothers decide to enter the labor market. Table 1.4 reports the quantile treatment effects of maternal work participation on child health from estimation of equation 1.3.9.

The estimates indicate that maternal work significantly benefits children at the bottom of the distribution ($\tau = 0.1$ to $\tau = 0.3$). In general, the magnitude of the quantile effects of maternal work participation decays with quantiles, providing evidence of larger effect at low quantiles than higher quantiles. The significant quantile estimates are between 0.62 standard deviations of height of a child in the first quantile and 0.30 standard deviations for one in the third quantile. These results suggest that the health effects of maternal labor participation are substantially heterogeneous and that it is children at bottom of the height-for-age distribution who experience more sizable ‘nutritional premiums’ due to maternal work; for the rest of the distribution, the effects are small. A comparison with the mean regression estimates indicates that the mean estimate understates the effect of maternal labor market participation among children at the bottom of the distribution.

The above result that maternal labor market participation translates into improved health of a child is interesting because the evidence from the data shows that the households of working mothers are generally poor, and have low per capita income (and low per capita expenditure) compared to households of their nonworking counterparts. One might have expected children of working mothers to have less stature than their counterparts for two reasons: Children born in poor households are vulnerable to poor health because they are raised in less favorable conditions where supply of nutrition and health inputs may not be
adequate. Work participation by a mother may also deprive her child of quality time and care, further depressing their health status. Nevertheless, the econometric estimations seem to indicate that maternal participation in the labor market generates a non-trivial income effect that offsets these adverse effects. Two explanations are possible for the positive effect of maternal work on child health. The first explanation is that mother’s desire to improve household welfare may induce positive changes in household consumption behavior. As earlier noted, the majority of working mothers in this sample are from poor households, suggesting that their decision to enter the labor market may have been driven by the desire to earn additional income and improve the welfare in their households. Further, households where both mother and father earn income are likely to spend their income differently than when only the man is earning. Working women are likely to spend their earnings on consumption of health improving inputs that are beneficial to the children. Indeed, studies in developing countries such as Duflo (2003) and Glick and Sahn (1998) find that female earnings generate greater child health outcomes than male earnings or nonmother income. Secondly, and consonant with the permanent income hypothesis, the marginal propensity to spend on consumption is likely to be higher in poor households than in richer households. Higher food expenditure share reported by households of working mothers (poor households) may be a direct income effect of maternal work participation, and it seems to support the permanent income hypothesis. Children from poor households are thus expected to experience a larger marginal effect of maternal income than those from relatively wealthy households.\(^\text{32}\)

The quantile regression estimates also show evidence of substantial heterogeneity in effects of other determinants of child nutritional status. For example, the impact of mother’s education is greatest among children at the bottom of the distribution and lowest among

\(^{32}\)Since poor households derive higher utility (relative to wealthy ones) from an additional dollar earned, the effect of larger share food expenditure is greater among children in the poor households.
those at the top. Similarly, more sizable health benefits of antenatal care visits by a mother are experienced by children at the upper tail of the distribution than those at the bottom. Heterogeneous effects of antenatal care could be due to inequalities in access and utilization of maternal health care services, caused by differences in the various socioeconomic conditions. For instance, Mohanty (2012) finds that less poor and educated women have a higher probability of utilizing antenatal care services in India. Nichter (1995) attributes inequalities in maternal care utilization to differences in social and religious beliefs. The results also show that mother’s height has a positive and significant effect on child’s stature in all the quantiles, but the effect is greater among children at the bottom of the distribution than those at the top. Furthermore, the health benefits of immunization are higher for children below the median standardized height compared to children at the top quantiles. Heterogeneity in the effects of immunization could be due to differences in awareness, and acceptance of immunization. The demand for vaccination remains low in many developing countries including India, despite the increased awareness created by the WHO, health workers and community leaders.

1.6 Conclusions

This essay estimates the quantile treatment effects of maternal participation in labor markets on child nutrition (the standardized height-for-age) using a nationally representative dataset from India. The study identifies the effect by using the exogenous local variations in social employment programs implemented by the government of India to stimulate job creation in communities. Both the mean and the quantile estimates are positive, suggesting that the income effect of maternal work is greater than the deleterious effects of reductions in quality and quantity of child care. However, the mean estimate is substantially smaller.

Nichter (1995) particularly reports that certain Hindu and Muslim groups have long held the belief that vaccination is a covert method of family planning, primarily targeting Muslims.
than some of the quantile estimates, meaning that large mean effects of maternal participation may mask the benefits for children from marginal households. The quantile estimates reveal large amount of heterogeneity in the benefits to the child from mother’s work. In particular, the results suggest that it is the children in the lower tail of the height-for-age distribution who experience sizable ‘nutritional premium’ due to maternal labor participation; the effects are small for the children in the rest of the distribution. Social programs such as the employment programs implemented by the government can significantly contribute to improved child nutrition by targeting mothers of less nourished children located at the lower part of the height-for-age distribution.
1.7 References


### 1.8 Tables and Figures

#### Table 1.1: Summary Statistics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Working Mother</th>
<th>Nonworking Mother</th>
<th>T Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child’s Height-for-Age (Z-Score)</td>
<td>Mean</td>
<td>Std.Dev</td>
<td>Mean</td>
</tr>
<tr>
<td></td>
<td>-1.407</td>
<td>2.850</td>
<td>-1.34</td>
</tr>
<tr>
<td>Child is Immunized</td>
<td>0.855</td>
<td>0.352</td>
<td>0.846</td>
</tr>
<tr>
<td>Mother’s Education (Years)</td>
<td>2.224</td>
<td>3.668</td>
<td>3.903</td>
</tr>
<tr>
<td>No. of Antenatal Visits</td>
<td>2.443</td>
<td>2.425</td>
<td>2.744</td>
</tr>
<tr>
<td>Mother’s Age at Child’s Birth</td>
<td>25.430</td>
<td>5.737</td>
<td>24.652</td>
</tr>
<tr>
<td>Early Breast Feeding is Good</td>
<td>0.724</td>
<td>0.447</td>
<td>0.740</td>
</tr>
<tr>
<td>Mother Uses Media channels</td>
<td>0.566</td>
<td>0.496</td>
<td>0.665</td>
</tr>
<tr>
<td>Birth Weight (1=Big, 0=Small)</td>
<td>0.792</td>
<td>0.406</td>
<td>0.799</td>
</tr>
<tr>
<td>Height of Mother (Centimeters)</td>
<td>149.793</td>
<td>6.333</td>
<td>150.827</td>
</tr>
<tr>
<td>No. of Children &lt; 14 Years</td>
<td>3.142</td>
<td>1.464</td>
<td>3.122</td>
</tr>
<tr>
<td>Mother Lives Near Her Parents</td>
<td>0.308</td>
<td>0.462</td>
<td>0.333</td>
</tr>
<tr>
<td>Has Access to Safe Water</td>
<td>0.764</td>
<td>0.425</td>
<td>0.826</td>
</tr>
<tr>
<td>Hygiene and Sanitation Index</td>
<td>0.711</td>
<td>0.162</td>
<td>0.734</td>
</tr>
<tr>
<td>Child is Male</td>
<td>0.500</td>
<td>0.500</td>
<td>0.522</td>
</tr>
<tr>
<td>Age of Child (Years)</td>
<td>2.714</td>
<td>1.551</td>
<td>2.466</td>
</tr>
<tr>
<td>Female Empowerment (Index)</td>
<td>0.430</td>
<td>0.200</td>
<td>0.384</td>
</tr>
<tr>
<td>Household is Poor</td>
<td>0.491</td>
<td>0.500</td>
<td>0.274</td>
</tr>
<tr>
<td>Other Working Members</td>
<td>1.234</td>
<td>0.713</td>
<td>0.912</td>
</tr>
<tr>
<td>Father Works (1=Participates)</td>
<td>0.944</td>
<td>0.231</td>
<td>0.751</td>
</tr>
<tr>
<td>Other Backward Class Caste</td>
<td>0.375</td>
<td>0.484</td>
<td>0.415</td>
</tr>
<tr>
<td>Scheduled Tribe (Caste)</td>
<td>0.285</td>
<td>0.451</td>
<td>0.232</td>
</tr>
<tr>
<td>Scheduled Caste (Caste)</td>
<td>0.232</td>
<td>0.422</td>
<td>0.073</td>
</tr>
<tr>
<td>South Region</td>
<td>0.199</td>
<td>0.399</td>
<td>0.133</td>
</tr>
<tr>
<td>Western Region</td>
<td>0.198</td>
<td>0.399</td>
<td>0.191</td>
</tr>
<tr>
<td>Eastern Region</td>
<td>0.014</td>
<td>0.116</td>
<td>0.032</td>
</tr>
<tr>
<td>Central Region</td>
<td>0.551</td>
<td>0.497</td>
<td>0.474</td>
</tr>
<tr>
<td><strong>Instrumental Variable</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employment Program in the Village</td>
<td>0.884</td>
<td>0.320</td>
<td>0.849</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>2838</td>
<td></td>
<td>7716</td>
</tr>
<tr>
<td>Variable</td>
<td>Coefficient</td>
<td>Standard Error</td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>-------------</td>
<td>----------------</td>
<td></td>
</tr>
<tr>
<td>Employment Programs in the Village</td>
<td>0.116***</td>
<td>0.043</td>
<td></td>
</tr>
<tr>
<td>Child is Immunized</td>
<td>0.051</td>
<td>0.043</td>
<td></td>
</tr>
<tr>
<td>Mother Education</td>
<td>-0.032***</td>
<td>0.004</td>
<td></td>
</tr>
<tr>
<td>Antenatal Care Visits</td>
<td>0.008</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>Mother’s Age at Child’s Birth</td>
<td>0.016***</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>No. of Children &lt; 14 Years</td>
<td>-0.056***</td>
<td>0.011</td>
<td></td>
</tr>
<tr>
<td>Mother Lives Near Parents</td>
<td>-0.074**</td>
<td>0.031</td>
<td></td>
</tr>
<tr>
<td>Maternal Height (cm)</td>
<td>-0.004</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>Birth Weight (1=Big, 0=Small)</td>
<td>-0.007</td>
<td>0.036</td>
<td></td>
</tr>
<tr>
<td>Early Breast Feeding is Good</td>
<td>0.019</td>
<td>0.033</td>
<td></td>
</tr>
<tr>
<td>Mother Uses Media channels</td>
<td>0.004</td>
<td>0.032</td>
<td></td>
</tr>
<tr>
<td>Child is Male</td>
<td>-0.026</td>
<td>0.029</td>
<td></td>
</tr>
<tr>
<td>Age of Child (years)</td>
<td>0.068***</td>
<td>0.009</td>
<td></td>
</tr>
<tr>
<td>Female Empowerment (Index)</td>
<td>0.577***</td>
<td>0.074</td>
<td></td>
</tr>
<tr>
<td>Household is poor</td>
<td>0.305***</td>
<td>0.033</td>
<td></td>
</tr>
<tr>
<td>Father Engages in Paid Work</td>
<td>0.667***</td>
<td>0.049</td>
<td></td>
</tr>
<tr>
<td>No. of other Working Members</td>
<td>0.130***</td>
<td>0.021</td>
<td></td>
</tr>
<tr>
<td>Access to Safe Water</td>
<td>-0.087**</td>
<td>0.036</td>
<td></td>
</tr>
<tr>
<td>Hygiene and Sanitation Index</td>
<td>-0.0548</td>
<td>0.088</td>
<td></td>
</tr>
<tr>
<td>South Region</td>
<td>0.954***</td>
<td>0.066</td>
<td></td>
</tr>
<tr>
<td>Western Region</td>
<td>0.710***</td>
<td>0.063</td>
<td></td>
</tr>
<tr>
<td>Eastern Region</td>
<td>-0.0002</td>
<td>0.119</td>
<td></td>
</tr>
<tr>
<td>Central Region</td>
<td>0.651***</td>
<td>0.059</td>
<td></td>
</tr>
<tr>
<td>Other Backward Caste</td>
<td>0.264***</td>
<td>0.043</td>
<td></td>
</tr>
<tr>
<td>Scheduled tribe</td>
<td>0.389***</td>
<td>0.046</td>
<td></td>
</tr>
<tr>
<td>Scheduled caste</td>
<td>0.869***</td>
<td>0.055</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-2.433***</td>
<td>0.366</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>10554</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LR chi2 (27)</td>
<td>1941***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudo $R^2$</td>
<td>0.161</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*** p<0.01, ** p<0.05, * p<0.1
<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal Work Participation</td>
<td>0.360</td>
<td>0.495</td>
</tr>
<tr>
<td>Child is Immunized</td>
<td>0.464***</td>
<td>0.074</td>
</tr>
<tr>
<td>Mother Education</td>
<td>0.0315***</td>
<td>0.009</td>
</tr>
<tr>
<td>Antenatal Care Visits</td>
<td>0.050***</td>
<td>0.016</td>
</tr>
<tr>
<td>Mother’s Age at Child’s Birth</td>
<td>0.006</td>
<td>0.007</td>
</tr>
<tr>
<td>No. of Children &lt; 14 Years</td>
<td>-0.065***</td>
<td>0.020</td>
</tr>
<tr>
<td>Mother Lives Near Parents</td>
<td>0.175***</td>
<td>0.067</td>
</tr>
<tr>
<td>Maternal Height (cm)</td>
<td>0.0361***</td>
<td>0.005</td>
</tr>
<tr>
<td>Birth Weight (1=Big, 0=Small)</td>
<td>0.119</td>
<td>0.081</td>
</tr>
<tr>
<td>Early Breast Feeding is Good</td>
<td>0.038</td>
<td>0.069</td>
</tr>
<tr>
<td>Mother Uses Media channels</td>
<td>0.115**</td>
<td>0.057</td>
</tr>
<tr>
<td>Child is Male</td>
<td>-0.077</td>
<td>0.059</td>
</tr>
<tr>
<td>Age of Child (years)</td>
<td>-0.219***</td>
<td>0.021</td>
</tr>
<tr>
<td>Female Empowerment (Index)</td>
<td>0.408****</td>
<td>0.144</td>
</tr>
<tr>
<td>Household is Below Poverty Line</td>
<td>-0.070</td>
<td>0.074</td>
</tr>
<tr>
<td>Father Engages in Paid Work</td>
<td>-0.069</td>
<td>0.100</td>
</tr>
<tr>
<td>No. of other Working Members</td>
<td>-0.026</td>
<td>0.045</td>
</tr>
<tr>
<td>Access to Safe Water</td>
<td>0.193***</td>
<td>0.074</td>
</tr>
<tr>
<td>Hygiene and Sanitation Index</td>
<td>0.252</td>
<td>0.188</td>
</tr>
<tr>
<td>South Region</td>
<td>-0.0002</td>
<td>0.155</td>
</tr>
<tr>
<td>Western Region</td>
<td>0.206*</td>
<td>0.118</td>
</tr>
<tr>
<td>Eastern Region</td>
<td>0.516***</td>
<td>0.188</td>
</tr>
<tr>
<td>Central Region</td>
<td>0.396***</td>
<td>0.117</td>
</tr>
<tr>
<td>Other Backward Caste</td>
<td>0.058</td>
<td>0.068</td>
</tr>
<tr>
<td>Scheduled tribe</td>
<td>-0.078</td>
<td>0.095</td>
</tr>
<tr>
<td>Scheduled caste</td>
<td>-0.264</td>
<td>0.196</td>
</tr>
<tr>
<td>Constant</td>
<td>-7.621***</td>
<td>0.737</td>
</tr>
</tbody>
</table>

Observations: 10,353
F Statistics: 911.000
R-Square: 0.045

*** p<0.01, ** p<0.05, * p<0.1
Table 1.4: Quantile Treatment Estimates (LATE)

<table>
<thead>
<tr>
<th>Variable</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal Work Participation</td>
<td>0.626</td>
<td>0.468</td>
<td>0.355</td>
<td>0.167</td>
<td>0.044</td>
<td>-0.051</td>
<td>-0.009</td>
<td>0.036</td>
<td>-0.106</td>
</tr>
<tr>
<td></td>
<td>(0.364)</td>
<td>(0.216)</td>
<td>(0.174)</td>
<td>(0.186)</td>
<td>(0.164)</td>
<td>(0.168)</td>
<td>(0.153)</td>
<td>(0.159)</td>
<td>(0.244)</td>
</tr>
<tr>
<td>Child is Immunized</td>
<td>0.243</td>
<td>0.648</td>
<td>0.836</td>
<td>0.620</td>
<td>0.580</td>
<td>0.537</td>
<td>0.295</td>
<td>0.256</td>
<td>0.185</td>
</tr>
<tr>
<td></td>
<td>(0.329)</td>
<td>(0.237)</td>
<td>(0.243)</td>
<td>(0.196)</td>
<td>(0.235)</td>
<td>(0.232)</td>
<td>(0.195)</td>
<td>(0.205)</td>
<td>(0.224)</td>
</tr>
<tr>
<td>Mother Education</td>
<td>0.076</td>
<td>0.117</td>
<td>0.073</td>
<td>0.059</td>
<td>0.054</td>
<td>0.046</td>
<td>0.034</td>
<td>0.015</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.023)</td>
<td>(0.022)</td>
<td>(0.020)</td>
<td>(0.010)</td>
<td>(0.022)</td>
<td>(0.022)</td>
<td>(0.022)</td>
<td>(0.025)</td>
</tr>
<tr>
<td>Antenatal Care Visits</td>
<td>0.0240</td>
<td>0.001</td>
<td>0.046</td>
<td>0.048</td>
<td>0.057</td>
<td>0.063</td>
<td>0.091</td>
<td>0.107</td>
<td>0.141</td>
</tr>
<tr>
<td></td>
<td>(0.061)</td>
<td>(0.049)</td>
<td>(0.038)</td>
<td>(0.038)</td>
<td>(0.038)</td>
<td>(0.038)</td>
<td>(0.041)</td>
<td>(0.036)</td>
<td>(0.043)</td>
</tr>
<tr>
<td>Mother’s Age at Child’s Birth</td>
<td>0.020</td>
<td>-0.001</td>
<td>0.0004</td>
<td>-0.007</td>
<td>0.011</td>
<td>0.015</td>
<td>0.021</td>
<td>0.020</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.017)</td>
<td>(0.014)</td>
<td>(0.015)</td>
<td>(0.015)</td>
<td>(0.015)</td>
<td>(0.015)</td>
<td>(0.013)</td>
<td>(0.016)</td>
</tr>
<tr>
<td>No. of Children &lt; 14 Years</td>
<td>-0.031</td>
<td>-0.026</td>
<td>-0.050</td>
<td>-0.061</td>
<td>-0.093</td>
<td>-0.048</td>
<td>-0.085</td>
<td>-0.081</td>
<td>0.074</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.067)</td>
<td>(0.048)</td>
<td>(0.048)</td>
<td>(0.056)</td>
<td>(0.058)</td>
<td>(0.053)</td>
<td>(0.057)</td>
<td>(0.064)</td>
</tr>
<tr>
<td>Mother Lives Near Parents</td>
<td>0.542</td>
<td>0.433</td>
<td>0.349</td>
<td>0.435</td>
<td>0.462</td>
<td>0.347</td>
<td>0.198</td>
<td>0.135</td>
<td>0.190</td>
</tr>
<tr>
<td></td>
<td>(0.312)</td>
<td>(0.204)</td>
<td>(0.158)</td>
<td>(0.150)</td>
<td>(0.165)</td>
<td>(0.151)</td>
<td>(0.137)</td>
<td>(0.150)</td>
<td>(0.182)</td>
</tr>
<tr>
<td>Height of Mother (Centimeters)</td>
<td>0.083</td>
<td>0.065</td>
<td>0.067</td>
<td>0.058</td>
<td>0.044</td>
<td>0.047</td>
<td>0.040</td>
<td>0.031</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.014)</td>
<td>(0.011)</td>
<td>(0.013)</td>
<td>(0.012)</td>
<td>(0.009)</td>
<td>(0.015)</td>
<td>(0.006)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>Birth Weight (1=Big, 0=Small)</td>
<td>-0.126</td>
<td>-0.426</td>
<td>-0.142</td>
<td>0.036</td>
<td>0.012</td>
<td>-0.112</td>
<td>-0.126</td>
<td>0.016</td>
<td>0.102</td>
</tr>
<tr>
<td></td>
<td>(0.340)</td>
<td>(0.193)</td>
<td>(0.169)</td>
<td>(0.164)</td>
<td>(0.188)</td>
<td>(0.195)</td>
<td>(0.165)</td>
<td>(0.138)</td>
<td>(0.201)</td>
</tr>
<tr>
<td>Early Breast Feeding is Good</td>
<td>0.122</td>
<td>0.117</td>
<td>0.281</td>
<td>0.480</td>
<td>0.275</td>
<td>0.066</td>
<td>0.108</td>
<td>0.065</td>
<td>0.307</td>
</tr>
<tr>
<td></td>
<td>(0.237)</td>
<td>(0.199)</td>
<td>(0.172)</td>
<td>(0.148)</td>
<td>(0.178)</td>
<td>(0.167)</td>
<td>(0.150)</td>
<td>(0.147)</td>
<td>(0.192)</td>
</tr>
</tbody>
</table>

*p < 0.01, c p < 0.05, b p < 0.1

Continues
<table>
<thead>
<tr>
<th>Variable</th>
<th>Quantile</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mother Uses Media channels</td>
<td></td>
<td>0.097</td>
<td>-0.088</td>
<td>-0.110</td>
<td>0.038</td>
<td>0.015</td>
<td>0.089</td>
<td>0.167</td>
<td>0.028</td>
<td>0.119</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child is Male</td>
<td></td>
<td>-0.187</td>
<td>-0.130</td>
<td>-0.025</td>
<td>-0.140</td>
<td>-0.011</td>
<td>-0.027</td>
<td>0.053</td>
<td>-0.053</td>
<td>-0.062</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age of Child (years)</td>
<td>0.434&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.303&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.108&lt;sup&gt;c&lt;/sup&gt;</td>
<td>-0.035</td>
<td>-0.124&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-0.276&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.450&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.647&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.892&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female Empowerment (Index)</td>
<td>0.373</td>
<td>0.566</td>
<td>0.481</td>
<td>0.302</td>
<td>0.253</td>
<td>0.207</td>
<td>-0.077</td>
<td>-0.084</td>
<td>-0.168</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household is Poor</td>
<td>-0.698&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-0.566&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.499&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.274&lt;sup&gt;c&lt;/sup&gt;</td>
<td>-0.082</td>
<td>-0.084</td>
<td>0.107</td>
<td>-0.007</td>
<td>-0.085</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Father Works</td>
<td>0.546</td>
<td>0.638</td>
<td>0.228</td>
<td>0.179</td>
<td>0.192</td>
<td>0.216</td>
<td>0.265</td>
<td>0.100</td>
<td>0.184</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other Working Members</td>
<td>0.064</td>
<td>0.022</td>
<td>0.144</td>
<td>0.044</td>
<td>-0.006</td>
<td>-0.013</td>
<td>-0.005</td>
<td>0.056</td>
<td>0.050</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Access to Safe Water</td>
<td>0.246</td>
<td>0.345</td>
<td>0.187</td>
<td>-0.013</td>
<td>-0.006</td>
<td>0.032</td>
<td>0.122</td>
<td>-0.081</td>
<td>-0.048</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hygiene and Sanitation</td>
<td>1.750&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.237&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.172&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.071&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.584</td>
<td>0.222</td>
<td>0.231</td>
<td>0.144</td>
<td>-0.073</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other Backward Caste</td>
<td>0.178</td>
<td>0.291</td>
<td>0.010</td>
<td>0.008</td>
<td>-0.168</td>
<td>-0.165</td>
<td>-0.111</td>
<td>-0.191</td>
<td>-0.031</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> p < 0.01, <sup>c</sup> p < 0.05, <sup>c</sup> p < 0.1; Table Continues
<table>
<thead>
<tr>
<th>Variable</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheduled Tribe</td>
<td>0.158</td>
<td>0.277</td>
<td>0.049</td>
<td>0.014</td>
<td>-0.013</td>
<td>-0.009</td>
<td>-0.193</td>
<td>-0.261</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>(0.482)</td>
<td>(0.261)</td>
<td>(0.277)</td>
<td>(0.219)</td>
<td>(0.214)</td>
<td>(0.204)</td>
<td>(0.209)</td>
<td>(0.231)</td>
<td>(0.300)</td>
</tr>
<tr>
<td>Scheduled Caste</td>
<td>-0.162</td>
<td>0.109</td>
<td>-0.188</td>
<td>-0.188</td>
<td>-0.300</td>
<td>-0.191</td>
<td>-0.169</td>
<td>-0.458</td>
<td>-0.262</td>
</tr>
<tr>
<td></td>
<td>(0.526)</td>
<td>(0.350)</td>
<td>(0.322)</td>
<td>(0.271)</td>
<td>(0.266)</td>
<td>(0.305)</td>
<td>(0.231)</td>
<td>(0.213)</td>
<td>(0.339)</td>
</tr>
<tr>
<td>South Region</td>
<td>-0.429</td>
<td>-0.516</td>
<td>-0.218</td>
<td>-0.061</td>
<td>0.031</td>
<td>0.125</td>
<td>0.307</td>
<td>0.057</td>
<td>-0.129</td>
</tr>
<tr>
<td></td>
<td>(0.470)</td>
<td>(0.399)</td>
<td>(0.343)</td>
<td>(0.320)</td>
<td>(0.314)</td>
<td>(0.307)</td>
<td>(0.304)</td>
<td>(0.364)</td>
<td>(0.395)</td>
</tr>
<tr>
<td>Western Region</td>
<td>-0.159</td>
<td>0.067</td>
<td>0.204</td>
<td>0.260</td>
<td>0.402</td>
<td>0.654</td>
<td>0.634</td>
<td>0.282</td>
<td>0.101</td>
</tr>
<tr>
<td></td>
<td>(0.436)</td>
<td>(0.417)</td>
<td>(0.300)</td>
<td>(0.283)</td>
<td>(0.274)</td>
<td>(0.271)</td>
<td>(0.273)</td>
<td>(0.334)</td>
<td>(0.384)</td>
</tr>
<tr>
<td>Eastern Region</td>
<td>0.198</td>
<td>0.237</td>
<td>0.778</td>
<td>1.138</td>
<td>1.354</td>
<td>1.761</td>
<td>1.981</td>
<td>2.022</td>
<td>1.291</td>
</tr>
<tr>
<td></td>
<td>(1.509)</td>
<td>(0.539)</td>
<td>(1.011)</td>
<td>(0.598)</td>
<td>(1.030)</td>
<td>(0.684)</td>
<td>(0.582)</td>
<td>(0.581)</td>
<td>(0.517)</td>
</tr>
<tr>
<td>Central Region</td>
<td>0.805</td>
<td>0.849</td>
<td>0.907</td>
<td>0.834</td>
<td>0.762</td>
<td>0.931</td>
<td>0.850</td>
<td>0.494</td>
<td>0.082</td>
</tr>
<tr>
<td></td>
<td>(0.439)</td>
<td>(0.343)</td>
<td>(0.289)</td>
<td>(0.254)</td>
<td>(0.257)</td>
<td>(0.216)</td>
<td>(0.248)</td>
<td>(0.338)</td>
<td>(0.350)</td>
</tr>
<tr>
<td></td>
<td>(3.8410)</td>
<td>(2.320)</td>
<td>(1.914)</td>
<td>(2.221)</td>
<td>(2.030)</td>
<td>(1.466)</td>
<td>(2.275)</td>
<td>(1.034)</td>
<td>(2.183)</td>
</tr>
</tbody>
</table>

\(^a\) p < 0.01, \(^c\) p < 0.05, \(^*\) p < 0.1
Figure 1.1: Kernel Density Estimates of Height-for-Age, Pooled Sample
Figure 1.2: Kernel Density Estimates of Height-for-Age, by Mother Category
Figure 1.3: Distribution of Height-for-Age of Children of Working and Nonworking Mothers
Chapter 2

SEMIPARAMETRIC ANALYSIS OF AGRICULTURAL TECHNOLOGY ADOPTION IN UGANDA: THE ROLE OF NONFARM EARNINGS

2.1 Introduction

Technological improvements in the agriculture sector are believed to be the most important pathway for reducing rural poverty in many agrarian economies such as those in Sub-Saharan Africa (SSA) (Bourdillon et al., 2003; Kijima, Otsuka, and Sserunkuuma, 2008; Mendola, 2007; Kassie, Shiferaw, and Muricho, 2011). For many of these countries, agriculture provides the leading source of employment and contributes large fractions of national income.\(^1\) In the case of Uganda, the agriculture sector contributes at least 40% of the Gross Domestic Product, about 85% of the export earnings, and employs over 70% of the national labor force (Government of Uganda, 2009b). Nearly 90% of the population of Uganda lives in rural areas, and directly derives its livelihood from subsistence farming (Government of Uganda, 2009a).

In recognition of its importance, development partners and governments of the SSA economies have invested in agricultural research and development to increase agricultural productivity and stimulate growth in these countries (Doss, 2006). In Uganda, the

---

\(^1\) Agriculture contributes a third of the regional “GNP” and employs at least two-thirds of the labor force (The World Bank, 2011).
government launched the Plan for Modernization of Agriculture (PMA), a holistic policy framework aimed at eradicating poverty by transforming farmers from subsistence farming to market oriented production. To achieve the PMA’s mission, the National Agriculture Research System (NARS) in collaboration with international research centers generated a wide range of improved technologies and management practices that have been disseminated to farmers through the National Agricultural Advisory Services (NAADS) and several other private service providers. Adoption of modern technologies such as the high yielding varieties (HYV) is expected to increase farm-level productivity and improve livelihoods of the farm households in developing countries (The World Bank, 2013). Several studies have reported that adoption of improved agricultural technologies enhances household well-being in developing countries (e.g. Bourdillon et al. (2003) in Zimbabwe; Mendola (2007) and Hossain et al. (2003) in Bangladesh; Ali and Abdulai (2010) in Pakistan; and Kijima, Otsuka, and Sserunkuuma (2008)and Kassie, Shiferaw, and Muricho (2011) in Uganda). In particular, Kijima, Otsuka, and Sserunkuuma (2008) and Kassie, Shiferaw, and Muricho (2011) report that adoption of upland rice and modern groundnut varieties may improve household income and reduce poverty in Uganda.

However, sustainable livelihood impacts of improved agricultural technologies in many SSA economies are hampered by the low rates of adoption of these technologies (Tripp and Rohrbach, 2001). For instance, only 28% of the land area allocated to maize in the SSA region is planted with improved maize varieties (Langyintuo et al., 2010).2 Furthermore, an average farmer in SSA applies only about 8 kg per hectare of fertilizers compared to 101 kg per hectare in South Asia (Morris et al., 2007) and over 145 kg per hectare in the developed world (The World Bank, 2010). The majority of rural farmers in SSA countries including Uganda are unable to purchase modern inputs because they lack liquid capital and have limited access to credit (Langyintuo et al., 2010; Simtowe et al., 2010; Ndjeunga and

---

2Maize is a major staple crop enterprise in SSA.
Markets for credit and insurance are either not available or dysfunctional in many of these economies (Gruhn and Rashid, 2001). The few available credit institutions are reluctant to give loans to agriculture (Gordon, 2000). As a result, credit institutions set high collateral requirements and charge high interest rates on credit, inhibiting farmers’ access to credit (Gruhn and Rashid, 2001).

Diversification into nonfarm income activities is an important strategy used by credit-constrained households in developing countries to obtain investment capital (Janvry and Sadoulet, 2001; Iiyama et al., 2008; Barrett, Reardon, and Webb, 2001; Reardon, Stamoulis, and Pingali, 2007). In Uganda, the share of total income from nonfarm activities for rural households increased from 46% in 2000 to 65% in 2006 (Uganda National Household Survey, 2006b). These shares are above the average of 35% reported for Africa (Haggblade, Hazell, and Reardon, 2010). Further, the level of household participation in rural nonfarm activities has significantly increased from 49% in 2003 (Kijima, Matsumoto, and Yamano, 2006) to about 59% in 2009 (Uganda Bureau of Statistics, 2010). Rural nonfarm opportunities are a more reliable source of household income, and often fetch higher returns to labor and capital than the agriculture sector (Reardon, Berdegue, and Escobar, 2001). In addition, the risk covariance between nonfarm and farm portfolios is low, making it possible for poor farmers to effectively insure against the risks and uncertainties in the agriculture sector and market failures (Reardon, Berdegue, and Escobar, 2001). The stream of income earned from nonfarm activities does not only enable farmers to smooth consumption (Kijima, Matsumoto, and Yamano, 2006; Janvry and Sadoulet, 2001) but may also provide them with liquid capital to purchase modern farm inputs (Reardon, 1997; Barrett, Reardon, and Webb, 2001).

Maize is a major staple crop enterprise in the SSA.
Despite the evidence of increasing importance of nonfarm income in Uganda, no empirical study to the best of my knowledge has analyzed its direct causal effect on the agricultural sector, the dominant sector of Uganda’s economy. The few published studies available have generally focused on investigating the impact of participation in nonfarm activities on rural poverty alleviation in general (Ellis and Bahiigwa, 2003; Kijima, Matsumoto, and Yamano, 2006; Matsumoto, Kijima, and Yamano, 2006). The main purpose of the study is to investigate if higher nonfarm earnings spur a greater likelihood of adoption of improved maize seed technologies in Uganda. If so, we can conclude that nonfarm earnings constitute an alternative source of investment capital for farmers who wish to adopt improved agricultural technologies but are frozen out of credit markets or cannot borrow the desired level of capital.

An important econometric challenge to the estimation of a causal effect of nonfarm income is its potential endogeneity, arising from unobservables that affect both household participation in nonfarm activities and their adoption decisions. Without proper instruments, the estimated effect may understate or overstate the true impact of nonfarm earnings. The effect may be biased if unobserved attributes of the farmers exert a positive impact on household nonfarm earnings and agricultural investment choices. For example, decision-makers that are more entrepreneurial may be more likely to (i) engage in nonfarm activities and (ii) adopt improved agricultural technologies. If so, the effect of nonfarm earnings would be biased upward because of the positive correlation with unobservable entrepreneurial skills. On the other hand, Hertz (2009) argues that households may, due to preferences and/or risk aversion, shift labor and capital resources from the risky agricultural sector to the nonfarm sector with less uncertain returns, creating negative/downward bias. For example, many youths in Uganda, especially those residing in the periphery of urban and peri-urban areas, prefer to operate transportation business to agriculture, and
thus attach less importance to farming. In this case, the potential effects of nonfarm earnings on adoption would be biased downwards if endogeneity of nonfarm participation is not accounted for during estimation.

Furthermore, the extant literature on agricultural technology adoption has heavily relied on distribution-dependent parametric methods (e.g., normal errors in the case of Probit) to estimate causal effects.\footnote{Selected examples from Africa include: Asfaw and Admassie (2004); Doss (2003); Olwande, Sikei, and Muthenge (2009); Simtowe et al. (2010).} Violation of distributional assumptions is well known to lead to misspecification error in likelihood based estimators of limited dependent variables models (Newey, 1985; Ichimura, 1993; Klein and Spady, 1993; Schafgans, 2004; Martins, 2001). This would render the inferences drawn from such models potentially incorrect and misleading for policy prescriptions. This essay primarily relies on a recently developed semiparametric estimator (Rothe, 2009) for dependent variables with binary outcomes such as the case in this study. Rothe’s estimator is consistent under mild regularity conditions and a linear single-index assumption, and unlike other well-known semiparametric single index models of binary data (Klein and Spady, 1993; Ichimura, 1993), it accommodates endogenous variables in a simple two-step process. Correcting for endogeneity of nonfarm earnings is crucial to estimating its causal effect on the adoption of improved maize seeds as discussed above. A parametric two step Probit model is also estimated to compute baseline estimates to compare with the estimates derived using the semiparametric estimator. The empirical analysis done in this essay shows that the methods yield qualitatively similar conclusions but the quantitative inferences are substantially different. Most notably, the parametric model understates the probability of adoption for the poor farmers (in terms of nonfarm income) but significantly overstates the likelihood of adoption for the richer farmers.

The rest of the paper is organized as follows. The second section describes the maize
subsector in Uganda and reviews the relevant literature related to the implications of non-
farm income on technology adoption in Africa. The third section outlines the conceptual
model and estimation methods, describing the parametric and semiparametric estimators
adopted in this paper. The fourth section describes the data used, and the selected charac-
teristics of the households in the sample. The fifth section presents the estimation results
and the specification tests for the parametric model. The last section concludes.

2.2 Background and Literature Review

2.2.1 Significance of Maize and Technological Progress in Uganda

Maize (corn) is one of the most important agricultural export commodities and the most
widely cultivated cereal crop in Uganda. The crop is cultivated by about 86% of Uganda’s
rural households (UBOS, 2006), accounting for about 46% of the total land under cereal
crops in the country. Recent data shows that the household marketed share of maize in-
creased from 14% in 2000 to 52% in 2006 (Uganda’s Plan for Modernization of Agricul-
ture, 2007; Uganda Bureau of Statistics, 2000). This is a substantial increase and demon-
strates that maize has become an important commercial crop in Uganda.

Overall, maize production has generally increased in the last two decades but with a
proportional increase in acreage. Annual production increased from 602,000 metric tons in
1990 to 1,272,000 metric tons in 2008 with an average of about 998,200 metric tons. Cor-
respondingly, the total land planted with maize increased from 401,000 hectares in 1992
to 887,000 hectares in 2008. The crop yield at farm level has remained low (average of
1.5 metric tons per hectare of grain yield) (Food and Agricultural Organization, 2010))
compared to the 3-7 metric tons per hectare from experimental plots at the national agri-
cultural research stations. Poor maize yields have been attributed to a number of technical
and economic factors such as poor management practices, shortage of good planting seeds,
infertile soils, prevalence of pests and diseases, low yield potential of local cultivars and poor farm implements, limited use of improved varieties and other purchased inputs such as fertilizers and other agrochemicals (Blackie, 1994).

There have been substantial technological advancements in Uganda’s maize subsector in the past two decades, especially the use of improved seed varieties. The National Agricultural Research System (NARS) of Uganda has generated several new maize varieties, among which are the Longe series (1 through 12), which were bred for disease resistance and high yields Uganda National Household Survey (2006a). However, the level of uptake of these new varieties is still low, only about 21%, (Uganda National Household Survey (2010) which has been attributed mainly to lack of capital to purchase seed. Many resource poor farmers in Uganda, including maize farmers, cannot afford the expensive modern inputs and have limited access to credit especially because of high interest rates and lack of collateral. Hence, farmers must find other sources of liquid capital, such as income from nonfarm activities, to purchase better agricultural inputs.

2.2.2 Literature Review

The literature on technology adoption highlights a number of different explanations for low adoption of improved agriculture technologies in developing countries, ranging from credit and liquidity constraints, information barriers, costs, uncertain benefits, risk and taste preferences, and differences in agro-ecological conditions. This section presents a review of selected recent studies on agricultural technology adoption in Sub-Saharan Africa. The literature is presented in three main strands.

One strand of literature has emphasized social interactions as an important determinant of technology adoption. For example, Bandiera and Rasul (2006) study the effect of social network of family and friends on adoption of sunflower in Mozambique. Their findings show that adoption decisions are more highly correlated within the network of family and
friends than within religion networks, and uncorrelated among individuals belonging to different religion networks. Similarly, Moser and Barrett (2003) find that learning effects, both from extension agents and from other farmers exert a significant influence over adoption decisions in Madagascar. More recently, Conley and Udry (2010) investigated the role of social learning in the diffusion of a new agricultural technology in Ghana. They collected data on farmers’ sources of information. Their findings suggest that farmers adjust their choice of inputs to align with those of their neighbors who had been successful in previous periods.

The second strand of literature focuses on the relationship between different forms of heterogeneity and technology adoption decisions. For example, in their analysis of fertilizer and improved seed adoption in Tanzania, Nkonya et al. (2005) find that farmer adoption decisions are significantly influenced by differences in biophysical, environmental, and socioeconomic conditions under which they operate. In particular, they report that adoption of improved seed was positively associated with the rate of nitrogen application, farm size, farmer education, and visits by extension agents. Further, Dercon and Christiansen (2007) analyze the impact of ex-post risk on adoption of fertilizers in Ethiopia, and find that downside risk in consumption exerts a negative and significant effect on the rate of fertilizer application. More recently, Suri (2011) specifies a generalized Roy model with comparative advantage and heterogeneity in returns to technology to analyze adoption of hybrid seed and fertilizers in Kenya. Findings from this study show heterogeneity in returns to hybrid maize, leading to a conclusion that high average returns to technologies mask the low returns realized by marginal farmers, who are the majority. The study further compares the estimated returns with adoption behavior, revealing that farmers with the highest returns did not use hybrid maize seed, those with lower returns adopted hybrid maize seed, while marginal farmers (with zero returns) switched in out of use of hybrid maize from year to year. These results demonstrate that aggregate adoption rates may remain low or stagnant.
despite high average returns to new maize technologies because either the marginal returns to adoption are low, or because the farmers with comparative advantage in adoption have already used the technology.

Another important strand of literature, and most relevant to this essay, relates adoption decisions to access to liquidity and credit. For example, Moser and Barrett (2003) analyze farmers’ decisions to adopt, expand, and dis-adopt high yielding rice varieties in Madagascar. They fit a dynamic Tobit model of technology adoption under incomplete financial and land markets, and find that seasonal liquidity constraints discouraged adoption by poorer farmers. Similarly, Croppenstedt, Demeke, and Meschi (2003) find that credit was the most important constraint to adoption of fertilizers in Ethiopia. It has been noted that subsistence farmers want to use advanced farm technologies but do not have liquid capital to purchase them (Duflo, Kremer, and Robinson, 2008). They have limited access to credit because it is either not available or they do not have collateral to get credit for farm investment (Hertz, 2009). Moreover, typical subsistence farmers are usually not able to save their farm earnings to purchase inputs later because they face other needs that compete for the limited financial resources. In their recent experiment conducted in Kenya, Duflo, Kremer, and Robinson (2011) find that farmers could only use farm revenue to purchase fertilizers immediately after harvesting. Their findings show that the proportion of farmers using fertilizer increased by at least 33% when farmers were offered the option to buy fertilizer immediately after the harvest.

This essay contributes to existing literature on technology adoption in SSA by exploring the linkage between nonfarm income and technology adoption under liquidity constraints. In the presence of imperfect credit markets, rural nonfarm income opportunities are expected to substitute for borrowed capital (Ellis and Freeman, 2004; Collier and Lal, 1984; Reardon, 1997), and can increase the collateral base of households (Reardon, Crawford,
and Kelly, 1994; Barrett, Reardon, and Webb, 2001). This translates into increased availability of liquid capital to farmers for financing purchase of capital intensive technologies. On the other hand, reliance on nonfarm income opportunities by rural farmers may reduce their adoption of modern technologies (especially labor intensive technologies) by shifting farm labor from agriculture to the nonfarm sector (McNally, 2002; Goodwin and Mishra, 2004). Nevertheless, increased demand for household labor for nonfarm work activities may increase adoption of capital intensive or/and time-saving technologies.

Empirical studies that have investigated the effect of nonfarm income on technology adoption in Africa report mixed findings. Holden, Shiferaw, and Pender (2004) explores the impact of improved access to nonfarm income on household welfare, agricultural production, and conservation investments in Ethiopia. Their results show that access to nonfarm income opportunities increases household income but reduces farmer incentives to invest in conservation, leading to rapid land degradation. Marenya and Barrett (2007) estimate a multivariate Probit model to quantify the determinants of adoption of natural resource management practices in Western Kenya, and find a positive and significant effect of non-farm income on use of inorganic fertilizers.\(^5\) Clay, Reardon, and Kangasniemi (1998) fit a farm investment model to analyze the effect of nonfarm income on farmers’ investment in land conservation and soil fertility in Rwanda. Their results indicate that nonfarm income significantly increased investment in land conservation but did not affect their decision to apply chemical fertilizers. Using a similar approach, Chikwama (2004) finds no evidence that existing rural wage opportunities contribute towards raising households’ farm investment in Zimbabwe. On the other hand, (Savadogo, Reardon, and Pietola, 1994, 1998) show that nonfarm income significantly increased adoption of animal traction in Burkina Faso.

The work in this essay differs from the above studies in three different ways. First, the study recognizes that nonfarm income and farm revenue are endogenous in a technology

\(^5\) They assume that nonfarm income is exogenous.
adoption model; second, a novel semiparametric estimator for endogenous binary models is used to relax the parametric distribution assumptions imposed in previous studies; third, it relies on a new dataset from Uganda.
2.3 Econometrics

2.3.1 Model Specification

The study examines adoption behavior of smallholder farmers in Uganda using the conventional random utility framework (Hausman and Wise, 1978; Hanemann, 1984). In this model, households are assumed to make rational production decisions by choosing technologies that maximize their expected utilities. This study assumes that households are faced with a set of two alternative maize (corn) technologies: the improved maize varieties and their traditional counterparts. Following literature (Hanemann, 1984; Baltasa and Doyle, 2001), the random utility function for a maize farm household facing a technology set can be specified as below:

\[ \pi_{ik} = \bar{\pi}(s_{ik}, t_i) + \varepsilon(s_{ik}, t_i) = x_{ik}\theta + \varepsilon_{ik}, i = 1, 2, ..., n \]  

(2.3.1)

Where \( \pi_{ik} \) denotes expected utility a farmer derives from planting maize seed variety \( k \) (\( k = 1 \) if improved variety, \( k = 0 \) if local variety), \( \bar{\pi}(s_{ik}, t_i) \) is the deterministic component of the utility function specified as a function of the observed attributes of the technology options \( s_{ik} \) and the socio-economic characteristics of a household \( t_i \); \( \varepsilon(s_{ik}, t_i) \) is the stochastic component of the utility function representing unobserved attributes affecting technology choice, heterogeneity in tastes and measurement errors; the matrix \( x_{ik} \) includes covariates \( s_{ik} \) and \( t_i \); \( \theta \) is the vector of parameters. Given the specification of \( \pi_{ik} \), a utility maximizing farm household would choose to adopt improved maize seed varieties if the expected utility derived from them is higher than the utility generated from their local

---

6Households’ preferences for specific technologies are driven by the expected benefits of using the technologies, and are influenced by both the observed and unobserved factors including attributes of the technologies and the tastes (Baltasa and Doyle, 2001).
(traditional) counterparts i.e. if $\pi_{i1} \geq \pi_{i0}$. The binary response model of adoption is thus specified in equation 2.3.2:

$$y_i = I\{\pi_{i1} - \pi_{i0} > 0\} = I\{x_{i1} \theta + \epsilon_{i1} - x_{10} \theta - \epsilon_{i0} > 0\} = I\{x_i \theta + u_i > 0\} \quad (2.3.2)$$

Where $u_i = \epsilon_{i1} - \epsilon_{i0}$ is a random error term with zero mean and is defined up to some scalar normalization.

The major focus of this study is to provide an understanding of how farmers adjust their adoption decisions ($y_i$) in response to changes in endogenous household nonfarm income. Nonfarm income in this case includes household revenue earned from wage employment and self employment, as well as the income transfers and remittances received from members of the household working outside home. The study also controls for the effect of the potentially endogenous household cash revenue generated from the sale of farm produce, and other exogenous variables. Potential endogeneity of the variables in the adoption equation can be accommodated by the following system of three equations:

$$y_i = I\{x_i \theta + u_i > 0\} = I\{x_{i1} \beta_1 + x_{2i} \beta_2 + x_{i0} \gamma + u_i > 0\} \quad (2.3.3)$$

$$x_{1i} = E[x_{1i} | z_i] + v_{1i} = \alpha_{1i} z_i + v_{1i}, \quad E[v_{1i} | z_i] = 0 \quad (2.3.4)$$

$$x_{2i} = E[x_{2i} | z_i] + v_{2i} = \alpha_{2i} z_i + v_{2i}, \quad E[v_{2i} | z_i] = 0 \quad (2.3.5)$$

where, $y_i$ is the binary decision for adoption, $x_i = (x_{1i}, x_{2i}, x_{i0})$ is a $1 \times (p \times q_1 \times q_2)$ dimensional vector of explanatory variables; $x_{1i}$ and $x_{2i}$ are measures of the endogenous variables—nonfarm earnings and the cash income from farm production, respectively; $z_i = (x_{i0}, z_{1i}, z_{2i})$ is a vector of instruments; $\theta = (\beta_1, \beta_2, \gamma)$, $\alpha_{1i}, \alpha_{2i}$ are unknown parameters to be estimated, where $\beta_1$ captures the causal effect of nonfarm income on adoption of improved seed; and $u_i, v_{1i}$ and $v_{2i}$ are the error terms. Equation (2.3.3) is the outcome equation or the technology adoption equation with a binary dependent variable as defined earlier. Equations
(2.3.4) and (2.3.5) are the first-stage equations for the endogenous variables nonfarm and farm earnings, respectively.

This study considers four instrumental variables to obtain unbiased estimates of farm and nonfarm income coefficients. The first instrument is the status of the local non-farm labor market captured by the share of non-farm income in the total household income at the village level as one of the instruments for nonfarm income. This variable was constructed by dividing the aggregate household non-farm income in a given village by the total income for all households in that village. A high share of non-farm income indicates high prevalence of nonfarm employment opportunities in the village, and translates into greater potential of households to diversify into non-farm income generating enterprises. It is expected that the existence of nonfarm opportunities in villages increases the probability of household participation in nonfarm work, leading to increased household nonfarm income. One may, however, argue that the share of nonfarm income likely affects household farm decisions via its negative effect on agricultural labor supply. To address this concern, the study controls for the amount of family labor supplied to agriculture (proxied by household size) and it is assumed that the only remaining pathway through which the share of nonfarm income influences farm production is through nonfarm income earned by the household.

The second set of instrumental variables are the 3-year lags of nonfarm income and farm revenue used in the nonfarm income equation. Past household income represents an important form of financial endowments, presenting households with an opportunity to invest in productive assets that generate income. This study hypothesizes that nonfarm income earned at least three years back directly affects current nonfarm income through its positive effect on investment in nonfarm activities (self-employment). However, it is unlikely that past income has a direct effect on current farm production decisions because few rural subsistence farmers, if any, in developing countries have savings accounts in formal institutions. It is also unlikely that subsistence farmers can keep cash savings for 3-5
years. Indeed, findings from a study of constraints to fertilizer adoption in Kenya by Duflo, Kremer, and Robinson (2011) show that farmers are unable to save money over even short periods of time, which is a major impediment to adoption. In practice, subsistence farmers including those in Uganda save money in the form of household assets, land, livestock and poultry. The value of household assets is included in the empirical model to control for household savings.

Third, the study uses existence of migrant networks in the district, defined as the percentage of households with at least one migrant, as an instrument for nonfarm income. According to the dynamic theory of migration, communities build migrant connections from interpersonal linkages involving migrants, former migrants, and non-migrants in the origin and destination areas (Massey et al., 1993). These networks facilitate migration because they lower the transaction costs and risks of movement, and provide information about economic opportunities elsewhere (Massey et al., 1993; McKenzie and Rapoport, 2007). Following the same logic, this study argues that existence of migrant networks in the districts in Uganda creates social capital which facilitates household access to nonfarm work opportunities in other communities. A possible threat to validity of this instrumental variable is that availability of migrant networks is likely to influence household agricultural decisions via its negative effects on local farm labor supply. The study, however, addresses this concern by controlling for family agricultural labor supply using household size in the estimators.

Fourth, the study uses contemporaneous weather shocks to instrument for household farm earnings. The weather shock variable was captured as an index, constructed as the mean of three indictors (dummies) of severe weather conditions in a village: drought, floods and landslides. The index takes values from 0 to 1, corresponding to favorable weather and severe weather conditions, respectively. Severe weather shocks during the
production season are expected to induce exogenous variations in farm earnings by reducing farm yields but have no direct effect on the adoption decisions in the current season or year.

The exogenous regressors \( (x_i) \) included in the model are drawn from recent literature of nonfarm labor supply (Chang and Yen, 2010; Holden, Shiferaw, and Pender, 2004; Isgut, 2004) and agricultural technology adoption (Conley and Udry, 2010; Munshi, 2004; Dercon and Christiansen, 2007; Phimister and Roberts, 2006; Fernandez-Cornejo, Hendricks, and Mishra, 2005; Ahituv and Kimhi, 2006; Gedikoglu, McCann, and Artz, 2011). Education of the household head measured in years of formal schooling, age of the household head, household access to agricultural extension and advisory services, and household size are included to capture the effects of human capital and risk tolerance (age) on the nonfarm income and technology adoption.\(^7\) The effect of ex-post savings and wealth on adoption are captured using the value of household assets. An index for lagged weather shock (1=severe weather, 0=favorable weather) is used to capture the effect of ex-post weather risk on ex-ante adoption decisions. The study also includes a dummy variable indicating whether a household had ever used improved farm technologies, to capture farmers’ adoption history. The effect of neighborhood experience on the adoption of improved maize seed was captured in the model by using the proportion of households in the village, except the \(i^{th}\) household, that had planted improved seed varieties. Consonant with social network theory, it is expected that households located in a village that has greater experience with improved seed are more likely to adopt them. Furthermore, the effect of competitiveness of maize enterprise in a household’s land allocation decisions is captured using the proportion

---

\(^7\)In rural areas in Uganda, household heads are the main decision makers in the households. The age of head of household can increase or decrease risk aversion.
of land size planted with maize. Finally, heterogeneous effects of adoption arising from location and agro-ecological characteristics are captured using regional dummies (northern, central, western, and eastern parts of the country).

### 2.3.2 A Semiparametric Binary Choice Model with Endogenous Covariates

The focus of the essay is to estimate the causal effect of nonfarm income on adoption of HYV maize crop seeds using a two-stage semiparametric model. A number of semiparametric methods for estimating binary choice models with endogenous regressors have been proposed (Newey, 1985; Blundell and Powell, 2004; Rothe, 2009). This essay implements Rothe’s two-stage semiparametric estimator, which is an extension of the semiparametric maximum likelihood estimator by Klein and Spady (1993). The estimator is \( \sqrt{n} \) consistent, asymptotically normal, allows a certain form of heteroskedasticity and, most importantly for the purpose of this essay, allows for endogeneity of continuous regressors. Monte Carlo simulations in Rothe (2009) indicate that the estimator exhibits better finite sample performance relative to competing semiparametric alternatives.

The key identifying assumption of the estimator is a distributional exclusion restriction, which requires that the dependence of the structural error terms \((u_i)\) on the explanatory variables and the instrumental variables is completely characterized by the residual vector (Blundell and Powell, 2004). By the distributional exclusion restriction, the conditional expectation of the binary dependent variable given the regressors and the reduced form error terms takes the form:

\[
E(y_i|x_i,z_i) = P(-u_i \leq x_i \theta | x_i, z_i) = P(-u_i \leq x_i \theta | v_i) = G(x_i \theta, v_i)
\]  

(2.3.6)

where \(G(x_i \theta, v_i)\) is the unknown cumulative distribution function (cdf) of \(-u_i\) conditional
on \( v_i \). If the link \( G(x_i \theta, v_i) \) were a known function, then the coefficient vector can be consistently estimated by maximizing the likelihood function:

\[
L_n(\theta) = \frac{1}{n} \sum_{i=1}^{n} (y_i \log G(x_i \theta, v_i) + (1 - y_i) (1 - \log G(x_i \theta, v_i))) \tag{2.3.7}
\]

In practice the link function is assumed to either belong to a known parametric family or is estimated nonparametrically. If, for example, we assume that the errors are jointly normally distributed then \( G(x_i \theta, v_i) \) becomes \( \Phi(x_i \theta + \rho v_i) \) where \( \Phi(\cdot) \) is the normal c.d.f. (Blundell and Powell, 2004). Following Klein and Spady (1993), Rothe (2009) proposes that the link function be estimated nonparametrically by the Nadaraya Watson estimator.

Let \( \hat{\omega} = (x_i \theta, \hat{v}_i) \) where \( \hat{v}_{si} = x_{si} - \hat{\alpha}_s z_i, s = 1, 2 \) are the first stage residuals used as control variables.\(^8\)

For given values of the parameter vector and the first stage residuals, the Nadaraya-Watson Kernel estimator of \( G(\hat{\omega}) \) is:

\[
\hat{G}(\hat{\omega}_i | \theta, \hat{v}_i) = \frac{\sum_j K_h(\hat{\omega}_j - \hat{\omega}_i) y_i}{\sum_j K_h(\hat{\omega}_j - \hat{\omega}_i)} \tag{2.3.8}
\]

where \( K_h(\xi) = K(\xi/h) / h \) is a kernel function, and is the bandwidth parameter.\(^9\) The semiparametric estimator is obtained by maximizing the quasi-maximum likelihood function (or the semiparametric likelihood function) obtained by substituting estimate \( \hat{G}(\hat{\omega}_i | \theta, \hat{v}_i) \) into the log likelihood function equation (2.3.7).\(^{10}\)

\(^8\)The first stage equation can be estimated using a fully nonparametric regression or assumed to satisfy certain parametric or semi parametric restrictions (Rothe, 2009). This essay uses OLS procedure in this application, as done in Rothe (2009).

\(^9\)Due to concerns over computational time, we consider only one bandwidth for index as well as the two control variables in our estimation. The “optimal” bandwidth is estimated jointly with the parameter vector \( \theta \). We use a nonlinear optimization subroutine by quasi-Newton method (nlpqn) of SAS/IML to maximize the quasi-maximum likelihood function.

\(^{10}\)As in the probit, location-scale normalization is needed to ensure identification of the parameter vector. For the probit, the location-scale normalization requires setting the first and second moments of the error term to zero and one, respectively. For the semiparametric estimator, the location-scale normalization is imposed by constraining the intercept to zero and one of the coefficients on continuous regressors to a constant.
In addition, the adoption model parameters were also estimated by maximizing the likelihood function (2.3.7) assuming that the errors are jointly normally distributed. This gives rise to the two-stage conditional maximum likelihood Probit estimator (Rivers and Vuong, 1988), which, like the semiparametric estimator, requires endogenous regressors to be continuous and includes the predicted first-stage residuals ($\hat{v}_1$ and $\hat{v}_2$) as control variables in the second stage.

2.4 Data and Summary Statistics

The study utilises the two waves of Uganda National Household Survey (UNHS) data collected by the Uganda Bureau of Statistics (UBOS) in 2005/06 and 2009/10. The study primarily relies on data from the 2009/10 wave but also utilizes lagged variables for income (used as instrumental variables), assets and weather shocks from 2005. The UNHS 2009/10 is a survey of a nationally representative sample of 3,123 households drawn from 322 enumeration areas (villages) distributed over 54 districts in Uganda.\footnote{The survey covered 72 enumeration areas in each of the four regions in the country: eastern, central, western and northern.} Data were collected for three main modules: the household module (socio-economic module), the agricultural module and the community module. The household module captured data on employment and income of each member of the household, education, household asset holdings, consumption expenditure, and access to services (and physical infrastructure). The agricultural survey collected data on agricultural production (agricultural inputs and technologies), land availability and land use, and shocks and uncertainties, agricultural extension services. Using the household identifiers in these two UNHS surveys, a household- matched dataset of 1218 maize farmer households was constructed. In this sample, about 22\% of the maize farmers planted improved maize seed varieties in 2009, and the proportion of farmers who...
had nonfarm income was about 83%. Nonfarm income accounts for about 67% of the total income of an average household in this sample.

Table 2.1 presents a summary of variables included in the econometric model, characterizing households in terms of adopters and non-adopters of improved maize seed. The descriptive statistics show that adopting households have better access to markets and advisory services and are more endowed with financial, physical and human capital than the non-adopters. In particular, adopters report significantly higher amounts of nonfarm income and a larger proportion of land planted with maize; they have more years of formal education, more interactions with agriculture extension workers, and better access to credit. On average, adopters report $1,269 in annual nonfarm income in the year 2009 vs. $1,074 for non-adopters.12 Adopters reported significant levels of past earnings from the farm and the non-farm subsectors relative to the non-adopters. All adopters vs. 74% of the non-adopters indicated that they had used improved seed varieties in the past. In addition, adopters resided in villages with significantly higher proportions of farmers using improved seed varieties than the non-adopters. Non-adopters experienced significantly severe past weather shocks relative to the adopters. Severity of contemporaneous weather shocks was, however, comparable between the communities of the two farmer categories. Further, adopters are closer to trading centers by about five km on average. Households that are located close to trading centers have better access to purchased inputs and are thus more likely to use these technologies. The summary statistics also show that the percentage of male adopters is higher (83.1%) than that of male non-adopters (69.7%). The statistics generally suggest that nonfarm income may catalyze adoption of improved seed varieties. However, simple comparison of the basic indicators of the two categories

---

12The median income is however much lower than the average income. The median nonfarm income is about $587 for adopters and $453 for non-adopters. Median farm revenue is $755 for adopters and $684 for non-adopters.
does not provide a causal interpretation of the effect because nonfarm income is potentially endogenous.

2.5 Estimation Results

Table 2.2 presents the results of the first stage equations and F-tests for the relevance of the instrumental variables used in this study. The results show significant coefficients on the instrumental variables in the expected direction as discussed above, indicating that the excluded variables satisfy the validity requirement. In particular, lagged nonfarm income and the village level share of nonfarm income have a positive and significant effect on current nonfarm income. On the other hand, occurrence of severe weather conditions significantly reduced the amount of farm income earned in a household. The F-tests suggest that the chosen instruments are relevant in explaining the exogenous changes in the endogenous variables. The Amemiya-Lee-Newey test of over-identification was also performed, and fails to reject the null hypothesis that the instruments are uncorrelated with the error term.\footnote{Chi-sq(2)= 0.464, P-value = 0.792.} Thus, the excluded variables jointly satisfy the over-identification requirement for the two income equations.

In the second step, the residuals generated from the first step were included as control variables in the parametric and semiparametric models of adoption. Table 2.3 presents the estimated coefficients of the parametric and semi parametric adoption models.\footnote{The scale restriction (see footnote 10) of the semiparametric model, the coefficient of the number of household members involved on farming is normalized to the corresponding Probit estimate.} The first-stage residuals are highly significant in the Probit model lending some support to the treatment of farm and nonfarm income as endogenous regressors. It is also important to note that while both the Probit and semiparametric models exhibit highly significant

13

14

coefficients, the standard errors for the semiparametric estimates are substantially smaller. These efficiency gains may be the result of the true link function being far from the normal distribution, in which case the semiparametric estimator dominates the Probit (Martins, 2001; Klein and Spady, 1993; Blundell and Powell, 2004; Rothe, 2009). The Lagrange multiplier test of normally distributed errors (Bera, Jarque, and Lee, 1984; Wilde, 2008) was performed to establish if the Probit functional specification is appropriate for the data. The test rejects the Probit specification with a p-value of 0.0049, indicating that Probit model may not capture the true adoption behavior of the sample of maize farmers.

Given that the specification test rejects the Probit functional form, the following discussions focus mostly on the semiparametrically estimated coefficients in the ensuing analysis. Starting with the key variable of interest, the findings show that an increase in nonfarm income is positively and significantly associated with an increase in the likelihood of adoption of improved maize seeds. These findings corroborate previous related studies that nonfarm income induces adoption of modern farm inputs in developing countries (e.g. (Savadogo, Reardon, and Pietola, 1998; Fernandez-Cornejo, Hendricks, and Mishra, 2005; Goodwin and Mishra, 2004; Phimister and Roberts, 2006; Marenya and Barrett, 2007; Hertz, 2009)). This study follows the procedure in Rothe (2009) and Blundell and Powell (2004) to evaluate the impact of change in the coefficient of nonfarm income on adoption. The procedure involves estimation of the average structural functions (ASF), which represents the marginal probability that the dependent variable takes a value of 1 for exogenously determined values of the regressors.\textsuperscript{15}

In Figure 2.1, the estimated ASF for the semiparametric model is plotted over the 1%

\textsuperscript{15}Marginal effects are difficult to compute in binary choice models with endogenous regressors. The alternative is the Average Structural Function (ASF), the average of the normal CDF transformed predictions (Bera, Jarque, and Lee, 1984; Wilde, 2008). Formally, ASF is the partial mean of the conditional distribution function with respect to the marginal distribution of the reduced form errors \(v_i\) while holding the index \((x_i \theta)\) constant: \(ASF = \int G(x_i \theta, v_i)F_v\), where \(F_v\) is the marginal distribution function of \(v\).
to 99% quintile range of the nonfarm income, given the sample average values for the remaining variables. As seen in the graph, the choice probabilities monotonically increase over the distribution of nonfarm income. The estimated probabilities imply that a farmer with annual nonfarm income of $4,363.276 (about the 95% percentile) is 32.8% more likely to adopt improved seed. Given the average adoption rate of 22.3%, probability estimates imply that one standard deviation increase in average nonfarm income increases the likelihood of adoption by 47%.\(^{17}\)

Other significant determinants of adoption include farm revenue, history of adopting technologically advanced inputs, neighborhood effects, extension education, age of the farmer, credit and previous weather shocks. In particular, the amount of farm revenue per hectare generated by farmers yields a significant impact on adoption.\(^{18}\) Having adopted one or more modern agricultural inputs in the past is a statistically significant and positive predictor of adoption of improved maize seeds. Presence of farmers who have used improved seed in the neighborhood also increases a farmer’s probability of adopting improved maize seed—suggesting that learning from other farmers plays a significant role in technology adoption. The findings also reveal that the probability of adoption increases with the number of interactions between the farmer and agricultural extension agents. This underscores the role played by the extension system and development projects in dissemination and promotion of improved technologies in Uganda.

The results further show that the probability of adoption significantly decreases with

---

\(^{16}\)We assume a standard deviation increase in nonfarm income for the average household which translates to total nonfarm income of $1,128.38+$3,234.88=$4,363.28 per columns 2 and 3 of the descriptive statistics table (Table 2.1).

\(^{17}\)The proportional marginal effect of 47% is obtained by dividing the probability of adoption (32.8%) for a farmer with an annual nonfarm income of $4,363.28 (see previous footnote) by the baseline adoption probability of 22.3%.

\(^{18}\)A better measure of farm income would have been farm profits which we unfortunately do not have because of lack of reliable input cost estimates from the survey.
the farmer’s age, probably due to aversion to risk. Experimental studies, such as that by Yesuf and Bluffstone (2007) in Ethiopia, have found that farmers become more risk averse as they age. The results also show that farmers who experienced severe weather shocks in the recent past were less likely to adopt improved seed, suggesting that weather risk discourages farmers to invest in risky farm technologies.

Surprisingly, the distance to the nearest trading center has a counterintuitive positive impact on the probability of adoption. A negative effect was expected because inputs markets are mainly located in town centers and farmers face transaction costs when buying inputs. Two possible explanations can be used to interpret the positive effect of distance on adoption. First, farmers farther away from trading centers receive lower farm gate price due to transport costs, implying that they may have a greater incentive to adopt improved seed to boost yields. Secondly, households located farther away from trading centers primarily derive their livelihoods from farming, and are therefore more likely to adopt improved seed compared to the households closer to trading centers with more nonfarm opportunities. The study findings also show a negative and significant relationship is found between receipt of credit and adoption of improved seed. The negative effect could be due to risk averse behavior of farmers. Because of high cost of credit, small scale farmers may be reluctant to allocate debt capital to the risky farm enterprises to avoid losing their collateral (Hertz, 2009). The three regional dummies show significant and positive coefficients suggesting that farmers in these agro-ecological zones are more likely to adopt improved seed relative to farmers in Northern region (the base region for the model). These findings suggest that regionally-specific attributes such as agro-ecological characteristics and locational factors wield a significant effect on farmer adoption decisions.
2.6 Conclusions

This essay investigates the causal effects of earned nonfarm income on farm level adoption of improved maize seed varieties in Uganda. It tests the stylized fact that earned nonfarm income is a source of liquid capital for credit constrained farmers. Departing from the conventional parametric paradigm, the study analyzes adoption behavior using both parametric and semiparametric estimators. The parametric specification test rejects the normality assumption of the Probit model, meaning that the estimates derived using a parametric model are inconsistent. The study findings provide evidence that nonfarm income is a critical determinant of adoption of improved maize seed. The findings suggest that in presence of credit constraints, nonfarm income, own farm revenue and remittance flows can induce investment in modern agricultural inputs. Strategic interventions aimed at promoting increased adoption and uptake of purchased agricultural technologies in Uganda should consider the important drivers of nonfarm income and agricultural earnings.
2.7 References


Dercon, S., and L. Christiansen. 2007. “Consumption Risk, Technology Adoption and


2.8 Tables and Figures

Figure 2.1: Average Structural Function
<table>
<thead>
<tr>
<th>Variable</th>
<th>Adopters</th>
<th>Non-adopters</th>
<th>T-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std.Dev.</td>
<td>Mean</td>
</tr>
<tr>
<td>Farm Revenue $10^3$ (US$)</td>
<td>1.104</td>
<td>1.318</td>
<td>1.794</td>
</tr>
<tr>
<td>Nonfarm Earnings $10^3$ (US$)</td>
<td>1.269</td>
<td>1.971</td>
<td>1.080</td>
</tr>
<tr>
<td>District Migrant Networks</td>
<td>0.227</td>
<td>0.096</td>
<td>0.242</td>
</tr>
<tr>
<td>Household Received Credit</td>
<td>0.864</td>
<td>0.343</td>
<td>0.865</td>
</tr>
<tr>
<td>Size of Household (No. of Persons)</td>
<td>7.081</td>
<td>3.070</td>
<td>6.795</td>
</tr>
<tr>
<td>Distance to the Trading Center (km)</td>
<td>3.685</td>
<td>6.077</td>
<td>4.483</td>
</tr>
<tr>
<td>Household Head is Male</td>
<td>0.831</td>
<td>0.376</td>
<td>0.697</td>
</tr>
<tr>
<td>Age of Head of Household (years)</td>
<td>44.985</td>
<td>13.700</td>
<td>48.428</td>
</tr>
<tr>
<td>Education of Household Head (years)</td>
<td>6.787</td>
<td>6.606</td>
<td>5.401</td>
</tr>
<tr>
<td>Agricultural Extension Visits</td>
<td>0.843</td>
<td>3.294</td>
<td>0.516</td>
</tr>
<tr>
<td>Nonfarm Share in the Village</td>
<td>0.527</td>
<td>0.202</td>
<td>0.194</td>
</tr>
<tr>
<td>Contemporaneous Weather Shock</td>
<td>0.031</td>
<td>1.491</td>
<td>1.313</td>
</tr>
<tr>
<td>Household Assets $10^3$ (US$)</td>
<td>12.583</td>
<td>95.484</td>
<td>5.656</td>
</tr>
<tr>
<td>Neighborhood Effects</td>
<td>0.417</td>
<td>0.249</td>
<td>0.251</td>
</tr>
<tr>
<td>Lagged Farm Revenue $10^3$ (US$)</td>
<td>0.067</td>
<td>0.119</td>
<td>0.046</td>
</tr>
<tr>
<td>Lagged Nonfarm Income $10^3$ (US$)</td>
<td>0.528</td>
<td>0.136</td>
<td>0.399</td>
</tr>
<tr>
<td>Lagged Weather Shock Index</td>
<td>0.269</td>
<td>0.308</td>
<td>0.350</td>
</tr>
<tr>
<td>Adoption History</td>
<td>1.000</td>
<td>0.743</td>
<td>0.437</td>
</tr>
</tbody>
</table>

Number of Observations 272 946

*** p<0.01, ** p<0.05, * p<0.1
### Table 2.2: First-Stage Regressions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Nonfarm Income Equation</th>
<th>Farm Revenue Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>Std. Error</td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.601</td>
<td>0.573</td>
</tr>
<tr>
<td>Lagged Nonfarm Income $\times 10^3$ (US$)</td>
<td>0.530***</td>
<td>0.104</td>
</tr>
<tr>
<td>Lagged Farm Revenue $\times 10^3$ (US$)</td>
<td>0.719</td>
<td>0.972</td>
</tr>
<tr>
<td>District Migrant Networks</td>
<td>-0.256</td>
<td>1.067</td>
</tr>
<tr>
<td>Nonfarm Share in the Village</td>
<td>2.809***</td>
<td>0.480</td>
</tr>
<tr>
<td>Household Assets $\times 10^3$ (US$)</td>
<td>0.016***</td>
<td>0.002</td>
</tr>
<tr>
<td>Contemporaneous Weather Shock</td>
<td>-0.155</td>
<td>0.522</td>
</tr>
<tr>
<td>Lagged Weather Shock Index</td>
<td>0.157</td>
<td>0.258</td>
</tr>
<tr>
<td>Household Received Credit</td>
<td>0.461*</td>
<td>0.253</td>
</tr>
<tr>
<td>Adoption History</td>
<td>0.118</td>
<td>0.215</td>
</tr>
<tr>
<td>Household Head is Male</td>
<td>0.002</td>
<td>0.206</td>
</tr>
<tr>
<td>Agricultural Extension Visits</td>
<td>-0.001</td>
<td>0.046</td>
</tr>
<tr>
<td>Age of Head of Household (years)</td>
<td>-0.001</td>
<td>0.006</td>
</tr>
<tr>
<td>Education of Household Head (years)</td>
<td>0.010</td>
<td>0.016</td>
</tr>
<tr>
<td>Size of Household (No. of Persons)</td>
<td>0.055*</td>
<td>0.029</td>
</tr>
<tr>
<td>Log of Distance (km)</td>
<td>-0.074</td>
<td>0.059</td>
</tr>
<tr>
<td>Neighborhood Effects</td>
<td>0.790**</td>
<td>0.400</td>
</tr>
<tr>
<td>Maize Farm Size (Share)</td>
<td>-0.367</td>
<td>0.498</td>
</tr>
<tr>
<td>Eastern Region</td>
<td>0.104</td>
<td>0.321</td>
</tr>
<tr>
<td>Central Region</td>
<td>-0.447</td>
<td>0.283</td>
</tr>
<tr>
<td>Western Region</td>
<td>-0.276</td>
<td>0.314</td>
</tr>
<tr>
<td>F-Statistic</td>
<td>11.370***</td>
<td></td>
</tr>
</tbody>
</table>

*** p<0.01, ** p<0.05, * p<0.1
Table 2.3: Second Stage Regressions

<table>
<thead>
<tr>
<th></th>
<th>Two-stage Probit Model</th>
<th></th>
<th>Semiparametric Model (Rothe)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>Std.Error</td>
<td>Coefficient</td>
<td>Std.Error</td>
</tr>
<tr>
<td>Nonfarm Earnings $\times 10^3$ (US$)</td>
<td>0.112*</td>
<td>0.066</td>
<td>0.176***</td>
<td>0.045</td>
</tr>
<tr>
<td>Farm Revenue $\times 10^3$ (US$)</td>
<td>0.130***</td>
<td>0.058</td>
<td>0.162***</td>
<td>0.029</td>
</tr>
<tr>
<td>Household Assets $\times 10^3$ (US$)</td>
<td>0.001</td>
<td>0.003</td>
<td>-0.001</td>
<td>0.004</td>
</tr>
<tr>
<td>Lagged Weather Shock Index</td>
<td>-0.041</td>
<td>0.133</td>
<td>-0.461***</td>
<td>0.185</td>
</tr>
<tr>
<td>Household Received to Credit</td>
<td>0.077</td>
<td>0.133</td>
<td>-0.247*</td>
<td>0.151</td>
</tr>
<tr>
<td>Size of Household (No. of Persons)</td>
<td>0.015</td>
<td>0.016</td>
<td>-0.030</td>
<td>0.020</td>
</tr>
<tr>
<td>Adoption History</td>
<td>0.276***</td>
<td>0.122</td>
<td>0.385***</td>
<td>0.147</td>
</tr>
<tr>
<td>Household Head is Male</td>
<td>0.310***</td>
<td>0.113</td>
<td>0.197</td>
<td>0.137</td>
</tr>
<tr>
<td>Agricultural Extension Visits</td>
<td>0.067***</td>
<td>0.021</td>
<td>0.068***</td>
<td>0.021</td>
</tr>
<tr>
<td>Age of Head of Household (years)</td>
<td>-0.013***</td>
<td>0.003</td>
<td>-0.007*</td>
<td>0.004</td>
</tr>
<tr>
<td>Education of Household Head (years)</td>
<td>0.005</td>
<td>0.008</td>
<td>0.010</td>
<td>0.010</td>
</tr>
<tr>
<td>Log of Distance (km)</td>
<td>0.034</td>
<td>0.035</td>
<td>0.105***</td>
<td>0.040</td>
</tr>
<tr>
<td>Maize Farm Size (Share)</td>
<td>-0.493</td>
<td>0.407</td>
<td>-0.493</td>
<td>N/A+</td>
</tr>
<tr>
<td>Neighborhood Effects</td>
<td>1.084***</td>
<td>0.201</td>
<td>1.015***</td>
<td>0.251</td>
</tr>
<tr>
<td>Eastern Region</td>
<td>1.176***</td>
<td>0.247</td>
<td>1.391***</td>
<td>0.186</td>
</tr>
<tr>
<td>Central Region</td>
<td>1.575***</td>
<td>0.264</td>
<td>1.648***</td>
<td>0.186</td>
</tr>
<tr>
<td>Western Region</td>
<td>1.108***</td>
<td>0.264</td>
<td>1.193***</td>
<td>0.203</td>
</tr>
<tr>
<td>Residuals From Nonfarm Equation</td>
<td>-0.119***</td>
<td>0.065</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residuals From Farm Equation</td>
<td>-0.158**</td>
<td>0.056</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Likelihood Value</td>
<td>-542.798</td>
<td></td>
<td>-468.574</td>
<td></td>
</tr>
</tbody>
</table>

*** p<0.01, ** p<0.05, * p<0.1

+ The coefficient of the share of maize was not estimated in the nonparametric model; see footnotes 10 and 14
Chapter 3

PARAMETRICALLY-GUIDED SEMIPARAMETRIC ESTIMATION OF BINARY-CHOICE MODELS: AN APPLICATION TO FERTILIZER ADOPTION IN UGANDA

3.1 Introduction

Discrete choice models are commonly used in many fields, including economics, political science, medicine, and psychology, to analyze situations where a decision or choice has to be made. Discrete choice models usually take the following general form

\[ y_i^* = v_i \beta + u_i \]  

(3.1.1)

where \( y_i^* \) is a latent variable which is operationalized by defining \( y_i = 1 \) if \( y_i^* \geq 0 \) and \( y_i = 0 \) otherwise, \( \beta \) is a \((q + 1) \times 1\) vector of unknowns, \( v_i \equiv (1, x_i) \) where \( x_i \) is a \( 1 \times q \) vector of explanatory variables, and \( u_i \) is the error term. The applied literature is dominated by parametric estimation where the distribution function (cdf) of \( u_i \), say \( F(u) \), is assumed to be normal (probit model) or logistic (logit model). Because parametric assumptions that are not consistent with the data could invalidate the results, some have considered nonparametric and semiparametric methods. In discrete choice models the estimated effects of the regressors are of interest as well as the estimated conditional mean. Thus instead of a fully nonparametric approach, semiparametric methods have been the focus to circumvent any distributional assumptions yet recover the desired estimates.
There has been significant research on semiparametric estimation of single-index models that contain parametric discrete choice models as a special case (see Klein and Spady (1993); Ichimura (1993); Horowitz and Hardle (1996); Cosslett (1987); Powell, Stock, and Stoker (1989); Stoker (1986)). A single-index model has the following form

\[ E(y|v) = F(v\beta) \]  

(3.1.2)

where \( F \) is an unknown (not necessarily a distribution) function, called the link function. The term \( v\beta \) is the index. Note that if \( F \) is the normal or logistic distribution function, the function in (3.1.2) is the binary probit or logit model and, if it is the identity function, equation (3.1.2) becomes the usual linear regression model. Among the advantages of single-index models is dimension reduction. The index \( v\beta \) is a scalar and thus single-index models do not suffer from the curse of dimensionality; if \( \beta \) were known it would be possible to estimate \( F \) as the nonparametric mean regression of \( y_i \) on \( z_i = v_i\beta \) which is a scalar. Therefore, in single-index models it is possible to estimate \( F \) at the nonparametric rate as if there is a single regressor and the coefficient vector \( \beta \) at the parametric rate \( O(n^{-1/2}) \).

In a completely different vein, a recent paper by Hjort and Glad (1995) proposes a semiparametric method for density estimation which starts with a parametric estimator and multiplies this parametric start with a correction factor (unknown density divided by the parametric start) which is estimated nonparametrically. The idea is based on bias reduction. If the parametric start captures a sufficient amount of the curvature of the unknown density, the correction factor will be close to a constant and less rough. Thus the bias associated with the nonparametric estimation of this correction factor will be less than that associated with the underlying density.

Unlike the semiparametric papers in the literature that estimate the link function nonparametrically, we propose to estimate the link function semiparametrically by employing
the Hjort and Glad (1995) bias reduction idea. Potentially relevant information is introduced in the form of a parametric function which is the parametric guide for the link function. The distinguishing characteristic of the proposed estimator is how the unknown link function $F$ is estimated using prior parametric information about its shape.

### 3.2 Ichimura’s Semiparametric Estimator of Binary Data

The single-index model of Ichimura (1993) is based on minimizing a nonlinear least squares (NLS) loss function. The NLS estimator of $b$ minimizes

$$
\frac{1}{n} \sum_{i=1}^{n} [y_i - \tilde{F}(x_i b)]^2
$$

(3.2.1)

where $\tilde{F}$ is the nonparametric estimator for the unknown link function, and $b \equiv (1, \beta_2, \ldots, \beta_q)^T$ is the $\beta$ vector after scale and location normalizations assuming the first regressor has a continuous distribution. $\tilde{F}$ is defined as

$$
\tilde{F}(x_i b) = \sum_{j \neq i} y_j K \left( \frac{x_i b - x_j b}{h} \right) / \sum_{j \neq i} K \left( \frac{x_i b - x_j b}{h} \right)
$$

(3.2.2)

where $K$ is the kernel function (usually a symmetric density function) and $h = h(n)$ is the smoothing parameter such that $h \to 0$ as $n \to \infty$. Ichimura (1993) denotes this model semiparametric least squares (SLS) and shows that $\tilde{b}$ is consistent and $\sqrt{n}(\tilde{b} - b_0) \xrightarrow{d} N(0, \Omega_{SLS})$ and gives a consistent estimator of $\Omega_{SLS} = \Gamma^{-1} \Sigma \Gamma^{-1}$. $\Gamma$ and $\Sigma$ can be consistently estimated by

$$
\hat{\Gamma} = \frac{1}{n} \sum_{i=1}^{n} \tilde{x}_i \tilde{x}_i^T \hat{F}'(x_i \tilde{b})^2,
$$

(3.2.3)

$$
\hat{\Sigma} = \frac{1}{n} \sum_{i=1}^{n} \tilde{x}_i \tilde{x}_i^T \hat{F}'(x_i \tilde{b})^2 [y_i - \hat{F}(x_i \tilde{b})]^2
$$

(3.2.4)

where $\tilde{x}_i = (x_{2i}, \ldots, x_{qi})$, $\tilde{b} = (1, \tilde{b}_T)^T$, and $\hat{F}'$ is the derivative of $\hat{F}$.
Ichimura (1993) also considers weighted SLS (WSLS) in which he weights the objective function (3.2.1) and the summands in $\tilde{F}$ by a weight function $W(x_i)$. Efficiency considerations play a role: the choice of the weight function does not affect the consistency and rate of convergence of the estimator of $b$ but does affect its efficiency. Optimally weighted (the weight function is a consistent estimator of $\text{Var}(y|x)^{-1}$) WSLS achieves the semiparametric efficiency bound.\textsuperscript{1}

Care must be taken to prevent the denominator of (3.2.2) from getting arbitrarily close to zero. Ichimura (1993) restricts the summands in (3.2.2) and (3.2.1) to those observations for which the density of the index is not too small.

### 3.3 Parametrically-Guided Single-Index Model

The proposed semiparametric estimator is motivated by the estimator of Hjort and Glad (1995); Glad (1998). Suppose one wishes to estimate the conditional mean function $E(y|x) = m(x)$. They first start with a parametric estimator $m(x, \hat{\beta})$ which could be, for instance, a simple linear regression or a more complex maximum likelihood estimation, and then multiply it with a correction factor $r(x) = m(x)/m(x, \hat{\beta})$ which is estimated nonparametrically. The idea is based on bias reduction: if the parametric start is close to the truth, the correction factor will be close to a constant and thus smoother and (bias-wise) easier to estimate than $m$ itself. Hence the bias associated with nonparametric estimation of this correction factor would be less than the bias from direct nonparametric estimation of the unknown regression function. Their estimator is

\[ \hat{m}(x) = m(x, \hat{\beta})\hat{r}(x). \]  

(3.3.1)

\textsuperscript{1}For this efficiency result, the weight function should depend on $x$ only through the index as in binary-choice models where $\text{Var}(y|x) = \text{Var}(y|xb)$. 

94
When the correction factor is estimated by the Nadaraya-Watson estimator their parametrically guided estimator is

$$\hat{m}(x) = \sum_{i=1}^{n} \frac{y_i}{m(x_i, \hat{\beta})} \frac{m(x_i, \hat{\beta})}{m(x_i, \hat{\beta})} K_h(x_i - x) / \sum_{i=1}^{n} K_h(x_i - x).$$

(3.3.2)

Glad (1998) shows that this estimator has the same large sample variance as the standard nonparametric estimators (Nadaraya-Watson and local linear) while bias reduction is possible if the parametric start belongs to a neighborhood around the true regression curve.

The proposed semiparametric estimator for the single-index models is based on the above idea. We introduce (potentially) relevant information in the form of a parametric function in attempts to reduce bias. The estimator starts with a parametric model for the link function, say $G(xb)$, where $G(\cdot)$ is a known function (for instance normal cdf) and multiplies it with the correction factor $r(xb) = F(xb)/G(xb)$ which is estimated nonparametrically. Note that the information that this estimator starts with is only related to the shape of the link function and not to the coefficient estimates. In this sense the estimator is using a fixed start vis--vis (Glad, 1998; Hjort and Glad, 1995). When the Nadaraya-Watson estimator is used to estimate the correction factor the proposed model is

$$\hat{F}(xb) = \sum_{j \neq i} \left\{ \frac{y_j G(x_i b)}{G(x_j b)} K \left( \frac{x_i b - x_j b}{h} \right) \right\} / \sum_{j \neq i} K \left( \frac{x_i b - x_j b}{h} \right).$$

(3.3.3)

This model will be referred to as parametrically-guided single-index model (PGSIM). Coefficient estimates are obtained by replacing $\hat{F}(xb)$ with our parametrically guided link function $\hat{F}(xb)$ (3.3.3) in the Ichimura loss function (3.2.1) which is then minimized with respect to the coefficient vector. A bias and variance comparison of the Nadaraya-Watson estimator and the PGSIM estimators of the link function is useful to see the bias reduction.

While both have the same variance

$$Var(\hat{F}(z)) = (nh)^{-1} f(z)^{-1} \sigma^2(z) R(K) + o_p((nh)^{-1})$$

(3.3.4)
where $f(z)$ is the density of $z$, $\sigma^2(z) = \text{Var}(y|z)$, and $R(K) = \int K^2(t)dt$, the bias of Nadaraya-Watson estimator is

$$
\text{Bias}(\hat{F}(z)) = \frac{h^2 \mu_2(K)}{2f(z)} \left( F''(z)f(z) + 2f'(z)F'(z) \right) + o_p(h^2) \quad (3.3.5)
$$

where $\mu_2(K) = \int t^2 K(t)dt$, whereas the bias of (3.3.3) is

$$
\text{Bias}(\hat{F}(z)) = \frac{h^2 \mu_2(K)}{2f(z)} \left( r''(z)G(z)f(z) + 2f'(z)r'(z)G(z) \right) + o_p(h^2). \quad (3.3.6)
$$

Thus, for the same $h$ and $K$, bias reduction is possible if the parametric start $G$ can be chosen such that

$$
|r''(z)G(z)f(z) + 2f'(z)r'(z)G(z)| < |F''(z)f(z) + 2f'(z)F'(z)|. \quad (3.3.7)
$$

If the parametric start is proportional to $F$, the correction factor $r$ is going to be a constant and thus $r' = r'' = 0$. If it is sufficiently close to $F$, roughness of $r$ will be less than roughness of $F$ and $r$ will have a smaller second derivative. Thus (3.3.7) defines a ‘neighborhood’ of $F$ where bias reduction is possible by choosing a parametric start from this neighborhood.\(^3\)

The above argument indicates that even though the estimator can not (asymptotically) outperform Ichimura (1993) (when optimally weighted in binary-choice model estimation) with respect to the coefficients as they attain the semiparametric efficiency bound, in finite samples there is potential to increase efficiency by using a parametric guide. On the other hand, the bias of (3.3.3) reduces to a smaller order than Ichimura (1993) probability estimates when the parametric start is proportional to $F$.

\(^2\)See Glad (1998) and Sam and Ker (2006) for the derivation of the bias and variance expressions.

\(^3\)(3.3.7) is a neighborhood for the link function and not for the coefficient estimates. For the latter, it is not straightforward to obtain a neighborhood like (3.3.7) as there is no closed form solution but it is reasonable to expect bias reduction for coefficient estimates as well since they would be obtained from first order conditions which are functions of less biased estimates of the link when (3.3.7) is satisfied.
In the semiparametric estimator of Ichimura care must be taken as the nonparametric density estimator in the denominator can get arbitrarily small. Ichimura restricts \( x \) to a set on which the above mentioned problem is avoided. With respect to the PGSIM estimator, the parametric start \( G(\cdot) \) should be nonzero throughout the support of the index in addition to restrictions so that the denominator does not get arbitrarily small.

Similar to the Ichimura estimator, the coefficient vector estimated based on our proposed link function \((\hat{b})\) is unbiased, \( \sqrt{n} \) consistent and asymptotically normal:

\[
\sqrt{n}(\hat{b} - b_0) \overset{d}{\rightarrow} N(0, \Psi)
\]

where \( \Psi \) can be consistently estimated by

\[
\hat{\Psi} = \left[ \frac{1}{n} \sum_{i=1}^{n} \tilde{x}_i \tilde{x}_i^T \hat{F}'(x_i \hat{b})^2 \right]^{-1} \left[ \frac{1}{n} \sum_{i=1}^{n} \tilde{x}_i \tilde{x}_i^T \hat{F}'(x_i \hat{b})^2 [y_i - \hat{F}(x_i \hat{b})]^2 \right] \left[ \frac{1}{n} \sum_{i=1}^{n} \tilde{x}_i \tilde{x}_i^T \hat{F}'(x_i \hat{b})^2 \right]^{-1}
\]

### 3.4 Empirical Application to Fertilizer Adoption in Uganda

As an empirical application of the proposed estimator, the study the determinants of fertilizer adoption in Uganda. Like many Sub-Saharan Africa (SSA) countries, economic growth in Uganda is hampered by low productivity of the agriculture sector, estimated at only 2% annual rate of growth compared to the 6% Government-set growth target (Uganda Bureau of Statistics, 2010). Poor soils and scarce use of modern technologies may be some of the major causes of low yields of agricultural enterprises. Furthermore, there is evidence of massive soil fertility depletion in the country, caused by the immense pressure

---

4 Asymptotic normality follows from the uniform consistency of \( \hat{b} \) (proof of uniform consistency is available upon request) and the theory of nonlinear least squares per which the asymptotic variance is identical to that of least squares based on the pseudo-regressors (see for example Hansen 2013, pp 187-89 for proof). A requirement for the existence of \( \Psi \) is the existence and boundedness of the first derivative of the link function. Our proposed link satisfies this requirement as a composition of bounded and smooth functions. We note that in practice, computing analytical derivatives of our link function estimator can be quite tedious. We therefore recommend the computation of the numerical derivative when estimating \( \Psi \).

5 The sector contributes at least 40% of the national GDP and about 85% of the export earnings, and employs over 70% of the national labor force (Government of Uganda, 2009).
on arable land due to increasing population density. For example Henao and Banaante (2006) estimate that the soils in Uganda lose about 66kg of nitrogen (N), phosphorus (P) and potassium (K) annually, making it one of the highest rates of soil nutrient depletion in SSA. Moreover, the cost of replenishing the depleted soil nutrients is high, about 20% of a farmer’s household income (Nkonya et al., 2005).

Inorganic fertilizer is one of the most important agricultural technologies that may have the potential to sustainably raise crop productivity and enhance food security and incomes of households. For instance, results from field trials conducted by Sasakawa Global 2000-Uganda show that application of fertilizer (at a rate of 90:40 kg per hectare of Nitrogen and Phosphorus) in maize yields about 4,312-6,054 kg per hectare compared to the 550kg per hectare for farmers who do not use fertilizers. Despite these proven beneficial effects and favorable Government policies and institutional frameworks, such as the National Agricultural Advisory Service (NAADS) that are designed to promote the use of modern inputs in Uganda, only a very small fraction of farmers use fertilizers. A study by Yamano and Arai (2010) found that only about 7-8% of Ugandan farmers used fertilizers in 2009, compared to about 17-31% reported by Suri (2011) in neighboring Kenya. Furthermore, an average farmer in Uganda applies about 2.1 kg per hectare, which is far below the 32.4 kg per hectare used by a farmer in Kenya (The World Bank, 2013). Moreover, organic alternative sources of soil nutrients such as manure and crop residues are labor intensive and do not have adequate nutrient levels (Yanggen et al., 1998; Morris et al., 2007). Abdoulaye and Sanders (2005) estimate that a farmer may need to apply about 6-10 tons of manure per

---

6Population pressure leads to land degradation because land scarcity forces farmers to cultivate the same piece of land every season, and also forces them to reclaim marginal land areas such as wetlands.

7Sasakawa Global 2000, is a program that aims to build the capacity of smallholder farmers to improve their livelihoods through adoption of modern farming methods, including the use of quality seed and small amounts of fertilizer, and value addition. In Uganda, SG 2000 has been active since 1997, covering over 24 districts in the country.
hectare to generate the adequate amount of nitrogen and phosphorous. In light of this background, this essay explores the empirical determinants of fertilizer use in Uganda using a novel semiparametric estimator.

While a number of studies have analyzed fertilizer adoption behavior of farmers in many developing countries, empirical work on the determinants of fertilizer adoption in Uganda is scarce. We are only aware of one study Okoboi and Barungi (2012) which finds that farmers do not use fertilizers primarily because of a lack of liquidity and knowledge on use of fertilizer. Our study uses a more recent, nationally representative dataset collected by the Uganda national Bureau of Statistics in 2010 and contributes to existing literature on adoption by testing whether non-debt liquidity (farm and nonfarm income) spurs fertilizer use for credit-constrained households. Furthermore, we depart from the conventional parametric approaches used by past studies, and use the semi-parametric estimator described above.

3.4.1 A Brief Literature Review on the Determinants of Fertilizer Adoption in SSA

Several explanations have been advanced by economists regarding the low level of fertilizer adoption in SSA countries, ranging from limited incentives to use fertilizers, to lack of physical and human capital to purchase and use fertilizers efficiently. Farmers in Africa generally have low incentives to use fertilizers because they are not certain of the returns to fertilizer application (Reardon et al., 1999). Fluctuations in prices and volatile crop yields caused by covariate shocks such as droughts, make it difficult for farmers to assess the expected returns to fertilizer use (Kelly, 2006). In addition, farmers lack adequate knowledge and skills in using fertilizers (Abdoulaye and Sanders, 2005; Kelly, 2006; Morris et al., 2007). Asfaw and Admassie (2004); Yamano and Arai (2010); Abdoulaye and Sanders (2005); Duflo, Kremer, and Robinson (2008) report that literacy or formal schooling and agricultural extension advice increase fertilizer adoption because they enhance farmer’s
awareness of and technical knowledge on the use of modern inputs. Farmer literacy is especially important because it enables farmers to synthesize, understand, and effectively utilize new technical information supplied by the extension agents (Kelly, 2006). Other disincentives to fertilizer adoption in SSA include high input to output price ratio, poor transportation network, limited and and untimely supply of fertilizers (Abrar, Morrissey, and Rayner, 2004; Yamano and Arai, 2010; Ndjeunga and Bantilan, 2005). In particular, poor transportation network and long distances to input suppliers further add to the high fertilizer cost, rendering fertilizers unaffordable for the subsistence farmers.\(^8\)

Further, liquidity constraints and lack of credit hamper fertilizer adoption. Indeed, Yamano and Arai (2010); Duflo, Kremer, and Robinson (2008) find that many subsistence farmers want to apply fertilizers but do not have liquid capital to purchase them. Many of the subsistence farmers have no access to credit or do not have collateral to get credit for farm investment (Hertz, 2009). Moreover, smallholder farmers are usually not able to save their farm income to purchase inputs later due to other competing financial needs (Duflo, Kremer, and Robinson, 2011).\(^9\) Some studies such as (Reardon, Stamoulis, and Pingali, 2007; Barrett, Reardon, and Webb, 2001) suggest that income from nonfarm activities may help farmers to overcome the credit constraints, and induce them to adopt purchased inputs. Researchers have also found that farmers’ decisions to use fertilizers are influenced by land size, albeit with mixed conclusions. For example, Nkonya, Schroeder, and Norman (1997) in Tanzania, and Freeman and Omiti (2003) in western Kenya report that households with smaller land per capita are more likely to use fertilizer relative to those with larger farm size per capita. In Zambia, however, Jha and Hojjati (1993) find that farmers with more

\(^8\)We note however, proximity to input shop does not necessarily imply increased quantity purchased (Ariga et al., 2008).

\(^9\)In their study, (Duflo, Kremer, and Robinson, 2011) find that the proportion of farmers using fertilizer increased by at least 33% when farmers were offered the option to buy fertilizer immediately after the harvest.
land per capita used fertilizer merely to maintain yields and address specific problems of infertility, and not to intensify (increase yields). Land size reflects the economic status of the household meaning that farmers with larger land size per capita may be wealthier and thus more likely to adopt fertilizers than their counterparts with smaller size of land. Richer farmers engage in more commercial and profitable cash enterprises and have better access to the infrastructure such as roads, extension, education, and markets giving them better access to fertilizers than the smaller farmers (Kelly, 2006).

3.4.2 Empirical Model and Data

We study fertilizer adoption behavior of smallholder farmers in Uganda using the conventional random utility framework (Hausman and Wise, 1978; Hanemann, 1984). It is assumed that households make rational production decisions and therefore only apply fertilizer if doing so maximizes their expected utilities. Households are faced with two production technologies: one with fertilizer and one without; following previous research (Baltasa and Doyle, 2001; Hanemann, 1984), the random utility function for a farm household facing the technology set can be specified as:

$$\pi_{ik} = \pi(s_{ik}, \tau_i) + \epsilon(s_{ik}, \tau_i) = x_{ik}\theta + u_i, \ i = 1, 2, ... n \quad (3.4.1)$$

where $\pi_{ik}$ denotes expected utility a farmer derives from fertilizer technology $k$ ($k = 1$ if fertilizer is applied, $k = 0$ otherwise), $\pi(s_{ik}, \tau_i)$ is the deterministic component of the utility function specified as a function of the observed attributes of the technology options ($s_{ik}$ ) and the socio-economic description of a household ($\tau_i$ ); $\epsilon(s_{ik}, \tau_i)$ is the stochastic component of the utility function representing the unobserved attributes affecting technology choice, heterogeneity in tastes and measurement errors; $x_{ik}$ is the arithmetic combination representing the covariates $s_{ik}$ and $\theta$ is the vector of parameters. Given the specification of $\pi_{ik}$ , a utility maximizing farm household would choose to apply inorganic fertilizer if the
corresponding expected utility is higher than that generated from the traditional technology (no fertilizer) i.e. if $\pi_{1} > \pi_{0}$. The binary response model of fertilizer adoption is thus specified as:

$$y_{i} = I \{ \pi_{1} > \pi_{0} \} = I \{ x_{i1} \theta + \varepsilon_{i1} - x_{i0} \theta - \varepsilon_{i0} > 0 \} = I \{ x_{i} \theta + u_{i} \}$$

where $u_{i} = \varepsilon_{i1} - \varepsilon_{i0}$ is random error term with zero mean and the coefficient vector $\theta$ is defined up to some scalar normalization.

The data we use are taken from the second wave (the 2009/10 wave) of the Uganda National Household Surveys (UNHS), a longitudinal survey of households in Uganda. The sample consists of 1912 farmer households. The proportion of farmers (in this sample) who applied fertilizers in 2009/2010 is about only 5%, which makes it relevant to analyze the factors hindering adoption of the input. We model adoption decision in terms of a binary response framework.

The control variables used in this essay are drawn from the empirical literature on agricultural technology adoption in developing countries. First, it has been noted that subsistence farmers want to use advanced farm technologies but do not have liquid capital to purchase them (Duflo, Kremer, and Robinson, 2008). Farmers have limited access to credit because markets for credit and insurance are either not available or dysfunctional (Gruhn and Rashid, 2001). Furthermore, the few credit institutions that are available are reluctant to give agricultural loans (Gordon, 2000). Moreover, typical subsistence farmers are usually not able to save their farm earnings to purchase inputs later because they face several other needs that compete for the limited financial resources. Farm and nonfarm income are therefore included in the empirical model to explore if they are used by credit-constrained farmers as an alternative financing mechanism to purchase fertilizer. To circumvent potential reverse causality bias introduced by farm and nonfarm income, we use lagged measures of these two variables from first wave (2005/06). In addition to lack of access to debt capital, riskiness of agricultural returns (primarily due to rainfall variation) has been identified...
in the literature as a critical impediment to wider adoption of improved agricultural technologies. We therefore capture the effects of ex-post weather shocks on fertilizer adoption by including a drought variable measured in the first wave (2005/06).

We also include in our model the average education of the household (measured as the mean of formal years of schooling of household members aged at least 15), age of the household head, household access to agricultural extension and advisory services, and household size to capture the effects of human capital and risk tolerance (age) on technology adoption. Further, the study captures household access to agricultural markets using the distance from the household to the closest trading center. Distance is expected to inversely affect the probability of fertilizer adoption through its positive effects on the cost of transportation and thus the effective prices of fertilizer. Heterogeneous effects of adoption arising from location and ago-ecological characteristics are captured using regional dummies (the northern, central, western, and eastern parts). Differences due to ago-ecological zones may influence fertilizer adoption decisions through their effect on farmer’s perception of soil quality and yield response. A dummy variable for male headed household is included to capture the effect of gender on adoption. We hypothesize that male-headed households have greater exposure to information about new technologies and that they are likely to be less risk averse than their female-headed counterparts.

A brief summary of the variables included in the model is provided in table 3.1. The data shows significant differences between users and nonusers of fertilizers with respect to all the variables except nonfarm revenue and agriculture extension services.
3.4.3 Empirical Results

In table 3.2 we present the empirical results for the parametric (probit model) and the semiparametric estimator (PGSIM).\textsuperscript{10} For the PGSIM, the optimal smoothing parameter \((h=.3)\) was estimated simultaneously with the parameter vector by minimizing the Ichimura loss function as suggested by Hardle, Hall, and Ichimura (1993). Nonetheless, we present the results from an alternative value of \(h (.32)\) near the optimal point; we note that the coefficient estimates are nearly identical for the two smoothing parameters. The stability of the coefficients suggest that we are near the optimal \(h\). The results show that the coefficient estimates for both approaches have the same sign. Moreover, both methods generate significant coefficients on the same variables except one, age of the household head: Age is significant in the probit model but not significant in the semiparametric estimator. However, all significant coefficient estimates derived using the semiparametric model are considerably larger in magnitude than the corresponding probit coefficients, suggesting that the parametric model might understate their effects on fertilizer adoption. For example the estimated coefficient of nonfarm income is more than twice larger for the PGSIM.

The significant determinants of adoption include extension services, male head of the household, past nonfarm income and farm revenue. As expected, the probability of adoption of fertilizers increases with nonfarm income as well as farm revenue, suggesting that both nonfarm and farm activities may be an important source of liquid capital for investing in fertilizers. Our findings corroborate previous research that diversification into nonfarm income activities is an important strategy used by credit-constrained households in developing countries to obtain needed capital for investment in agricultural technologies (Janvry

\textsuperscript{10} As in the probit, location-scale normalization is needed to ensure identification of the parameter vector. For the probit, the location-scale normalization requires setting the first and second moments of the error term to zero and one, respectively. For the semiparametric estimator, the location-scale normalization is imposed by constraining the intercept to zero and one of the coefficients on continuous regressors to a constant. The coefficient on the log of distance to a trading center is normalized to the corresponding Probit estimate to facilitate comparison of the results.
and Sadoulet, 2001; Iiyama et al., 2008; Barrett, Reardon, and Webb, 2001; Reardon, Stamoulis, and Pingali, 2007).

The results also indicate that male farmers have a higher probability of adopting fertilizers than their female counterparts. Previous studies in Africa show that women farmers are constrained by limited access to most of the important production resources such as land, liquid capital and labor (De Groote and Coulibaly, 1988). The results further show that the probability of adoption significantly increases with the number of interactions between the farmer and the agricultural extension agents. Agricultural extension and advisory services are important in enhancing farmer’s knowledge and skills in fertilizer application.

3.5 Conclusions

In this essay, a new semiparametric estimator for binary-choice single-index models, which uses parametric information in the form of a known (parametric) link function and nonparametrically corrects it, is applied to empirical data. An extensive simulation study is conducted and the new estimator is compared with the semiparametric estimator of Ichimura (1993) and the parametric probit. The performance of the estimator is robust to the correctness of the parametric guide since all that is required from the guide is to smooth the function to be estimated nonparametrically so as to achieve bias reduction, not that it be correct all the time.

The method of Hardle, Hall, and Ichimura (1993) is followed regarding the choice of the smoothing parameter for the Ichimura and proposed estimator and the objective functions are optimized with respect to the smoothing parameter as well as the unknown coefficients. Results show that it works well and gives quite reasonable results.

The empirical results show that the conclusions drawn from probit model and the
PGSIM estimates are qualitatively similar but are different in magnitude. The semiparametric estimator generates substantially larger coefficients on the significant covariates compared to the probit model. We find that extension and advisory services, nonfarm income and farm revenue significantly increase adoption, and that the probability of adoption is higher among male headed households.
3.6 References


3.7 Tables and Figures

<table>
<thead>
<tr>
<th>Variable</th>
<th>Nonfertilizer Users</th>
<th>Fertilizer Users</th>
<th>T-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Past Farm Revenue (US$)</td>
<td>87.75</td>
<td>248.66</td>
<td>-6.70</td>
</tr>
<tr>
<td></td>
<td>(191.01)</td>
<td>(537.52)</td>
<td></td>
</tr>
<tr>
<td>Past Nonfarm Income (US$)</td>
<td>543.16</td>
<td>811.99</td>
<td>-1.24</td>
</tr>
<tr>
<td></td>
<td>(1847.21)</td>
<td>(1599.29)</td>
<td></td>
</tr>
<tr>
<td>Famr Size (Hectares)</td>
<td>1.00</td>
<td>1.36</td>
<td>-2.57</td>
</tr>
<tr>
<td></td>
<td>(1.29)</td>
<td>(1.43)</td>
<td></td>
</tr>
<tr>
<td>No. of Extension Visits</td>
<td>0.39</td>
<td>1.38</td>
<td>-1.44</td>
</tr>
<tr>
<td></td>
<td>(1.97)</td>
<td>(3.92)</td>
<td></td>
</tr>
<tr>
<td>Age of Household Head (Years)</td>
<td>47.83</td>
<td>44.84</td>
<td>1.82</td>
</tr>
<tr>
<td></td>
<td>(15.05)</td>
<td>(13.63)</td>
<td></td>
</tr>
<tr>
<td>Average Education of Household (Years)</td>
<td>4.40</td>
<td>5.73</td>
<td>-2.39</td>
</tr>
<tr>
<td></td>
<td>(5.09)</td>
<td>(4.61)</td>
<td></td>
</tr>
<tr>
<td>Gender of Head (1=Male, 0=Female)</td>
<td>0.70</td>
<td>0.86</td>
<td>-3.21</td>
</tr>
<tr>
<td></td>
<td>(0.46)</td>
<td>(0.35)</td>
<td></td>
</tr>
<tr>
<td>Household size (No. of Persons)</td>
<td>6.62</td>
<td>7.77</td>
<td>-3.35</td>
</tr>
<tr>
<td></td>
<td>(3.10)</td>
<td>(3.68)</td>
<td></td>
</tr>
<tr>
<td>Distance to Trading Center (KM)</td>
<td>4.30</td>
<td>2.60</td>
<td>2.03</td>
</tr>
<tr>
<td></td>
<td>(7.83)</td>
<td>(2.70)</td>
<td></td>
</tr>
<tr>
<td>Past Drought Shock (1=Yes, 0=No)</td>
<td>0.55</td>
<td>0.49</td>
<td>1.91</td>
</tr>
<tr>
<td></td>
<td>(0.49)</td>
<td>(0.50)</td>
<td></td>
</tr>
<tr>
<td>Central Region</td>
<td>0.19</td>
<td>0.34</td>
<td>-3.52</td>
</tr>
<tr>
<td></td>
<td>(0.39)</td>
<td>(0.48)</td>
<td></td>
</tr>
<tr>
<td>Eastern Region</td>
<td>0.26</td>
<td>0.37</td>
<td>-2.26</td>
</tr>
<tr>
<td></td>
<td>(0.44)</td>
<td>(0.49)</td>
<td></td>
</tr>
<tr>
<td>Northern Region</td>
<td>0.28</td>
<td>0.22</td>
<td>1.25</td>
</tr>
<tr>
<td></td>
<td>(0.45)</td>
<td>(0.42)</td>
<td></td>
</tr>
<tr>
<td>Western Region</td>
<td>0.27</td>
<td>0.07</td>
<td>4.19</td>
</tr>
<tr>
<td></td>
<td>(0.44)</td>
<td>(0.25)</td>
<td></td>
</tr>
<tr>
<td>Number of Observations</td>
<td>1825</td>
<td>87</td>
<td></td>
</tr>
</tbody>
</table>

*Standard Deviations in Parentheses*
Table 3.2: Results from PGSIM and Probit Procedures

<table>
<thead>
<tr>
<th>Variable</th>
<th>Probit (1)</th>
<th>PGSIM (2a)</th>
<th>PGSIM (2b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(Past of Farm Revenue in US$)</td>
<td>0.137***</td>
<td>0.233***</td>
<td>0.234***</td>
</tr>
<tr>
<td></td>
<td>(0.041)</td>
<td>(0.076)</td>
<td>(0.079)</td>
</tr>
<tr>
<td>ln(Past Nonfarm Income in US$)</td>
<td>0.057***</td>
<td>0.112***</td>
<td>0.112***</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.044)</td>
<td>(0.045)</td>
</tr>
<tr>
<td>Farmsize (Hectares)</td>
<td>0.021</td>
<td>0.028</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td>(0.038)</td>
<td>(0.079)</td>
<td>(0.082)</td>
</tr>
<tr>
<td>No. of Extension Visits</td>
<td>0.039***</td>
<td>0.071***</td>
<td>0.072***</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.028)</td>
<td>(0.032)</td>
</tr>
<tr>
<td>Age of Household Head (Years)</td>
<td>-0.008**</td>
<td>-0.006</td>
<td>-0.006</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.007)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Mean Household Education (Years)</td>
<td>0.007</td>
<td>0.016</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.020)</td>
<td>(0.021)</td>
</tr>
<tr>
<td>Gender of Head (1=Male, 0=Female)</td>
<td>0.320***</td>
<td>0.793***</td>
<td>0.799***</td>
</tr>
<tr>
<td></td>
<td>(0.147)</td>
<td>(0.300)</td>
<td>(0.308)</td>
</tr>
<tr>
<td>Household Size (No. of Persons)</td>
<td>0.018</td>
<td>0.034</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.034)</td>
<td>(0.035)</td>
</tr>
<tr>
<td>Past Drought Shock (1=Yes, 0=No)</td>
<td>0.017</td>
<td>0.042</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>(0.113)</td>
<td>(0.219)</td>
<td>(0.226)</td>
</tr>
<tr>
<td>Eastern Region</td>
<td>-0.052</td>
<td>-0.223</td>
<td>-0.225</td>
</tr>
<tr>
<td></td>
<td>(0.137)</td>
<td>(0.261)</td>
<td>(0.27)</td>
</tr>
<tr>
<td>Northern Region</td>
<td>-0.179</td>
<td>-0.468*</td>
<td>-0.473*</td>
</tr>
<tr>
<td></td>
<td>(0.157)</td>
<td>(0.307)</td>
<td>(0.316)</td>
</tr>
<tr>
<td>Western Region</td>
<td>-0.862***</td>
<td>-1.425***</td>
<td>-1.436***</td>
</tr>
<tr>
<td></td>
<td>(0.202)</td>
<td>(0.336)</td>
<td>(0.349)</td>
</tr>
<tr>
<td>ln(Distance to Trading Center in KM)</td>
<td>-0.134**</td>
<td>-0.134</td>
<td>-0.134</td>
</tr>
<tr>
<td></td>
<td>(0.072)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Log Likelihood (LL)</td>
<td>-309.850</td>
<td>-308.723</td>
<td>-309.058</td>
</tr>
<tr>
<td>Bandwidth (h)</td>
<td>N/A</td>
<td>0.30</td>
<td>0.32</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>1912</td>
<td>1912</td>
<td>1912</td>
</tr>
</tbody>
</table>

*Standard Errors in Parentheses; *** p<0.01, ** p<0.05, * p<0.1*
Bibliography


119


